

# II CONGRESSO DE TECNOLOGIA DA INFORMAÇÃO

II Congresso de Tecnologia da Informação IFSUL Passo Fundo

Passo Fundo - RS - Brasil

## ANAIIS

### **Organização Editorial**

Alex Sebben da Cunha - IFSUL Passo Fundo

Alexandre Tagliari Lazzaretti - IFSUL Passo Fundo

### **Coordenação**

João Mário Lopes Brezolin - IFSUL Passo Fundo

### **Realização**

Instituto Federal Sul-Riograndense - Campus Passo Fundo

### **Apoio**

Alura Cursos

Atua – Sistemas de Informação

Compasso – Uma Empresa UOL Diveo

DioLinux

Germânias Hotel

KLM Informática

Meta Tecnologia

PoloSul.org

Prefeitura Municipal de Passo Fundo

Sociedade Brasileira de Computação – SBC

## Prefácio

---

Sejam bem-vindos a segunda edição do Congresso de Tecnologia da Informação do IFSUL Campus Passo Fundo, ocorrido nos dias 30, 31 de outubro e 01 de novembro de 2019. O congresso é um evento realizado pelos cursos de informática do campus Passo Fundo.

O evento busca motivar o estudante para exercer a atividade científica e tecnológica, uma vez que proporciona um espaço de divulgação e discussão de temas relevantes e atuais na área de informática desenvolvidos em sua instituição. É nossa aspiração estimular no estudante o interesse pela observação, construção do conhecimento, troca de experiências, destacando assim o comprometimento da instituição com o ensino, pesquisa e a extensão. Além disso, busca-se oferecer uma diversificada programação, com palestras, oficinas e minicursos.

O tema do simpósio este ano é "**Computação Aplicada**", enfatizando o aspecto prático das contribuições que a informática pode proporcionar nas diferentes áreas do conhecimento

Em relação aos trabalhos submetidos, foram totalizados 21 trabalhos. Sendo aceitos os 16 trabalhos com melhor média resultante das avaliações dos pares. Dentre os trabalhos os 3 primeiros colocados receberam uma premiação no evento de encerramento.

A realização do simpósio congresso não seria possível sem o trabalho incessante da comissão organizadora do eventos, principalmente do coordenador, Prof. João Mário Lopes Brezolin, e do auxílio de todos os professores e técnicos administrativos do Instituto Federal Sul-Riograndense Campus Passo Fundo.

Cabe um agradecimento aos alunos, alunos bolsistas e técnicos administrativos que auxiliaram na realização das tarefas demandadas pelo evento. Além disto, um agradecimento especial aos membros do comitê de avaliação dos trabalhos científicos e a todos que participaram do evento, muito obrigado!

Um ótimo evento a todos!

Alexandre Tagliari Lazzaretti  
Carlos Alberto Petry  
Roberto Wiest

Comissão de Trabalhos Científicos

## **Coordenação**

---

### **Coordenação geral:**

Prof. Dr. João Mário Lopes Brezolin

### **Comissão de trabalhos científicos:**

Prof. Dr. Alexandre Tagliari Lazzaretti

Prof. Me Carlos Alberto Petry

Prof. Dr. Roberto Wiest

### **Comissão das palestras, minicursos e oficinas:**

Prof. Me. André Fernando Rollwagen

Prof. Me. Élder F. F. Bernardi

Prof. Me. Jorge L. B. Bavaresco

Prof. Me. Lisandro Lemos Machado

### **Aluno bolsista:**

Daniel Nicolau Saito

## Comitê de Programa

---

Adilso Souza	IFSUL - Passo Fundo
Alexandre Tagliari Lazzaretti	IFSUL - Passo Fundo
Alexandre Zamberlan	Universidade Franciscana
Alexandre Zanatta	Universidade de Passo Fundo
André Schaeffer	Universidade Federal da Fronteira Sul
Anubis Rossetto	IFSUL - Passo Fundo
Élder Bernardi	IFSUL - Passo Fundo
Carlos Rolim	UFRGS
Carlos Silva	Instituto Federal Farroupilha
Carmen Brezolin	IFSUL - Passo Fundo
Dionatrã Kirchoff	PUC
Edimar Manica	IFRS
Emerson Oliveira Jr.	Instituto Federal do Rio Grande do Sul
Evandro Kuszera	UTFPR - Dois Vizinhos
Fahad Kalil	Faculdade Meridional - IMED
Gabriel Santin	IFSUL - Passo Fundo
Guilherme Antonio Borges	Instituto Federal Sul-Rio-Grandense
Joao Mário Brezolin	Instituto Federal Sul-Riograndense
José de Figueiredo	Instituto Federal Sul-rio-grandense
José Maurício Fernandes	Embrapa Trigo
Juliano Weber	IFFarroupilha
Marco Trentin	UPF
Marcos José Brusso	UPF
Rômulo de Oliveira	PUCRS
Ricardo Schmidt	University of Passo Fundo
Roberto Wiest	IFSUL - Passo Fundo
Suellen Sotille	IMED
Telmo De Cesaro Júnior	IFSUL - Passo Fundo
Tiago Rocha	Instituto Federal do Rio Grande do Sul
Valderi Leithardt	Univali
Vanessa Lago Machado	Instituto Federal Sul-rio-grandense
Willingthon Pavan	Universidade de Passo Fundo

## Sumário

---

<b>Ferramenta Educativa para Levantamento de Requisitos de Software</b>	<b>7</b>
André Rollwagen (IFSUL - Passo Fundo - Brazil), Vinícius Maciel (IFSUL - Algeria), Joseane Amaral (IFSUL - Passo Fundo - Brazil)	
<b>Uma alternativa de software para mosaicação de imagens digitais no meio agrícola</b>	<b>17</b>
Gustavo Visentini (Universidade de Passo Fundo - Brazil), Willingthon Pavan (Universidade de Passo Fundo - Brazil), Carlos Holbig (Universidade de Passo Fundo - Brazil), José Maurício Cunha Fernandes (Embrapa Trigo - Brazil)	
<b>Desenvolvimento de uma fresadora CNC utilizando software de modelagem 3D e hardware de código aberto</b>	<b>27</b>
Amauri Lopes (IFSul - Brazil)	
<b>Sistemas ERP em Ferramentarias de Micro e Pequeno Porte em Santa Catarina</b>	<b>37</b>
Júlia de Farias (Centro Universitário Municipal de São José - Brazil), Osmar Da Cunha Filho	
<b>DESENVOLVIMENTO DE UMA PRÓTESE DE MÃO ROBOTIZADA</b>	<b>47</b>
Henrique de Lima (Instituto Federal Sul-Rio-Grandense (IFSul) - Brazil)	
<b>Uso de Sistema de Recomendação baseado em personas para potencializar a aprendizagem dos alunos de Algoritmos</b>	<b>57</b>
Gabriel Bonato (IFSUL Passo Fundo - Brazil), Joao Mário Brezolin (Instituto Federal Sul-Riograndense - Brazil), Carmen Brezolin (IFSUL Passo Fundo - Brazil), - Brazil)	
<b>Um estudo sobre a Classificação de Risco em Pronto Atendimento utilizando o método KNN</b>	<b>67</b>
Antonio A.G. Silveira (IFES - Brazil), Jefferson Andrade (Instituto Federal do Espírito Santo - Brazil)	
<b>Confiabilidade nos Navegadores de Internet</b>	<b>74</b>
Lisandro Machado (Instituto Federal Sul-Rio-Grandense - Brazil), Luis Fernando Pooter (Instituto Federal Sul-rio-grandense - Brazil)	
<b>Progressive Web Apps: um estudo de caso para o 1º Batalhão Rodoviário da Brigada Militar</b>	<b>84</b>
Daniel Nicolau Saito (IFSUL - Brazil), Thauany Martins (IFSUL - Brazil), Marcelo Fernandes (Instituto Federal Sul-rio-grandense (IFSUL) - Brazil), Adilso Souza (Instituto Federal Sul-RioGrandense Campus Passo Fundo - Brazil), Anubis Rossetto (Instituto Federal Sul-RioGrandense Campus Passo Fundo - Brazil), Guilherme Gehring (IFSUL Passo Fundo - Brazil)	

<b>O uso de aplicativo educacional na Educação Profissional e Tecnológica: o que os dados revelam</b>	<b>94</b>
<p>Angélica Tommasini (IFRS - Campus Sertão/RS - Brazil),  Ana Sara Castaman (IFRS- Brazil),  Márcio Vieira (IFRS-Campus Sertão - Brazil),  Ângela Helena Peretti (IFRS - Campus Sertão - Brazil),  Luciane Inocente (IFRS - Campus Sertão - Brazil),  Lis Ângela De Bortoli (IFRS - Sertão - Brazil),  Daniela Hoffmann (Secretaria de Educação de Getúlio Vargas/RS - Brazil)</p>	
<b>Aprenda Talian: Plataforma colaborativa para aprendizado do dialeto Talian</b>	<b>100</b>
<p>Leandro José Grando (IFSUL - Brazil),  Anubis Rossetto (IFSUL - Campus Passo Fundo - Brazil)</p>	
<b>A difícil vida dos testadores em plataformas de crowdtesting</b>	<b>110</b>
<p>Fernando Luiz Costella (UPF - Brazil),  Alexandre Zanatta (Universidade de Passo Fundo - Brazil),  Cristiano Cervi (Universidade de Passo Fundo (UPF) - Brazil)</p>	
<b>Comparação de Performance entre o PostgreSQL e o MongoDB</b>	<b>118</b>
<p>Emanuele Solagna (Universidade de Passo Fundo (UPF) - Brazil),  Alexandre Tagliari Lazzaretti (IFSUL - Passo Fundo - Brazil)</p>	
<b>Caracterizando os Desenvolvedores de Ecossistemas de Software</b>	<b>128</b>
<p>Marina Moreira (Pontifícia Universidade Católica do Rio Grande do Sul - Brazil),  Caio Steglich (PUCRS - Brazil),  Sabrina Marczak (PUCRS - Brazil),  Fernando Figueira Filho (Universidade Federal do Rio Grande do Norte - Brazil),  Marcelo Gattermann Perin (Independent Researcher - Brazil),  Cleidson de Souza (UFPA - Brazil)</p>	
<b>Uso de Redes Neurais Convolucionais para a identificação e classificação de Insetos</b>	<b>138</b>
<p>Joao Mário Brezolin (Instituto Federal Sul-Riograndense - Brazil),  Bruno Bearsi Paixão (IFSUL - Brazil)</p>	
<b>Predição da Nota do ENEM 2018 usando Técnicas de Regressão</b>	<b>147</b>
<p>Pedro Lessa (Instituto Federal do Espírito Santo - Campus Serra - Brazil),  Karin Komati (Ifes Campus Serra - Brazil)</p>	

# Ferramenta Educativa para Levantamento de Requisitos de Software

André F. Rollwagen<sup>1</sup>, Joseane Amaral<sup>1</sup>, Vinícius Macial<sup>1</sup>

<sup>1</sup>Instituto Federal de Educação Ciência e Tecnologia Sul-Rio-Grandense (IFSUL)  
Estrada Perimetral Leste – 99.064-440 – Passo Fundo – RS – Brasil

{andre.rollwagen,joseane.amaral}@passofundo.ifsul.edu.br,  
vinicius.maciavel.pf@hotmail.com

***Abstract.** Digital technologies like computers, tablets, smartphones and internet of things are inserted in society helping us daily. In education is no different, educational institutions, teachers and students are meeting these innovations. But the majority of software was not intended as teaching tools. In this context this project was developed in order to provide a tool that can be applied in the Systems Analysis and Software Engineering disciplines so that the students have the experience of using a software for requirements gathering that brings them practicality and ease in this process.*

***Resumo.** Tecnologia digitais como computadores, tablets, smartphones e internet das coisas estão inserida na sociedade nos auxiliando diariamente. Na educação não é diferente, instituições de ensino, professores e alunos estão indo ao encontro destas inovações. Porém a grande maioria dos softwares não foram planejados como ferramentas de ensino. Neste contexto desenvolveu-se este projeto, a fim de fornecer uma ferramenta que possa ser aplicada nas disciplinas de Análise e Projeto de Sistemas e Engenharia de Software, para que os educandos possam experienciar a utilização de um software para levantamento de requisitos que lhes traga praticidade e facilidade neste processo.*

## 1. Introdução

A maioria de softwares para gestão de projetos e requisitos de sistemas está inserida no meio comercial, dificultando a utilização de tais ferramentas no meio acadêmico. Sendo assim, propôs-se o desenvolvimento de um software para levantamento de requisitos com foco educacional para ser utilizado nas disciplinas de Análise e Projeto de Sistemas e Engenharia de Software, nos cursos técnicos e superiores de Informática de uma instituição de ensino localizada na cidade de Passo Fundo.

Tendo em vista que o projeto desenvolvido trata de um tema comum ao meio comercial, é difícil encontrar aplicativos que realizem este processo de forma gratuita e sejam desenvolvidos sob a perspectiva educacional. Assim, provou-se justificável o desenvolvimento desta pesquisa e da criação do software com foco educacional, para suprir a deficiência na área de levantamento de requisitos para fins de aprendizagem.

Para alcançar tal objetivo, durante o desenvolvimento do projeto foram feitas análises de softwares na área de gerenciamento e levantamento de requisitos como OSRMT [Brito 2010], SIGERAR [Grande 2010] e OpenReq [Grings 2008]; estudadas as teorias de criação de softwares educacionais [Campos 2001, Benitti 2005], bem como as teorias de aprendizagem de Vygotsky (2010) utilizadas como base para este projeto. Após o estudo teórico foi realizado levantamento de requisitos para o desenvolvimento do software educacional e por fim o mesmo foi aplicado nas disciplinas de Análise e Projeto de Sistemas e Engenharia de Software, finalizado o projeto com a análise dos resultados.

## **2. Referencial Teórico**

Neste capítulo são apresentados os principais conceitos necessários para a realização deste projeto, abrangendo teorias de aprendizagem, análise e projeto de sistemas, softwares educacionais e as ferramentas utilizadas para o desenvolvimento do software educacional.

### **2.1. Teorias da Aprendizagem**

Em geral, as teorias de aprendizagem são a forma como a humanidade sistematiza uma determinada área de conhecimento, podendo assim ver, explicar e resolver problemas nessa esfera. Para Moreira:

Uma teoria de aprendizagem é, então, uma construção humana para interpretar sistematicamente a área de conhecimento que chamamos de aprendizagem. Representa o ponto de vista de um autor/pesquisador sobre como interpretar o tema aprendizagem, quais as variáveis independentes, dependentes e intervenientes. Tenta explicar o que é aprendizagem e porque funciona como funciona [Moreira 1999, p12].

Segundo Moreira (1999), Vygotsky foi um pensador que pregava a influencia dos contextos sociais e culturais no processo de aprendizagem. Juntamente a isso, também focalizava em todos os mecanismos através dos quais se davam o desenvolvimento cognitivo. Neste âmbito, para entender o processo cognitivo do ser humano deve-se considerar o ambiente social que o indivíduo vive.

Outro ponto abordado por Moreira (1999) é o conceito de mediação, que para Vygotsky é uma atividade típica da cognição humana e ocorre através de instrumentos e signos. O instrumento é tudo aquilo que pode ser usado para realização de alguma tarefa. Já um signo, é algo que significa alguma coisa. Estes signos são classificados em indicadores, icônicos e simbólicos.

Para o desenvolvimento deste projeto foram levadas em consideração as teorias de aprendizagem de Vygotsky, que entende que o processo de aprendizagem dos alunos ocorre de forma cognitiva e por meio de instrumentos e signos onde, neste caso o instrumento será a ferramenta que deverá intervir no ensino e ampliar a qualidade do ensino das matérias.

## **2.2. Análise e Projeto de Sistemas**

No desenvolvimento de um software a engenharia de requisitos é fundamental para o sucesso do projeto. Esta etapa pode ser considerada como uma ponte entre o projeto e a construção de um sistema e está fortemente ligada às necessidades do projeto, produto e também das pessoas envolvidas. Para Pressman:

A engenharia de requisitos fornece o mecanismo apropriado para entender aquilo que o cliente deseja, analisando as necessidades, analisando a viabilidade negociando uma solução razoável, especificando a solução sem ambigüidades, validando a especificação e gerenciando as necessidades à medida que são transformadas em sistema operacional [Pressman, 2011, p 127].

Para Sommerville (2016), os requisitos de um sistema e a descrição dos serviços refletem as necessidades do cliente. O processo de verificar e entender estes requisitos é chamado de engenharia de requisitos. Segundo o autor, existem várias divisões de requisitos, porém para este projeto, serão aprofundados os requisitos funcionais e não funcionais. Os requisitos funcionais descrevem o que um sistema deve fazer, abordando as funções do software detalhadamente, suas entradas, saídas e exceções, assim:

Requisitos Funcionais são declarações de serviços que o sistema deve fornecer, como o sistema deve reagir a entradas específicas e como o sistema deve se comportar em determinadas situações. Em alguns casos os requisitos funcionais também podem estabelecer explicitamente o que o sistema não deve fazer [Sommerville 2011, p. 59].

O autor atenta para os problemas que podem ser causados devido a má especificação de requisitos. A equipe de desenvolvimento pode interpretar requisitos mal definidos de forma ambígua, o que resulta em funcionalidades que não refletem exatamente aquilo que o cliente estava esperando, que por sua vez, acaba gerando um novo levantamento de requisitos para que as mudanças no sistema possam ser desenvolvidas e o mesmo passe a atender a expectativa do cliente.

Requisitos não funcionais não necessariamente estão ligados a funções específicas do sistema. Eles podem estar relacionados as propriedades do software como confiabilidade, tempo de resposta e espaço de armazenamento. Por isso muitas vezes não referem-se apenas ao sistema, mas também a todo o processo de desenvolvimento do mesmo [Sommerville 2016].

## **2.3. Softwares Educacionais**

Softwares educacionais são ferramentas desenvolvidas para auxiliar usuários em tarefas com fins de aprendizagem. Um software pode ser classificado como educacional, segundo Giraffa (2009), se o mesmo for contextualizado no processo de ensino-aprendizagem seguindo a metodologia estipulada pelo docente.

Baseado em Campos (2001), pode-se dizer que um software educacional contempla as características da educação que levam à formação global do aluno,

que necessita aprender a aprender e pensar para melhor intervir, inovar e questionar, trabalhando com as funções da cognição.

Já Benitti (2005), relata que para desenvolvimento de softwares educacionais leva-se em consideração quatro etapas principais: concepção, elaboração, finalização e viabilização. As etapas tem enfoque específico e bem definido, uma fase só é iniciada após a finalização da anterior.

#### **2.4. Ferramentas de Desenvolvimento**

Para o desenvolvimento deste projeto elencou-se as ferramentas consideradas fundamentais para sua realização. Assim, a linguagem base para codificação do software foi *Personal Home Page* (PHP) [Niederauer 2017], utilizando o *framework Codeigniter* [Codeigniter 2019]. Para a parte gráfica optou-se pelo *Bootstrap*, um *framework* de código aberto para desenvolvimento *front-end*, que permite responsividade, e abrange HTML, CSS e *JavaScript* [Silva 2015].

### **3. Metodologias**

A realização deste projeto foi dividida em duas etapas. A primeira consiste no desenvolvimento do software, na qual foi realizado o estudo das necessidades e a implementação do mesmo. Na segunda fase disponibilizou-se o software para uso pelos educandos e aplicou-se um questionário envolvendo a utilização, para obter retorno dos alunos sobre a ferramenta.

Durante a etapa de desenvolvimento foram levantados os principais requisitos que o sistema deveria conter para um bom funcionamento. Após iniciou-se o processo de programação lógica do sistema e a partir deste ponto foram definidas a arquitetura e linguagem de programação, já apresentados no referencial teórico. Após o término da codificação, foram realizadas reuniões com o orientador para definição de melhorias de acordo com os conteúdos da matéria de Análise e Projeto de Sistemas.

Na segunda fase o software foi aplicado em sala de aula, tendo como participantes da pesquisa as turmas do Curso Técnico em Informática e a turma do Tecnólogo em Sistemas para Internet. Para medir a eficácia do trabalho desenvolvido, foi realizada uma pesquisa exploratória qualitativa através de um questionário com seis perguntas de múltipla escolha sobre o software e sua utilização, e mais uma descritiva para críticas e sugestões. A pesquisa foi finalizada com a análise dos dados coletados com os estudantes sobre a eficácia do software após utilização em sala de aula.

### **4. Resultados**

#### **4.1. Desenvolvimento do Software Educacional**

Ao finalizar a pesquisa bibliográfica, iniciou-se o processo de desenvolvimento da ferramenta onde se buscou unir os conhecimentos tecnológicos de programação às teorias de aprendizagem, visando assim a criação de um software educacional para levantamento de requisitos.

O software foi desenvolvido para que cada aluno tenha uma conta de usuário, podendo desta forma visualizar e acessar seus projetos. Dentro de cada projeto o educando pode cadastrar os dados básicos que devem ser preenchidos no documento de especificação de requisitos, que é solicitado pelo professor em aula. Dentre estes dados, se destacam a descrição do projeto e a descrição do mini-mundo, que são uma tentativa de representação do propósito e das condições em que o software deverá ser utilizado. Este cadastro pode ser visto na figura 1.

The screenshot shows a web form titled 'Cadastro' with a sub-section 'Informações Básicas'. The form contains the following fields:

- Nome:** A text input field containing 'Projeto de Aula' and a label 'Obrigatório'.
- Descrição Propósito do Sistema:** A text area containing the text: 'O sistema tem por sua finalidade ter o controle de cadastros de consultório médico, permitindo o usuário gerar: ficha de cadastros de pacientes, emissão de receitas, e recibos de pagamento de consultas e relatório de inadimplência.'
- Descrição Mini-Mundo:** A text area containing the text: 'O sistema permitirá o usuário a realizar o cadastro de pessoas, em que terá que informar os dados como: nome, telefone, endereço, cidade, UF, CEP, número e complemento, CPF, R.G, órgão expedidor, data de nascimento, idade, e sua função no sistema que poderá ser:'.
- Responsável:** A dropdown menu with 'Vinicius' selected.

At the bottom right of the form are two buttons: a red 'Cancelar' button and a green 'Salvar' button.

**Figura 1 - Cadastro do Projeto**

Após a descrição inicial do projeto é possível partir para o módulo de requisitos, onde o sistema foi dividido entre requisitos funcionais e não funcionais. No cadastro dos requisitos funcionais o aluno pode informar os campos “nome”, “descrição”, “categoria” e a qual projeto pertence. Já nos requisitos não funcionais também existem os campos “nome” e “descrição” e também os campos “obrigatório” e “permanente”. Outra particularidade dos requisitos não funcionais é a possibilidade de cadastrá-los ligados a um requisito funcional, criando assim os requisitos dependentes, que são de grande importância na especificação do projeto, conforme mostra a Figura 2.

The screenshot shows two parts of the application interface:

**Top Part: Cadastro de Requisitos não Funcionais**

- Nome:** Text input 'RNF1' with label 'Obrigatório'.
- Descrição:** Text area with 'Um médico não terá acesso aos dados do paciente de outro médico'.
- Obrigatório:** Dropdown menu with 'Sim' selected.
- Permanente:** Dropdown menu with 'Sim' selected.
- Tipo de requisito:** Dropdown menu with 'Especificação' selected and label 'Obrigatório'.
- A green 'Salvar' button is at the bottom.

**Bottom Part: Requisitos do projeto**

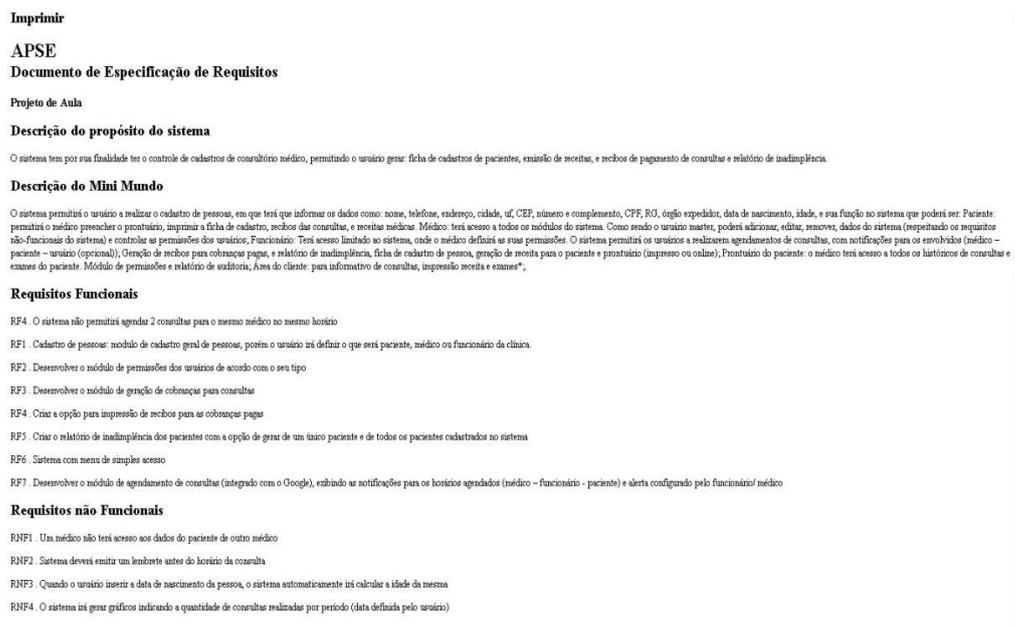
A table with the following data:

Código	Nome	Categoria	Obrigatório	Permanente	Ações
54	RNF1	Estrutural	s	s	[Icon]

Below the table, it says 'Mostrando 1 de 1 de 1 registros' and navigation buttons: 'Primeiro', 'Anterior', '1', 'Próximo', 'Último'.

**Figura 2 - Cadastro de Requisitos Não Funcionais**

Para que o sistema cumprisse seu papel como um software educacional e pudesse proporcionar aos alunos praticidade durante a utilização, foram criados gatilhos nos cadastros de projeto, requisitos funcionais e não funcionais. Estes gatilhos se apresentam em forma de botões que podem ser acionados e disponibilizar aos alunos os conteúdos estudados em aula. Ao final do processo de cadastro do projeto o educando pode gerar através do software o documento de especificação de requisitos, que exhibe os dados no modelo do documento solicitado pelo professor em aula, conforme mostra a Figura 3.



**Figura 3 - Documento de especificação de requisitos**

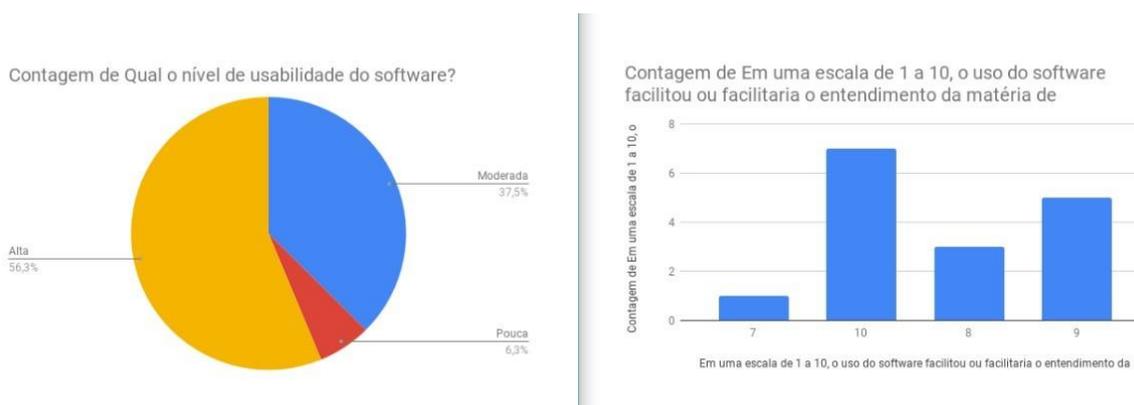
A geração deste documento é a concretização do objetivo do software: auxiliar os alunos em uma tarefa que anteriormente era realizada de forma manual, podendo contar com as dicas do conteúdo educacional integradas ao sistema.

## 4.2. Resultados da Pesquisa

Posterior ao desenvolvimento do software o mesmo foi aplicado em sala de aula. O sistema foi apresentado aos alunos de duas turmas do Instituto Federal Sul-rio-grandense (IFSUL), uma no curso Técnico em Informática, na matéria de Análise e Projeto de Sistemas e na turma do Tecnólogo em Sistemas para Internet na disciplina de Engenharia de Software. Nas duas ocasiões o sistema foi disponibilizado aos alunos e solicitado que realizassem um teste cadastrando um projeto para que depois fosse gerado o documento de especificação de requisitos.

Ao finalizar a demonstração do software foi aplicada uma pesquisa qualitativa, com seis questões de múltipla escolha e uma descritiva para medir a qualidade e viabilidade do projeto. Foram obtidos no total 16 respostas, sendo nove do curso Técnico em Informática e sete do Tecnólogo em Sistemas para Internet.

A primeira questão indagou sobre a usabilidade do software com as seguintes opções de resposta: “alta”, “moderada”, “pouca” e “nenhuma”. A mostra de dados foi de 56,3% para a resposta “alta”, 37,5% para “moderada” e 6,3% para “pouca” usabilidade. A segunda pergunta questionou sobre o auxílio do software em relação ao entendimento da disciplina e isto foi feito disponibilizando-se uma escala de notas de um a dez. Os resultados se apresentaram da seguinte forma: uma resposta contendo nota sete, três notas oito, cinco notas nove e sete notas dez. Os resultados podem ser acompanhados na figura 4.

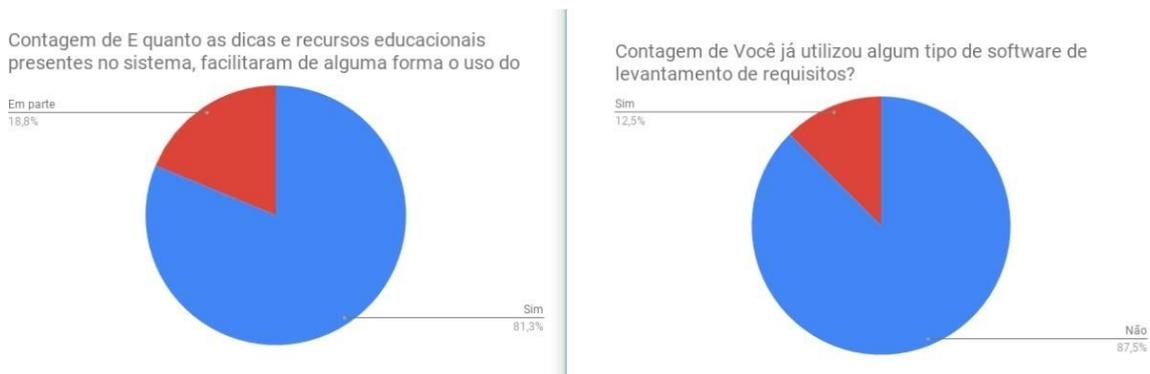


**Figura 4 - Questões 1 e 2**

Nas questões três e quatro abordaram-se os conceitos de software educacional. Na terceira pergunta, o questionamento referia-se a possibilidade de identificar os conceitos da matéria de levantamento de requisitos durante o uso do software, dando aos entrevistados as opções “sim”, “não” e “em parte”. O resultado foi de 100% para a opção “sim”. A quarta questão analisou se as dicas e recursos educacionais do sistema auxiliaram os educandos no processo de levantamento dos requisitos de seu projeto. Nesta indagação as opções foram “sim”, “não” e “em parte”, e o resultado foi de que 100% dos entrevistados opinaram “sim”, compreendendo que as dicas auxiliaram no processo.

A quinta pergunta examinou se as dicas educacionais auxiliaram no uso do software e continha as mesmas opções da quarta questão, que geraram um resultado de 87,3% de respostas “sim” e 18,8% da opção “em parte”. Esta mostra de dados pode ser acompanhada pela figura 5.

A sexta e última das questões de múltipla escolha verificou se os alunos já haviam utilizado algum tipo de software para documentação de levantamento de requisitos, sendo que as alternativas foram “sim” e “não”. O resultado obtido foi 12,5% para “sim” e 87,5% para “não”, mostrando que a maioria dos educandos não teve contato prévio com softwares dessa área. Estes dados podem ser vistos na figura 5.



**Figura 5 - Questões 5 e 6**

Para finalizar a pesquisa foi criada uma questão descritiva que permitiu aos entrevistados a possibilidade de mencionar alguma crítica ou sugestão relacionada ao software. Dentre estas respostas foram separadas duas críticas construtivas que serão abordadas na sessão de trabalhos futuros.

A pesquisa foi respondida por todos os alunos que estavam presentes nas aulas em que a pesquisa foi aplicada. Os resultados do trabalho realizado obtiveram índices satisfatórios.

## 5. Conclusão

Este projeto objetivou o desenvolvimento de uma ferramenta capaz de auxiliar os alunos a realizar o levantamento de requisitos de um projeto de forma prática, podendo ter contato com os conteúdos vistos em sala de aula. Assim, constatou-se que o software educacional permite a geração do documento de especificação de requisitos, destacando a possibilidade de acesso aos conteúdos específicos da área. Portanto, a proposta foi executada com sucesso.

O processo de levantamento de requisitos deu-se de forma satisfatória, sendo realizado por meio de pesquisa bibliográfica e análise de softwares comerciais na área de gerenciamento de requisitos, trazendo idéias e funcionalidades para o sistema. Durante o desenvolvimento o software foi sofrendo modificações provenientes das orientações fornecidas pelo professor responsável pelas disciplinas da área. Essas etapas também tiveram êxito, e possibilitaram aos educandos uma experiência prática durante a utilização do software educacional.

Conforme descrito nos resultados obtidos, houveram sugestões e críticas que foram analisadas e adicionadas ao plano de desenvolvimento futuro do projeto. Uma crítica originou-se de um entrevistado que possui deficiência visual, o qual apontou que o software não dispunha de *tags placeholder* nos campos, o que dificultou ao software de leitura de telas traduzir a informação ao usuário. O aluno também relatou que os botões de "dúvidas" eram acionados, porém não ficavam disponíveis para o leitor da tela. Neste contexto, verificou-se dificuldades de acessibilidade no software para usuários com deficiência visual.

Outra sugestão referia-se ao campo “nome” dos requisitos, onde habitualmente nas aulas o professor instrua o aluno que adicionasse no nome do requisito uma inscrição RF (requisito funcional) ou RNF (requisito não funcional) seguido do número do requisito. O educando entrevistado recomendou a geração automática desse identificador. Assim, percebeu-se a importância desses requisitos para melhorar a qualidade desta ferramenta. Mas devido ao tempo necessário para modificações, essas sugestões estão destinadas a trabalhos futuros.

Ao finalizar este projeto concluiu-se que o objetivo principal de criar um software educacional para auxiliar os alunos a realizar o levantamento de requisitos foi alcançado. Pode-se constatar através da pesquisa, que o software cumpriu de forma eficaz o seu propósito, resolvendo o problema inicial do projeto, que era a falta de uma ferramenta educacional para o levantamento de requisitos. Com análise dos resultados obtidos por meio do questionário aplicado aos educandos, constatou-se que o software auxiliou os alunos no entendimento dos conceitos pertinentes ao levantamento de requisitos, e também na aplicação dos mesmos de forma prática e lúdica, facilitando a aprendizagem nas disciplinas de Análise e Projeto de Sistemas e Engenharia de Software.

Como trabalhos futuros pretende-se implementar *tags placeholder* nos campos das telas do sistema, com a finalidade de permitir aos softwares de leitura de telas traduzir as informações ao usuário com deficiência visual. Também pretende-se implantar o sistema em um servidor do IFSUL, deixando-o disponível aos estudantes para utilização, facilitando assim o processo de ensino-aprendizagem nas disciplinas relacionadas.

## 6. Referencias

- Benitti, F. B. V., Seara, E. F. R., Schlindwein, L. M. (2005) "Processo de desenvolvimento de Software educacional: proposta e experimentação", CINTED-UFRGS. Novas Tecnologias da Educação. V. 3, Nº 1, Maio.
- Brito, A., Neves, L., Ronny, M., Bezerra, S., Yoshidome, E. (2010) "Uma Análise Avaliativa de Ferramentas de Software Livre no Contexto da Implementação do Processo de Gerência de Requisitos do MPS.BR". Workshop em Engenharia de Requisitos, Cuenca, Ecuador, abril 12-13.
- Campos, G. H. (2001) "Qualidade em Software Educacional", <http://www.casadaciencia.com.br/artigos>, julho.
- Codeigniter. (2019) "British Columbia Institute of Technology", [https://www.codeigniter.com/user\\_guide](https://www.codeigniter.com/user_guide), março.
- Giraffa, L. M. M. (2009) "Uma odisséia no ciberespaço: O software educacional dos tutoriais aos mundos virtuais", Revista Brasileira de Informática na Educação", v.17, No 1. (ISSN: 1414-5685; online: 2317-6121).
- Grande, J., Martins, L. (2010) "SIGERAR: Uma ferramenta para gerenciamento de requisitos". Dissertação de Mestrado, UNIMEP- Universidade Metodista de Piracicaba.

- Grings, C., Sayão, M., (2008) "OpenReq: uma Ferramenta para Auxílio à Gerência de Requisitos".
- Moreira, M. A. (1999) "Teoria de aprendizagem", São Paulo: E.D.U.
- Niederauer, J. (2017) "PHP para quem conhece PHP", 5 a ed. São Paulo. Novatec.
- Pressman, R. S. (2011) "Engenharia de software", 7<sup>a</sup> ed. São Paulo: MCGRAW-Hill.
- Silva, M. S. (2015) "Bootstrap 3.3.5 Aprenda a usar o framework Bootstrap para criar layouts CSS complexos e responsivos". Novatec.
- Sommerville, I. (2011) "Engenharia de Software", 9<sup>a</sup> ed. São Paulo: Pearson - Prentice Hall.
- Sommerville, I. (2016) "Software Engineering", 10<sup>a</sup> ed. Pearson Education.
- Tchounikine, P. (2011) "Computer Science and Educational Software Design. A Resource for Multidisciplinary Work in Technology Enhanced Learning", 1st Edition, XIII, 180 p.
- Vygotsky, L. S. (2010) "Aprendizagem e desenvolvimento na Idade Escolar". In: Linguagem, desenvolvimento e aprendizagem". 11<sup>a</sup>. Edição. São Paulo: Ícone, p. 103-116.

# Uma alternativa de software para mosaicagem de imagens digitais no meio agrícola

Gustavo César Visentini<sup>1</sup>, Willingthon Pavan<sup>1</sup>, Carlos A. Holbig<sup>1</sup>, Jose Mauricio Fernandes<sup>1,2</sup>

<sup>1</sup>Universidade de Passo Fundo (UPF)  
Passo Fundo – RS – Brazil

119692@upf.br, pavan@upf.br, holbig@upf.br

<sup>2</sup>Embrapa Trigo  
Passo Fundo – RS – Brazil

mauricio.fernandes@embrapa.br

**Abstract.** *The increase in the use of Drones in the agricultural environment has been a great tool in the monitoring of cultivars. To do this, automating image processing in a simple way, so that it is not necessary to purchase or use specific GIS software (Geographic Information Systems) becomes important for the diagnosis and decision-making of the farmer in a timely manner. In this way this work presents the development of a Desktop tool for the mosaic of RGB and NDVI images. The tool was developed in the C++ language using the OpenCV framework.*

**Resumo.** *O aumento na utilização de Drones no meio agrícola tem se mostrado uma ótima ferramenta no monitoramento de cultivares. Para isso, automatizar o processamento destas imagens de maneira simples, de forma que não seja necessário a compra ou utilização de softwares específicos de GIS (Geographics Information Systems), torna-se importante para o diagnóstico e tomada de decisão do agricultor em tempo hábil. Desta forma este trabalho apresenta o desenvolvimento de uma ferramenta Desktop para a mosaicagem de imagens RGB e NDVI. A Ferramenta foi desenvolvida na linguagem C++ utilizando o framework OpenCV.*

## 1. Introdução

Segundo Rong [Rong et al. 2009] a construção de um mosaico de imagens é um processo que consiste em sobrepor uma imagem sobre a outra, levando em consideração pontos em comum, gerando assim um mosaico, o que amplia o campo de visão da região de interesse.

A construção de mosaicos e o seu uso na topografia vem se mostrando importante há bastante tempo [Loch 2001]. Desde os primeiros métodos de mosaicagem muitas técnicas e equipamentos fotográficos foram criados e melhorados [Ferreira et al. 2013]. Deste modo a utilização de equipamentos e o processamento de dados para mapeamento aéreo faz uso de múltiplas estratégias e tecnologias, incluindo uma série de passos e trabalhos a serem efetuados para poder chegar ao produto final de um mapeamento, um mosaico em alta resolução.

Atualmente no mercado existem vários softwares para desempenhar tal papel, sendo em sua maioria de uso proprietário. Como exemplo pode-se citar o Photoscan

da empresa Agisoft [Photoscan 2019]. A lista de funcionalidades inclui foto-triangulação aérea e terrestre com exportação dos dados para outros softwares fotogramétricos, geração de nuvens de pontos e modelos poligonais triangulares 3D, MDT (Modelo Digital de Terreno) e MDS (Modelo Digital de Superfície), ortofoto e processamento de imagens multiespectrais, sendo estas amplamente empregadas no meio agrícola.

Por outro lado, Tarallo [Tarallo et al. 2011] discute formas alternativas aos softwares proprietários para criar mosaicos de imagens digitais com um resultado satisfatório. Segundo o autor, utilizando a biblioteca de visão computacional OpenCV [OpenCV 2019], é possível criar uma ferramenta de mosaicagem automática.

No cenário descrito por Tarallo, a forma manual de mosaicar imagens consiste em obter o MDT (modelo digital do terreno), fazer a ortorretificação de imagens e colocação manual de bandeiras marcadoras, usadas para que o software possa reconhecer e construir o mosaico deste terreno. Utilizando esta técnica, a construção de mosaicos pode ser demorada e trabalhosa.

A solução proposta, consiste em utilizar uma técnica denominada SIFT (*Scale Invariant Feature Transform*). O método SIFT consiste em transformar uma imagem em uma coleção de vetores de características locais (descritores de características) e cada um desses vetores são invariantes à escala, rotação e parcialmente invariante às mudanças de iluminação e ponto de vista. Esta técnica é considerada um mosaico não controlado, onde as imagens são combinadas por características comuns entre elas, sem a utilização de marcadores pré-indicando sua disposição no mosaico.

O primeiro estágio do SIFT é a Detecção de extremos no Espaço Escala. Neste primeiro estágio é feita a procura por todas as escalas e locais de uma imagem, onde para isto é utilizada uma função conhecida como diferença da matriz Gaussiana, servindo como identificação dos potenciais pontos de interesse que são invariantes à escala e orientação. Esta é a parte que mais demanda poder computacional do algoritmo. A segunda etapa é a localização dos pontos-chaves (*KeyPoints*), onde para cada local na imagem que foi selecionado como possível *KeyPoints* é determinada sua posição e escala e por fim, eles realmente são selecionados baseados em medidas de sua estabilidade. A terceira etapa é chamada de Definição da Orientação. Uma ou mais orientações são atribuídas para cada ponto-chave localizado, baseada em direções do gradiente. Todas as operações posteriores são realizadas sobre os dados da imagem que foram considerados ponto-chave e que foram transformados em relação à orientação, escala e localização, proporcionando invariância a estas transformações. Por fim, a etapa Descritor dos Pontos-Chave é responsável por mensurar os gradientes na escala selecionada na região ao redor de cada ponto-chave, sendo criados histogramas de orientações para compor o descritor que após a finalização é utilizado para fazer a mesclagem e união das imagens em um único panorama.

Outros autores demonstram outras técnicas para a mosaicagem não controlada semelhantes ao SIFT, o método SURF (*Speeded Up Robust Features*). Este é um detector de características invariante à escala de rotação. Semelhante ao SIFT, ele também é baseado na teoria do espaço de escala, mas no entanto, utiliza a matriz Hessiana de toda a imagem para estimar os máximos locais em diferentes espaços de escala [Yang et al. 2011].

Existem outros métodos discutidos em alguns trabalhos, sendo um deles os mosaicos baseados em característica de baixo nível. Este método não exige que as imagens tenham uma grande área de sobreposição entre elas, pois com a detecção de alguns pontos já é possível fazer a mosaicagem. As características detectadas por este método são bordas, canto, pixel, cor, dentre outras [Ghosh 2012]. Este método inclui o detector de cantos FAST, o qual, detecta cantos em comum nas imagens.

A utilização de mosaicos torna-se importante quando é necessário fazer o mapeamento de grandes áreas, mas para tal, precisa-se capturar inúmeras fotos devido a vários fatores como tamanho de pixel e altura de onde ela foi capturada, uma única imagem não consegue cobrir toda área. Outras aplicações que pode-se citar é o aumento do campo de visão de um UAV (Veículo aéreo não Tripulado) em tempo real, utilizando mais do que uma câmera, gerando uma Única imagem só destas câmeras, em um grande mosaico.

Com a constante evolução da computação e das técnicas empregadas na construção de mosaicos alguns padrões foram sendo mais usados. Deste modo, segundo Wolf e Dewitt [Wolf and Dewitt 2000] a criação de mosaicos provenientes de fontes aéreas geralmente podem ser divididos em três classes: Mosaicos Controlados, Mosaicos Não Controlados e Mosaicos Semi-controlados. Neste trabalho foi utilizada a técnica de Mosaicos Não Controlados, sendo que a detecção de pontos chaves para a posterior sobreposição, utilizou de recursos do *framework* OpenCV.

## 2. Material e Métodos

O software denominado Auto Mosaicing tem como objetivo mosaicar imagens digitais de maneira simples e rápida, resultando em um mosaico de imagens digitais com qualidade suficiente para detectar como por exemplo, falhas de plantação, má desenvolvimento da planta, níveis de biomassa, entre outros. Desenvolvido e compilado para a plataforma MacOS, ele também poderá ser compilado para outras plataformas como Windows e distribuições Linux diversas, pois seu código, desenvolvido em C++ permite esta portabilidade, bastando apenas o usuário ter instalado e configurado em seu computador um compilador GCC com o *framework* OpenCV instalado e suas bibliotecas ligadas ao projeto.

### 2.1. Obtenção do mosaico com a stitching API

A criação de um mosaico envolve basicamente quatro etapas de processamento: Registro, Re-projeção, Stitching e Mesclagem. Embora o estado da arte indique avanços nessa área de pesquisa nos últimos anos, o mosaico de imagens continua a ser um desafio devido a fatores que podem mudar de um cenário para outro, aonde provavelmente não poderá ser utilizado um mesmo modelo para todos os cenários [Ghosh 2014]. A seguir, cada uma destas etapas são brevemente descritas:

- **Registro:** Para registrar um conjunto de imagens, é necessário estimar as transformações geométricas que alinham as imagens de acordo com uma imagem de referência dentro deste conjunto. O conjunto pode consistir de duas ou mais imagens tiradas de uma única cena em momentos diferentes, a partir de pontos de vistas diferentes.
- **Re-projeção:** Refere-se ao alinhamento das imagens para um sistema de coordenadas comum utilizando as transformações geométricas calculadas.

- **Stitching:** O objetivo desta etapa é sobrepor as imagens alinhadas em uma composição maior, combinando valores de pixels das partes sobrepostas e retendo pixels aonde não ocorre sobreposição.
- **Mesclagem:** Diferenças fotométricas globais frequentemente podem ocorrer entre imagens em uma sequência, resultando na visibilidade da emenda entre elas. Esta etapa é utilizada para minimizar esse efeito e homogenizar a aparência global do mosaico.

Toda documentação do *framework* está disponível no site do OpenCV [OpenCV 2019], juntamente com exemplos de implementações. Na Figura 1 é apresentado o código para criação de um mosaico utilizando a classe *Stitcher*.

Na linha 14 da Figura 1, é declarada uma variável do tipo *Mat* responsável por armazenar o mosaico final, sendo este, um tipo de variável disponibilizado pela classe *Stitcher*. Na linha 15 é criado um objeto "create", passando para ele o atributo "mode", aonde pode ser PANORAMA (*imagens lado a lado*) ou SCANS (*imagens dispersas para todos os lados*). O segundo atributo "try\_use\_gpu" é responsável por habilitar ou não o processamento em GPU (*Graphics Processing Unit*) se disponível no hardware utilizado para o processamento.

Na linha 16 da Figura 1 o objeto "status" recebe o retorno do processamento do objeto "stitch", sendo que este por sua vez, tem como entrada um Array de imagens e uma variável para armazenar o resultado do processamento. Nesta etapa, o código realiza todas as etapas descritas anteriormente de forma automática.

## 2.2. Otimização da mosaicagem

Seguindo o estudo realizado por Ferreira [Ferreira et al. 2013], seguir uma metodologia de coleta das imagens, torna o processo de mosaicagem mais eficiente. Desta maneira o autor mostra que duas imagens vizinhas devem conter uma sobreposição chamada de *Overlapping* de 70% em média. Esta técnica é segmentada em duas, sendo *Sidelap* o percentual de sobreposição lateral de duas imagens e *Frontlap* a sobreposição frontal ou traseira de duas imagens. Para atingir o valor de 70% de *Overlapping*, uma média do *Sidelap* e do *Frontlap* é calculada para tal.

Para implementar o estudo de Ferreira, foi utilizada uma técnica chamada ROI (rectangular region of interest), ou seja, retângulo da área de interesse. Esse ROI é criado utilizando exatamente a área em que existe uma sobreposição (*Overlapping*) entre duas imagens.

Sugerindo que a captura das imagens é feita de maneira em que o *Overlapping* corresponda a 70% e sua posição no vetor de imagens já está definido, o algoritmo terá conhecimento de aonde calcular a área de interesse (ROI) de cada imagem. Como observado na figura 2, quando carregadas as imagens para o vetor, o algoritmo calcula e estima que o mosaico deve ser formado por 6 imagens dispostas em uma matriz 2x3, de forma que a primeira imagem possui dois pontos de sobreposição (*sidelap*: img 4 e *frontlap*: img 2) e assim sucessivamente até a última imagem.

```

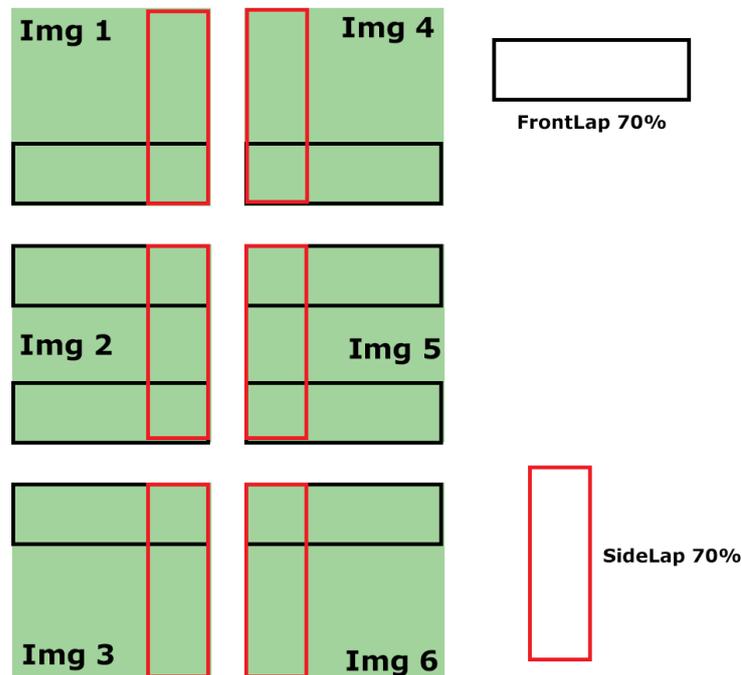
1  #include <iostream>
2  #include <fstream>
3  #include "opencv2/imgcodecs.hpp"
4  #include "opencv2/highgui.hpp"
5  #include "opencv2/stitching.hpp"
6  using namespace std;
7  using namespace cv;
8  bool try_use_gpu = false;
9  Stitcher::Mode mode = Stitcher::SCANS;
10 vector<Mat> imgs;
11 string result_name = "result.jpg";
12 int main()
13 {
14     Mat pano;
15     Ptr<Stitcher> stitcher = Stitcher::create(mode, try_use_gpu);
16     Stitcher::Status status = stitcher->stitch(imgs, pano);
17     if (status != Stitcher::OK)
18     {
19         cout << "Can't stitch images, error code = " << int(status)
20             << endl;
21         return -1;
22     }
23     imwrite(result_name, pano);
24     return 0;
25 }

```

Figura 1. Trecho do código responsável pela mosaicação das imagens

### 3. Resultados e Discussão

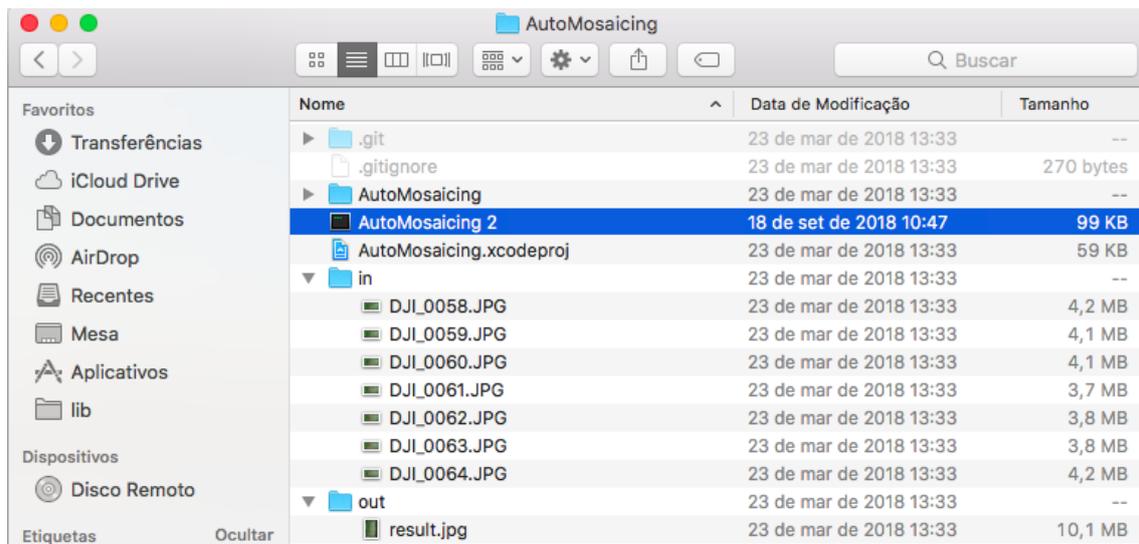
O software desenvolvido foi compilado para a plataforma MacOS, sem interface gráfica, sendo sua execução feita por meio do terminal. A pasta do aplicativo contém o arquivo binário compilado e duas pastas, IN e OUT, aonde deverão ser colocadas as imagens de entrada (IN) e o resultado será armazenado na pasta (OUT).



**Figura 2.** Áreas de cálculo do ROI em cada uma das imagens, levando em consideração sua posição. (\*O tamanho dos retângulos não corresponde à 70% da imagem, é meramente uma ilustração de sua posição dentro da imagem.

Para implementar a técnica de ROI, foi desenvolvida uma função no algoritmo, onde seu principal objetivo é delimitar a área em que existe um *Overlapping* verdadeiro. Desta maneira a função responsável por detectar as características entre as imagens sabe que somente nesta área deve-se fazer a procura, poupando processamento, tempo de execução e deixando o algoritmo mais assertivo para gerar um mosaico de imagens.

Observando o código criado (Figura 4), na linha 3, é definido qual a porcentagem utilizada no *Overlapping*, nas linhas 6 e 7, são criados os vetores chamados *rois*, onde serão armazenadas as posições das linhas e colunas da imagem que contém o *Overlapping*. Após, utilizando um laço de repetição, é percorrida imagem por imagem do vetor de imagens que anteriormente foi carregado para o programa. Neste laço, por meio de testes condicionais, é identificado qual a posição da imagem no mosaico, e então é criada uma área de interesse de 70% e armazenado em um vetor de *ROIS*.



**Figura 3. Diretório do Software com o executável e as pastas de entrada e saída de imagens.**

Após o término deste processo, o algoritmo comporta-se da mesma maneira que foi demonstrada anteriormente, com exceção da linha 3 do algoritmo apresentado na Figura 5. Além do vetor de imagens e uma variável para armazenar o resultado do processamento, é incluído o vetor de ROIS anteriormente processado. Com isso os passos seguintes do algoritmo são verificar se existe algum retorno de erro do processamento e gravar o resultado final na pasta de destino.

```

1
2  int VALUE = 70;
3  int OVERLAP = imgs.at(0).rows*VALUE/100;
4  int IMGSIZE = imgs.at(0).rows*50/100;
5
6  vector<vector<Rect>> rois;
7  vector<Rect> roisVect;
8  //Images of the columns
9  for (int j = 0; j < imgs.size(); j++) {
10     //First top image Rect(x,y,width,height)
11     if (j == 0) {
12         Rect rect1 = Rect(0, 0, imgs.at(j).cols,IMGSIZE);
13         cout << rect1.size() << endl;
14         roisVect.push_back(rect1);
15         rois.push_back(roisVect);
16         roisVect.clear();
17
18         //Last image
19     } else if (j == imgs.size() - 1) {

```

**Figura 4. Técnica de retângulo da área de interesse, onde o algoritmo determina a região da imagens aonde deve-se procurar os pontos em comum para a mosaicagem.**

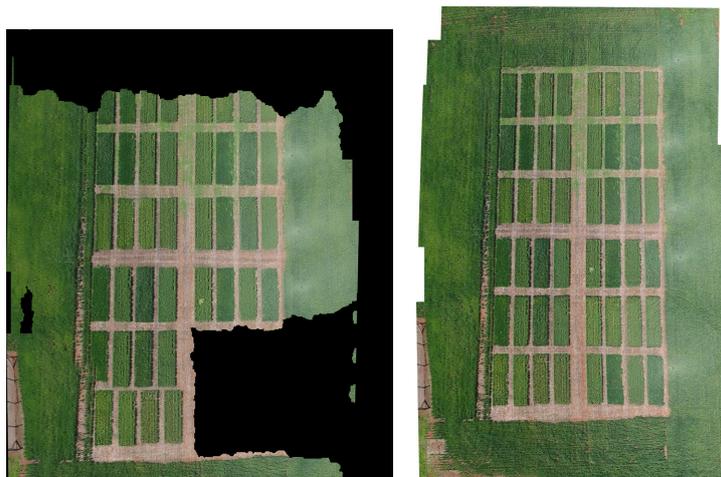
```

1  Mat pano;
2  Ptr<Stitcher> stitcher = Stitcher::create(mode, try_use_gpu);
3  Stitcher::Status status = stitcher->stitch(imgs, rois, pano);
4  cout << endl << status;
5  cout << endl << pano.size();

```

**Figura 5. Técnica de retângulo da área de interesse, onde o algoritmo determina a região da imagens aonde deve-se procurar os pontos em comum para a mosaicagem.**

Para testar o software desenvolvido foram utilizadas imagens coletadas em um campo experimental com cultivares de trigo. A captura destas imagens seguiu o pré-requisito de possuir o *Overlapping* médio de 70%. Na Figura 6 são apresentados dois processamentos utilizando exatamente as mesmas imagens em ambos os mosaicos, sendo que a Figura 6(a) não está utilizando a técnica de ROI e a Figura 6(b), utiliza a técnica de ROI. Ao total foram utilizadas 10 imagens para processar cada mosaico.



(a) Mosaico gerado sem utilizar a função de ROI. (b) Mosaico gerado utilizando a função ROI.

**Figura 6. Diferentes mosaicos gerados utilizando o software desenvolvido.**

#### 4. Conclusões

A utilização da técnica de pontos de interesse (ROI) tornou-se um fator importante para este trabalho, visto que o algoritmo anteriormente testado não obteve sucesso na mosaicagem justamente por não ter um ponto de partida para a etapa de registro da mosaicagem. Para este trabalho, seus resultados podem ser considerados satisfatórios, pois as imagens utilizadas aqui foram posteriormente analisadas e processadas, gerando resultados condizentes com a utilização de outros aplicativos e softwares proprietários de mosaicagem.

Como trabalhos futuros, destaca-se a necessidade de desenvolvimento de uma interface gráfica, o que poderá ajudar pessoas leigas no assunto a utilizarem o software, além de um aprimoramento na função que faz o cálculo da área de interesse (ROI).

#### Referências

- Ferreira, A., Roig, H., Marotta, G., and Meneses, P. (2013). Utilização de aeronaves remotamente pilotadas para extração de mosaico georreferenciado multiespectral e modelo digital de elevação de altíssima resolução espacial. *Simposio Brasileiro de Sensoriamento Remoto*, 1:9308–9315.
- Ghosh, D. (2012). Quantitative evaluation of image mosaicing in multiple scene categories. pages 1–6.
- Ghosh, Kaabouch, F. (2014). Robust spatial-domain based super-resolution mosaicing of cubesat video frames: Algorithm and evaluation. page 68.
- Loch, C. (2001). Importância do monitoramento global e integrado do planejamento municipal. In *VI CONGRESSO BRASILEIRO DE SENSORIAMENTO REMOTO*, volume 2, pages 523–530.
- OpenCV (2019). Documentation open source computer vision.
- Photoscan, A. (2019). Photogrammetric processing of digital images and generates 3d spatial data.

- Rong, W., Chen, H., Liu, J., Xu, Y., and Haeusler, R. (2009). Mosaicing of microscope images based on surf. In *2009 24th International Conference Image and Vision Computing New Zealand*, pages 271–275.
- Tarallo, A. D. S., Hiraga, A. K., de Paiva, M. S. V., Jorge, L. A. d. C., and da Silva, F. A. (2011). Construção automática de mosaico de imagens agrícolas aéreas sequenciais. *Agricultura de precisão: um novo olhar*, pages 106–109.
- Wolf, P. R. and Dewitt, B. A. (2000). Elements of photogrammetry: with applications in gis. *McGraw Hills Comp., USA*.
- Yang, L., Wu, X., Zhai, J., and Li, H. (2011). A research of feature-based image mosaic algorithm. In *2011 4th International Congress on Image and Signal Processing*, volume 2, pages 846–849.

# Desenvolvimento de uma fresadora CNC utilizando software de modelagem 3D e hardware de código aberto

Amauri Lopes<sup>1</sup>, Rodrigo Otávio Oliveira<sup>1</sup>, Anselmo Rafael Cukla<sup>1</sup>

<sup>1</sup> Instituto Federal de Educação, Ciência e Tecnologia Sul-Rio-Grandense (IFSUL)  
Câmpus Passo Fundo

Estrada Perimetral Leste, 150 – Passo Fundo – RS – Brasil

{amaurilopes1203, roo1039}@gmail.com,  
anselmo.cukla@passofundo.ifsul.edu.br

**Abstract.** *This work approaches a methodology of construction of a CNC milling machine. A 3D design software was used to develop the conceptual design. Simulations were performed using the finite element method to perform the milling machine design. For command and control, the Arduino integrated circuit board was used. In this platform, the Marlin® firmware was applied and configured to be used for milling process. Based on the conceptual design, it was possible to develop a prototype CNC milling machine to validate the design. Also, in this article, an example for image conversion for engraving and application of the CNC milling machine is discussed.*

**Resumo.** *Este artigo descreve uma metodologia de construção de uma fresadora CNC. Utilizou-se um software desenho 3D para realizar o desenvolvimento do projeto conceitual. Ainda, foram realizadas simulações, utilizando o método de elementos finitos para realizar o dimensionamento da fresadora. Para o comando e controle, utilizou-se a placa de circuito integrado Arduino, no qual foi aplicado o firmware Marlin® sendo este configurado para ser utilizado para processo de fresamento. Com base no projeto conceitual, foi possível o desenvolvimento de um protótipo de uma fresadora CNC para validar o projeto. Ainda, nesse artigo, se aborda um exemplo para a conversão de imagens para realização de gravações e método de aplicação da fresadora CNC.*

## 1. Introdução

A exigência humana por ter uma maior produção em um curto período de tempo (eficiência), tanto em produtos quanto em serviços, tem origem desde do início da humanidade e teve um grande crescimento após a Primeira Revolução Industrial. Porém, somente no século anterior que foram criados projetos de automatização, tendo um trabalho de precisão e com elevada eficiência, possibilitando uma produção em série sem perder a qualidade do produto [Andrade *et al.* 2017], como exemplo, a produção em série de carros, a automatização dos equipamentos, entre outros.

As fresadoras CNC, são máquinas controladas por comandos numéricos computadorizados, tendo uma estrutura capaz de criar peças desenvolvidas computacionalmente a partir de mecanismos automatizado, utilizados para usinagem.

Essas máquinas automatizadas são amplamente usadas em indústrias para a usinagem de peças, sendo possível obter uma maior repetibilidade na fabricação [Polastrini 2016].

Atualmente, no mercado existem fresadora, mesmo sendo de pequeno porte, com preço elevado, sendo utilizada para usinagem de madeiras, plásticos e placas de circuito impresso. Para diminuir custo, utiliza-se plataformas eletrônicas, como o Arduino que possui uma plataforma de código aberto acessível [Arduino 2019], para realizar o controle da máquina e estruturas em perfis de aço ou em alumínio, desenvolvidas em softwares de desenho para criar a máquina.

O presente artigo descreve uma metodologia de desenvolvimento do projeto de uma fresadora CNC de baixo custo, baseada no uso de *hardware* e *firmware* de código aberto. O projeto apresenta a construção com base em simulações em 3D e sua programação empregando circuito eletroeletrônico com ferramentas *Open Source*. Finalmente, é apresentado um exemplo de aplicação de fresado sobre uma superfície em acrílico.

Este artigo está organizado da seguinte forma: No capítulo 1 a Introdução, no capítulo 2 é tratada a Metodologia de construção da máquina CNC; já no capítulo 3 é tratado a Implementação de uma aplicação utilizando a máquina CNC, e finalmente; nos capítulos 4 e 5 os Resultados e Conclusões.

## 2. Metodologia

A metodologia empregada no presente trabalho consiste basicamente em três etapas, conforme mostrado na Figura 1:



Figura 1. Ordem das etapas da metodologia empregada.

### 2.1. Projeto elétrico

Para o desenvolvimento da máquina fresadora CNC se faz necessário primeiramente a seleção e dimensionamento dos componentes eletrônicos que serão utilizados na máquina conforme mostrado na Figura 2, para posteriormente projetar a estrutura mecânica mais adequada.

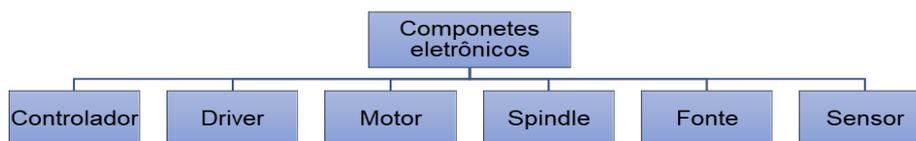


Figura 2. Componentes eletrônicos empregados.

Para realizar o controle do equipamento foi utilizado Arduino, que é uma placa capaz de ler entradas de sinais e transformá-los em sinais de saída, permitindo também a aplicação de comandos, que podem ser enviados para o microcontrolador da placa [Arduino 2019].

Para facilitar as conexões entre os *drivers*, motores e o Arduino, foi empregado uma *Ramps*, que seria uma placa projetada para ser encaixada ao Arduino, permitindo, de forma simplificada, o encaixe e ligação de todos os componentes eletrônicos de uma

máquina CNC, além proporcionar a adição de componentes como telas lcd, adaptadores para cartão de memória, entre outros [Reprap 2019].

O designer modular da *Ramps* é formado principalmente por plugs para *driver* de passo e um *shield* para encaixe do Arduino MEGA. Para facilitar a sua utilização, permite que se possa ser feita a substituição de peças, atualização e expansão da placa de forma simples [Reprap 2019].

Os motores de passo são dispositivos eletromecânicos que utilizam pulsos elétricos para realizar a movimentação angular do seu eixo, este movimento é chamado de passo e controlado por meio de um *driver* [Ferrari 2011]. Utilizou-se no projeto três motores de 5 kg.cm, do tipo unipolar, com tensão de 7V e corrente de 1.2 A.

Um *driver* para motores de passo é um circuito empregado para amplificar os sinais digitais de baixa potência, permitindo o controle de um atuador, proporcionando utilizar maiores níveis de energia em relação ao que a interface de saída pode oferecer, conforme esquematizado na Figura 3 [Pazos 2002].

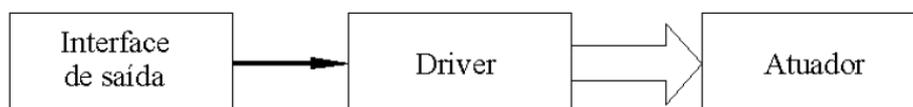


Figura 3. Esquema de funcionamento de um *driver*.

Nesse projeto foi empregado três *drivers* DRV8825 para controlar os motores de passo. Além de ter um baixo custo, possui proteção contra sobrecarga de corrente e temperatura, aceitando uma alimentação de até 45 V e corrente por fase de até 1.5 A, permitindo também uma resolução de 1/32 por passo. O esquema de ligação do *driver* é apresentado na Figura 4, apresentado a ligação entre o *driver*, o controlador e a fonte de alimentação [Pololu,2019].

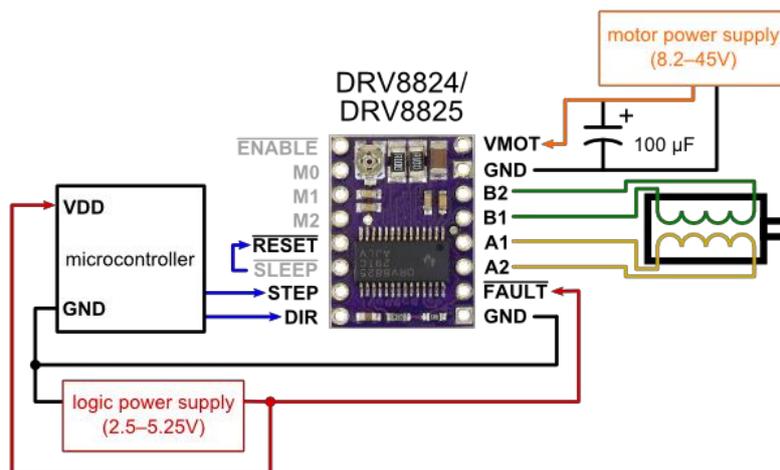
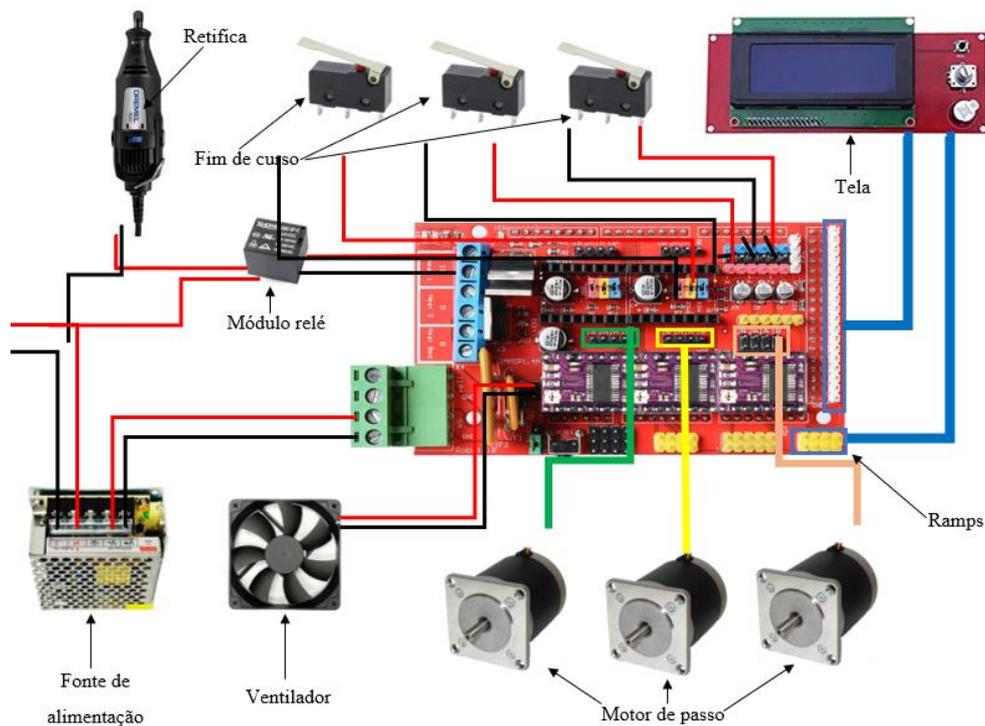


Figura 4. Ligação do *driver* de motor de passo.

Com base nesses componentes, foi possível a criação do esquema elétrico mostrado na Figura 5, sendo adicionado ainda um ventilador 80x80 mm para realizar o arrefecimento dos *drivers* e um tela LCD 2004 20x04, que permite realizar o controle da máquina por esse componente. Para realizar o fresamento, utilizou-se um módulo de relé de 5V para acionar a retífica, utilizando para tal uma Dremel® de 85W.



**Figura 5. Circuito elétrico utilizado na fresadora.**

## 2.2. Projeto mecânico

No desenvolvimento de uma máquina, o projeto mecânico deve cumprir algumas condições de acabamento e tolerâncias dimensionais sem necessitar da habilidade do operador [Stoeterau 2004]. Para se obter um desempenho competitivo para estas condições é de essencial importância a escolha e dimensionamento de forma correta dos componentes mecânicos que serão aplicados no projeto, sendo estes apresentados na Figura 6.



**Figura 6. Componentes mecânicos empregados.**

Selecionados os componentes mecânicos, realizou-se o desenho e dimensionamento do projeto utilizando o programa SolidWorks® empregando as ferramentas de desenho e simulação disponíveis no aplicativo.

O projeto da estrutura, pode ser realizado por aplicativos de desenho auxiliado por computador (CAD) que contam com funções de análise de elementos finitos (FEA) para análise de tensões, permitindo o desenvolvimento de projetos mais detalhado. Para a análise da estrutura deste projeto, foi realizado apenas a análise estática das forças, ainda considerando que se os movimentos da máquina forem lentos e as acelerações desprezíveis [Norton 2013].

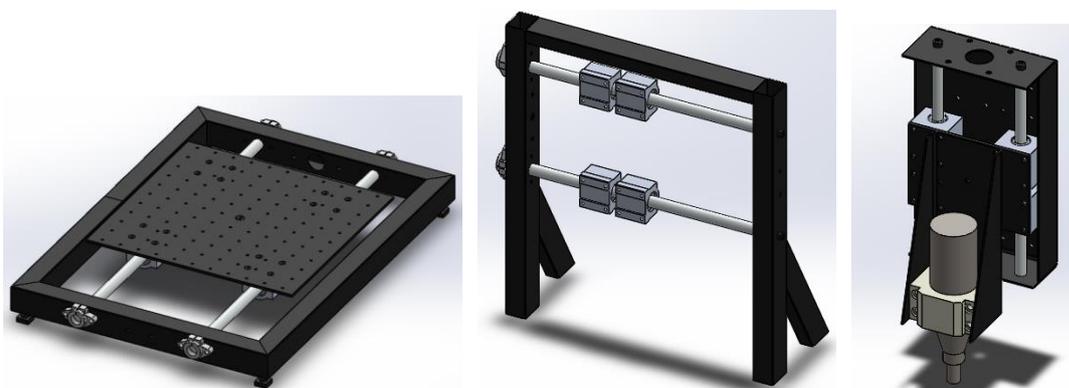
A estrutura da máquina deve “ser capaz de resistir aos movimentos de corte, tração, compressão, flexão, torção ou encurvadura sem que estes fenômenos se façam notar na peça final obtida, garantindo sempre a segurança do utilizador” [Moreira 2015].

A estrutura mecânica, construída em perfis retangulares de aço, compõe-se basicamente em três partes:

**Eixo Y:** corresponde como a base da máquina fresadora, contendo a mesa na qual será fixada a peça e permitirá a movimentação da peça no eixo Y, conforme mostrado na Figura 7.a.

**Eixo X:** esta parte será soldada sobre o eixo Y e permitirá a movimentação do cabeçote da fresadora no eixo X, podendo ser observado na Figura 7.b, além de servir de base para a fixação da parte eletrônica.

**Eixo Z:** este conjunto de componentes constitui o cabeçote da fresadora, sendo fixado essa estrutura no eixo X e nesse local onde será fixado a retífica, conforme abordado na Figura 7.c.



**Figura 7. Estrutura da fresadora: a) Eixo Y, b) Eixo X e c) Eixo Z.**

As forças de usinagem são essenciais na para análise do funcionamento da máquina, sendo estas utilizadas para determinar a rigidez da fresadora. Normalmente para tal dimensionamento se desconsidera a força da gravidade, pois o erro causado por tal força é sistemático e pode ser corrigido por meio da programação empregada.

Com base na força de usinagem pode ser realizada a simulação da estrutura, para verificar sua rigidez, não sendo extremamente precisa pois foi realizado algumas considerações, como exemplo, a maioria dos componentes tem contato unido, mas tendo grande importância na avaliação da estrutura e dimensionamento da mesma.

Para realizar a simulação, não foi adicionado o sistema de transmissão de movimento e alguns dos parafusos para facilitar a simulação no SolidWorks®. Para simular, fixou-se as pernas da base da máquina e realizou-se a aplicação uma força de 2 kgf na ponta da ferramenta. O posicionamento da força, foi no local a onde ocorre maior flexão, para descobrir o deslocamento máximo. A malha utilizada para a simulação foi mista com base em curvatura e tamanho máximo de elementos de 30 mm, a simulação pode ser vista na Figura 12.a. da seção 4.

### 2.3. Interface de programação

Para o movimento correto de uma máquina CNC é preciso que seja adicionado instruções adequadas. Para isso se tem programas utilizados para tais ações de controle do equipamento, determinando assim, que ações a máquina deve fazer para realizar o processo de usinagem [Pérche 2013]. Sendo apresentado nesta seção as partes da programação utilizadas no presente trabalho.

O Marlin® é um *firmware* de código aberto, utilizado em controladores, principalmente, para impressoras 3D, permitindo o gerenciamento de todas as tarefas do equipamento [Marlin 2019]. Além de ser um *firmware* gratuito, proporciona uma fácil configuração e permite o uso e controle de aquecedores, motores de passo, sensores, displays, teclados, leds, botões entre outros.

A linguagem de controle do Marlin® é baseada no código G. Essa linguagem pode ser adicionada por meio do monitor serial do programa Arduino ou através do programa CNC Gcode Controller, quando conectado com o computador. Também é possível a adição de arquivos em código G por meio de cartões de memória na máquina.

O *firmware* do Marlin® permite realizar modificações, tais como adição de placas gráficas, controle da retifica, arquivos por meio de cartão de memória, entre outras funções. Por ter tantas funções, esse *firmware* se torna complexo, dificultando realizar configurações específicas, no entanto, na página do Marlin® se tem as etapas passo a passo das configurações principais para funcionamento da máquina. Por esse diferencial e por permitir o gerenciamento da máquina sem a necessidade de um computador foi utilizado o Marlin®.

### 3. Implementação

Para a criação de gravações pode-se empregar o programa Inscap®®, no qual com base em uma imagem ou a partir de um desenho criado, pode-se realizar a gravação da mesma ou o fresamento dos contornos da imagem, no qual é utilizado para fresamento de placas de circuito impresso, realizando a vetorização da imagem e a convertendo em código G. Para exemplificar, pode-se fazer a conversão de uma imagem de placa de circuito a ser impresso como da Figura 8 em código G.

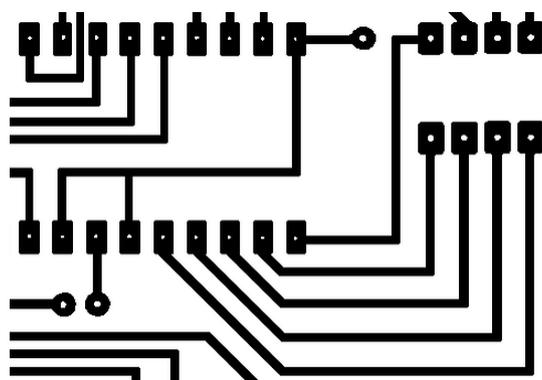
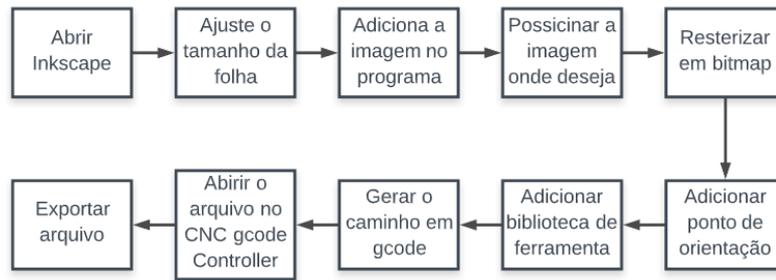


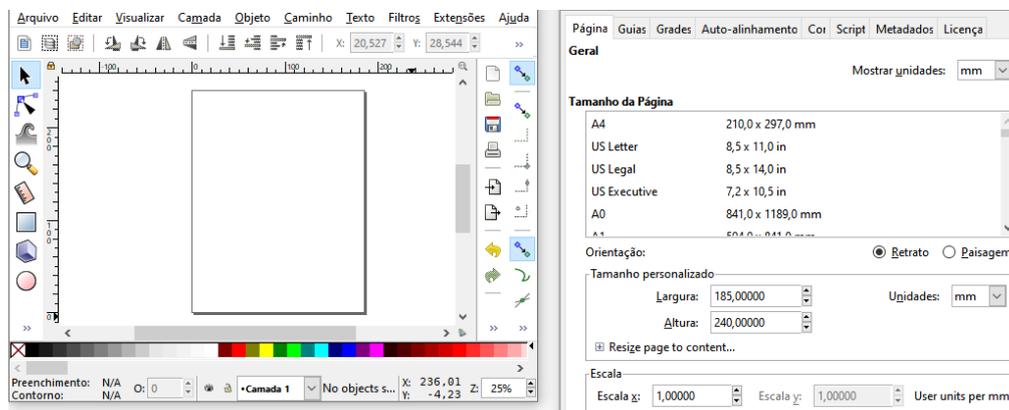
Figura 8. Exemplo de imagem de circuito impresso.

O esquema da sequência de etapas que se deve realizar pode ser visto no diagrama da Figura 9, sendo utilizado os programas Inscap para criar o código G e o CNC Gcode Controller para ajustar os comandos para ser aplicado na máquina.



**Figura 9. Diagrama das etapas de criação do caminho da gravação.**

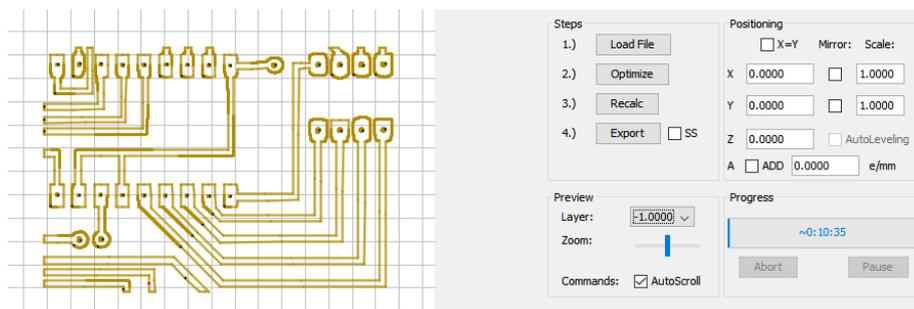
Ao iniciar um novo arquivo deve-se configurar o tamanho da folha conforme a Figura 10, para ser o mesmo que o da área de trabalho da fresadora. Posteriormente, ao adicionar a imagem, deve-se ajustar no ponto na área de trabalho no qual se deseja gravar.



**Figura 20. Dimensão da área de trabalho.**

Ajustado a imagem, deve-se realizar a rasterização em bitmap, para o programa poder encontrar os pontos para indicar o caminho de contorno da gravação. Em sequência, adiciona-se o ponto de orientação e escolhe na biblioteca de ferramentas o tipo de fresa que irá utilizar. Com esses dados informados, pode-se criar o caminho e gerar a construção do código G.

Em seguida, inicializa-se o programa CNC Gcode Controller e abra o arquivo gerado. Por fim, exporte o arquivo aberto para ser configurado automaticamente para funcionar na fresadora CNC desenvolvida. Ao reabrir o arquivo, o código G criado apresentará a imagem que será gravada conforme apresentado na Figura 11, sendo esta imagem o caminho que a fresadora realizará ao fresar a peça.

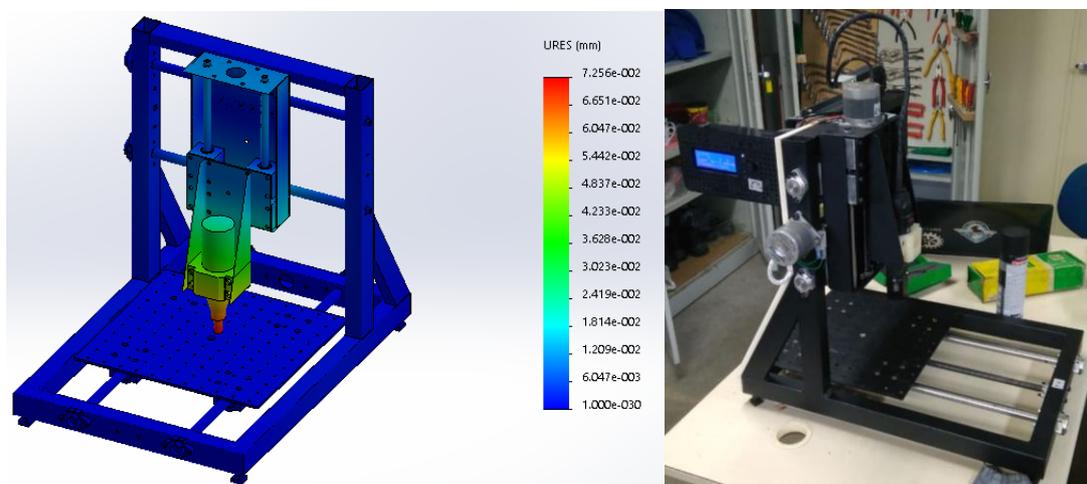


**Figura 31. Simulação do código G gerado.**

#### 4. Resultados e aplicações

Utilizando os métodos de elementos finitos na parte de simulação do SolidWorks®, pode verificar a precisão da máquina, sendo esta apresentada na Figura 12.a, no qual pode-se obter uma deflexão máxima de 0,0726 mm na ponta da ferramenta, sendo aceitável para realizar a usinagem de matérias macios como madeiras, plásticos, etc.

Com base no projeto desenvolvido teoricamente pode-se construir a fresadora CNC apresentada na Figura 12.b, tendo para sua construção um grande aproveitamento de materiais reciclados como chapas, motores, caixas de plásticos entre outras que são responsáveis por manter a estrutura do projeto com um baixo custo.



**Figura 42. Fresadora CNC: a) Simulação de análise estática e b) fresadora construída.**

Existem diversas aplicações que a máquina pode ser utilizada, dentre elas pode-se citar o uso para realizar gravações, para isso gera-se um código G, para criar o caminho da ferramenta para realizar a gravura. No projeto utilizou-se o programa CNC Gcode Controller para realizar o controle da máquina e também para realizar a adição e edição do código G criado e testes, antes de ser enviado para a máquina.

Com o código G criado, pode-se enviá-lo para a máquina por meio da porta serial do Arduino, ou por meio de cartão de memória. Após enviado o código a máquina inicia o processo de gravura, fazendo o fresamento da peça como pode ser observado na Figura 13.a.

O resultado da gravação pode ser visto na Figura 13.b, apresenta o desfecho do processo de fresamento, no qual se tem a palavra “IFSUL” escrita em acrílico. No entanto, como pode ser visto na figura, a espessura das linhas variou por causa que no processo não foi utilizado nivelamento automático.



**Figura 53. a) Funcionamento da fresadora e b) Exemplo de aplicação.**

## **5. Conclusões**

Os resultados obtidos pela máquina indicam que, mesmo com a utilização de matérias reutilizáveis, sistemas computacionais de código aberto, como o Marlin®, permitem a criação de uma máquina de fácil configuração e ainda tendo grande precisão.

O projeto proporcionou a eficácia e a praticidade gerada pela criação do projeto mecânico, facilitando a criação de desenhos a partir do programa SolidWorks®, que permitiu a criação de forma fácil das peças fabricadas como o suporte para a retífica, os suportes para os fusos, entre outros.

Pode-se concluir que o presente projeto proporciona uma forma simples de criação de uma fresadora CNC, sendo recomendada tanto para uso educacional, para entender mais a fundo o uso de comandos numéricos, quanto para uso para fins pessoais, como gravação de imagens, placas entre outros.

## **Referências**

Andrade, B. H. S., Gonçalves, G. P., Carvalho, R. R. M., Costa, M. M. and Rodrigues, F. I. B. (2017). “Projeto e construção de mesa a comando numericamente computadorizado com mini retífica”. In: Congresso Técnico Científico da Engenharia e Agronomia, Pará, Brasil.

Arduino (2019). “What is Arduino?”, <https://www.arduino.cc/en/Guide/Introduction>, Outubro.

Ferrari, C. D. (2011). “Implementação de um sistema de comando numérico computadorizado (CNC) por retrofitting em torno mecânico universal”. FATEC, São Paulo, Brasil.

Marlin (2019). “What is Marlin?”, <http://marlinfw.org/docs/basics/introduction.html>, Outubro.

Moreira, A. F. L. (2015). “Máquina-Ferramenta CNC para Gravação - Desenvolvimento de Protótipo de Baixo Custo”, Dissertação (Mestrado em Engenharia Mecatrônica) – Universidade do Minho, Braga, Portugal.

Pazos, F. (2002). “Automação de sistema e robótica”, Axel Books, Rio de Janeiro, Brasil.

Pérche, C. F. de P. (2013). “Desenvolvimento de um sistema de prototipagem para placas de circuito impressos”, Trabalho de Conclusão de Curso – Universidade Federal de Itajubá, Itabira, Brasil.

Polastrini, F. H. (2016). “Desenvolvimento de uma máquina CNC de baixo custo com software e hardware abertos”, Trabalho de Conclusão de Curso – Instituto Federal Minas Gerais, Minas Gerais, Brasil.

Pololu. (2019). “Stepper Motor Drivers”, <https://www.pololu.com/category/120/stepper-motor-drivers>, Outubro.

Reprap. (2019). “Arduino Mega Pololu Shield”. [https://reprap.org/wiki/Arduino\\_Mega\\_Pololu\\_Shield](https://reprap.org/wiki/Arduino_Mega_Pololu_Shield), Outubro.

Stoeterau, R. L. (2004). “Introdução ao Projeto de Máquina-Ferramentas Modernas”, Universidade Federal de Santa Catarina, Santa Catarina, Brasil.

# Sistemas ERP em Ferramentarias de Micro e Pequeno Porte em Santa Catarina

Júlia Ilze de Farias<sup>1</sup>, Osmar da Cunha Filho<sup>1</sup>

<sup>1</sup>Centro Universitário Municipal de São José – São José – SC – Brasil

{juliailze825, osmarcf.usj}@gmail.com

**Abstract.** *This paper presents a research results about ERP's implementation and use inside micro and small machine shops of Santa Catarina. Data were collected through semi-structured interview and questionnaires sent to 70 industries, obtaining 11 responses in total. The answers show that simplicity of interface, supplier/consulting proximity and process control are benefits noted, although improvements in production functionality and lower costs are necessary.*

**Resumo.** *Este trabalho apresenta o resultado de uma pesquisa sobre implantação e uso de softwares ERP dentro do estado de Santa Catarina entre empresas de fabricação de ferramentas de micro e pequeno porte. Os dados foram coletados através de entrevistas semi-estruturadas e questionários distribuídos entre 70 empresas, sendo obtidas 11 respostas ao todo. As respostas obtidas demonstram que simplicidade de interface, proximidade com o fornecedor/consultor e maior controle de processos são benefícios notados, contudo são necessárias melhorias nas funcionalidades relacionadas à produção e preços mais acessíveis.*

## 1. Introdução

A indústria catarinense movimenta um 5,1% do PIB na indústria nacional e possui um setor metal mecânico aquecido e movimentado quase totalmente por micro e pequenas indústrias [PORTAL DA INDÚSTRIA, 2017], [FIESC, 2017].

Dentro do setor metal mecânico há o segmento da fabricação de ferramentas, moldes e matrizes, que se destaca nacionalmente em qualidade e quantidade. A produção, lucratividade e competitividade dessas empresas dependem de um cuidadoso controle e gestão dos recursos. Sabendo disso, algumas dessas empresas têm se arriscado a incorporar sistemas ERP.

Os sistemas ERP são softwares que permitem a gestão, integração e controle das diversas áreas de uma fábrica através de um banco de dados consistente e unificado. Embora pareça muito vantajoso, a aquisição e uso deste tipo de recurso possui desafios diferenciados quando são observados do ponto de vista de um pequeno negócio.

Existe vasta bibliografia a respeito desse tipo de sistema, inclusive voltada para empresas de micro e pequeno porte, porém muitos desses estudos se limitam a análises financeiras, visões sobre vantagens e desvantagens, correspondências entre trabalho de consultoria e implantação, deixando de lado questões relevantes como funcionalidade, usabilidade e saúde do trabalhador. Há também a necessidade de encontrar estudos mais específicos quanto ao nicho de mercado que o software atende, sabendo que no universo dos negócios, cada setor possui características próprias e portanto, necessidades.

Este trabalho tem por objetivo relatar o comportamento e visão das micro e pequenas ferramentarias de Santa Catarina a respeito da utilização de softwares ERP observando aspectos financeiros, funcionais, ergonômicos e de consultoria. A secção 2 descreve em termos gerais o que são sistemas ERP, a secção 3 explica os critérios de seleção do porte das empresas, na 4 são fornecidos dados a respeito da indústria catarinense e por fim, os resultados de pesquisa estão dispostos na secção 5.

## **2. Sistemas ERP**

Segundo Padilha e Marins [apud CHOPRA e MEINDL, 2003], ERP, *Enterprise Resource Planning*, significa Planejamento dos Recursos da Empresa. Este tipo de sistema promove a integração de todo o negócio fornecendo dados consistentes e um fluxo de informações único e contínuo. Isso se dá através da base de dados unificada, que provê rastreamento e visibilidade da informação de qualquer setor da empresa e da cadeia de suprimento, permitindo a tomada de decisões estratégicas.

No Brasil, esses sistemas são chamados de Sistemas Integrados de Gestão Empresarial, e realizam o controle, suporte aos processos operacionais, produtivos, administrativos e comerciais das empresas. As informações registradas são base para consultas que possam refletir o mais próximo possível a realidade do negócio [PADILHA e MARINS, 2005].

### **2.1 Benefícios e desafios de um sistema ERP**

A automação, a integração de processos, o compartilhamento de práticas e dados de negócio e a disponibilização da informação em tempo real são alguns dos benefícios procurados quando uma empresa toma a decisão de adotar um pacote de software ERP. Por possuir essas características, sistemas dessa natureza são chamados de sistemas de gestão, sendo adquiridos para agilizar processos empresariais e extrair informações mais precisas da empresa [RODRIGUES e ASSOLARI, 2007]. Contudo, existem outros pontos que podem representar uma desvantagem estratégica e precisam ser considerados. Zwicker e Souza (2003) alertam sobre a resistência dos funcionários à mudança, sobre a dependência do fornecedor, sobre a necessidade de adequação entre outros diversos problemas.

## **3. Caracterização de empresas de micro e pequeno porte**

Segundo o IBGE (2003), não há unanimidade sobre os critérios de definição dos portes das empresas. Instituições financeiras oficiais e órgãos representativos do setor, ou ambos, costumam usar valores de faturamento como referência para essa delimitação, enquanto que institutos de pesquisa e apoio empresarial adotam o número de pessoas ocupadas. Os conceitos para enquadramento adotados são heterogêneos, pois a finalidade e objetivos das instituições não são os mesmos: regulamentação, crédito, estudos, etc.

O IBGE (2016) apesar de analisar a situação das empresas brasileiras por região, taxa de crescimento, mobilidade e diversos outros aspectos, utiliza a mesma classificação de porte por número de pessoas ocupadas. Limas (2009) explica que esse critério desconsidera as diferenças entre atividades com processos produtivos distintos, uso intensivo de tecnologia da informação (TI) e presença de mão-de-obra qualificada, ignorando a possibilidade da realização de alto volume de negócios com pequena mão-de-obra.

O SEBRAE (2013) realiza a classificação das empresas através do número de pessoas e do setor de atividade econômica. Abaixo, segue a tabela de classificação comparativa entre os dois órgãos citados:

**Tabela 1- Número de pessoas ocupadas por porte da empresa.**

Critério	Micro	Pequena	Média	Grande
IBGE	De 1 a 9 pessoas	De 10 a 49	De 50 a 249	Acima de 249
SEBRAE	Indústria: até 19 pessoas. Comércio e Serviço: até 9	Indústria: de 20 a 99 pessoas. Comércio e Serviço: 10 a 49	Indústria: de 100 a 499 pessoas. Comércio e Serviço: de 50 a 99 pessoas.	Indústria: acima de 499 pessoas. Comércio e Serviço: acima de 99 pessoas.

#### 4. Cenário da indústria em Santa Catarina

Segundo o Portal da Indústria (2017), o estado de Santa Catarina contava com aproximadamente 42.961 indústrias, gerando um PIB de R\$ 59,1 bilhões, o que equivale a 5,1% da indústria nacional. O porte das indústrias é formado em 72,3% de micro, 22,5% de pequenas, 4,3% de médias e apenas 0,9% de grandes empresas, sendo que os micros e pequenos negócios empregam 39,4 % do total de pessoas que trabalham na indústria.

O setor metalmeccânico e metalúrgico é responsável 7,9% dos estabelecimentos industriais do estado e é formado principalmente por empresas de micro e pequeno porte, representando 90,6% e 7,9% do total respectivamente. Atualmente, existem aproximadamente 51.800 pessoas trabalhando nesse segmento, das quais 7,3% trabalham com artigos de cutelaria, de serralheria e ferramentas [FIESC, 2017].

O segmento ferramental em Santa Catarina é formado por 464 indústrias, onde 49% estão distribuídas na região norte do estado. Segundo Geres (2018), “essas indústrias se dedicam à confecção de moldes e matrizes para a produção serial de todo tipo de peças, atendendo outras indústrias, sob encomenda”. O surgimento de muitas dessas empresas ocorreu com a terceirização ou enxugamento de empresas maiores do setor que decidiram praticar a reengenharia. Outra razão para o crescimento do setor no estado está na desativação de grandes indústrias durante a década de 1990, dando oportunidade para trabalhadores iniciarem seus próprios negócios e comesçassem a atender a demanda deixada.

#### 5. Descrição da pesquisa

A pesquisa realizada tem natureza básica e abordagem quali-quantitativa. Para a coleta de dados foram realizadas entrevistas por pauta, com o auxílio de um roteiro, e questionários quando essas não eram possíveis. Os entrevistados eram em sua totalidade profissionais que lidavam diretamente com o uso do software dentro de setores de planejamento e controle de produção ou compras. A seleção das empresas seguiu os seguintes critérios: empresas dentro do estado de Santa Catarina na CNAE (Classificação Nacional de Atividades Econômicas) 25.43-8, fabricantes de ferramentas, que tivessem no máximo 99 funcionários, disponibilidade e permissão para o uso de dados.

Foram convidadas a participar 70 empresas dentro dos critérios adotados, sendo contatadas inicialmente por telefone e por e-mail. As empresas que decidiram responder o questionário receberam um convite via e-mail com o link de acesso ao formulário. Ao todo, existiram 11 respostas, uma taxa 15,7%, sabendo-se que a taxa de retorno média para pesquisas do gênero oscila entre 3% e 5% [MATTAR apud LIMAS, 2009]. Entre as empresas que responderam, 4 informaram ter até 19 funcionários e 7, de 20 a 99, seguindo as classificações de micro e pequena indústria de acordo com o SEBRAE.

### 5.1 Entrevistas e questionário

A entrevista e o questionário foram estruturados para que fosse possível identificar a condição das empresas selecionadas em relação à implantação e uso de softwares ERP. Houve 4 aspectos importantes abordados: porte da empresa; implantação e satisfação com o retorno do investimento; funcionalidade, usabilidade e ergonomia de software, e consultoria e assistência. O questionário foi organizado em 15 questões, enquanto a entrevista abordou as mesmas perguntas de forma semi-estruturada, permitindo que o usuário relatasse livremente circunstâncias do seu dia-a-dia.

### 5.2 Adoção de sistema ERP

Mendes e Escrivão Filho (2002) descrevem o setor das pequenas empresas como limitados em recursos, o que compromete o investimento em tecnologia e gera preocupação no momento de implantar um ERP. Na presente pesquisa, 8 entre os 11 entrevistados declararam fazer uso de algum sistema ERP, sendo que apenas 5 demonstraram adoção exclusiva desse sistema de informação. Os demais usuários faziam uso complementar de planilhas eletrônicas, estavam em processo de migração entre sistemas ou desistência.

Os usuários que não usavam ou que não faziam uso restrito apontaram como principal causa o custo elevado para aquisição ou manutenção do software confirmando, assim, o que é amplamente dito por autores da área. O uso complementar de planilhas eletrônicas vem para sanar lacunas funcionais do software que serão explicados no tópico 5.5.

A mortalidade do ERP para as empresas que estavam em migração ou desistência foi em média dois anos, período esse, marcado pelo impacto dos custos segundo Padilha e Marins (2005).

O gráfico abaixo apresenta a seguinte distribuição por tempo de uso entre aqueles que responderam fazer uso de ERP:

Gráfico 1- Tempo de uso de software ERP.



### **5.3 Benefícios do ERP**

Segundo Souza & Zwicker (2003), ao adotar um sistema ERP as empresas esperam vários benefícios: integração, incremento de possibilidades de controle sobre os processos, atualização tecnológica, acesso a informações em tempo real e redução de custos em informática.

As empresas entrevistadas citaram as seguintes melhorias:

- Maior controle dos processos;
- Agilização de processos e de retorno ao cliente;
- Organização de informações;
- Detalhamento de informações de processo e produção;
- Permite o gerenciamento de tarefas e levantamento de prioridades;

Contudo, de forma menos unânime, foram citados estes outros pontos:

- Adequação com as atividades e necessidades da empresa;
- Integração com a produção;
- Visibilidade de ações na produção.

As ferramentarias que demonstraram maior satisfação relataram um relacionamento bastante próximo com os fornecedores e consultores do software. Em alguns casos, os clientes informaram que empresas de tecnologia da informação de pequeno porte, situadas nos arredores das fábricas, conseguiram reconhecer as necessidades, desenvolver soluções e customizar apropriadamente o sistema criado, reduzindo o impacto da implantação do sistema e fortalecendo vínculos de negócio.

### **5.4 Desafios do ERP**

As dificuldades apontadas no momento da implantação foram:

- Resistência dos funcionários à mudança;
- Adequação com as atividades e exigências do negócio;
- Dúvidas quanto ao uso do software;
- Complexidade excessiva, levando a necessidade de contratação de mais mão-de-obra.

Percebe-se que questões relacionadas à adequação do software não são totalmente compartilhadas pelas as empresas. Entre os entrevistados, houve 1 que afirmou não ter precisado fazer adaptações, apenas inclusões de novos pacotes; no entanto, 6 encontraram necessidades de customização e 1 fazia uso de um software voltado para o comércio, que além de inadequado para o ramo da empresa, possuía alto custo. Os altos custos para a compra de pacotes mais completos e para contratação de consultoria estão entre os fatores para a insatisfação.

### **5.5 Funcionalidade**

As principais lacunas dos softwares ERP, aquelas que criaram necessidade de customização ou que causam insatisfação, informadas pelos usuários eram aquelas relacionadas à produção:

- Necessidade de incluir acesso ao plano mestre por meio do controle de produção;
- Falta da possibilidade de setorização do controle de produção;
- Falta de interação com a produção;
- Necessidade de grandes adaptações para realização do planejamento de produção.

Outros pontos citados foram a necessidade de particionamento ou setorização do controle de estoque, excesso de validações (muitos cliques para uma única tarefa) e a mentalidade de que o negócio deve se adaptar ao sistema levando à falta de flexibilidade, o que será melhor debatido no tópico seguinte.

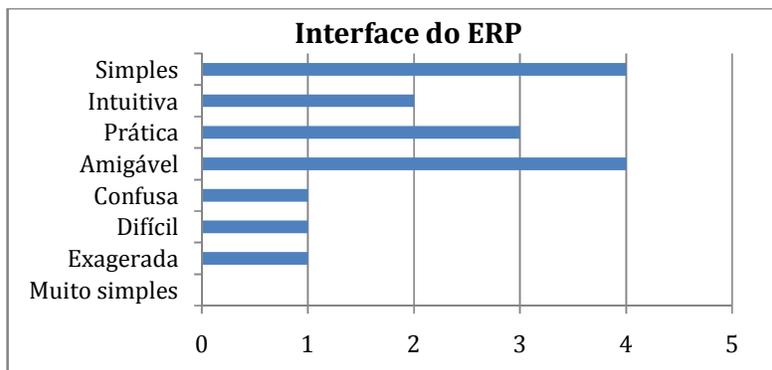
## 5.6 Usabilidade e ergonomia de software

Segundo Mendes e Escrivão Filho (2002), o segmento das pequenas empresas prioriza o preço entre os critérios de seleção de um software. Após esta etapa, as funcionalidades são verificadas, porém de forma pouco aprofundada. Os autores informaram que entre as dez empresas entrevistadas na pesquisa, apenas uma havia realizado um estudo de aderência em vários sistemas e poucas mencionaram realizar a adequação das funcionalidades.

Sendo as funcionalidades pouco estudadas por pequenas empresas, a usabilidade do software e a ergonomia ficam muito mais distanciadas das preocupações daqueles que adquirem o sistema. Quando alguns problemas relacionados à resistência dos funcionários e reclamações quanto às condições de trabalho aparecem, raramente a tecnologia utilizada é revista.

Nesta pesquisa, foram sugeridos alguns adjetivos para que os usuários identificassem características correspondentes à interface do sistema. Os resultados estão distribuídos no gráfico abaixo:

Gráfico 2- Características da interface do ERP.

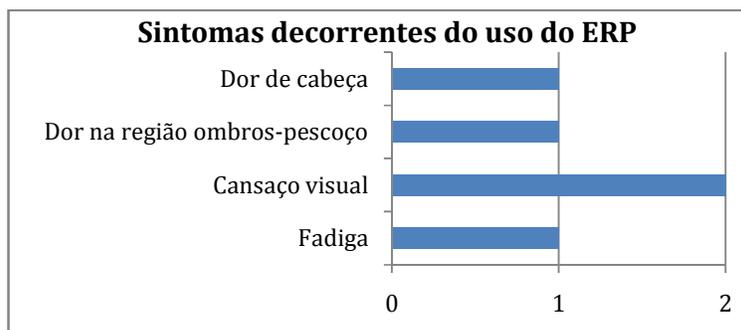


Algumas das empresas entrevistadas ressaltaram a importância da simplicidade e informaram que telas preenchidas de grande quantidade de ícones causam cansaço visual. Características como fonte, tamanho das letras ou cores não foram relatados.

Reunindo estas impressões e somando à insatisfação com o excesso de validações que alguns sistemas pedem (citado o tópico 5.5) observa-se que existem queixas relacionadas à carga de trabalho [BASTIEN e SCAPIN, 1993]. Sistemas ERP possuem grande complexidade pela variedade de atividades que desenvolvem, portanto se a densidade informacional e o número de ações não forem reduzidos, facilmente ocorrerão erros e desconforto para o usuário.

Outro ponto abordado pela pesquisa foi a condição do usuário após o dia de trabalho. Entre os 8 usuários de softwares ERP, 2 relataram sentir cansaço associado ao uso do software após algumas horas de uso e 3 relataram ter tido algum dos sintomas dispostos no gráfico abaixo:

**Gráfico 3- Sintomas decorrentes do uso do software ERP.**



É válido observar que ambientes como ferramentarias possuem ruídos, iluminação, ventilação e outros fatores ambientais característicos e que esses aspectos não foram observados na pesquisa, assim como questões ergonômicas relacionadas ao uso do computador, portanto são necessários novos estudos que relacionem todas essas variáveis.

### **5.7 Consultoria na implantação e utilização**

Certamente, o trabalho de consultoria e suporte ao usuário é essencial para a satisfação das empresas que adotam um sistema ERP. Ainda na fase de adoção, quando a consultoria não está presente e a compra do software é feita com base em motivos isolados ao invés de uma noção clara do negócio, a probabilidade do fracasso na implantação aumenta consideravelmente. Mendes e Escrivão Filho (2002) ressaltam que apesar do grande investimento, muitas “empresas tomam as decisões de forma apressada, impulsionadas por uma onda, sem a realização de estudo mais criterioso”.

Durante o processo de coleta de dados, todos os usuários alegaram ter recebido alguma forma de treinamento para a operação conforme o gráfico abaixo:

**Gráfico 4 - Treinamento para operação.**



Como falado anteriormente no tópico 5.3, as empresas de micro e pequeno porte catarinenses têm procurado contato com outras empresas no mesmo perfil dentro do setor de TI. É perceptível que a visita periódica do consultor transmite segurança ao usuário e antecipa a resolução de problemas. Além disso, os custos de contratação deste tipo de consultoria são menores quando comparados aos produzidos por grandes empresas e quase sempre vem desacompanhado da apreensão pela falta de continuidade

do relacionamento, relembando o caso das empresas que estão há mais de cinco anos com o mesmo consultor-fornecedor.

Os clientes que receberam outros tipos de treinamento tiveram que lidar com frustrações no processo de aprendizagem, que muitas vezes foi por tentativa e erro, relataram morosidade nas atualizações e de forma consequente, perceberam maior resistência dos funcionários.

## **6.Considerações finais**

A pesquisa teve por objetivo apresentar dados sobre micro e pequenas ferramentarias quanto à adoção, implantação e uso de sistemas ERP limitando-se ao estado de Santa Catarina. Os pontos abordados foram os fatores econômicos qualitativamente, as funcionalidades e ergonomia de software e o trabalho de consultoria.

Foi apresentado o cenário entre os entrevistados quanto à adoção do sistema, os benefícios encontrados pelos clientes e também as principais dificuldades. Verificou-se que as principais insatisfações quanto às funcionalidades do software são aquelas que se relacionam à produção: inclusão do plano mestre, interação com o chão de fábrica, facilitação do planejamento.

Após esta etapa, foram levantadas as características de interface de software que mais encantam os clientes e a simplicidade esteve entre as mais citadas e necessárias, pois previne o cansaço visual. Os usuários também relataram suas sensações após um dia de trabalho, mas estes aspectos devem ser estudados de forma mais profunda, pois há possibilidade do ambiente fabril ter influência considerável sobre essas circunstâncias.

A importância da consultoria foi abordada sendo citado o perfil das empresas que demonstraram ter tido mais sucesso na região. Também foi estabelecido um relacionamento entre a ausência do consultor e a frustração no caminho da aprendizagem e adaptação do sistema.

Ao longo deste trabalho, foi visto que o alto custo de aquisição, manutenção, extensão, customização e consultoria de software afetam bastante a micro e pequena ferramentaria, nem sempre tornando atraente a compra, levando ao uso compartilhado com planilhas eletrônicas ou sistemas criados para outro ramo.

Novas pesquisas com maior grupo amostral e que abordem mais profundamente os temas deste trabalho cooperarão para a visualização dos anseios do segmento ferramental e são necessárias para o desenvolvimento local da indústria e serviços de tecnologia da informação.

## **Referências bibliográficas**

Rodrigues, M. ; Assolari, L. M. A. A tecnologia da informação ERP e seus benefícios na gestão de processos e crescimento dos negócios. In: ENCONTRO DA ANPAD, 31., 2007, Rio de Janeiro. **Anais...**Rio de Janeiro: ANPAD, 2007.

Mendes, J. V.; Escrivão Filho, E. Sistemas integrados de gestão ERP em pequenas empresas: um confronto entre o referencial teórico e a prática empresarial. *Gestão & Produção*, v.9, n.3, p.277-296, 2002.

Limas, C. E. A. Sistemas integrados de gestão – ERP: benefícios esperados e problemas encontrados na implantação em pequenas empresas brasileiras. 2009. 100f.

Dissertação (Mestrado em Engenharia de Produção) – Universidade Tecnológica Federal do Paraná, Ponta Grossa, 2009.

Padilha, T. C. C.; Marins, F. A. S. Sistemas ERP: características, custos e tendências. Revista Produção, v. 15, n. 1, p. 102-113, 2005.

Zwicker, C.; Souza, A. Sistemas ERP: conceituação, ciclo de vida e estudos de casos comparados. In: SOUZA, C. A.; SACCOL, A. Z. (organizadores). Sistemas ERP no Brasil (*Enterprise Resource Planning*): Teoria e Casos. São Paulo: Atlas, 2003, 368 p.

Bastien, C.; Scapin, D. Ergonomic Criteria for the Evaluation of Human Computer Interfaces. INRIA, 1993

Portal da Indústria. Perfil da indústria: Santa Catarina. Disponível em: <<http://perfilindustria.portaldaindustria.com.br/estado/sc>>. Acesso em: 19 de set. de 2019.

ENODATA. Principais cidades de Santa Catarina de fabricação de ferramentas (C-2543-8/00). Disponível em: <<https://www.econodata.com.br/consulta-cnae/C2543800-FABRICACAO-DE-FERRAMENTAS/SANTA-CATARINA/cidades-SC>>. Acesso em: 19 de set. de 2019.

Geres, M. A cidade dos moldes. Indústria e Competitividade, Santa Catarina, n. 16, jul. 2018. Disponível em : <[http://fiesc.com.br/sites/default/files/2018-08/IC\\_FIESC\\_16\\_OK\\_DUPLAS.pdf](http://fiesc.com.br/sites/default/files/2018-08/IC_FIESC_16_OK_DUPLAS.pdf)>. Acesso em: 19 de set. de 2019.

SEBRAE. Anuário do trabalho na micro e pequena empresa: 2013. 6. ed. São Paulo, 2013, 288 p. Disponível em <[http://www.sebrae.com.br/Sebrae/Portal%20Sebrae/Anexos/Anuario%20do%20Trabalho%20Na%20Micro%20e%20Pequena%20Empresa\\_2013.pdf](http://www.sebrae.com.br/Sebrae/Portal%20Sebrae/Anexos/Anuario%20do%20Trabalho%20Na%20Micro%20e%20Pequena%20Empresa_2013.pdf)>. Acesso em: 19 de set. de 2019.

FIESC. Santa Catarina em Dados: Metalmeccânica e metalurgia, 2017. Disponível em: <<https://www.observatoriofiesc.com.br/metalmeccanica-metalurgia>>. Acesso em: 19 de set. de 2019.

IBGE. Demografia das empresas e estatísticas de empreendedorismo: 2016. Rio de Janeiro, 2018, 119 p. Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv101612.pdf>>. Acesso em: 19 de set. de 2019.

IBGE. As micro e pequenas empresas comerciais e de serviços no Brasil 2001. Rio de Janeiro, 2003. Disponível em:< [https:// biblioteca.ibge.gov.br/visualizacao/livros/liv1898.pdf](https://biblioteca.ibge.gov.br/visualizacao/livros/liv1898.pdf)>. Acesso em: 19 de set. de 2019.

Mattar, F. N. Pesquisa de Marketing: edição compacta. São Paulo: Atlas, 1996.

Chopra, S.; Meindl, P. Gerenciamento da Cadeia de Suprimentos - Estratégia, Planejamento e Operação. Prentice Hall, 2003

## Apêndice A – Questionário/ Entrevista

- A. Nome da empresa:
- B. Número de funcionários:
  - 1. Você permite o registro e uso dos dados obtidos neste questionário?
  - 2. Você prefere manter sigilo do nome da empresa?
  - 3. Faz uso de software ERP? Por quê?
  - 4. Caso não, como é feito o controle dos recursos? Existe conhecimento do que é esse tipo de software?
  - 5. Caso sim, qual é o nome do software?
  - 6. Caso sim, há quanto tempo?
  - 7. Houve resistência de alguma parte no momento em que o software foi implementado? Houve dificuldade de aprendizagem para o uso?
  - 8. No que diz respeito a custos de implantação e manutenção, quais são suas impressões? A aquisição deste recurso tem correspondido ao investimento?
  - 9. Os recursos que o software proporciona estão adequados a sua atividade? Por quê?
  - 10. Houve necessidade de adaptação/customização? Onde?
  - 11. Sobre a interface do software, clique nos adjetivos que combinam melhor:  
Simples demais, simples, intuitiva, amigável, confusa, difícil, prática, exagerada, outros.
  - 12. O que você sente após algumas horas utilizando o software? Atenção a expressões como: cansaço, dor de cabeça, dificuldades para manter atenção.
  - 13. Existe uma equipe de suporte ou assistência do próprio software que ajude a empresa?
  - 14. Você encontra orientação para as dúvidas com facilidade?
  - 15. O que você melhoraria no software? Tiraria ou acrescentaria?

# DESENVOLVIMENTO DE UMA PRÓTESE DE MÃO ROBOTIZADA

**Gregório Salvador<sup>1</sup>, Gustavo Vaccari<sup>1</sup>, José C. Barbon Jr<sup>1</sup>, Laura Bortoncello<sup>1</sup>,  
William Oliveira<sup>1</sup>, Henrique Zardin de Lima<sup>2</sup>, Maria Emília Ferreira Gonzalez  
Ribeiro<sup>1</sup>, Anselmo Cukla, Bianca Deon Rossato<sup>1</sup>, Alexsander Furtado Carneiro<sup>1</sup>.**

<sup>1</sup>Engenharia Mecânica – Instituto Federal Sul-Rio-Grandense (IFSul)  
CEP – 99064-440 – Passo Fundo – RS – Brasil

<sup>2</sup>Ciência da Computação – Instituto Federal Sul-Rio-Grandense (IFSul)  
CEP – 99064-440 – Passo Fundo – RS – Brasil

{gregoriolussani@hotmail.com, gustavoandre.vaccari@hotmail.com,  
jbarbonjunior@gmail.com, laura.bortoncello@hotmail.com,  
william\_a\_f\_o@hotmail.com, henriquezardindelimal@gmail.com,  
mariafgonzalez505@gmail.com, anselmo.cukla@passofundo.ifsul.edu.br,  
biancadrossato@gmail.com, alexsander.carneiro@passofundo.ifsul.edu.br}

## **Abstract.**

*This article intends to present the mechanical and electrical design of a robotic human hand prosthesis, as well as the computational algorithms to perform the initial motion function tests. The intention is that this prosthesis will be controlled by the muscle contraction of the user's own arm using electromyographic sensors and evolutionary control algorithms. The mechanical system was designed to perform elementary human hand movements and was manufactured by 3D printing, consisted mainly of polymer. This research proposes mechanical solutions to develop an economic prosthesis that can be reached by people of different social layers. It also aims as a result an improvement of the life quality of these individuals who suffered amputations.*

## **Resumo.**

*O presente artigo tem o intuito de apresentar o projeto mecânico e elétrico de uma prótese robótica de mão humana, bem como os algoritmos computacionais para realizar os testes iniciais das funções de movimento. Pretende-se que esta prótese seja controlada pela contração muscular do próprio braço do usuário utilizando-se de sensores eletromiográficos e algoritmos de controle evolutivos. O sistema mecânico foi projetado para realizar movimentos elementares de uma mão humana e foi fabricado por impressão 3D, constituindo-se principalmente de polímero. Esta pesquisa propõe soluções mecânicas para desenvolver uma prótese econômica que possa atender a pessoas de diferentes níveis sociais, também visa como resultado a melhora da qualidade de vida desses indivíduos que sofreram amputações.*

## **1. Introdução**

Nos últimos anos, os avanços da tecnologia permitiram que muitas áreas das engenharias e a medicina tenham proporcionado o aumento da qualidade de vida de pessoas com deficiências físicas, permitindo assim aos pacientes maior acessibilidade a estas tecnologias e a diminuição dos custos envolvidos no desenvolvimento próteses mecânicas que tenham evasão mínima (Vasconcelos, 2017). Baseado em pesquisas como (Bennett et al., 2015; Calderon et al., 2017; Johansen et al., 2016), este trabalho tem como objetivo propor e construir um protótipo funcional de uma prótese biomédica de uma mão robotizada, para que futuramente seja utilizada em pessoas que tenham o punho amputado. O protótipo visa à característica antropomórfica, ou seja, imitar a mão humana em aparência e funcionalidade ao realizar movimentos elementares e, assim, que seja capaz de manusear objetos pequenos (Calderon et al., 2017; Raurale and Chatur, 2014). Esse tipo de tecnologia poderia contribuir com o paciente na melhora da qualidade de vida e no aumento da própria autoestima (Vasconcelos, 2017). O primeiro protótipo experimental visa construir uma prótese de baixo custo e com capacidades de movimentos controladas externamente mediante comandos computadorizados. Cabe ressaltar que nas primeiras etapas não serão realizados ensaios em humanos. A estrutura mecânica do braço é composta essencialmente por polímeros confeccionados em impressão 3D, já o sistema de transmissão de movimentos se dá mediante linhas elásticas de alta resistência. O sistema eletromecânico utiliza servomotores, condicionadores de sinais e uma plataforma Arduino. O sistema computacional comanda os movimentos dos dedos através dos servomotores, que o sistema utiliza uma trajetória preestabelecida no microcontrolador. Neste trabalho, no capítulo 2 é apresentada a metodologia de como deve ser basicamente estruturado o braço protético. No capítulo 3, apresenta-se uma aplicação e resultados iniciais (simulação e experimental) do funcionamento das partes mecânicas, elétricas e computacionais. Finalmente, no capítulo 5 são apresentados os resultados e conclusões. E por último, as bibliográficas utilizadas.

## **2. Arquitetura do braço protético**

O sistema mecânico da prótese foi baseado em modelos anteriormente aplicados a outras pesquisas, tais como Bennett *et al.* (2015), Johansen *et al.* (2016). Já para o desenho 3D do primeiro protótipo, utilizou-se o mesmo que em Gross (2018). Quanto ao sistema elétrico e computacional, estes foram desenvolvidos especificamente para esta finalidade, como apresentado seguidamente.

## 2.1 Sistema mecânico

O primeiro protótipo a ser utilizado no projeto foi baseado no modelo 3D do braço mecânico de Gross (2018). Este modelo apresenta muita similaridade quanto aos movimentos e anatomia de uma mão e punho humana. Na Figura 1, apresenta-se o protótipo 3D.

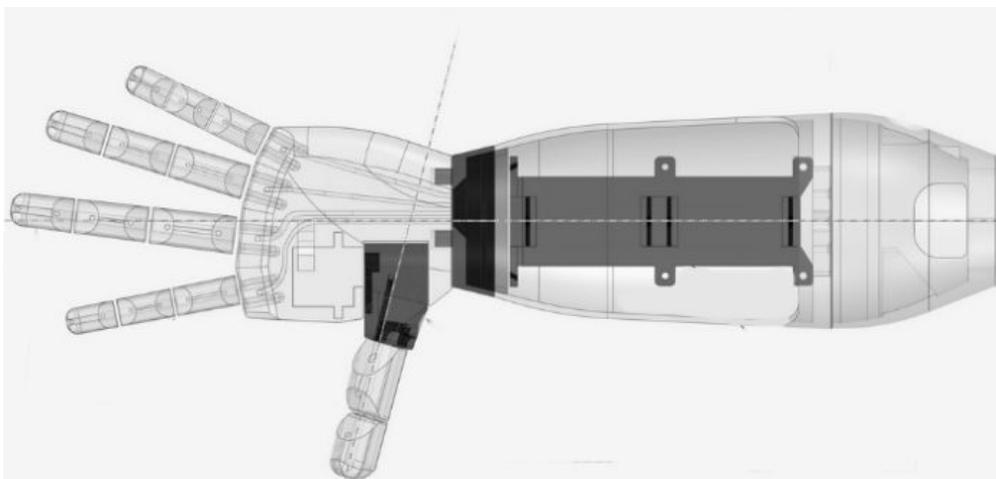


Figura 1 – Modelo 3D do primeiro protótipo da mão robótica (Gross, 2018)

Esta Prótese Robotizada pode ser dividida em duas partes, mão e punho. A mão possui todas as articulações dos dedos, que realizam movimentos de abertura e fechamento (contrações) dos mesmos mediante linhas elásticas que são acionadas mediante servomotores. No caso do “dedão” ou dedo polegar, possui um servomotor embutido para o acionamento de fechamento até a palma da mão. A contração das articulações do polegar também é feita mediante cordão compostos por fibra elásticas.

No braço, estão os 4 motores de acionamento dos dedos da mão, ainda aqui foram alocados os *drivers* de força e o controlador eletrônico Arduino. Para o acionamento dos dedos, existe uma polia que traciona a linha elástica e fecha os dedos. Devido a configuração mecânica dos dedos, estes retornam a sua posição original de estendido (ou mão aberta) quando a linha é solta. Na Figura 2 é apresentado o primeiro protótipo funcional que foi impresso no IFSUL, campus Passo Fundo.

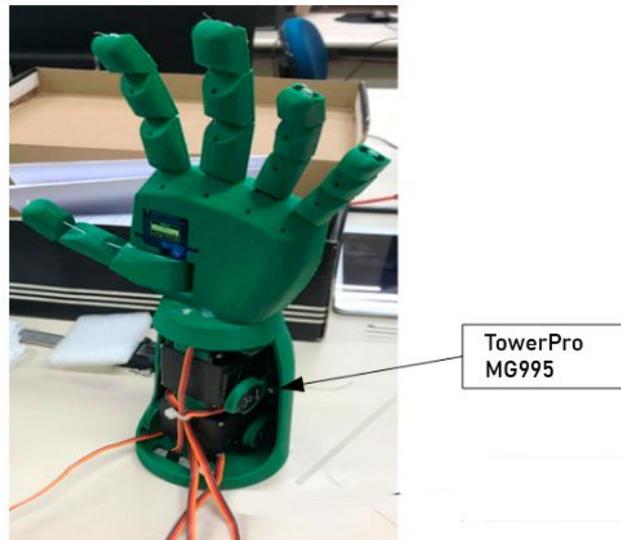


Figura 2 – Protótipo eletromecânico da prótese

## 2.2 Sistema elétrico

O sistema elétrico do primeiro protótipo da prótese está composto basicamente por motores, uma plataforma controladora de prototipagem eletrônica de hardware livre denominada Arduino e o *driver* de potência para o acionamento dos motores. O Sistema de acionamento dos dedos está composto por 4 servos motores do tipo TowerPro MG995, e um motor Micro Servo Motor 9g Tower Pro. A plataforma de processamento de dados e controle é um Arduino UNO R3; já o *driver* de potência e acionamento dos motores é um *Shield* v5.0, foi utilizada uma fonte regulável para a alimentação do sistema a uma tensão de 5Vcc. Neste trabalho, os Servomotores MG995 são nomeados como  $s_1 \cdots s_4$ , sendo eles os encarregados de movimentação dos dedos através de linhas elásticas. Isto difere no caso do dedo polegar que é movimentado com um Micro Servo SG90 ( $s_5$ ), que é acionado diretamente.

Detalhamento das funcionalidades dos motores: o motor  $s_1$  controla o movimento do dedo mínimo e anelar, o  $s_3$  do médio,  $s_4$  do indicador e o  $s_2$  controla o polegar juntamente com o Micro Servo que é nomeado como  $s_5$ . Todos os motores são ligados diretamente ao *driver Shield* v5.0 e controlados via Arduino UNO. Os motores são alimentados com uma tensão de 5 V (que pode ser como máximo 7,2 V) a uma corrente de 0,35 A ( $I_{max} = 2,5 A$ ) cada. Na Figura 3, é apresentado um diagrama esquemático das conexões dos componentes elétricos do sistema em funcionamento no braço. Já a Figura 4 representa os esquema de conexões dos servomotores no Arduino com a shield, sendo os fios vermelhos o de alimentação, os pretos da saída gnu e o amarelo de entrada de informações.

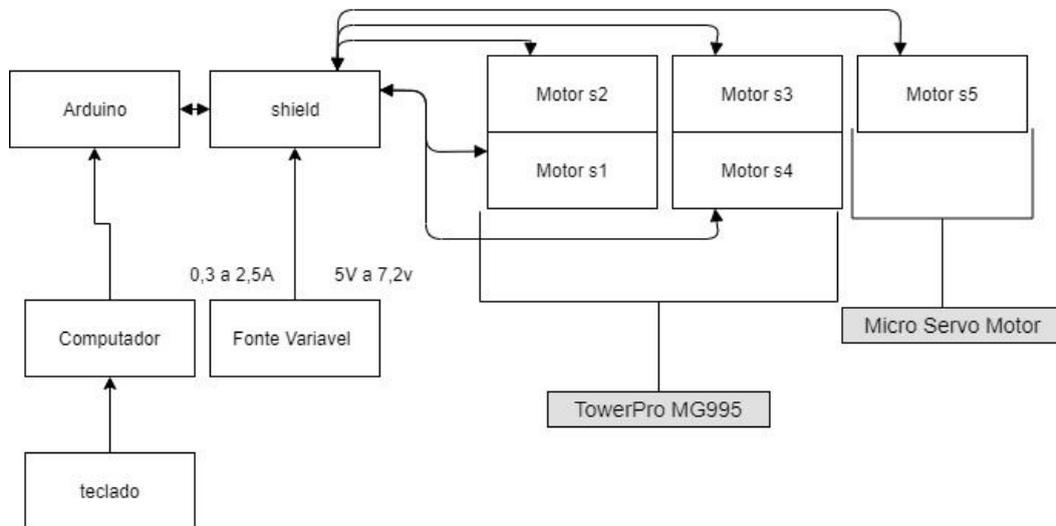


Figura 3 – Diagrama esquemático de conexão elétrica

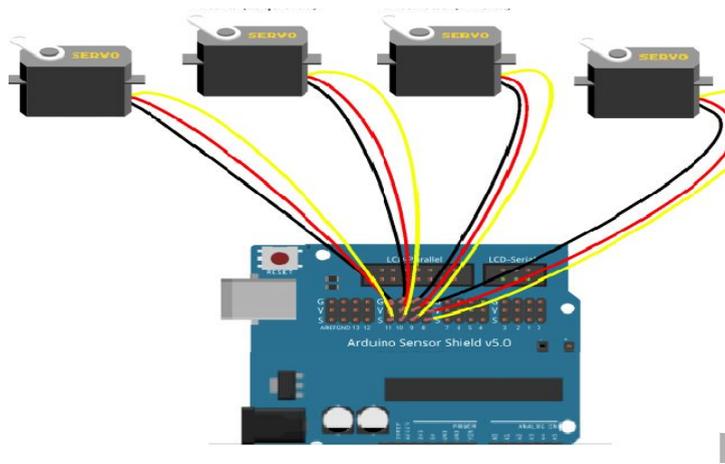


Figura 4 – Esquemático de conexão entre o Arduino e os servomotores

### 2.3 Sistema computacional

O sistema computacional foi projetado para comandar os movimentos fundamentais da prótese, que são executados através dos servomotores, cujos comandos são inicialmente processados pelo Arduino. A linguagem de programação utilizada é a própria do Arduino IDE®, baseado em C/C++. Para controlar os movimentos dos Servomotores ( $s_1 \cdots s_5$ ), empregou-se a biblioteca “servo.h”, como é ilustrado na Figura 4.

```
comandos_certos | Arduino 1.8.10
Arquivo Editar Sketch Ferramentas Ajuda

comandos_certos
#include <Servo.h> //INCLUSÃO DA BIBLIOTECA NECESSÁRIA

const int pinoServo6 = 6;
const int pinoServo8 = 8; //PINO DIGITAL UTILIZADO PELO SERVO
const int pinoServo10 = 10;
```

Figura 5 – Biblioteca Usada

A execução dos movimentos da prótese foi programada para que, inicialmente, os motores sejam ativados e conduzidos à posição zero, que coincide com a posição de “mão aberta”. A partir desta funcionalidade inicial, são realizadas outras funções preestabelecidas da prótese, tais como as funções de fechar, abrir, apontar e chamar. Estas funções são combinações de movimentos de motores em forma conjunta ou individual para a realização do movimento.

Para a execução dos movimentos, o Arduino recebe os comandos de movimento via teclado, utilizando a função “Serial.read()”, que recebe a informação e armazena na variável “tecla” permitindo a execução dos comandos de: abrir, fechar, apontar e chamar. Na Figura 5, é apresentado o fluxograma da lógica do programa principal armazenado no Arduino e, na Figura 6, é apresentado o exemplo do programa com a condição de Abrir a mão .

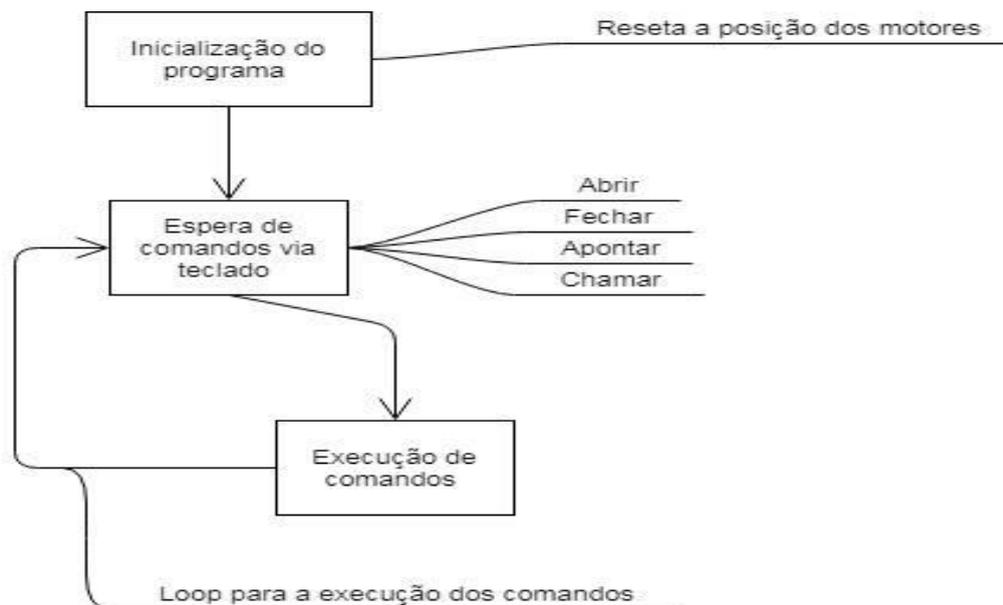


Figura 6 – Fluxograma da lógica do programa principal do braço protético

```
comandos_certos | Arduino 1.8.10
Arquivo Editar Sketch Ferramentas Ajuda

comandos_certos

tecla = Serial.read();

if(tecla == 'a'){
  Serial.println(tecla);
  pos1=0;
  pos2=150;
  medio.write(pos1);
  indicador.write(pos1);
  polegar2.write(pos2);
  anelarminimo.write(pos2);
} else if(tecla=='f'){
  Serial.println(tecla);
  pos1=150;
  pos2=0;
  medio.write(pos1);
  indicador.write(pos1);
  polegar2.write(pos2);
}
```

Figura 7 – Programa com a condição de Abrir a mão

### 2.3 Sinais eletromiográficos

Os sinais que movimentam a nossa mão são sinais elétricos que geram as contrações musculares e permitem fechar, abrir, indicar ou pegar objetos. Estes movimentos são o resultado de comandos gerados no cérebro e propagado ao longo dos nervos periféricos. A grande vantagem para esta aplicação, é que estes sinais podem ser captados mediante específicos, que são conhecidos como sinais EMG (Eletromiográficos) (Schultz and Kuiken, 2011). O EMG pode ser utilizado para comandar o acionamento dos motores dando completo controle do usuário sobre a prótese utilizando unicamente os seus próprios sinais elétricos. Na Figura 8, é apresentado um exemplo das possibilidades de utilização destes tipos de sinais para a reabilitação de pessoas que tenham sofrido a perda de algum membro superior.

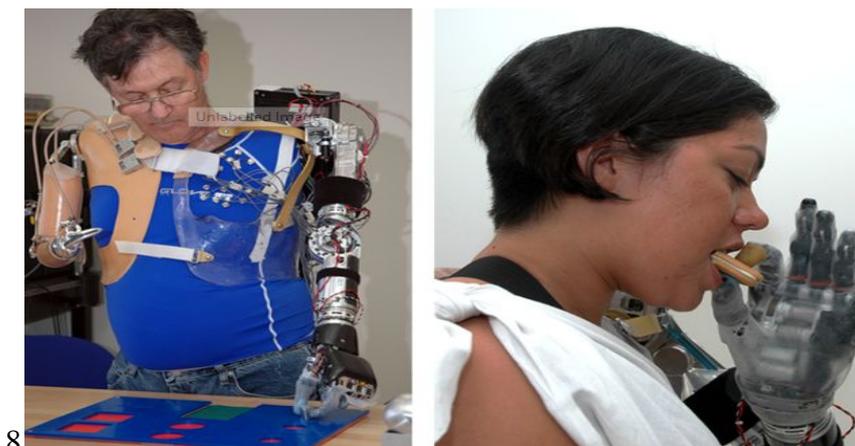


Figura 8 – Pacientes utilizando prótese ativas que são comandadas mediante sinais EMG (Schultz and Kuiken, 2011)

### 3. Resultados iniciais

#### 3.1 Sistema Eletromecânico

Os testes iniciais da prótese funcional, baseada no projeto mecânico de (Gross, 2018), apresentaram grandes similaridades com os movimentos naturais de uma mão humana. Os 4 movimentos iniciais que foram programadas na mão, foram implementadas satisfatoriamente (abrir, fechar, apontar e chamar), embora foi necessário o ajustes de parâmetros do driver dos motores, e ajustes de programação sobre a movimentação dos motores. Na Figura 9a e 9b, apresentam movimentos realizados pela prótese que representam movimentos do dia a dia, tal como pegar uma laranja.



Figura 9 A/B– Exemplo de movimentos da prótese.

A primeira apresentação pública do projeto de pesquisa se deu durante o evento “IFSUL Mostra A Sua Cara”, no dia 25/09/2019. O braço demonstrou as suas funcionalidades para os alunos de diferentes escolas da região, pais, professores e os funcionários do IFSUL. Nessa ocasião, embora inicialmente tenha havido problemas funcionais com um servo-motor que estava desregulado, a aceitação do público foi alta, motivando aos pesquisadores a melhorar ainda mais o projeto.

#### 3.2 Sistema computacional

Nos primeiros testes do algoritmo computacional, foi preciso realizar os ajustes com os motores. Primeiramente, o acionamento individual e, posteriormente, o acionamento de um grupo de motores para realização de movimentos sincronizados. Outros testes foram realizar ajuste do percurso de movimentação do motor, para não forçar o motor ao fechar mecanicamente os dedos. Feita a calibração dos motores e dos movimentos individuais dos dedos, foram programadas as funções de movimentos, que devem ser

comandadas via teclado (tais como, fechar, abrir, etc.). Nesse sistema, não foi utilizado nenhum tipo de sensores para acompanhar os movimentos da prótese, sendo assim, um sistema sem realimentação dos parâmetros de movimentos.

A partir dos resultados iniciais de movimentação da prótese, foi concebida a ideia da criação de uma interface gráfica com acesso via rede sem fio. Tendo como princípio para a criação da interface o uso do .NET para rodar em qualquer dispositivo com acesso a internet, tornando o acesso ao recurso mais fácil. Este conceito ainda está em etapa de implementação.

#### **4. Discussões e Conclusões**

Devido ao fato de que a pesquisa ainda está nas primeiras etapas do plano de trabalho, os resultados alcançados foram satisfatórios. A alta aceitação das pessoas no evento “IFSUL Mostra A Sua Cara”, no dia 25/09/2019, indica que a nossa pesquisa é promissora e de alto impacto social.

Percebeu-se que é necessária que seja implantada em breve a realimentação do sistema de controle. Dita realimentação será feita mediante a leitura dos deslocamentos dos dedos durante a trajetória, permitindo, assim, um controle mais preciso das funções de movimentos. Para isso, é necessário a implementação de sensores de posição.

Sem dúvida, o principal diferencial deste trabalho será a implementação de um sistema computacional sofisticado, capaz de realizar um processamento de alto nível, e que execute algoritmos evolutivos e sejam utilizada inteligência artificial para a aprendizagem dos movimentos de cada usuário.

#### **5. Bibliografia**

Bennett, D.A., Dalley, S.A., Truex, D., Goldfarb, M., 2015. A Multigrasp Hand Prosthesis for Providing Precision and Conformal Grasps. *IEEE/ASME Transactions on Mechatronics* 20, 1697–1704. <https://doi.org/10.1109/TMECH.2014.2349855>

Calderon, C.A., Ramirez, C., Barros, V., Punin, G., 2017. Design and Deployment of Grasp Control System applied to robotic hand prosthesis. *IEEE Latin America Transactions* 15, 181–188. <https://doi.org/10.1109/TLA.2017.7854610>

Gross, R., 2018. 3D Printable Humanoid Robotic Hand [WWW Document]. My MiniFactory. [URL https://www.myminifactory.com/object/3d-print-humanoid-robotic-hand-34508](https://www.myminifactory.com/object/3d-print-humanoid-robotic-hand-34508) (accessed 10.4.19).

Johansen, D., Cipriani, C., Popović, D.B., Struijk, L.N.S.A., 2016. Control of a Robotic Hand Using a Tongue Control System—A Prosthesis Application. *IEEE Transactions on Biomedical Engineering* 63, 1368–1376. <https://doi.org/10.1109/TBME.2016.2517742>

Raurale, S.A., Chatur, P.N., 2014. Identification of real-time active hand movements EMG signals for control of prosthesis robotic hand, in: 2014 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC). Presented at the 2014 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC), pp. 482–487. <https://doi.org/10.1109/ICCPEIC.2014.6915412>

Schultz, A.E., Kuiken, T.A., 2011. Neural Interfaces for Control of Upper Limb Prostheses: The State of the Art and Future Possibilities. *PM&R* 3, 55–67. <https://doi.org/10.1016/j.pmrj.2010.06.016>

Vasconcelos, T.B. de, 2017. Avaliação da qualidade de vida de pacientes amputados transtibiais unilaterais antes e após a protetização. *Fisioterapia Brasil* 12, 291–297. <https://doi.org/10.33233/fb.v12i4.927>

# Uso de Sistema de Recomendação baseado em personas para potencializar a aprendizagem dos alunos de Algoritmos

Gabriel Albrecht Bonato<sup>1</sup>, João Mário Lopes Brezolin<sup>1</sup>,  
Carmen Vera Scorsatto Brezolin<sup>1</sup>

<sup>1</sup>Instituto Federal de Educação, Ciência e Tecnologia Sul-rio-grandense (IFSUL)  
99064-440 – Passo Fundo – RS – Brazil

`gabriel.bonato@outlook.com,`

`{joao.brezolin, carmen.scorsatto}@passofundo.ifsul.edu.br`

**Abstract.** *The Algorithms, in Computing courses, usually is the first contact that the student has with the concepts of logic and programming languages. Due to the complexity of its contents, the students face various difficulties that often can lead them to absenteeism. Thus, considering the individual learning process of each student, it was observed that the implementation of a personalized recommendation system could help on their learning process. Also, presents the validation of the recommendation functionality.*

**Resumo.** *A disciplina de Algoritmos, nos cursos de Computação, em geral, é o primeiro contato que o aluno tem com os conceitos de lógica e linguagens de programação. Devido a complexidade dos seus conteúdos, os alunos deparam-se com diversas dificuldades que muitas vezes podem levá-los a evasão do curso. Assim, considerando o processo individual de aprendizado de cada aluno, observou-se que a implementação de um sistema de recomendação personalizado poderia auxiliar no processo de aprendizagem dos alunos. Além disso, apresenta a validação da funcionalidade de recomendação.*

## 1. Introdução

O problema da evasão discente afeta todos os níveis educacionais. Segundo Filho et al. (2007), a evasão de Cursos Superiores presente não é uma exclusividade brasileira, pelo contrário, é um problema que vem sendo detectado em diversos países. Segundo o autor, “as perdas de estudantes que iniciam mas não terminam seus cursos são desperdícios sociais, acadêmicos e econômicos”. Desse modo, observa-se que os investimentos nas Instituições de Ensino Superior não trazem os resultados esperados.

Bardagi e Hutz (2009), definem a evasão “a saída do aluno do curso de origem, sem concluí-lo”. Essa desistência, pode ser causada por diversos motivos: descontentamento com horários, falta de cursos noturnos, impossibilidade de trabalhar e estudar, mau relacionamento entre professor e aluno, pouca integração social à universidade, expectativas não correspondidas e falta de informações sobre curso e profissão, mau desempenho, reprovações e atraso, e problemas financeiros.

Vale destacar, que a evasão ocorre para todas as áreas de conhecimento, porém os índices mais altos encontram-se na área das Ciências Exatas [Morães and Pombeiro 2016], De acordo com Rodrigues (2013), os alunos chegam à

graduação com pouca base matemática, devido às deficiências existentes nos níveis fundamental e médio, o que diminui o interesse dos alunos pela área, causando-lhes frustrações por não conseguirem realizar seus sonhos, podendo até mesmo causar impactos irreversíveis a eles.

Muitas vezes a evasão discente dos cursos de Computação ocorre devido as dificuldades com o raciocínio lógico requerido nas disciplinas de programação [Alves and Jaques 2014], em especial as de Algoritmos. Por serem disciplinas introdutórias nesses cursos, surge a necessidade de encontrar soluções que venham a melhorar o processo de aprendizagem dos alunos nelas [Gossman 2017]. Nesse sentido o presente trabalho apresenta a proposta de um sistema personalizado que leva em consideração o perfil do aluno no momento para fazer a recomendação de materiais de aprendizagem.

## **2. Referencial teórico**

O estudo do processo de aprendizagem permeia o trabalho de diversos autores. Segundo Vygotsky (1998), o aprendizado é um elemento propulsor do desenvolvimento do sujeito que o capacita não somente para a realização das operações solicitadas para o aprendizado, mas o instrumentaliza para utilizar esses conhecimentos nas mais variadas situações. Para explicar as dimensões do desenvolvimento Vygotsky desenvolveu o conceito de zona de desenvolvimento proximal (ZDP). Na concepção dele, a ZDP consiste na “região” que representa a distância entre aquilo que o sujeito sabe e aquilo que ele pode vir a aprender.

Para haver aprendizagem, segundo Ausubel (2003), o conhecimento deve fazer sentido para o sujeito; o aprendiz deve correlacionar o novo conhecimento com os conceitos preexistentes em sua estrutura cognitiva. O forte embasamento conceitual por parte do discente poderá garantir um processo de aprendizagem consistente. Visando criar “âncoras” que transformem a aprendizagem mecânica em significativa, Ausubel aponta para o uso de organizadores prévios, uma estratégia para direcionar a estrutura cognitiva para a aprendizagem significativa. Os organizadores são materiais de aprendizagem que preparam o aluno para a aquisição de novos conhecimentos .

Como forma de se definir as diferenças entre o modo de aprendizado de cada indivíduo, também pode-se citar a teoria de Estilos de Aprendizagem (EA). De acordo com Ferreira (2018), “diversos estudos nas áreas de psicologia cognitiva e pedagogia apontam que cada indivíduo possui diferentes maneiras de captar, processar, analisar e organizar informações durante o processo de aprendizado, o que fundamenta o conceito de Estilos de Aprendizagem”. Do mesmo modo, um EA classifica a forma como os alunos recebem e processam as informações em escalas relativas [Felder et al. 1988].

Com base nesse contexto, o presente trabalho tem como objetivo desenvolver um sistema de recomendação personalizado de materiais de aprendizagem que auxilie alunos a alcançarem uma aprendizagem significativa e assim obterem êxito nas disciplinas de Algoritmos. Um sistema de recomendação (SR) busca prever o interesse do usuário por produtos ou serviços, com base nas suas preferências. Para tanto, o sistema utiliza recursos como a avaliação de conteúdos e o monitoramento das ações do usuário durante a navegação para prever que itens podem ser sugeridos a ele [Ricci et al. 2010].

Segundo Reategui e Cazella (2005):

“Para que se possa recomendar itens a um usuário, é necessário se ter conhecimento sobre quem é este usuário. É necessário capturar e armazenar seus dados pessoais e comportamentais relativos. Portanto, é essencial que se possa identificar o usuário no momento em que ele acessa o sistema onde foram implantadas as rotinas de recomendação”.

Esses sistemas usam técnicas de *Machine Learning* para que seja possível prever as preferências dos usuários. As ações e as informações do usuário permitem ao sistema estabelecer o processo de inferência lógica. De acordo com Ferreira (2018), esses métodos podem ser classificados em indução, quando as premissas não necessariamente são verdadeiras, mas que têm como resultado uma hipótese; em dedução, parte de premissas das quais já se há um conhecimento definido, e a partir de cálculos em cima disso, tem-se como consequência um resultado verdadeiro; e em abdução, aplicadas no sentido de analisar informações sobre as quais há um conhecimento incompleto, afim de completá-lo por meio de hipóteses baseadas nas informações já conhecidas.

Em aplicações voltadas a área educacional os sistemas de recomendação buscam potencializar o aprendizado dos alunos. Segundo Rezende (2018): “os sistemas de recomendação (SR) extraem e utilizam informações sobre o contexto e perfil dos alunos, dentre elas suas preferências, objetivos, estilos de aprendizagem e dados dos dispositivos tecnológicos utilizados”.

Existem diversas pesquisas envolvendo o uso de sistemas de recomendação. Ferro et al (2011) desenvolveu um sistema no ambiente virtual Moodle, com as funcionalidades de Perfil do Usuário, em que os usuários alunos deveriam indicar as áreas de sua preferência; Perfil do Material Didático, em que os usuários professores ou administradores deveriam indicar as áreas que compõem o Perfil do Material Didático ao submeter um material didático; Biblioteca, onde os alunos ao entrarem e visualizarem algum conteúdo, teriam a área do material didático adicionada ao seu perfil, caso ainda não fizessem parte dele, ou passariam a ter maior importância, caso já fizessem parte dele; Módulo visualizador de recomendações, em que as recomendações seriam exibidas no início da tela, de modo que não fosse necessário muitos cliques para que os usuários conseguissem visualizá-la; e ainda a Avaliação de Recomendação, em que o usuário, ao clicar sobre a recomendação, poderá avaliá-la, e após a avaliação poderá fazer o download do material didático.

Além disso, Vieira e Nunes (2012) desenvolveram o DICA, que se trata de um sistema de recomendação de objetos de aprendizagem baseado em conteúdo. Inicialmente, através do uso da técnica de Filtragem Baseada em Conteúdo, ele identifica a proximidade entre os temas dos Objetos de Aprendizagem, selecionando 20 itens com temas mais próximos. Após isso, ele verifica a quantidade de vezes que um termo ocorre no documento e, em seguida, identifica a relevância de um termo em um conjunto de documentos e, dessa forma, é realizada a comparação para apresentar quais Objetos possuem o maior número de termos em comum, dentre os 20 selecionados anteriormente.

Ainda, Reis e Barrére (2014), propuseram uma arquitetura em que conteúdos previamente categorizados com relação ao nível de conhecimento, pelo professor, sejam recomendados aos alunos, possibilitando a eles, inclusive, indicar conteúdos. Conforme o sistema for utilizado, informações de perfil, contexto e qualidade de experiência serão utilizados para que seja realizada a composição de uma recomendação adaptada. Para

validação dessa arquitetura, eles desenvolveram um protótipo denominado *Context Aware Learning*.

Conforme Reategui e Cazella (2005), há três técnicas de filtragem de informações que podem ser aplicadas aos SRs: Filtragem Baseada em Conteúdo, a qual se baseia no conteúdo dos itens do sistema e também nas informações do perfil do usuário; Filtragem Colaborativa, não exige análise do conteúdo dos itens, pois a filtragem é baseada na troca de experiências entre usuários; e Filtragem Híbrida, é a técnica em que se une vantagens das duas técnicas de filtragem citadas anteriormente.

O sistema de recomendação proposto, utiliza a técnica de filtragem híbrida, pois identifica similaridades entre o perfil de cada usuário, a qual pode ser definida em filtragem baseada em conteúdo, e analisará avaliações dos usuários, que é uma filtragem colaborativa. Além disso, para possibilitar o ensino personalizado através dessas ferramentas, os EAs são essenciais, de modo a permitirem a modelagem do perfil de cada aluno [Oliveira et al. 2018].

Os usuários são uma fonte importante de conhecimento para o desenvolvimento de sistemas. Eles podem contribuir de diferentes modos nesse processo, seja atuando diretamente no processo de Engenharia de Software, através de sugestões de modificações, ou ainda indiretamente, por meio de avaliações do sistema [Almaliki et al. 2015]. Tendo em vista que cada pessoa possui uma personalidade e que é necessário um grande número de pessoas envolvidas para a especificação dos sistemas, o custo torna-se elevado. Assim, as Personas surgem como uma forma de identificar as preferências dos usuários, mediante a criação de pessoas e personalidades fictícias [Shahri et al. 2016].

Conforme Almaliki, Ncube e Ali (2015), o uso dessa ferramenta torna mais simples o processo de design do sistema, pois fornece uma forma mais eficaz de comunicação entre os profissionais que trabalharão no processo da engenharia do sistema, limita a moldagem das necessidades e desejos dos usuários pela conveniência e minimiza projetos autorreferenciais dos profissionais que trabalharão no processo, ajuda os engenheiros a limitarem-se a um conjunto limitado de personalidades, auxilia na revisão das necessidades para fins de validação e, também, serve como fonte de inspiração durante todo o processo. Para tanto, é necessário segmentar os usuários de acordo com os padrões encontrados entre eles, sendo que para cada um dos segmentos, deve ser elaborada uma persona, dando-lhe nome, idade, gênero, imagem e cenários, o que possibilita uma maior facilidade a entender as necessidades dos usuários que o sistema deverá suprimir.

O conceito de personas foi utilizado no sistema para permitir ao sistema criar recomendações mais personalizadas para os alunos. Além de permitir a busca de materiais de aprendizagem baseado nas avaliações gerais dos usuários, o sistema realizará a recomendação com base nas informações e preferências deles, ou seja, será desenvolvido de modo que se adéque aos perfis de usuário dos alunos do IFSUL, câmpus Passo Fundo.

### **3. Materiais e Métodos**

De modo a persistir as informações dos alunos e os dados sobre suas preferências foi necessário utilizar o conceito de Big Data. O volume e o formato divergente dos dados armazenados dificultam o processamento realizado em sistemas relacionais tradicionais [IBM et al. 2011]. Nesse sentido, os bancos de dados NoSQL perfazem uma alternativa,

pois estabelecem uma abordagem baseada em disponibilidade, desempenho e escalabilidade dos dados [Toth 2011]. O MongoDB é um exemplo de Sistema de Gerenciamento de Banco de Dados (SGBD) NoSQL cujo modelo e estratégias de persistências têm enfoque na alta taxa de transferência de dados, seja para leitura como gravação [Banker 2011].

O modelo desse banco de dados é orientado a documentos que permitem representar propriedades diversas de um mesmo conjunto de dados [Banker 2011]. A abordagem orientada a documentos desses bancos permite a representação de estruturas hierárquicas complexas. O MongoDB foi utilizado porque os dados das avaliações dos usuários tendem a possuir formatos bastante divergentes uns dos outros, o que tornaria o processamento lento se utilizado banco de dados convencional.

Para possibilitar o desenvolvimento de uma *Graphic User Interface* (GUI), e permitir que ela seja acessada pelos mais variados dispositivos que tenham acesso a Internet, usou-se a tecnologia de Progressive Web Apps (PWA). As PWAs também oferecem alguns benefícios como responsividade, são interativas como um aplicativo, estão sempre atualizadas com a última versão disponível, podem ser encontradas através de mecanismos de busca, podem ser instaladas no dispositivo e ainda podem ser acessadas através de um endereço na internet [Hume 2017].

O Ionic Framework fornece grande número de componentes desenvolvidos através das tecnologias HTML, CSS e JavaScript, para o desenvolvimento de uma poderosa interface de usuário. Além disso, também conta com um sistema de linha de comando que auxilia o desenvolvedor em tarefas como criação de projeto e construção do software [Wilken and Bradley 2016]. O Ionic foi utilizado para desenvolver a interface responsiva do sistema, oferecendo possibilidade de acesso, por se tratar de uma PWA, em vários dispositivos.

Para haver comunicação entre a GUI e o sistema provedor de recomendações, foi necessário a utilização de Serviços WEB. Esses serviços permitem a interação entre duas máquinas em uma rede, conforme a W3C [Moraes et al. 2013]. No caso do sistema proposto, foi utilizado o estilo arquitetural Representation State Transfer (REST), o qual se trata de um conjunto de normas e padrões a serem seguidos na composição de um Serviço WEB [Fernandes Ribeiro and Elias Francisco 2016]. O Framework Spring Boot foi utilizado no desenvolvimento, para que permita, por meio das configurações automáticas de funcionalidades e gerenciamento, a simplificação da criação da interface de API REST do sistema.

O Apache Mahout é um framework de Machine Learning que tem enfoque em três áreas de Machine Learning, são elas mecanismo de recomendação, clustering e classificação [Anil et al. 2010]. Em alguns cenários a quantidade de dados a serem processados é tão grande a ponto de tornar impossível o processamento em uma única máquina. Através da utilização do Apache Hadoop, o Mahout possibilita a escalabilidade do processamento a um conjunto de máquinas [Anil et al. 2010]. O Mahout foi utilizado para realizar a implementação do sistema de recomendação proposto.

A YouTube API é uma ferramenta que permite pesquisa por vídeos através da informação de diversos parâmetros suportados por ela, que permitem a filtragem por esses critérios [Qin et al. 2010]. Essa ferramenta foi utilizada no sistema proposto para a realização de pesquisas por materiais de aprendizagem disponíveis na plataforma de

compartilhamento de vídeos YouTube, os quais são classificados e exibidos ao usuário. A interação de todas essas ferramentas será detalhada na próxima seção.

A disponibilização do sistema será feita juntamente aos demais materiais da disciplina de Algoritmos I, no ambiente Moodle.

#### 4. Funcionamento do Sistema

O sistema perfaz uma arquitetura Cliente-Servidor, na qual o Cliente (front-end), é responsável por realizar as requisições de dados ao Servidor (back-end). O servidor, por sua vez, faz o processamento da requisição e retorna uma resposta ao Cliente com os dados solicitados. Esse processo é descrito esquematicamente na Figura 1:

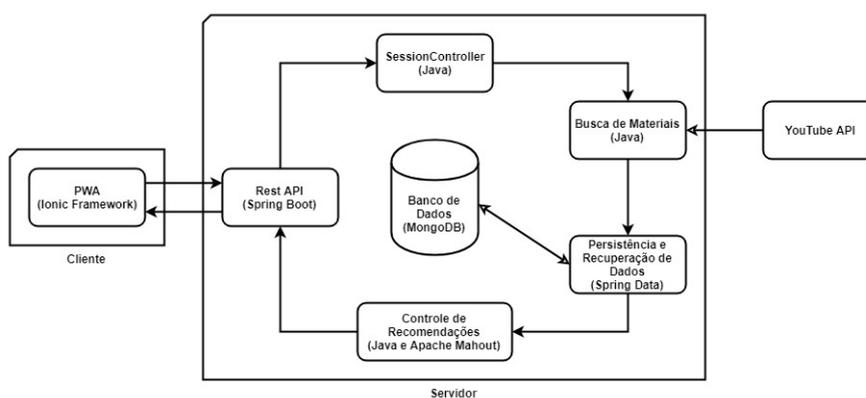
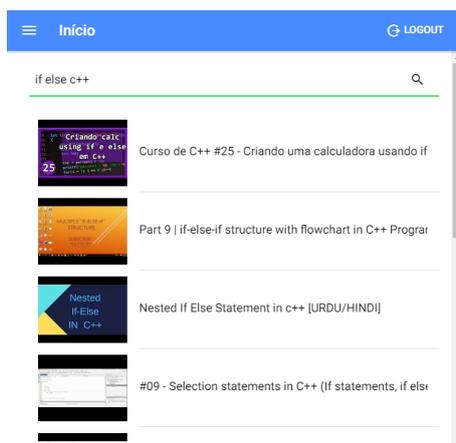


Figura 1. Arquitetura do Sistema

Essa organização permite que as requisições de usuários diferentes sejam processadas pelo servidor. O módulo cliente é uma PWA desenvolvida através do framework Ionic.

Como o sistema é composto por uma arquitetura Cliente-Servidor, é utilizado o modelo REST para o Serviço WEB, que utiliza dos métodos HTTP para transmissão de pacotes de dados. Assim, foi implementado no Servidor uma interface REST que recebe as requisições do Cliente. Desse modo, quando o usuário pesquisar por Materiais de Aprendizagem no Cliente, será realizada uma requisição ao Servidor, na qual serão enviados os termos da busca. Ao receber essa requisição, ocorrerá um controle de sessão, para que o sistema consiga verificar se o usuário está logado ou não. Essa informação será utilizada durante a recomendação dos materiais de aprendizagem.

Na primeira interface apresentada ao usuário é exibida uma tela de pesquisa e busca por materiais. O usuário tem a opção de realizar ou não o login. Caso o usuário não se autentique no sistema, o cálculo das recomendações será com base nas avaliações gerais dos usuários. Assim, o sistema realiza uma busca de Materiais de Aprendizagem através da tecnologia YouTube API com os termos enviados pelo usuário. Então, é realizada uma busca na base de dados NoSQL para recuperar as avaliações dos Materiais de Aprendizagem retornados pela YouTube API. Caso contrário, o sistema irá utilizar as informações da persona do usuário para realizar a recomendação de materiais. Ao realizar uma pesquisa por materiais, será exibida uma lista de itens recomendados, conforme a Figura 2:

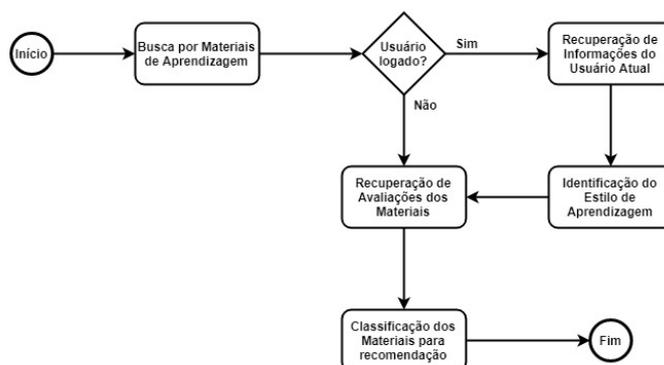


**Figura 2. Tela de Listagem dos Materiais de Aprendizagem Recomendados**

Para usuários logados, os Materiais de Aprendizagem retornados e suas respectivas avaliações, e as informações de perfil do usuário são encaminhados a um Controle de Recomendações. Com o auxílio do framework Apache Mahout, ele é responsável por filtrar as avaliações de usuários que possuam o perfil semelhante ao atual e também por avaliações anteriores similares, o que auxilia na identificação de um EA. Após a elaboração do conjunto de dados, é realizada a classificação dos Materiais de Aprendizagem para a realização da recomendação.

Para os usuários não autenticados, os materiais de aprendizagem são retornados com suas respectivas avaliações, e então é realizada a classificação dos materiais com maior número de avaliações positivas. Ao selecionar um Material de Aprendizagem, o usuário será direcionado a uma página para a visualização desse conteúdo, na qual ele poderá visualizar o material de aprendizagem, como também, avaliá-lo.

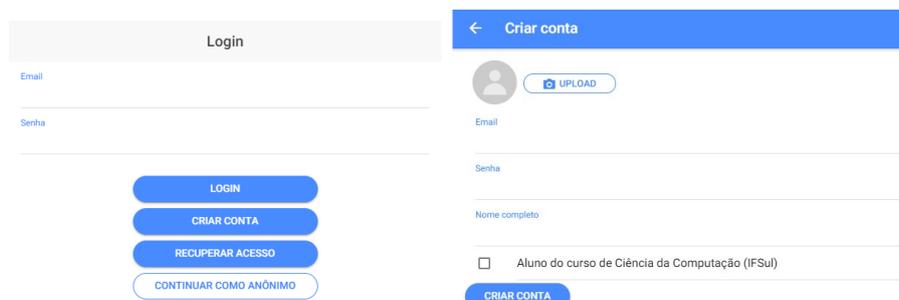
Por fim, os Materiais de Aprendizagem classificados pelo Controle de Recomendações, são retornados a interface de API REST, que é responsável por enviar a resposta ao Cliente, o qual irá exibir a listagem de recomendações ao usuário. Na Figura 3, pode-se visualizar o fluxo do Controle de Recomendações.



**Figura 3. Diagrama de Fluxos do Controle de Recomendações**

Para o desenvolvimento do sistema, foi utilizado o conceito de Personas, de modo a simplificar a engenharia de requisitos e possibilitar que as recomendações sejam reali-

zadas para vários segmentos de usuários. As personas utilizadas para o desenvolvimento são: a persona matriculada no curso de Ciência da Computação, do IFSUL; as personas para cada nível de conhecimento; e as personas para cada Estilo de Aprendizagem. E para que o sistema realize recomendações específicas para os usuários de cada segmento, é necessário que ele esteja logado. Na Figura 4, a exibição da tela de login e de criação de conta, respectivamente.



**Figura 4. Tela de Login e de Criação de Conta**

Assim, os alunos terão recomendações específicas se estiverem matriculados na disciplina, informação que pode ser definida no momento da criação de um usuário ou na alteração dos dados do usuário. E, além disso, as informações de nível de conhecimento e estilo de aprendizagem serão identificadas pelo sistema através das avaliações dos materiais realizadas pelos alunos. Quando um aluno possui várias avaliações em comum com outro, existe grande probabilidade de eles possuírem nível de conhecimento e estilo de aprendizagem aproximados. Dessa forma, quanto maior o número de materiais avaliados em comum entre dois alunos, maior a relevância dos materiais avaliados por um deles na recomendação de materiais para o outro.

## 5. Conclusões

Pode-se perceber que a evasão discente é um problema recorrente e que há diversas razões para a ocorrência dela, sendo que para os cursos de Computação um dos principais motivos são as dificuldades enfrentadas pelos alunos com relação aos conteúdos das disciplinas de Algoritmos. Dessa forma, verifica-se que o trabalho proposto visa minimizar as dificuldades enfrentadas por eles, desenvolvendo um sistema para a recomendação de materiais de aprendizagem baseado nas personas a fim de potencializar a aprendizagem significativa dos discentes.

A utilização do sistema pelos alunos da disciplina de Algoritmos I, do curso de Ciência da Computação do IFSUL, campus Passo Fundo, possibilitou a validação preliminar das recomendações, na qual foi possível verificar a realização das recomendações dos materiais de aprendizagem, as quais se dão através da identificação dos utentes segmentados em cada persona.

Em trabalhos futuros será realizada a validação do sistema de recomendação, por meio da métrica de avaliação *Mean Average Precision* (MAP), de modo a verificar que as

recomendações estão sendo realizadas corretamente.

## Referências

- Almaliki, M., Ncube, C., and Ali, R. (2015). Adaptive software-based feedback acquisition: A persona-based design. In *2015 IEEE 9th International Conference on Research Challenges in Information Science (RCIS)*, pages 100–111.
- Alves, F. P. and Jaques, P. (2014). Um ambiente virtual com feedback personalizado para apoio a disciplinas de programação. In *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, volume 3, page 51.
- Anil, R., Owen, S., Dunning, T., and Friedman, E. (2010). *Mahout in Action*. Manning Publications Co. Sound View Ct. 3B Greenwich, CT 06830.
- Banker, K. (2011). *MongoDB in Action*. Manning Publications Co., Greenwich, CT, USA.
- Felder, R. M., Silverman, L. K., et al. (1988). Learning and teaching styles in engineering education. *Engineering education*, 78(7):674–681.
- Fernandes Ribeiro, M. and Elias Francisco, R. (2016). Web Services REST Conceitos, análise e implementação. 2525-3859.
- Gossman, T. (2017). A utilização de algoritmos nos processos de ensino e de aprendizagem de programação de computadores em cursos técnicos.
- Hume, D. A. (2017). *Progressive Web Apps*. Manning Publications.
- IBM, Zikopoulos, P., and Eaton, C. (2011). *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. McGraw-Hill Osborne Media, 1st edition.
- Moraes, J., Breda, M., Gil, P., and Medaglia, R. (2013). Web Services. *Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Rio Grande do Sul, Brasil*.
- Morães, M. J. F. and Pombeiro, O. J. (2016). Evasão nos cursos de graduação em computação de curitiba. *Anais do EVINCI-UniBrasil*, 1(4):2088–2103.
- Oliveira, H. L., Vivas, A., Assis, L., Pitangui, C., de Almeida, A. M. F., and Dorça, F. (2018). Detecção automática e dinâmica de estilos de aprendizagem para sistemas adaptativos inteligentes para a educação utilizando a meta-heurística vitis vinífera. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 29, page 1323.
- Qin, S., Menezes, R., and Silaghi, M. (2010). A recommender system for youtube based on its network of reviewers. In *2010 IEEE Second International Conference on Social Computing*, pages 323–328.
- Ricci, F., Rokach, L., and Shapira, B. (2010). *Recommender Systems Handbook*, volume 1-35, pages 1–35.
- Shahri, A., Hosseini, M., Almaliki, M., Phalp, K., Taylor, J., and Ali, R. (2016). Engineering software-based motivation: A persona-based approach. In *2016 IEEE Tenth International Conference on Research Challenges in Information Science (RCIS)*, pages 1–12.

Toth, R. M. (2011). *Abordagem NoSQL—uma real alternativa*. Sorocaba, São Paulo, Brasil: Abril, 13.

Wilken, J. and Bradley, A. (2016). *Ionic in action: Hybrid mobile apps with Ionic and AngularJS*. Manning Publications.

# Um estudo sobre a Classificação de Risco em Pronto Atendimento utilizando o método KNN

Antonio A. G. Silveira<sup>1</sup>, Jefferson O. Andrade<sup>1</sup>

<sup>1</sup>Instituto Federal do Espírito Santo (IFES)

antonioadolfo@gmail.com, joandrade@ifes.edu.br

**Abstract.** *This paper presents a study using data mining in the environment of a Health Care Unit. The data of the risk classifications present in this study belong to patients who were attended at the UPA Serra Sede, in SERRA (ES), between 2014 and 2019. The KNN (K Nearest Neighbor) algorithm was used to perform a new risk classification through some measured patient vital signs. As a result of this work, it has been shown that nurses responsible for risk ratings have a certain classification pattern for patients.*

**Resumo.** *O presente artigo apresenta um estudo utilizando a mineração de dados no ambiente de uma Unidade de Saúde de Pronto Atendimento. Os dados das classificações de risco presentes neste trabalho pertencem a pacientes que foram atendidos na UPA de Serra Sede, em SERRA (ES), no período entre o ano de 2014 e 2019. Foi utilizado o algoritmo KNN (K Nearest Neighbor) para realizar uma nova classificação de risco através de alguns sinais vitais medidos dos pacientes. Como resultado desse trabalho, foi mostrado que os enfermeiros responsáveis pelas classificações de risco possuem um determinado padrão de classificação para os pacientes.*

## 1. Introdução

As Unidades de Pronto Atendimento (UPA) funcionam 24 horas por dia e podem resolver grande parte das urgências e emergências, servindo assim como uma importante peça na redução das filas de atendimentos nos prontos-socorros de hospitais. A avaliação da complexidade dos casos de cada paciente que busca atendimento nessas unidades é feita através de um modelo de Classificação de Risco.

A classificação de risco é responsável por ordenar o atendimento nos serviços de pronto atendimento, dividindo os pacientes em grupos e tem como principal objetivo reduzir os riscos aos pacientes, uma vez que determinadas informações dos pacientes são levadas em consideração e a fila deixa de obedecer somente a ordem de chegada.

Segundo o Grupo Brasileiro de Classificação de Risco [GBCR14], atualmente os modelos de classificação mais utilizados no mundo são: Modelo Australiano - Australian Triage Scale (ATS), Modelo Canadense - Canadian Triage and Scale (CTAS), Modelo Americano - Emergency Severity Index (ESI), Modelo de Manchester - Manchester Triage System (MTS).

Este último, o Modelo de Manchester, é um dos mais utilizados no Brasil [GBCR14], sendo utilizado pela UPA de Serra Sede, que é o objeto de estudo deste projeto e fica localizada no município da Serra-ES.

## 2. Classificação de Risco

A utilização da classificação de risco nas unidades de pronto atendimento é estabelecida pela portaria 2048/2002 do Ministério da Saúde que cita a organização do processo de trabalho quanto a classificação de risco da seguinte forma:

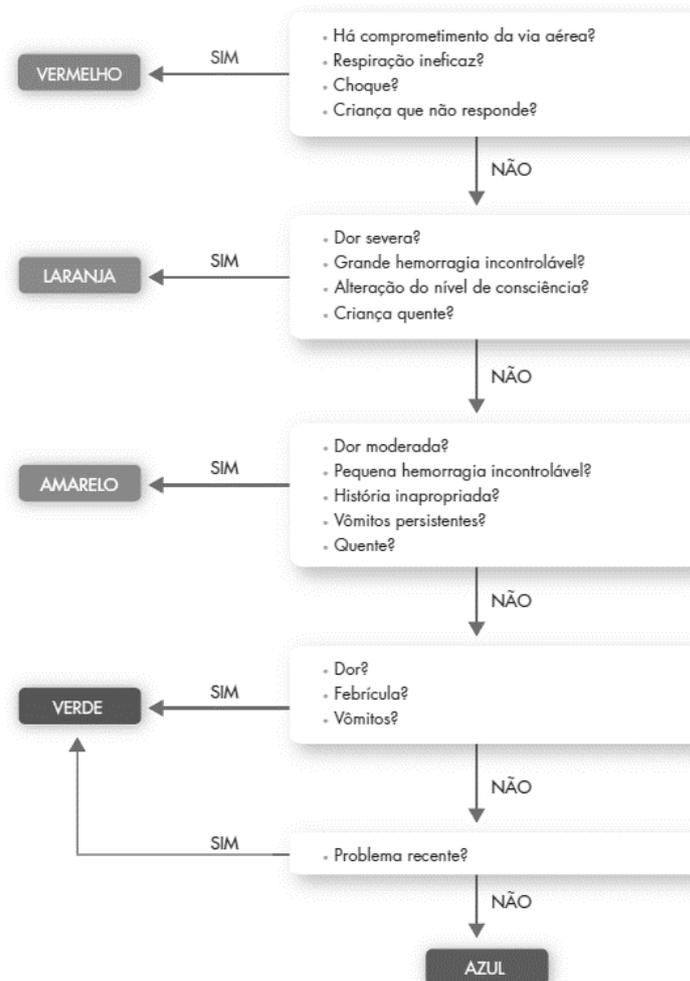
*O processo de triagem classificatória deve ser realizado por profissional de saúde, de nível superior, mediante treinamento específico e utilização de protocolos pré-estabelecidos e tem por objetivo avaliar o grau de urgência das queixas dos pacientes, colocando-os em ordem de prioridade para o atendimento. A esta triagem classificatória é vedada a dispensa de pacientes antes que estes recebam atendimento médico. Após a triagem, os pacientes são encaminhados aos consultórios médicos [Ministério da Saúde 2002].*

No Modelo de Manchester utilizado em nosso estudo, as classificações são indicadas por cores que separam o nível de urgência de atendimento para cada paciente, propondo tempo que cada um pode esperar de acordo com a Tabela 1.

**Tabela1. Quantidade de tempo que prioridade pode esperar pelo atendimento**

Cor	Prioridade	Descrição	Tempo
Vermelha	Emergência	Necessitam de atendimento imediato	0 min
Laranja	Muito Urgente	Necessitam de atendimento praticamente imediato	10 min
Amarelo	Urgente	Necessitam de atendimento rápido, mas podem aguardar	50 min
Verde	Pouco Urgente	Podem aguardar atendimento ou serem encaminhados para outros serviços de saúde	120 min
Azul	Não Urgente	Podem aguardar atendimento ou serem encaminhados para outros serviços de saúde	240 min

De acordo com o protocolo de Manchester, os pacientes respondem a determinadas perguntas exibidas na Figura 1, que conforme a resposta, sua prioridade é definida. [GBCR14]



**Figura 1 – Fluxograma resumido do protocolo Manchester**

### 3. A Mineração de Dados

A Mineração de dados é o processo que busca pela extração de informações úteis, padrões e tendências, frequentemente desconhecidos, a partir de uma grande quantidade de dados armazenada em banco de dados [BT99]

O processo de minerar grandes quantidades de dados pode encontrar padrões existentes, regras de associação e detectar relacionamentos entre as variáveis, fazendo com que seja possível também classificar os dados.

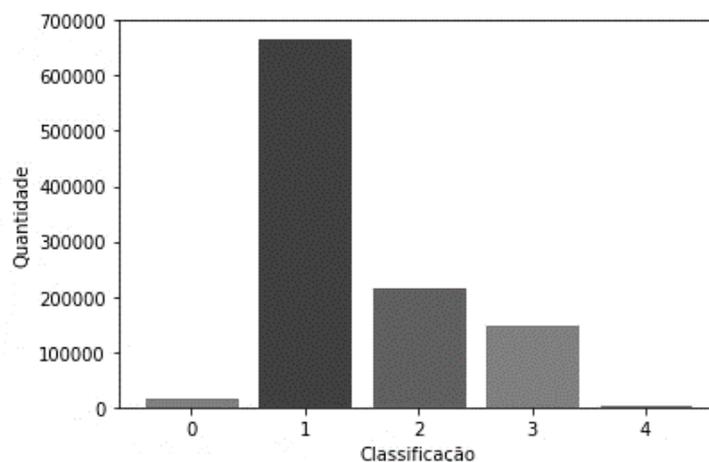
Os registros, exemplificados na Figura 2, foram coletados no banco de dados do sistema e salvos em um arquivo formato CSV, foi utilizada a versão 3.7 da linguagem Python, juntamente com as bibliotecas NumPy e Pandas para o desenvolvimento deste estudo.

Prioridade	Glicemia	Temperatura	Oximetria	Glasgow	Dor	FC
Amarelo	204.0	36.5	98.0	15.0	5.0	90.0
Amarelo	110.0	37.2	98.0	15.0	6.0	76.0
Amarelo	100.0	36.2	97.0	15.0	5.0	72.0
Amarelo	91.0	38.9	96.0	15.0	5.0	96.0
Amarelo	82.0	36.1	93.0	15.0	4.0	79.0
Amarelo	82.0	36.1	93.0	15.0	4.0	79.0

**Figura 2 – Pequena amostra dos dados coletados no banco de dados**

No estudo foram analisadas 1.044.615 classificações de risco, conforme Figura 3, realizadas através do protocolo de Manchester e as variáveis escolhidas para simulação de uma nova classificação foram os seis sinais vitais medidos durante a classificação: glicemia, temperatura, oximetria, escala Glasgow, escala de dor e frequência cardíaca.

Para menor esforço computacional, as cores foram transformadas nos seguintes códigos: 0 – Azul, 1 – Verde, 2 – Amarelo, 3 – Laranja, 4 - Vermelho



**Figura 3 – Quantidade de classificações realizadas e suas respectivas cores.**

#### **4. Uma nova classificação utilizando o KNN**

A Classificação é um dos resultados da mineração de dados e possibilita a associação de objetos a determinadas classes. [BL97]

Métodos baseados em distâncias adotam a ideia de que objetos que pertencem a uma mesma classe possuem uma relação de proximidade no espaço de atributos considerados [Faceli et al. 2011].

O algoritmo kNN(k-Nearest Neighbors) é um método de classificação que utiliza a distância para classificar um objeto. Ao classificar um objeto desconhecido, o método busca pelos K vizinhos mais próximos a ele em um conjunto de dados previamente conhecido e a classificação do objeto será de responsabilidade dos votos dos seus K vizinhos.

A distância utilizada pelo algoritmo foi a euclidiana e o valor de K=3 foi escolhido para o estudo com a intenção de que não aconteça empate na apuração dos votos de cada objeto, é importante também ressaltar que os sinais vitais com atributos ausentes foram substituídos pela média geral de cada coluna.

Os dados foram divididos em 67% para a base de treinamento e 33% para a base de teste e o resultado exibido na Figura 4 apresentou uma acurácia de 85%.

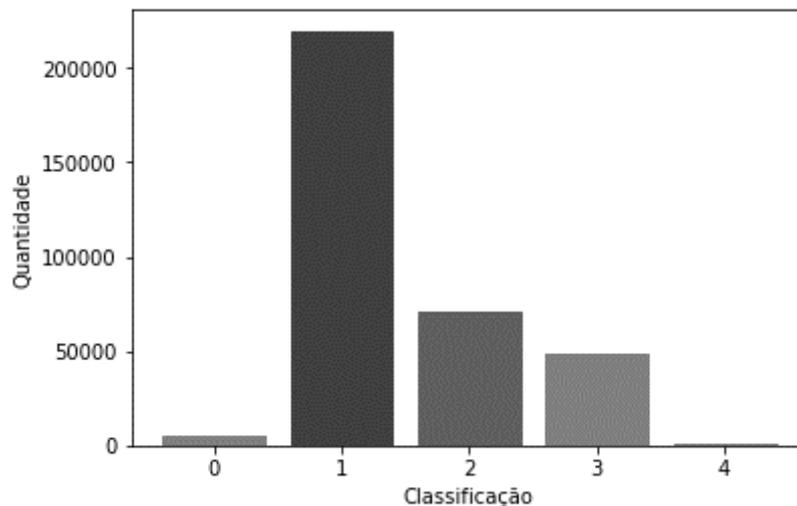
```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33, random_state=42)
knn = KNeighborsClassifier(n_neighbors=3)
knn.fit(X_train, y_train)
# predict the response
pred = knn.predict(X_test)
print(pred)
print(y_test)
accuracy = accuracy_score(y_test, pred)
print('\nThe accuracy of our classifier is %f%%' % accuracy)

[3 2 2 ... 2 2 2]
[3 2 2 ... 2 3 2]

The accuracy of our classifier is 0.854042%
```

**Figura 4 – Resultado apresentado no estudo**

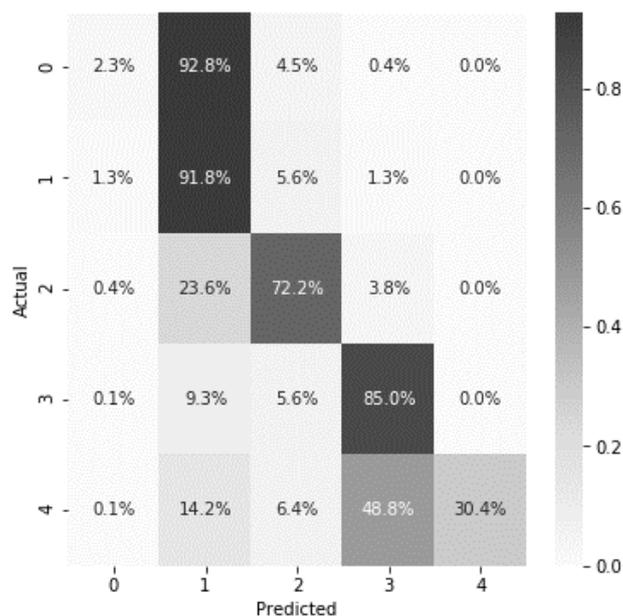
Na base de teste contendo 344.723 registros, em que a classificação foi realizada pelo algoritmo do método KNN, o resultado apresentado na Figura 5 é satisfatório uma vez que é semelhante à classificação realizada pelo Modelo de Manchester exibido na Figura 2.



**Figura 5 – Classificação pelo método KNN**

#### 4. Matriz de Confusão

Através da Matriz de Confusão representada na Figura 6 é possível fazer uma melhor análise dos dados. A matriz serve como métrica para o modelo, uma vez que exhibe os valores reais e os valores preditos pelo modelo KNN.



**Figura 6 – Matriz de Confusão**

Ao interpretar a matriz podemos verificar que 92.8% dos classificados como Azul no Modelo de Manchester foram classificados como Verde no método KNN. Que 91,8% dos classificados como Verde no Modelo de Manchester foram classificados também como Verde no método KNN. Que 72,2% dos classificados como Amarelo no Modelo de Manchester foram classificados também como Amarelo no método KNN. Que 85% dos classificados como Laranja no Modelo de Manchester foram classificados também como Laranja no método KNN. E que os classificados como Vermelho se dividiram mantendo 30,4% no Vermelho, porém mudando 48% para a cor Laranja.

## 5. Resultados Obtidos

É visível e previsto cada vez mais a tecnologia se integrando com outras áreas de conhecimento, não seria diferente para a área da saúde, que com o armazenamento do prontuário eletrônico se tornou possível a realização de diversos estudos.

Através do estudo realizado neste artigo foi possível verificar um resultado satisfatório quanto a mineração dos dados propostos. Considerando que o algoritmo KNN possuía apenas seis informações sobre o início do atendimento e que em 85% das vezes fez a mesma classificação que um profissional da área, fica visível que com um estudo mais aprofundado talvez seja possível chegar em um modelo de classificação mais rápido e eficiente baseado em uma grande quantidade de outros atendimentos.

Não é intenção desse estudo desprezar a análise clínica realizada pelo profissional de saúde, que tem conhecimento técnico para agir, independentemente de qualquer sinal vital medido no momento da classificação de risco, mas sim, propor uma combinação de conhecimentos para que constantemente mais variáveis do atendimento sejam analisadas para que mais vidas sejam salvas.

Para evolução do estudo, é possível trabalhar com outros métodos mais complexos de classificação, como as árvores, e principalmente a análise textual dos dados. Na análise textual, informações são extraídas com base na frequência dos termos encontrados em um determinado texto, conforme exibido na Figura 7.



# Confiabilidade nos Navegadores de Internet

Lisandro Lemos Machado<sup>1</sup>, Luis Fernando Pooter<sup>1</sup>

<sup>1</sup>Instituto Federal Sul-rio-grandense (IFSUL), campus Passo Fundo  
Estrada Perimetral Leste, 150, CEP 99064-440, Passo Fundo – RS

lisandro.machado@passofundo.ifsul.edu.br, ifsulluis@gmail.com

**Abstract.** *This work aims to investigate the security of the main internet browsers regarding their confidentiality in operating the information inserted in them. A comparison was made between the main browsers, seeking to verify possible abnormalities between them in terms of their security. The data flows of the browsers were monitored and the data sent was compared. Finally, it is concluded that with these methodologies of tests and comparisons, the result obtained allows to highlight possible Internet browsers more reliable to be used in organizational environments.*

**Resumo.** *Esse trabalho tem como objetivo investigar a segurança dos principais navegadores de internet quanto a sua confidencialidade em operar as informações neles inseridas. Foi realizada uma comparação entre os principais navegadores, buscando verificar possíveis anormalidades entre eles no que se refere a sua segurança. Foram monitorados os fluxos de dados dos navegadores e realizada a comparação de dados enviados. Por fim, conclui-se que com estas metodologias de testes e comparações, o resultado obtido permite salientar possíveis navegadores de internet mais confiáveis para serem utilizados em âmbitos organizacionais.*

## 1. Introdução

Em uma rede corporativa é necessário um controle das informações com um alto nível de prioridade. Nesse contexto, o analista de rede deve decidir qual ferramenta seus usuários deverão utilizar, priorizando a segurança de cada uma e a proteção das informações que as mesmas manipulam. Os navegadores de internet fazem parte destas ferramentas e são indispensáveis para acessar a rede mundial de computadores, pois são o ponto de entrada para ela, devendo se definir qual será o mais seguro a ser utilizado.

Diante disso, o problema que impulsiona a presente pesquisa é a confidencialidade oferecida por estes navegadores de internet em manipular as informações pessoais de cada indivíduo. Assim, esta análise busca responder as seguintes perguntas: Existem navegadores mais confiáveis que outros? Os navegadores possuem diferenças em sua forma de lidar com as informações?

Para responder estas questões, foram realizadas, pesquisas e testes utilizando os navegadores de internet como parte central deste estudo, o principal objetivo foi analisar a confiabilidade dos navegadores quanto a segurança das informações. Assim, foram estabelecidos critérios específicos a serem seguidos como etapas para chegar ao objetivo principal de forma organizada e correta.

## **2. Segurança da Informação e navegadores**

Assim como em uma conversação presencial com uma pessoa, em uma comunicação na internet existem informações das quais o seu sigilo é necessário, porém a dificuldade em manter o conteúdo em sigilo torna-se bem mais difícil. A informação na internet, pode ser interceptada, roubada e alterada se a rede não possuir critérios de segurança.

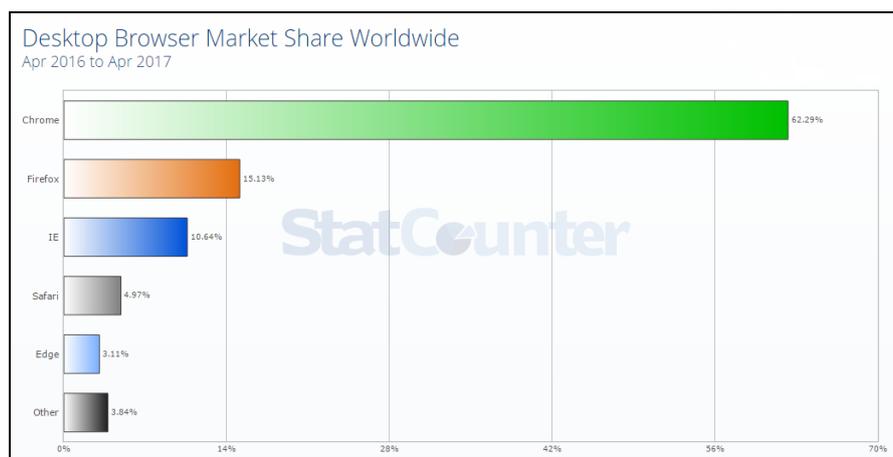
Segundo define Moraes “a segurança da informação pode ser definida como um processo de proteger a informação do mau uso tanto acidental como intencional, por pessoas internas ou externas à organização, incluindo empregados, consultores e hackers”. (2010, p. 19). É possível compreender que estamos vulneráveis a qualquer indivíduo, dentro ou fora de uma rede privada, capaz de praticar atos que resultem no comprometimento das principais propriedades da segurança da informação.

A confidencialidade dos dados é o princípio da segurança que impulsiona esta pesquisa, pelo fato da sua importância, tanto para o usuário e quanto para a empresa, onde nas mãos erradas as informações privadas podem resultar em grandes prejuízos. Conforme Moraes “o princípio da confidencialidade é proteger a informação em sistemas, recursos e processos para que eles não possam ser acessados por pessoas não autorizadas” (2010, p. 28). Ao acessar uma página web em que é solicitado um dado pessoal, o navegador inicia um processo de autenticação do site, enquanto este realiza o mesmo processo no navegador para verificar se o mesmo é autêntico ou se estamos autorizados a acessar a página web de acordo com a política de controle. O navegador então solicita à página web, uma chave de encriptação para codificar os dados e envia-los apenas no formato encriptado (Goodrich e Tamassia, 2013).

### **2.1 Navegadores de Internet**

Assim como um barco se movimenta pelos oceanos, transportando sua tripulação para um determinado destino, os navegadores de internet têm por função permitir ao usuário navegar pelo mar de informações da *web*. Conforme Comer (2006, p. 322), “um navegador web consiste em um aplicativo que um usuário invoca para acessar e exibir uma página web”. A concorrência se faz presente na categoria dos navegadores, com vários possuindo a mesma função, porém com suas próprias características e serviços.

Existem serviços com o objetivo de qualificar os navegadores mais utilizados por meio de métricas para contabilizar a frequência de uso. O StatCounter é um serviço de análise da *web*, que fornece estatísticas independentes sobre as tendências de uso da internet, baseando-se em mais de 15 bilhões de páginas vistas por mês por pessoas de todo o mundo por meio dos 2,5 milhões de sites membros do grupo. (STATCOUNTER, 2017). A Figura 1 apresenta análise realizada de Abril de 2016 à Abril de 2017.



**Figura 1 - Market Share dos navegadores para desktops do mundo**

Com estes dados é possível destacar os navegadores *web* mais utilizados.

## 2.2 Trabalhos Correlatos

Em se tratando da utilização de navegadores de internet, segundo Coelho (2013), as motivações dos usuários variam conforme a utilização, tais como, estudos, trabalho, acesso a notícias entre outros. Em sua pesquisa, realiza uma análise nos fatores sociais de influência na escolha dos navegadores de internet pelos usuários, nos quais 79,9% destes escolhem sua ferramenta através da reputação/popularidade do navegador. Outro fator de escolha do navegador é o tecnológico, onde 93,5% do total de amostras correspondem a velocidade de carregamento das páginas, seguidos da facilidade de navegação (89,2%) e em terceiro a segurança (83,8%). Algumas perguntas voltadas ao grau de confiabilidade dos usuários nos navegadores de internet permitiram destacar o Google Chrome com 79,9% do total da amostra, seguido pelo Firefox com 69,8% (COELHO, 2013).

Confiar na privacidade que os navegadores de internet oferecem aos usuários é uma questão que deveria exigir mais cuidados. Existem vários métodos na internet que prometem segurança e serviços oferecendo sigilo. Toledo e Fedel (2014) apontam que o usuário não possui o conhecimento sobre coletas de informações, não sendo estas, claras, objetivas e algumas vezes sem o seu consentimento. Segundo Ishitani (*apud* Toledo e Fedel, 2014), este tipo de coleta de informações pode ser considerado como invasão de privacidade, pois o usuário não tem conhecimento de que está sendo rastreado muito menos consentiu permissão para executar estes serviços. Os autores ainda expõem novamente a invasão e roubo de informações pessoais como ponto chave para concluir o trabalho. Conforme Toledo e Fedel (2014, p. 50) “diversas corporações faturam bilhões com venda de informações de seus usuários, seja para empresas ou para espionagem governamental”. Tendo em vista os relatos mencionados, é possível utilizar os mesmos como fonte de pesquisa e ainda como meio de impulsionar este projeto que pretende analisar a confiabilidade dos navegadores.

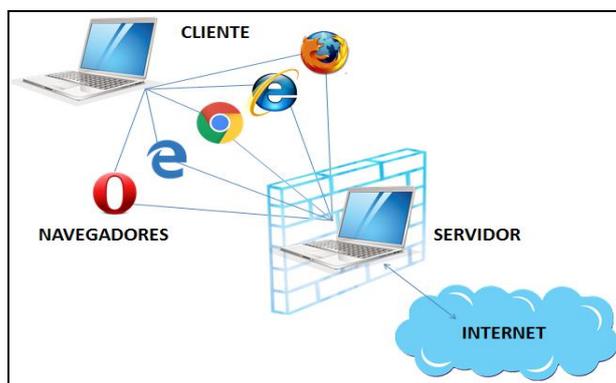
## 3. Análise da confiabilidade de navegadores da internet

Conforme definido no objetivo geral do projeto, a intenção foi a de analisar a confiabilidade dos navegadores quanto à segurança da informação. Com base no levantamento realizado no tópico anterior, os testes e análises foram executados nos

cinco navegadores mais qualificados no ano de 2017 e suas respectivas versões da época: Google Chrome, Mozilla Firefox, Internet Explorer, Microsoft Edge e Opera. O Safari não foi avaliado nesta análise, pois o mesmo não se encontra disponível oficialmente para Windows.

### 3.1 Ambiente

Assim, a estrutura elaborada para implementação dos testes de navegadores possui dois notebooks comunicando-se um com o outro por meio de serviços Cliente/Servidor, sendo que o Servidor disponibiliza acesso à internet para o cliente, monitorando o fluxo de dados entre os cinco navegadores de internet, anteriormente mencionados, instalados nele, permitindo a realização dos testes, conforme mostra a Figura 2.



**Figura 2 - Estrutura de implementação**

A máquina Servidor, é um notebook da marca Acer, modelo E1-531-2608, com processador Intel Celeron 1000M de 1,8 GHz, memória RAM de 4 GB e armazenamento de 500 GB com o sistema operacional Kali Linux versão 2017.2, que apresenta mais de trezentas ferramentas de segurança e de testes de invasão já instaladas. O Kali Linux, conforme Broad e Bindner, “é o *live disk* mais recente de uma distribuição de segurança disponibilizada pela *Offensive Security*” (2014, p. 24), sendo utilizado por especialistas em testes de intrusão e auditorias de segurança.

A máquina Cliente possui acesso à internet a partir do Servidor. É um notebook da marca DELL pertencente ao modelo Inspiron i15-5566-A10P com processador Intel Core i3 de 2,0 GHz, memória RAM de 4 GB e armazenamento de 1 TB. O sistema operacional utilizado foi o Windows 10. Este sistema foi escolhido pelo fato de ser o sistema operacional mais utilizado segundo dados do StatCounter, que apresenta no período de agosto de 2016 a agosto de 2017 o sistema operacional Windows com 83,41% de usuários frente a 11,16% do segundo colocado, o OS X.

Com os sistemas em funcionamento foi realizada a configuração das interfaces de rede nas duas máquinas para obter e disponibilizar o acesso à internet via cabo *crossover*. Na máquina Cliente, com o sistema operacional Windows10, foi instalado a versão mais recente e atualizada, em 2017, dos cinco navegadores selecionados. No Internet Explorer a versão do navegador mais recente utilizada nesta análise é a 11.608.15063.0, o Google Chrome encontra-se na versão 61.0.3163.100, o Mozilla Firefox foi instalado com a 55.0.3, Opera com a versão 48.0.2685.35 e o Microsoft Edge 40.15063.0.0.

Ao iniciar a etapa de testes, a primeira verificação e configuração nos navegadores foi realizada após a instalação dos mesmos. Para que a análise ocorresse da forma mais justa e correta possível, foi procurado e desativado qualquer *plug-in*, extensão ou complemento já instalado no navegador, estruturando um ambiente limpo para a realização dos testes, sem perigo de interferência destes programas. A seguir, foram determinadas ferramentas próprias para realização da verificação desejada.

### 3.2 Metodologia de testes

Como forma de se atingir o objetivo determinado, foram destacas ferramentas e recursos que possibilitaram a análise da confiabilidade dos navegadores de internet diante do ambiente configurado, sendo eles: Browsercope, Html5Test e Wireshark.

O Browsercope é um projeto de código aberto para perfis de navegadores da *web* com licença Apache 2.0, lançado em setembro de 2009. Esta ferramenta *online*<sup>1</sup> é específica para realização de testes de segurança, executando um total de dezessete testes, analisando o comportamento dos navegadores. Entre alguns de seus testes estão a verificação de aplicativos importantes, como o Java, verificação de quesitos referentes a política de segurança, invasões de diversas formas entre outros. Ele não necessita da instalação de nenhum *plug-in* para o seu funcionamento, sendo possível realizar a comparação dos aplicativos *web* e verificar a situação referente a cada um (BROWSCOPE, 2017). Com cada teste especificado no Browsercope, deu-se início a execução dos testes individuais em cada navegador utilizando esta ferramenta.

O *html5test*<sup>2</sup> é outra ferramenta para realizar testes nos navegadores *web* com o objetivo de identificar os mais adeptos a tecnologia do HTML5. Sua análise percorre serviços como performance, integração, conectividade, multimídias e ainda uma avaliação na segurança dos navegadores (HTML5 TEST). Este serviço entrega um resultado total da soma de cada funcionalidade que ele requisita e que o navegador possui. Cada especificação é adicionada como um ponto, podendo chegar a quinhentos e cinquenta e cinco pontos se todos os critérios forem cumpridos.

Para este trabalho o único grupo utilizado é o que aborda os testes de segurança e confiabilidade. Este grupo de análise verifica dez funcionalidades que os navegadores devem possuir como APIs, protocolos de segurança e métodos de autenticação. Se todas as especificações forem atendidas, receberá o total de trinta e dois pontos e informando ao lado de cada uma se o navegador possui (*Yes*) ou não (*No*).

O Wireshark é o sniffer de rede escolhido para realizar os testes de tráfego de informações, sendo indispensável para monitoramentos e abstração do que está ocorrendo na rede interna. Conforme Morimoto, o Wireshark “é um sniffer bastante completo que permite capturar o tráfego de rede, fornecendo uma ferramenta poderosa para detectar problemas e entender melhor o funcionamento de cada protocolo”. (2011, p. 446). O Wireshark é capaz de detalhar cada pacote capturado, fornecendo informações referentes ao mesmo. Ao clicar no pacote desejado, é possível realizar um acompanhamento entre origem e destino dos pacotes.

O primeiro passo deste teste foi iniciar o Wireshark no Servidor para capturar o tráfego e logo em seguida abrir o navegador a ser verificado no Cliente. Cada teste foi

---

<sup>1</sup> Disponível em <http://www.browserscope.org/>

<sup>2</sup> Disponível em [html5test.com](http://html5test.com)

executado utilizando apenas um navegador por vez. Ao encerrar uma sequência de acessos, parava-se a execução do Wireshark, salvando o *script* gerado com todas as informações referentes ao tráfego e acessos do Cliente. O principal critério a ser avaliado neste teste é a quantidade de pacotes que o navegador de internet gera enquanto está conectado à internet. Foram realizados os seguintes procedimentos:

- Teste de Minuto Sem Navegação: se iniciou o navegador *web*, sem realizar nenhuma ação durante o período de um minuto. Após este período encerrava-se o navegador juntamente com a captura de tráfego.
- Teste de Minuto Com Navegação: ao iniciar o Wireshark, o navegador é executado e direcionado para o endereço da página *web* selecionada. Os três sites testados separadamente foram <http://inf.passofundo.ifsul.edu.br/>, <https://www.americanas.com.br/> e [https://moodle.passofundo.ifsul.edu.br](https://moodle.passofundo.ifsul.edu.br/). Estes sites foram selecionados por conter diferentes métodos de segurança, sendo que um utiliza o protocolo HTTP, o outro HTTPS e por fim um HTTPS com certificação, sem realizava nenhuma ação, aguardando ao término do minuto.

#### 4. Resultados Obtidos

Com os testes devidamente executados os resultados foram analisados tendo como base os objetivos traçados pela pesquisa. Conforme apresentado na Tabela 1, a grande maioria dos requisitos solicitados pelo Browserscope foram cumpridos por parte dos navegadores. Apenas três testes retornaram resultados negativos para seus respectivos navegadores e unicamente o teste referente a compatibilidade do navegador com HSTS (HTTP Strict Transport Security) novo padrão de segurança SSL que força a utilização do HTTPS demonstrou instabilidade, não completando o teste, o que comprova uma falha desta verificação.

**Tabela 1 - Resultado do teste utilizando o Browserscope**

TESTE	CHROME	FIREFOX	IE	OPERA	EDGE
postMessage API	Pass	Pass	Pass	Pass	Pass
JSON.parse API	Pass	Pass	Pass	Pass	Pass
toStaticHTML API	Fail	Fail	Pass	Fail	Fail
httpOnly cookie API	Pass	Pass	Pass	Pass	Pass
X-Frame Options	Pass	Pass	Pass	Pass	Pass
X-Content-Type-Options	Pass	Pass	Pass	Pass	Pass
Block Reflected XSS	Pass	Pass	Pass	Pass	Pass
Block Location Spoofing	Pass	Pass	Pass	Pass	Pass
Block JSON hojacking	Pass	Pass	Pass	Pass	Pass
Block XSS in CSS	Pass	Pass	Pass	Pass	Pass
Sandbox Attribute	Pass	Pass	Pass	Pass	Pass
Origin header	Pass	Fail	Fail	Pass	Pass
Strict Transport Security	Running	Running	Running	Running	Running
Block cross-origin CSS attac	Pass	Pass	Pass	Pass	Pass
Content Security Policy	Pass	Pass	Fail	Pass	Pass
Cross Origin Resource Shari	Pass	Pass	Pass	Pass	Pass
Block visited link sniffing	Pass	Pass	Pass	Pass	Pass

O teste que verifica se o navegador suporta a API `toStaticHTML` foi o que obteve menos aprovações. Nesta avaliação, a API para desinfecção de entradas não confiáveis foi encontrada somente no Internet Explorer. Contudo este mesmo navegador foi o único que falhou no teste de Política de Segurança do Conteúdo, que verifica a

compatibilidade do navegador com Content Security Policy, que reduz as superfícies de ataques de injeção de conteúdo, tais como XSS. Por fim, outro teste que obteve falha por parte dos navegadores foi o Origin Header, sendo que o Internet Explorer e o Firefox não passaram neste quesito. Este recurso ameniza ataques do tipo *cross-site request forgery* que explora a confiança que um *website* tem no navegador do usuário.

A Tabela 2 mostra os resultados obtidos com a HTML5 Test. Com o resultado, é possível visualizar cada teste realizado pelo recurso bem como as pontuações individuais de cada requisito e ainda a soma total que cada navegador conquistou.

**Tabela 2 - Resultado do teste utilizando o HTML5 Test**

TESTE	PONTOS	CHROME	FIREFOX	IE	OPERA	EDGE
Web Cryptography API	5	YES	YES	YES	YES	YES
Content Security Policy 1	3	NO	NO	NO	YES	YES
Content Security Policy 2	2	YES	NO	NO	YES	YES
Cross-Origin Resource Sharing	4	YES	YES	YES	YES	YES
Subresource Integrity	2	YES	YES	NO	YES	NO
Cross-document messaging	2	YES	YES	YES	YES	YES
Web Authentication / FIDO 2	3	NO	NO	NO	NO	PARTIAL
Credential Management	3	YES	NO	NO	YES	YES
Sandboxed iframe	4	YES	YES	YES	YES	YES
iframe with inline contents	4	YES	YES	NO	YES	NO
<b>TOTAL</b>	32	26	21	15	29	23

No teste de Autenticação na *Web* utilizando o protocolo FIDO 2, para comunicação entre um autenticador externo e outro cliente, praticamente todos os navegadores reprovaram, apenas o Microsoft Edge obteve parcialmente êxito neste quesito, não alcançando completamente os requisitos exigidos. O Opera foi o navegador que mais somou pontos nesta análise, falhando apenas na verificação descrita a pouco. Seguido pelo Chrome, Edge, Firefox e com a menor pontuação o Internet Explorer, reprovando em mais da metade dos testes executados.

Com o Wireshark, foi realizada a verificação e comparação da quantidade de pacotes gerados em cada navegador com os mesmos acessos e tempo para cada teste. O primeiro teste, o de minuto sem navegação, é exibido na Tabela 3. Com ele é possível verificar a diferença dos números de pacotes gerados de um navegador para outro durante a sua inicialização no tempo de um minuto sem nenhum tipo de navegação.

**Tabela 3 - Quantidade de pacotes capturados durante um minuto sem navegação**

NAVEGADORES	N° PACOTES (por minuto)
Internet Explorer	606
Mozilla Firefox	418
Google Chrome	554
Microsoft Edge	967
Opera	520

O número de pacotes transferidos durante a utilização de cada navegador em teste pode ser observado na coluna da direita desta tabela. No teste sem navegação ou autenticação, onde apenas é inicializado o navegador, é possível observar números elevados na comparação destes programas.

Na execução do segundo teste fica visível a comparação entre cada navegador diante do acesso às páginas selecionadas. Ainda que todas as informações estejam apresentadas em uma única tabela, cada teste foi executado separadamente. A Tabela 4 demonstra o número de pacotes capturados em cada navegador *web*, durante o tempo de um minuto ao acessar cada página descrita, realizando *login* em cada uma delas.

**Tabela 4 - Quantidade de pacotes capturados durante um minuto com navegação**

NAVEGADORES	MOODLE	INF	AMERICANAS
Internet Explorer	1900	5600	15781
Mozilla Firefox	1349	4997	11786
Google Chrome	2026	4489	13590
Microsoft Edge	2483	6072	18820
Opera	1896	4590	13498

Como apresentado no teste anterior, o resultado apresentado indica as mesmas variações na quantidade de pacotes transferidos durante o teste entre os navegadores.

## 5. Conclusão

De acordo com os objetivos traçados para a realização da presente pesquisa, cada navegador selecionado pôde ser devidamente reconhecido, através de suas características comparadas. Os critérios e métodos para a realização da avaliação dos navegadores se basearam nas ferramentas Browserscope, HTML5 Test e o Wireshark, que tornaram possível obter e analisar dados através de testes realizados.

Com a ferramenta Browserscope fora uma falha na verificação de um dos requisitos em que todos os navegadores não o concluíram, foi possível obter resultados equivalentes. O quesito que mais apontou falha nos navegadores foi o teste de verificação da API `toStaticHTML`, neste apenas o Internet Explorer passou. Contudo este navegador, juntamente com o Firefox, foram os únicos a reprovar em dois testes. A verificação do cabeçalho Origin header, reprovou estes dois navegadores. O teste de verificação da compatibilidade do navegador com a política de segurança de conteúdo obteve uma única falha, em que o Internet Explorer reprovou.

O HTML5 Test se demonstrou uma ferramenta mais categórica em suas verificações, detectando mais falhas nos navegadores dentre os exigidos para qualificação. Seu sistema de pontuação para cada verificação auxilia na compreensão da relevância de cada teste ao atribuir uma nota máxima ou parcial. Nela, todos os navegadores apresentaram irregularidade em pelo menos um critério. O Opera conquistou o menor número de falhas, reprovando em apenas um teste, neste em questão todos os outros navegadores reprovaram. Em seguida estão o Chrome e o Edge, com problemas em duas verificações. No Firefox o índice aumentou para quatro reprovações. E por fim, o Internet Explorer com o maior número de testes reprovados. Ao total são seis falhas, somando apenas quinze pontos, quatorze a menos se comparado ao Opera, navegador que totalizou vinte e nove pontos dos trinta e dois possíveis.

Utilizando o Wireshark surgiram resultados instigantes para a análise. Em todos os testes de captura e análise de tráfego, os navegadores Edge e Internet Explorer respectivamente, se sobressaíram na quantidade de pacotes que trafegaram na rede quando os mesmos foram utilizados. A quantidade de tráfego dos dados no Microsoft Edge ficou, em todos os testes realizados, muito acima dos demais navegadores. Ainda, é seguido pelo seu antecedente, o Internet Explorer, que também se coloca em segundo lugar, quanto ao número de pacotes percorridos, em todos os testes. Opera, Chrome e Firefox possuem uma quantidade de fluxo semelhantes, não distanciando assim, um dos outros, e alterando suas posições no *ranking* a cada teste realizado. Contudo, a quantidade total de pacotes capturados, por si só, não condiz necessariamente com a falta ou excesso de segurança na informação. Porém, é instigante que alguns navegadores demandem de um número consideravelmente maior de protocolos do que outros, para processarem as mesmas informações e requisitos.

Em conjunto, o resultado dos testes realizados em todas as ferramentas, levando em consideração os números referentes ao tráfego de pacotes, tendo em vista que, outros navegadores realizam os mesmos processos com menor quantidade na transferência das informações, é possível destacar navegadores menos propensos a confidencialidade.

Os navegadores Opera, Firefox e Chrome, obtiveram resultados semelhantes nos testes. Entre estes, nenhum apresentou que demonstrem problemas na segurança e confidencialidade das informações. Todavia os navegadores Edge e Internet Explorer apresentaram resultados intrigantes nos testes realizados. O mais recente navegador do mercado, Edge obteve resultados positivos nos testes utilizando o Browserscope e HTML5 Test, se igualando ao demais navegadores. Porém ao analisar o tráfego na rede utilizando o Edge, a quantidade de pacotes que são transmitidos com sua utilização é comprovadamente maior do que qualquer outro navegador.

O Internet Explorer evidenciou mais problemas de segurança desde o teste com o Browserscope, no qual apresentou falhas incomuns aos demais navegadores, passando ainda pelo teste HTML5 Test, no qual foi o navegador que mais apontou falhas na sua verificação, somando a menor pontuação entre todos. Por fim, mesmo o Edge possuindo o maior número de pacotes transferidos, o Internet Explorer, não fica muito distante neste teste, mantendo-se em segundo lugar em todas as verificações executadas.

É possível concluir se pode distinguir estes programas quanto a confidencialidade. Fica evidente que o problema exposto possui soluções cabíveis a serem empregadas para que em ambientes corporativos, os responsáveis pela segurança optem por ferramentas que visando segurança e sigilo para suas informações. A verificação de cada pacote que trafega na rede é uma análise importante a ser mais aprofundada futuramente, pois se trata de um processo que demanda tempo para verificação devido ao grande fluxo de pacotes na rede, porém com a sua realização poderá se buscar identificar o porquê da diferença na geração de dados entre os navegadores.

Cabe ressaltar, que por se tratar de uma área ampla, conforme o tema abordado na presente pesquisa, a segurança de redes sempre apresentará novas possibilidades a serem estudadas e desenvolvidas. Especificamente em ambientes corporativos, a obtenção e manutenção de um ambiente seguro deve ser constantemente perseguida e periodicamente revisada. Esse processo passa fundamentalmente pela compreensão e

análise dos mais variados elementos envolvidos nesse meio, como foi o caso dos navegadores de internet aqui pesquisados.

## 6. Referências

- BROAD, James; BINDNER, Andrew. (2014) “Hacking com Kali Linux: Técnicas práticas para testes de invasão” Novatec.
- BROWERSCOPE. (2017) “Browserscope FAQ”, <http://www.browserscope.org/faq>, Agosto.
- COELHO, Milena de Lima Cavalcanti. (2013) “Impacto dos perfis tecnológicos na escolha e no comportamento de uso de navegadores de internet”. 60 f. Monografia (bacharelado) – Administração, Universidade de Brasília.
- COMER, Douglas E. (2006) “Interligação de Redes com TCP/IP”. Elsevier.
- GOODRICH, Michael T.; TAMASSIA, Roberto. (2013) “Introdução à Segurança de Computadores”. Bookman.
- GOOGLE. (2017) “Google Terms of Service”, <https://www.google.com.br/chrome/browser/privacy/utos-text.html>, Maio.
- HTML5 TEST. (2017) “How well does your browser support html5?”, <https://html5test.com>, Novembro.
- MELIUZ. (2017) “As 50 Melhores Lojas Online do Brasil”, <https://www.meliuz.com.br/blog/50-melhores-lojas-brasil>, Setembro.
- MORAES, Alexandre Fernandes de. (2010) “Segurança em Redes – Fundamentos”. Érica.
- MORIMOTO, Carlos Eduardo. (2011) “Redes: Guia Prático”. Sul Editores.
- NETMARKETSHARE. (2017) “Desktop Top Browser Share Trend, April 2016 to April 2017”, <https://www.netmarketshare.com>, Abril.
- STATCOUNTER GLOBAL STATS. (2017) “Desktop Browser Market Share Worldwide, April 2016 to April 2017”, <http://gs.statcounter.com/browser-market-share/desktop/worldwide/#monthly-201604-201704-bar>, Abril.
- TOLEDO, André Yuri; FEDEL, Gabriel de Souza. (2014) “Privacidade na WEB: uma análise sobre a navegação privativa dos navegadores”. Tecnologia em Segurança da Informação, Faculdade de Tecnologia de Americana.

# Progressive Web Apps: um estudo de caso para o 1º Batalhão Rodoviário da Brigada Militar

Daniel Nicolau Saito<sup>1</sup>, Thauany Corrêa Martins<sup>1</sup>, Marcelo Felipe Guarani Fernandes<sup>1</sup>, Guilherme Gehring<sup>1</sup>, Adilso Nunes de Souza<sup>1</sup>, Anubis G. M. Rossetto<sup>1</sup>

<sup>1</sup>Instituto Federal Sul-rio-grandense (IFSUL) Câmpus Passo Fundo  
Estrada Perimetral Leste, 150 – CEP: 99064-440 – Passo Fundo – RS –

daniel.nicolau.saito@yahoo.com.br

{thauanycorreamartins12, guaranimarcelo, guilherme.gehring}@gmail.com

{adilso.souza, anubisrossetto}@passofundo.ifsul.edu.br

**Abstract.** *Progressive Web Apps (PWAs) are a set of techniques for developing mobile web-based applications that run in a browser without having to be installed. One of PWA's key features is its ability to operate independently of connectivity, enabling data synchronization with background servers. This paper presents the development of two applications using the PWA: BRBM App and BRBM Vehicles. Given the need for the second application to operate offline, two technologies for data storage and synchronization were studied and one of them selected for development. Details about the application architecture and the results obtained in the evaluation of the BRBM Vehicles app using the Lighthouse tool are presented.*

**Resumo.** *Progressive Web Apps (PWAs) são um conjunto de técnicas para desenvolver aplicações web direcionadas para dispositivos móveis, rodando em um navegador, sem a necessidade de serem instaladas. Uma das principais características da PWA é a capacidade da aplicação operar independente de conectividade, possibilitando a sincronização de dados com servidores em segundo plano. Este trabalho apresenta o desenvolvimento de duas aplicações empregando a abordagem PWA utilizando o Framework Ionic: BRBM App e Veículos BRBM App. Tendo em vista a necessidade da segunda aplicação operar offline, foi realizado o estudo de duas tecnologias para armazenamento e sincronização de dados offline e selecionada uma delas para o desenvolvimento. São apresentados detalhes sobre a arquitetura das aplicações e resultados obtidos na avaliação da aplicação Veículos BRBM App utilizando a ferramenta Lighthouse.*

## 1. Introdução

Os aplicativos utilizados na plataforma móvel são basicamente de dois tipos: instaláveis ou Web. Os aplicativos instaláveis são aqueles que o usuário faz uma busca na loja da plataforma, escolhe a aplicação, faz download e instala. Esses aplicativos podem ser desenvolvidos utilizando as ferramentas específicas da plataforma (Android, iOS ou Windows), chamados nativos, ou *frameworks* que possibilitam a criação de um único código fonte e a geração da aplicação para diferentes plataformas, denominados multiplataforma.

Também existem os aplicativos híbridos que usam as linguagens da Web para construção e precisam de um *container* para executar [Lopes 2016]. Nesse trabalho vamos nos referir aos aplicativos que são instalados no dispositivo móvel como nativos. Já os aplicativos Web são sites independentes de plataforma, acessados e executados por um navegador.

Pode-se considerar que aplicativos *Web mobile* perdem dos aplicativos nativos em termos de desempenho e acesso a recursos do dispositivo. Por isso, o desenvolvimento de aplicativos móveis nativos é muitas vezes a opção escolhida para superar as limitações que a Web como plataforma impõe em aparelhos móveis. Entretanto, os aplicativos móveis nativos têm suas próprias deficiências, por exemplo o alto uso de recursos e tempo de desenvolvimento, sendo necessário manter e distribuir aplicativos em lojas das diferentes plataformas. Para usar o aplicativo nativo, os usuários precisam passar por muitas etapas que incluem a inscrição na loja da plataforma, verificar disponibilidade de memória, fazer download, instalar e finalmente abrir o aplicativo para usar. Associado a isso, alguns aplicativos ainda exigem um processo de cadastro de usuário antes do seu uso efetivo. [Cselle 2012] fez um estudo onde revela que, em média, um aplicativo perde 20% de seus usuários para cada etapa entre o primeiro contato do usuário e iniciar o uso. Além disso, muitos usuários consideram este processo difícil e tedioso. Nesse sentido, os aplicativos nativos podem gerar desvantagens para empresas e desenvolvedores.

Percebe-se, portanto, pontos frágeis em ambas as plataformas. No ambiente Web tais fragilidades podem ser contornadas pelo surgimento da chamada Progressive Web Application (PWA), cujo objetivo é combinar as capacidades e experiências de aplicativos nativos com o alcance da web [Biørn-Hansen et al. 2017]. Ao possuir o alcance da web, a PWA compatibiliza o melhor da web com o melhor dos aplicativos nativos, sendo possível seu uso em todas as plataformas, reduzindo o custo e os recursos e aumentando a acessibilidade.

A partir da necessidade verificada pelo 1º Batalhão Rodoviário da Brigada Militar, este trabalho desenvolveu duas aplicações empregando a abordagem de PWA com o Framework Ionic. O 1º Batalhão Rodoviário da Brigada Militar (BRBM), com sede em Passo Fundo, atua em 3.914 quilômetros de malha viária abrangendo cerca de 229 municípios do norte do Rio Grande do Sul.

As aplicações foram nomeadas como *BRBM App* e *Veículos BRBM App*. A aplicação *BRBM App* teve como objetivo desenvolver um sistema informatizado para auxiliar nos processos administrativos do 1º BRBM, envolvendo o controle de diversas informações de pessoal. Já aplicação *Veículos BRBM App* foi desenvolvida para tratar exclusivamente do controle de saída e retorno de veículos. Entre os requisitos do Batalhão, existem funcionalidades que serão executadas por policiais em atuação nas ruas/estradas e que com muita frequência não tem conexão estável com a Internet. Assim, empregar o conceito *offline* na *Veículos BRBM App* foi a solução encontrada para atender a esta demanda.

Muito embora as duas aplicações desenvolvidas empreguem o conceito de PWA, na aplicação *BRBM App* optou-se por utilizar um banco de dados relacional para persistência dos dados considerando que a aplicação envolve a necessidade de armazenar dados complexos, isso é, muitos dados e de diferentes contextos/tipos, que exigem também consultas complexas. Neste sentido, um banco de dados relacional se aplica como me-

lhor solução. Foi adotado o sistema de gerenciamento de banco de dados PostgreSQL que gerou quarenta tabelas para armazenar os dados da aplicação. Para acesso aos dados foi desenvolvida uma API na linguagem PHP que permite requisições HTTP para acesso e atualização dos dados mediante usuário autenticado. Dado o requisito de tratar com dados sempre atualizados e que serão usados por diversos usuários, essa aplicação não opera *offline*, porém tira proveito das demais características de uma PWA.

De outro lado, um requisito importante para a aplicação *Veículos BRBM App* era justamente a possibilidade de operação *offline*. Essa aplicação tem por objetivo registrar as saídas dos veículos da corporação. Possibilitar a operação *offline* pressupõe armazenar localmente as informações geridas pela aplicação quando não possuir conexão e ao retornar, deve ocorrer a sincronização dos dados com a base de dados remota. Todavia, o processo de sincronização envolve uma série de problemas complexos e de difícil resolução. Por exemplo, o desenvolvimento de algoritmos de resolução de conflitos e revisões e a prevenção de perda de dados são requisitos que precisam ser implementados. Porém, a grande maioria das soluções em termos de banco de dados, delega ao desenvolver o tratamento de armazenar localmente as informações e da sincronização dos dados com base de dados remota [Dal’Pizzol and Rossetto 2018].

Duas ferramentas que trazem solução para tratar este contexto *offline* são: CouchDB/PouchDB e Firebase Firestore. Tais ferramentas possuem APIs que permitem salvar os dados localmente quando em estado *offline* e quando a conexão é restabelecida fazer a replicação e sincronização de maneira consistente com o banco de dados remoto. Estas ferramentas incorporam banco de dados NoSQL que são aplicados para estruturas de dados mais simples. Assim, para a aplicação *Veículos BRBM App* optou-se em utilizar uma dessas duas ferramentas. A partir do estudo realizado, uma tecnologia foi selecionada para o desenvolvimento da aplicação.

## 2. Referencial Teórico

Esta seção apresenta detalhes sobre as PWAs, a operação *offline*, bem como sobre as ferramentas com suporte *offline* Firebase Firestore e PouchDB/CouchDB e o *framework* utilizado para desenvolvimento.

### 2.1. Progressive Web Apps

Progressive Web Application – PWA corresponde à um conjunto de técnicas e tecnologias que podem oferecer uma experiência mais produtiva, adaptativa e prática aos usuários, sem a necessidade de instalar um aplicativo nativo no dispositivo, sem os inconvenientes problemas de conexão dos sistemas web tradicionais, mas construindo um desempenho com aprimoramento progressivo da experiência do usuário.[Russel and Berriman 2015] cunharam o termo Progressive Web Apps visando denominar uma aplicação web que mantém as características da Web enquanto entrega uma experiência de usuário próxima à dos aplicativos nativos. Os autores ainda enumeram as características para uma PWA:

- Progressiva: deve implementar melhoramento progressivo, possibilitando sua operação independentemente das funcionalidades suportadas ou não pelo navegador;
- Responsiva: deve funcionar igualmente em qualquer tamanho de tela;

- Independente de conectividade: deve ser melhorada progressivamente com Service Workers para que possa operar *offline*;
- App-like: deve adotar um modelo de aplicação *Shell* + Conteúdo para criar navegações e interações que simulem as de um aplicativo nativo;
- Atualizada: deve atualizar-se automática e transparentemente por meio de Service Workers;
- Segura: devem ser servidas através de HTTPS para garantir a integridade das informações;
- Descobrível: é identificável através de manifestos W3C permitindo que mecanismos de busca a encontrem;
- Reengajável: pode acessar as interfaces de reengajamento do sistema operacional, por exemplo, as notificações *push*;
- Instalável: deve poder ser adicionada à *homescreen* do dispositivo através de diálogos do navegador, permitindo que o *app* seja mantido sem a necessidade de uma *app store*;
- Linkável: deve aproveitar-se do poder social das URLs para compartilhar e ser compartilhada.

## 2.2. Operação Offline

O conceito *offline* está relacionado a possibilidade de armazenamento primário de informações localmente - no próprio dispositivo do usuário - e sua sincronização com servidores e outros dispositivos na presença de conexão com a Internet. A capacidade *offline* de PWAs oferece grandes benefícios e conveniência aos usuários em cenários em que o sinal de Internet seja lento, intermitente ou inexistente [DA-14 2017] . Dessa forma, o usuário dispõe de uma interface mais rápida, uma vez que as operações de escrita e leitura são realizadas localmente, reduzindo drasticamente o tempo de resposta. Outro ganho é a economia de bateria, já que a redução de requisições externas diminuem os ciclos de processamento e, em cenários de desastre, a aplicação ainda é operável até que a conexão seja restabelecida para a sincronização dos dados com os servidores [Holt 2015].

Um aspecto importante que precisa ser compreendido é que o cenário *offline* não é um erro - como geralmente é tratado pelos desenvolvedores -, mas um fato do ambiente *mobile* [Feyerke 2013] e que uma abordagem *offline* é a única maneira de se obter uma experiência ubíqua, ou seja, permitindo a operação da aplicação em qualquer cenário de conectividade [Holt 2015].

### 2.2.1. CouchDB/PouchDB

CouchDB é um banco de dados NoSQL criado em 2005 [CouchDB 2017] e atualmente mantido pela Apache Software Foundation [PouchDB 2018]. Já o PouchDB é uma implementação JavaScript do CouchDB que tem por objetivo emular a API do CouchDB com uma fidelidade de execução no navegador. Assim, o CouchDB é usado no lado do servidor e o Pouch no próprio aplicativo, por meio de API. Isso possibilita a sincronização de ambos quando o aplicativo estiver *online*. Segundo [Lawson 2017] , o grande poder do CouchDB é a sincronização. O autor reforça que o CouchDB não é um banco de dados, mas um motor de sincronização, ou seja, ele é uma forma eficiente de transferir dados de

um lugar a outro, enquanto gerencia conflitos e revisões. Portanto, a persistência dos dados a nível local se dá com o uso do PouchDB enquanto seu mecanismo de sincronização envia os dados a um servidor CouchDB.

Na estrutura de armazenamento do Pouch, quando comparado com um banco relacional, a equivalência no que diz respeito a tabela é com o próprio *database*, de modo que para cada conjunto de dados, por exemplo de cidades, é criado um *database* no Pouch, já um registro (linha) é um *document* e uma coluna é um *field*. Todo documento tem dois identificadores denominados *\_id* e *\_rev*, sendo o *\_id* o identificador e o *\_rev* um número de revisão para cada alteração no documento.

### 2.2.2. Firebase Firestore

O Cloud Firestore é uma banco de dados de nuvem NoSQL flexível e escalonável para armazenar e sincronizar dados para o desenvolvimento no lado do cliente e do servidor [Google 2018]. O foco é o desenvolvimento para dispositivos móveis, Web e servidores a partir do Firebase e do Google Cloud Platform. Ele mantém seus dados em sincronia em aplicativos cliente por meio de *listeners* em tempo real e oferece suporte *offline* para dispositivos móveis e Web, permitindo criar aplicativos responsivos que funcionem independentemente da latência da rede ou da conectividade com a Internet. Com o passar dos anos a plataforma cresceu muito, se tornando uma ferramenta que vem sendo bastante utilizada em função da quantidade de serviços oferecidos, bem como, da facilidade de implementação. O Cloud Firestore estrutura os dados por meio de coleções e documentos, sendo possível adotar a abordagem de criar subcoleções. Os documentos sempre pertencem a uma coleção e armazenam dados na forma chave e valor. Um identificador (*id*) pode ser gerado automaticamente para cada documento.

### 2.2.3. Ionic

O Ionic Framework pertence a abordagem híbrida baseada na Web, onde componentes da interface do usuário são estruturados usando tecnologias da web, incluindo HTML, CSS e Javascript.

O desenvolvimento híbrido fundamenta em implementar sistemas que funcionem em inúmeras plataformas, tais como: IOS, Android, Desktop e Progressive Web Apps (PWA). O lema em frameworks multiplataforma é: “*Write once, run anywhere*”, ou seja, com uma única base de código, tem-se aplicações que funcionem através de um navegador e possuem acesso à plataforma quase à nível nativo. Com o crescimento da abordagem PWA, o Framework Ionic foi um dos primeiros a incorporar a possibilidade de gerar a aplicação para PWA. Como o Ionic é construído na Web, os aplicativos Ionic podem funcionar em qualquer lugar em que a Web é executada, inclusive como uma PWA.

## 3. Análise das ferramentas de sincronização

Com o intuito de melhor conhecer e avaliar as duas ferramentas para armazenamento e sincronização de dados selecionadas, foram desenvolvidos protótipos da aplicação *Veículos BRBM App* em cada uma delas. A seguir são apresentadas vantagens e desvantagens verificadas em cada ferramenta.

O desenvolvimento com o CouchDB/PouchDB exigiu inicialmente a sua instalação, que foi realizada de forma local. Já o Firestore é um BaaS (BackEnd As A Service), ou seja um serviço disponibilizado em que toda a estrutura do *backend* como configuração de servidor, integração com banco de dados, sistema de *push notification* e outros serviços, que fazem parte do *backend*, estão completamente prontos para se integrar com o aplicativo. Assim, foi necessário apenas criar uma conta na plataforma.

Com relação a autenticação, o Firebase Auth oferece diversos métodos de autenticação, como e-mail/senha, provedores de terceiros, como o Google ou Facebook, Login Anônimo, Confirmação via SMS, ou o uso direto do seu sistema de contas. O gerenciamento de usuários se torna simples e seguro. No projeto a necessidade era de criação de usuário por meio de e-mail/senha e o processo de autenticação foi bem simplificado de ser desenvolvido. No CouchDB o processo de autenticação é bem mais complexo, gerando grandes dificuldades na implementação. Também se verificou que uma desvantagem do CouchDB é não oferecer mecanismo de controle de acesso a nível de documentos. Nesse caso, para cada usuário, é necessário um banco de dados específico, criado dinamicamente pela aplicação.

Em relação ao processo de sincronização de dados, ambas as ferramentas funcionam adequadamente. Foi possível observar que os dados atualizados são sincronizados em todos os dispositivos conectados em segundos. Além disso, os dados permanecem disponíveis quando o aplicativo fica *offline*, o que oferece uma ótima experiência do usuário, independentemente da conectividade de rede.

Um aspecto que gera uma desvantagem ao Firebase é o fato de haver limites para o acesso gratuito. Porém, avaliando a necessidade da aplicação desenvolvida, verificou-se que o uso ficará dentro dos limites estipulados para o acesso gratuito, sendo 10 mil autenticações/mês, 1 GB de armazenamento, 20 mil gravações de documentos/dia, 50 mil leituras de documento/dia e 20 mil exclusões de documento/dia.

Por fim, após a avaliação desses diferentes critérios, como também podem ser observados na Tabela 1, optou-se por empregar a plataforma Firebase Firestore para o desenvolvimento da aplicação *Veículos BRBM App*, considerando ainda algumas outras vantagens: já possui uma estrutura pronta, rápida implementação, segura, facilmente escalável, configuração intuitiva, múltiplas ferramentas disponibilizadas; e grande base de usuários, com muita documentação disponível para sanar possíveis dificuldades. No caso do CouchDB/PouchDB seria necessário um servidor para fazer a sua instalação e configuração, além de preocupação de se manter um servidor, tanto em relação a custo, quanto de tempo para sua configuração, manutenção e futuras atualizações.

## **4. Arquitetura da solução PWA**

Nesta seção são descritos os requisitos das aplicações *BRBM App* e *Veículos BRBM App* e apresentada a arquitetura proposta para a solução.

### **4.1. BRBM App**

A aplicação *BRBM App* gerencia os dados básicos dos módulos que compõem o sistema gerencial do Batalhão Rodoviário, tais como: cadastro de pessoal controlando os dados de todo o efetivo, órgãos, setores e cidades, controle de concessão de medalha, cursos e indisponibilidades, registro das solicitações de fardamento sendo possível cada envolvido

**Tabela 1. Comparativo entre as ferramentas**

<b>Crítérios</b>	<b>CouchDB/PouchDB</b>	<b>Firebase Firestore</b>
Instalação/ Configuração	Precisa ser instalado e configurado	Estrutura pronta; Configuração intuitiva
Autenticação	Criação de Usuário ( login e senha ); com difícil implementação	Diversos mecanismos disponíveis; implementação facilitada
Controle de acesso aos dados	não oferecer mecanismo de controle de acesso em nível de documentos	Pode restringir o acesso às coleções
Mecanismos de sincronização	Funciona adequadamente	Funciona adequadamente
Documentação	Documentação Restrita	Ampla Documentação
Custo	Totalmente Gratuito	Limites para acesso gratuito, porém bem amplos

manter atualizado o registro do seu manequim, previsão de férias, registro de materiais em cautela, controle de frotas.

Neste aplicativo são disponibilizados inúmeros relatórios gerenciais que auxiliam os gestores do BRBM na tomada de decisão, facilitando o acesso as informações com a possibilidade de aplicação de filtros específicos para cada relatório.

#### **4.2. Veículos BRBM App**

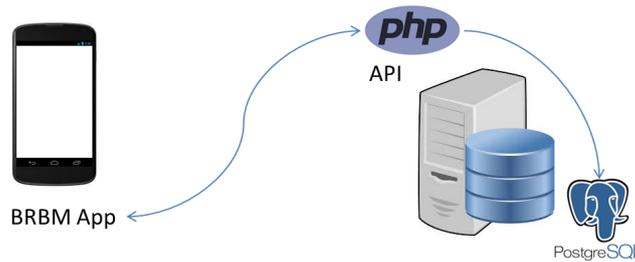
Essa aplicação tem por objetivo registrar as saídas dos veículos da corporação. Assim, o usuário (policial) que retirar um veículo deve fazer o registro de saída informando data e hora de saída, motivo e quilometragem inicial. Ao entregar o veículo faz o registro de entrega com data e hora de entrega e quilometragem final. Desta forma, é possível manter um histórico de todas as movimentações dos veículos. A aplicação pode funcionar *offline*, utilizando a persistência do *browser* para o armazenamento das informações até que haja conexão para que, então, essas informações sejam enviadas para a Nuvem.

#### **4.3. Arquitetura**

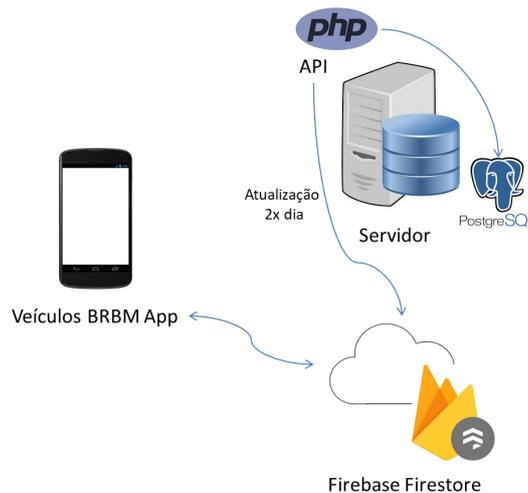
Com os requisitos das aplicações e definição da ferramenta a ser utilizada para o desenvolvimento da aplicação *Veículos BRBM App*, foram então definidas e detalhadas as interações entre as aplicações e as bases de dados.

A aplicação *BRBM App* persiste os dados em uma base de dados PostgreSQL, que fica em um servidor remoto, por meio de uma API desenvolvida em PHP e que usa JWT (JSON Web Token). O JWT é um padrão que define como transmitir e armazenar objetos JSON de forma compacta e segura entre diferentes aplicações. Os dados nele contidos podem ser validados a qualquer momento pois o *token* é assinado digitalmente, garantindo assim a segurança dos dados. A Figura 1 exemplifica essa interação, destacando que a PWA pode estar ou não hospedada no mesmo servidor do banco de dados.

A aplicação *Veículos BRBM App* também é uma PWA, porém, como já mencionado anteriormente faz uso da plataforma Firebase Firestore que traz recursos para gerenciar a persistência quando operando *offline* e posterior sincronização de dados quando tiver a conexão re-estabelecida. Nesse caso, a Figura 2 ilustra como ocorre a interação entre a aplicação e o Firestore e também deste último com a base de dados PostgreSQL.



**Figura 1. Interação da aplicação *BRBM App* e o banco de dados**



**Figura 2. Interação aplicação *Veículos BRBM App* e banco de dados**

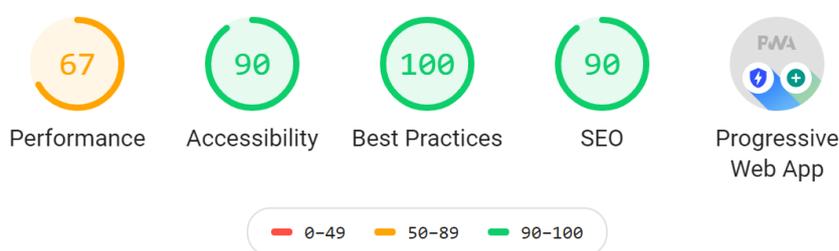
Cabe destacar alguns aspectos sobre os dados armazenados e atualizados em cada um dos bancos de dados. A base de dados PostgreSQL persiste dados de quarenta tabelas. Todos esses dados são mantidos (inclusão, alteração, exclusão) pela aplicação *BRBM App*, com exceção da tabela *controle\_veiculo* que é apenas fonte de consulta para a aplicação. Essa tabela (*controle\_veiculo*) tem seus dados gerados justamente pela aplicação *Veículos BRBM App*. Assim, a integração das bases de dados PostgreSQL e Firebase Firestore é realizada através de um *script* PHP que é programado para ser executado em dois horários diários. Esse *script* executa as seguintes ações: a) atualiza dados no Firestore: dados usados apenas para consulta, tais quais são as coleções pessoas, órgãos e veículos; b) Envia dados dos *controle\_veiculos* do Firestore para PostgreSQL (dados marcados como não atualizados).

## 5. Avaliação da aplicação *Veículos BRBM App*

No sentido de verificar se os aspectos relacionados a uma PWA foram contemplados, utilizou-se a ferramenta Lighthouse para avaliar a *Veículos BRBM App*. A Lighthouse [Lighthouse 2019] é uma ferramenta desenvolvida pelo Google que roda como extensão do Google Chrome para avaliar aplicativos da Web em relação à sua qualidade de construção, coletando métricas de desempenho e *insights*, com base em uma série de auditorias. A partir disto, a qualidade é então medida em cinco categorias, sendo elas: Performance (Desempenho), Accessibility (Acessibilidade), Best Practices (Boas práticas), SEO (Otimização para mecanismos de busca) e Progressive Web App. O Lighthouse

afere uma pontuação única entre 0 e 100 para cada uma dessas cinco categorias. Ela pode ser usada como uma ferramenta de verificação para os requisitos de uma Progressive Web App.

A Figura 3 apresenta as pontuações alcançadas pela aplicação, com destaque a categoria PWA onde a aplicação atendeu todos os requisitos para ser caracterizada como tal. Três condições devem ser asseguradas para considerar uma aplicação web móvel como uma PWA, a saber: (i) se ela é servida por HTTPS (esse é um requisito para evitar ataques intermediários), (ii) se ela é fornecida com um manifesto de aplicativo da web, declarando seus metadados (nome, ícones, URL de base) e (iii) se usa service workers, ou seja, um conjunto de APIs para permitir que os desenvolvedores programaticamente armazenem em cache e pré-carreguem recursos e dados.



**Figura 3. Avaliação da Veículos BRBM App usando a ferramenta Lighthouse**

## 6. Considerações Finais

Este trabalho apresentou o desenvolvimento de duas aplicações empregando a abordagem PWA, bem como o resultado de um comparativo realizado entre duas ferramentas para armazenamento e sincronização de dados: CouchDB/PouchDB e Firebase Firestore. O comparativo entre tais ferramentas considerou diversos aspectos tais como instalação e configuração, autenticação, controle de acesso aos dados, mecanismo de sincronização, documentação e custo. Por fim, a ferramenta Firebase Firestore foi utilizada no desenvolvimento.

A solução desenvolvida alcançou os requisitos definidos inicialmente, tratando-se de uma aplicação com características bastante peculiares uma vez que está integrada com uma aplicação mais robusta que gerencia muitos dados. Dessa forma, arquiteturas e plataformas específicas para cada solução foram adotadas. O desenvolvimento com a abordagem PWA forneceu um acesso fácil para usuários, pois inclui o *design* responsivo da web que pode ser acessado por meio de navegador de *desktop* e dispositivos móveis. Além disso, com base nos testes feitos usando a ferramenta *Lighthouse*, a aplicação desenvolvida obteve resultados satisfatórios.

## Referências

- Biørn-Hansen, A., Majchrzak, T. A., and Grønli, T.-M. (2017). Progressive web apps: The possible web-native unifier for mobile development. In *WEBIST*, pages 344–351.
- CouchDB (2017). <http://couchdb.apache.org/>. Acesso em: 11 dez 2017.

- Cselle, G. (2012). Every step costs you 20% of users. <http://blog.gaborcselle.com/2012/10/every-step-costs-you-20-of-users.html>. Acesso em: 2 out 2017.
- DA-14 (2017). Eight reasons to consider progressive web app development. <https://da-14.com/blog/eight-reasons-consider-progressive-web-app-development>. Acesso em: 2 out 2017.
- Dal’Pizzol, M. and Rossetto, A. G. (2018). Offline-first multidatabase progressive web apps com pouchdb/couchdb. In *I Congresso de Tecnologia da Informação - IFSUL Passo Fundo*, pages 58–67. Acesso em: 6 out 2019.
- Feyerke, A. (2013). Say hello to offline first. <http://hood.ie/blog/say-hello-to-offline-first.html>. Acesso em: 2 out 2017.
- Google (2018). Firebase firestore. <https://firebase.google.com/docs/firestore>. Acessado em: 15 jun 2018.
- Holt, B. (2015). Offline-first apps with pouchdb. <https://www.youtube.com/watch?v=7L7esHWAjSU>. Acesso em: 31 out 2017.
- Lawson, N. (2017). Pouchdb and couchdb: an interview with nolan lawson. <https://blog.couchdb.org/2017/04/04/pouchdb-couchdb-an-interview-with-nolan-lawson>. Acesso em: 15 out 2017.
- Lighthouse (2019). Lighthouse. <https://developers.google.com/web/tools/lighthouse/?hl=pt-br/>. Acesso em: 12 aug 2019.
- Lopes, S. (2016). *Aplicações mobile híbridas com Cordova e PhoneGap*. Editora Casa do Código.
- PouchDB (2018). Introduction to pouchdb. <https://pouchdb.com/guides>. Acesso em: 15 out 2018.
- Russel, A. and Berriman, F. (2015). Progressive web apps: escaping tabs without losing our soul. Acesso em: 12 out 2017.

# O uso de aplicativo educacional na Educação Profissional e Tecnológica: o que os dados revelam

Angélica Tommasini<sup>1</sup>, Ana Sara Castaman<sup>1</sup>, Márcio Luis Vieira<sup>1</sup>, Ângela Helena Peretti<sup>1</sup>, Luciane Inocente<sup>1</sup>, Lis Ângela de Bortoli<sup>1</sup>, Daniela Hoffmann<sup>2</sup>

<sup>1</sup> Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS)– Sertão – RS – Brasil

<sup>2</sup> Secretaria Municipal de Educação – Prefeitura Municipal de Getúlio Vargas – Getúlio Vargas - RS - Brasil

angelicatommasinil@gmail.com, ana.castaman@sertao.ifrs.edu.br, marcio.vieira@sertao.ifrs.edu.br, angelahperetti2001@gmail.com, lucianeinocente@gmail.com, lis.debortoli@sertao.ifrs.edu.br, danihoffmann1980@gmail.com

**Abstract.** *This study aims to report the use of the Kahoot educational application in Vocational and Technological Education (EPT). Therefore, the methodology used refers to a qualitative and quantitative approach study, based on bibliographic and case study procedures. Thus, a structured questionnaire was applied to a class enrolled in the course of Water and Soil Management, in the Technical Course in Integrated Farming High School, from the Federal Institute of Education, Science and Technology of Rio Grande do Sul, Campus Sertão. With the obtained results, it is concluded that the use of the application added even more in the construction of the knowledge with the students of the EPT, because through the used resource the pedagogical work was strengthened and the learning of the studied content facilitated interactively.*

**Resumo.** *Este estudo tem por objetivo relatar o emprego do aplicativo educacional Kahoot na Educação Profissional e Tecnológica (EPT). Para tanto, a metodologia empregada remete a um estudo de abordagem qualitativa e quantitativa, pautado em procedimentos bibliográficos e de estudo de caso. Assim, aplicou-se um questionário estruturado em uma turma matriculada na unidade curricular de Manejo de Água e Solo, no curso Técnico em Agropecuária Integrado ao Ensino Médio, do Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul, Campus Sertão. Com os resultados obtidos, conclui-se que o uso do aplicativo agregou ainda mais na construção do conhecimento junto aos estudantes da EPT, pois por meio do recurso utilizado fortaleceu-se o trabalho pedagógico e facilitou a aprendizagem do conteúdo estudado de maneira interativa.*

## 1. Introdução

Atualmente, com o impacto das tecnologias digitais, observa-se uma grande influência dos aplicativos para fins educacionais. Para [Soad 2017], o uso dos aplicativos

educacionais vêm possibilitando aos estudantes e educadores maior flexibilidade e comodidade no andamento do trabalho pedagógico.

O *kahoot* sendo um aplicativo educacional atrativo, de fácil acesso e gratuito pode colaborar enquanto um recurso utilizado pelos docentes em seu trabalho pedagógico na Educação Profissional e Tecnológica (EPT). Assim, este estudo de caso, de técnica bibliográfica e de campo, tem por objetivo relatar o emprego do aplicativo educacional *Kahoot*, na Educação Profissional e Tecnológica. Refere-se a resultados parciais de uma pesquisa que vem sendo desenvolvida no Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS) - *Campus* Sertão no projeto de pesquisa “Estratégias de Aprendizagem para construção do conhecimento na Educação Profissional e Tecnológica”.

O estudo está dividido em três momentos acrescidos de introdução mais considerações finais. No primeiro momento, conceitua-se acerca do aplicativo *kahoot* e seu emprego na educação. No segundo momento, apresenta-se a metodologia utilizada no estudo e, no terceiro e último momento, analisa-se e discute-se os dados da pesquisa.

## **2. *Kahoot*: um recurso no processo de ensino e aprendizagem**

Na contemporaneidade, observa-se cada vez mais o surgimento e a inserção de novos recursos de ensino no trabalho pedagógico visando à facilitação da construção das aprendizagens. Estas aliadas às tecnologias digitais têm servido de aporte na área educacional. A exemplo, o emprego da sala de aula invertida, com a construção de conhecimentos, a partir de vídeos aulas, materiais digitais e gamificação. A gamificação caracteriza-se pelo uso de elementos dos jogos, fora do contexto desses, como na educação [Coli *et.al* 2017].

Inúmeros autores têm publicado experiências de elaboração e aplicação de jogos em sala de aula, bem como da gamificação. Os jogos educativos digitais, além de inovadores, podem alcançar o maior número de participantes, uma vez há diversos jogos para *downloads* gratuitos [Andrade & Rezende 2018]. Logo, o aplicativo *Kahoot* tem constituído-se como um forte dispositivo para o trabalho pedagógico no processo de ensino e aprendizagem, inclusive, na modalidade de ensino da EPT.

Destacam [Correia & Santos 2017] que o aplicativo educacional *kahoot* é uma plataforma gratuita, disponível a partir de uma computador, *smartphone* ou *tablet*, de fácil acesso, que permite o acréscimo de ações de pergunta/resposta. Portanto, compreende-se que o *kahoot* por ser um jogo digital, propicia espaços de aprendizagem colaborativa, entre docentes e estudantes criando, trocando e valorizando os conhecimentos construídos durante o jogo [Andrade & Rezende 2018]. Ou seja, dos conteúdos estudados na unidade curricular.

O *Kahoot* é uma ferramenta que pode ser utilizada em uma sala de aula invertida ou *Flipped Classroom*, em que o educando tem a possibilidade de obter o conhecimento fora do espaço escolar. Percebe-se que, estas estratégias proporcionam maior independência e compreensão dos educandos no decorrer das atividades propostas dentro e fora da sala [Toscano & Gil 2018]. [Toscano & Gil 2018] enfatizam que para o andamento da aula, o docente necessitará organizar os conteúdos/materiais e fornecer

aos estudantes, orientações no decorrer das atividades propostas, já que este previamente deverá estudar a temática proposta, fazendo uma prévia para as aulas.

Diante da relevância do uso deste aplicativo para a educação, analisa-se nas próximas seções a aplicação do mesmo na EPT. Para tanto, a seguir serão discutidos os caminhos metodológicos deste estudo.

### **3. Metodologia**

Os rumos metodológicos que orientam este estudo consistiram em explorar acerca do emprego do aplicativo educacional *kahoot* no trabalho pedagógico, especificamente, na unidade curricular de Manejo de Água e Solo, no curso Técnico em Agropecuária Integrado ao Ensino Médio, do Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul, *Campus Sertão*, avaliando sua cooperação no processo de ensino e aprendizagem. A pesquisa caracteriza-se por uma abordagem qualitativa e quantitativa. [Roesch 2009] destaca que uma pesquisa qualitativa almeja a compreensão do contexto, possibilitando melhores condições de avaliação. Ainda, a presente investigação de objetivos exploratórios está pautada em um estudo bibliográfico e por um estudo de caso.

O estudo foi realizado nos meses de março e abril de 2019. A turma é formada por 30 estudantes matriculados, porém a população participante da pesquisa corresponde a 26 alunos da referida turma.

Para o andamento e condução da atividade atentou-se para alguns passos: a) a seleção do conteúdo discutido em aula; b) a escolha dos materiais; c) revisão da bibliografia; d) a evolução da atividade. A coleta de dados se desenvolveu por meio de um questionário com perguntas fechadas (dicotômicas e múltiplas escolhas).

O método de coleta de dados aplicado foi questionário viabilizado em ambiente *web*, com preenchimento *online*, sendo remetido para o *e-mail* de todos os integrantes da turma pesquisada. Para a análise dos dados, apresentou-se as estatísticas e refletiu-se acerca dos dados, a partir do referencial teórico que serviu de base para elaboração do estudo.

### **4. Resultados e discussões: o que os dados revelam**

O aplicativo educacional *kahoot* foi usado no andamento das aulas da unidade curricular de Manejo de Água e Solo, a qual totaliza 80 horas e consta no primeiro ano, na matriz curricular do curso Técnico em Agropecuária Integrado ao Ensino Médio. Tem por objetivo monitorar, planejar e organizar a exploração, o manejo e conservação do solo, de acordo com a sua formação, características e propriedades. O docente responsável pela unidade curricular atua na mesma há 04 anos, na instituição estudada. Porém, no último ano tem utilizado o *kahoot* em seu trabalho pedagógico.

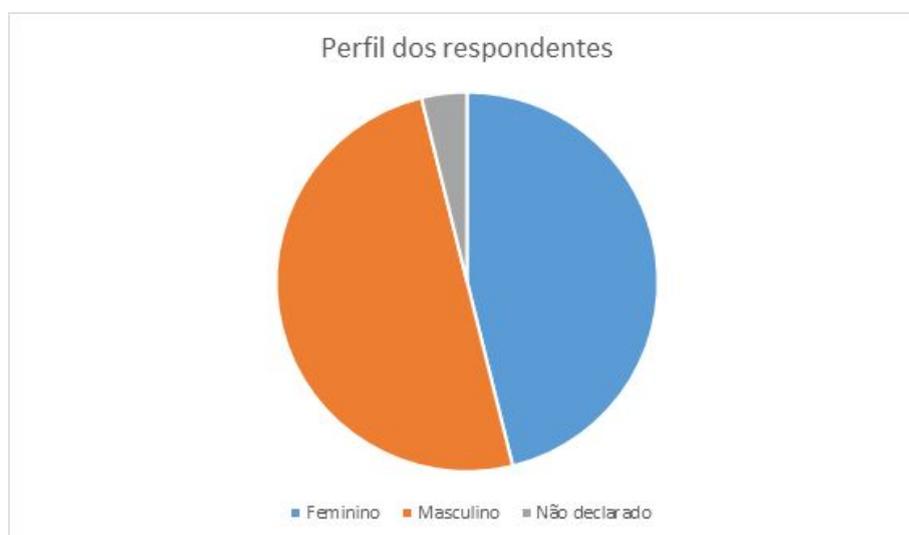
Destaca-se que objetivo desta atividade não foi apenas de adentrar na temática proposta, mas disponibilizar auxílio/subsídios para construção de uma aprendizagem intencional e significativa, que faça sentido ao educando.

Logo após ocorrer a aula expositiva-dialogada com assistência do livro didático que trata do conteúdo do componente curricular estudado, aplicou-se o *kahoot*. Para

conduzir a atividade orientou-se que os educandos realizassem o acesso à plataforma do jogo (*kahoot*) com aparelho celular de modo individual. Para o trabalho pedagógico foram criadas 20 perguntas com tempo de 30 segundos para resolução de cada pergunta. No decorrer do jogo, os educandos manifestaram interesse e atenção ao responder às questões propostas com rapidez.

Após a aplicação do *quiz* (jogo *kahoot*), os estudantes foram chamados a realizarem um questionário em ambiente *web*<sup>1</sup> para verificar a eficácia da atividade. O questionário continha 14 questões fechadas, sendo encaminhado para 27 alunos presentes na data da aplicação do *kahoot*. Contudo a discussão e análise dos dados baseiam-se na população de 26 respondentes, pois obteve-se o retorno destes.

O perfil dos respondentes é composto por sua maioria do gênero masculino (13), seguido pelo gênero feminino (12) e um estudante não declarou.



**Figura 1: Gráfico gênero dos estudantes**

Todos os educandos possuem *smartphone*, sendo que 88,46% possuem internet pessoal. Os demais utilizam a internet via *wi-fi*. Destaca-se que 80,77% possuem computador, *notebook* ou *tablet*. A partir desta análise percebe-se que os educandos estão cada vez mais conectados à era tecnológica, ou seja, as tecnologias digitais. Portanto, o docente necessita estar atento e utilizando destas inovações para beneficiar o seu trabalho pedagógico. O aparelho de celular hoje é um significativo instrumento para autonomia de adolescentes e jovens, pois possibilita a comunicação ao receber ou enviar mensagens criando um pleno processo de interatividade [Weigelt 2013].

Os dados revelam ainda que, os educandos desconheciam o aplicativo *kahoot*. Somente 15,88% dos entrevistados já tinha algum conhecimento sobre o aplicativo e 84,62% o desconheciam. Quando perguntados se já tinham algum contato com uma técnica semelhante ao *kahoot*, 69,23% já tinham alguma experiência semelhante e 30,77% nunca tiveram contato. Analisa-se que mesmo diante da inserção dos estudantes

---

<sup>1</sup> Para esta atividade foi enviado via plataforma do *google* o questionário

às tecnologias digitais, ainda não empregadas ações pedagógicas com tal recurso, já que muitos estudantes não conhecem este aplicativo.

É imprescindível implementar o uso das tecnologias digitais no processo de ensino e aprendizagem. O docente confronta-se com um mundo tecnológico e necessita buscar novas possibilidades frente a esta realidade no espaço educacional. No entanto, verifica-se que as tecnologias apesar de poder facilitar o processo de ensino, também têm se tornado um barreira para muitos docentes, já que muitos não as utilizam em sala de aula [Silva *et.al* 2016]. Como observa-se os dados a seguir.

Dos educandos entrevistados 15,88% já tinha vivenciado alguma técnica de aprendizagem em outros componentes curriculares e 84,62% nunca tiveram essa vivência. Quando questionados sobre a experiência do uso do aplicativo durante as aulas, os estudantes responderam que 92,15% achou muito interessante, e 3,85% achou indiferente. Dos estudantes pesquisados, 65,38% entenderam que a atividade teve desenvolvimento adequado, 11,54% indiferente e 23,8% acharam pouco adequado. Portanto, os aplicativos educacionais podem constituir-se como um instrumento no trabalho pedagógico. Ao desfrutar dos dispositivos móveis, o estudante estará aperfeiçoando diversas habilidades e competências, além de estimular o desenvolvimento do raciocínio, planejamento e organização [Mousquer & Rolim 2011].

Quanto a colaboração para o raciocínio, os dados revelam que 88,46% acham que a atividade contribuiu e 11,54% que é uma ação indiferente. Quando perguntados se o aplicativo *kahoot* pode ser uma ferramenta que facilita o processo de aprendizagem, 88,46% marcaram que sim, 7,69% que era indiferente e 3,85% apontaram que era pouco. Na indagação quanto ao uso do *kahoot* na incorporação do conhecimento, 88,46% entendem que muito, 7,69 que indiferente e 3,85% pouco.

O aplicativo educacional possibilita ao educando uma aprendizagem por descobertas e superação de desafios com desenvolvimento de habilidades de forma prazerosa e divertida [Andrade & Rezende 2018]. Ainda, o uso da plataforma *kahoot* proporciona um ambiente desafiador, instigando o estudante para a aprendizagem.

Ao analisar os dados da pesquisa percebe-se que o emprego do aplicativo *kahoot* no espaço escolar pode fortalecer o processo de ensino e aprendizagem, envolvendo estudantes e docentes de modo inovador por meio de jogos. O *feedback* recebido pelos educandos é um dado valioso para analisar o trabalho pedagógico que o docente utiliza em sala de aula bem como, a construção da aprendizagem por parte dos estudantes.

## 5. Conclusão

Este estudo buscou analisar a importância do uso de aplicativos para fins educacionais. Em especial, o *locus* de análise foi do aplicativo *kahoot* como recurso para o processo de ensino e aprendizagem na educação profissional e tecnológica, em turmas do curso Técnico em Agropecuária integrado ao ensino médio, na unidade curricular de Manejo de Água e Solo.

Os frutos deste estudo revelam que o uso do aplicativo educacional *kahoot* tem sido bem aceito pelos estudantes. Ainda, tem apoiado o processo de ensino e

aprendizagem na EPT, e quando bem aplicados tornam-se grandes colaboradores para a construção das aprendizagens significativas.

## Referências

- Andrade, C. D. F.; Rezende, A. L. A. (2018). As potencialidades do uso do kahoot nas práticas pedagógicas no processo de alfabetização. Disponível em:[http://www.editorarealize.com.br/revistas/conedu/trabalhos/TRABALHO\\_EV117\\_MD1\\_SA19\\_ID10236\\_09092018111954.pdf](http://www.editorarealize.com.br/revistas/conedu/trabalhos/TRABALHO_EV117_MD1_SA19_ID10236_09092018111954.pdf). Acesso em: 14 Out 2019.
- Coil, D. A.; Ettinger, C. L.; Eisen, J. A. (2017). Gut Check: the evolution of an educational board game. *PLOS Biology*, 15(4).
- Correia, M.; Santos, R. (2017). A aprendizagem baseada em jogos online: uma experiência de uso do kahoot na formação de professores. Disponível em: [https://repositorio.ipsantarem.pt/bitstream/10400.15/2433/1/siie-cied\\_2017\\_artigo%20onas%20atas%20pt.pdf](https://repositorio.ipsantarem.pt/bitstream/10400.15/2433/1/siie-cied_2017_artigo%20onas%20atas%20pt.pdf) . Acesso em: 1 out. 2019.
- Mousquer, T.; Rolim, C. O. (2011). A utilização de dispositivos móveis como ferramenta pedagógica colaborativa na Educação Infantil. In: Anais do II Simpósio da Tecnologia da Informação. Santo Ângelo, 2011. Disponível em: <http://www.santoangelo.uri.br/stin/Stin/trabalhos/11.pdf> Acesso em: 01 out. 2019.
- Roesch, S. M. A. (2009). Projetos de estágio e de pesquisa em administração: guia para estágios, trabalhos de conclusão, estudos de caso. 3 ed. São Paulo: Atlas.
- Silva, I. C.; Prates, T. S.; Ribeiro, L. F. S.; (2016). As Novas Tecnologias e aprendizagem: desafios enfrentados pelo professor na sala de aula. Disponível em: <https://periodicos.ufsc.br/index.php/emdebate/article/download/1980-3532.../33788>. Acesso em: 01 out. 2019.
- Soad, G. W. (2017). Avaliação de qualidade em aplicativos educacionais móveis. Disponível em: [www.teses.usp.br/teses/disponiveis/55/55134/tde.../GustavoWilliansSoad\\_revisada.pdf](http://www.teses.usp.br/teses/disponiveis/55/55134/tde.../GustavoWilliansSoad_revisada.pdf) Acesso em: 01 out. 2019.
- Toscano, C. S. S; Gil, H. M. P .T; (2018). O contributo da aplicação digital do kahoot!no processo de ensino/aprendizagem no âmbito da prática supervisionada no 1º ciclo de ensino básico. Disponível: [https://repositorio.ipcb.pt/bitstream/10400.11/6103/1/R.%20E.%20-%20FINAL\\_.pdf](https://repositorio.ipcb.pt/bitstream/10400.11/6103/1/R.%20E.%20-%20FINAL_.pdf) . Acesso em: 01 out. 2019.
- Weigelt, D. (2013). Os jovens e o celular: O poder da comunicação móvel. Disponível em:<http://portalintercom.org.br/anais/sul2013/resumos/R35-0425-1.pdf> Acesso em: 01 Out. 2019.

# Aprenda Talian: Plataforma colaborativa para aprendizado do dialeto Talian

Leandro José Grando<sup>1</sup>, Anubis Graciela de Moraes Rossetto<sup>1</sup>

<sup>1</sup>Instituto Federal Sul-rio-grandense (IFSUL) Câmpus Passo Fundo  
Estrada Perimetral Leste, 150 – CEP: 99064-440 – Passo Fundo – RS – Brasil

leandro-grando@hotmail.com, anubisrossetto@gmail.com

**Abstract.** *Aiming to spread Talian dialect knowledge through mobile technologies, this paper presents the development of an app, called Aprenda Talian, available for Android and iOS, that allows translations of Talian words and expressions with texts and audios. Thus, the application provides learning of this dialect to mobile users and provides the opportunity for inclusion for those who have Talian as their main language. Since a dialect is created in a decentralized and constant way, the application encompasses the possibility for users to contribute by adding or modifying translations. To manage contributions, a site with an API to communicate with the app and with restricted access to people moderating contributions is also presented. User evaluation results indicate that the application can contribute to learning the Talian dialect.*

**Resumo.** *Com o intuito disseminar o conhecimento do dialeto Talian por meio de tecnologias móveis, esse trabalho apresenta o desenvolvimento de um aplicativo intitulado Aprenda Talian, disponível para Android e iOS, que possibilita traduções de palavras e expressões do português para o Talian, com textos e áudios. Dessa forma, o aplicativo proporciona o aprendizado desse dialeto aos utilizadores de dispositivos móveis e oportuniza a inclusão para os que tem o Talian como língua principal. Por conta de um dialeto ser criado de forma descentralizada e constante, o aplicativo engloba a possibilidade dos usuários contribuírem, acrescentando ou modificando traduções. Para gerenciar as contribuições, também é apresentado um site com API para se comunicar com o aplicativo e com acesso restrito às pessoas que moderam as contribuições. Os resultados de avaliação com usuários indicam que o aplicativo pode contribuir para o aprendizado do dialeto Talian.*

## 1. Introdução

A cultura é tida como um dos principais fatores que distinguem a identidade de um povo, ela pode ser observada nos costumes, na arte, na arquitetura, na religiosidade, nas tradições e na fala de um determinado grupo social. Ela é a soma de padrões de comportamentos humanos. A compreensão de diversas culturas nos permitem entender subjetivamente e objetivamente, como se integrar numa determinada sociedade, como foram formadas suas leis, seus costumes, seu entretenimento, sua política, sua religiosidade e sua lógica mercantil. Dada a importância da cultura, a sua preservação é necessária.

O Talian é um dialeto (variedade linguística) que se originou no sul do Brasil devido a miscigenação entre italianos e brasileiros, após um movimento de migração de

italianos para o Brasil iniciado por volta de 1930 [Corrêa 2019]. Na época, a Itália não possuía uma língua oficial (o Italiano Standard) e cada região possuía um dialeto que a distinguia das demais. Assim, os italianos que chegavam ao Brasil falavam diferentes dialetos mas o que predominava era o dialeto Vêneto, de onde vinham a maioria dos imigrantes. Com a mistura do dialeto vênето com o português, seus desdobramentos e influências, criou-se um novo dialeto diferente do dialeto da região de Vêneto na Itália, cunhou-se assim o termo Talian para referenciá-lo. A grande quantidade de brasileiros descendentes de italianos que falam Talian fez com que em 2014, o Talian fosse reconhecido como patrimônio nacional imaterial brasileiro. O Talian é preservado sendo passado de geração a geração, através da fala, de rádios que fazem parte da sua programação em Talian, de livros, matérias de jornais, teatros e outros. Porém, esses meios de difusão estão a cada dia atingindo menos os jovens, e consequentemente fazendo com que o Talian aos poucos caia em desuso.

Apesar dos dispositivos móveis trazerem inúmeros benefícios, alguns grupos de pessoas, principalmente os mais idosos, ainda não fazem o uso constante desses dispositivos. Isso cria um distanciamento na comunicação entre os utilizadores e os não utilizadores. Esse distanciamento é ainda maior para um grupo de brasileiros que tem como seu principal idioma o dialeto Talian (estima-se que haja 500 mil brasileiros que falam Talian).

Considerando que os mais jovens consomem bastante conteúdo por meio da internet e de smartphones, foi escolhido desenvolver uma aplicação multiplataforma que auxilie no aprendizado básico do Talian, dessa forma, possibilitando a compreensão desse dialeto aos utilizadores de dispositivos móveis, trazendo a inclusão para os que tem o Talian como língua principal e assim criando uma ponte entre a cultura dos primeiros imigrantes italianos e as futuras gerações.

Baseado nisso, esse trabalho apresenta o desenvolvimento de um aplicativo mobile híbrido (para Android e iOS) intitulado “Aprenda Talian” desenvolvido com o Framework Ionic [Framework 2019]. O aplicativo conta com traduções de palavras e expressões do português para o Talian, organizadas em categorias. Além do aplicativo, também foi desenvolvido um site com acesso restrito onde os administradores podem cadastrar categorias, palavras, traduções e áudios, além de moderar contribuições de usuários.

O artigo está organizado da seguinte forma: a Seção 2 apresenta os trabalhos relacionados. A Seção 3 detalha o aplicativo Aprenda Talian, bem como apresenta detalhes do site de gerenciamento. A Seção 4 apresenta os resultados obtidos a partir de uma avaliação com usuários. Por fim, apresenta as considerações finais com sugestões de trabalhos futuros.

## **2. Trabalhos Relacionados**

Nessa seção são analisados aplicativos de aprendizagem de idiomas que se assemelham de alguma forma com o aplicativo proposto nesse trabalho. A análise levou em consideração dois aplicativos, o “Aprenda Pomerano” e o “Expressões Idiomáticas”.

O aplicativo “Aprenda Pomerano” [Pure 2019] se propõe a ensinar a língua Pomerano aos seus usuários, foi desenvolvido para dispositivos com o sistema operacional Android e, assim como no aplicativo proposto neste trabalho, disponibiliza traduções de

palavras e frases, organizadas por categorias, permitindo ao usuário ouvir a pronúncia correta de cada tradução.

Durante a análise, não foi encontrada nenhuma forma de contribuir com o conteúdo do aplicativo, também percebeu-se que as traduções e áudios ficam armazenados localmente no dispositivo onde o APP está instalado, não sendo assim necessário conexão com a internet para o seu uso e também impossibilitando que o seu conteúdo seja atualizado em tempo real, fazendo com que seja necessário criar uma atualização do aplicativo para que seu conteúdo mude. No momento em que esse trabalho foi escrito, o *Aprenda Pomerano* não estava disponível na loja de aplicativos oficial do Android, a Google Play, mesmo assim, foi possível adquirir o aplicativo de forma não oficial no site “apkpure.com” [Pure 2019].

O aplicativo *Expressões Idiomáticas* se propõe a ensinar ao usuário o significado de expressões em inglês utilizadas por americanos, encontrando uma expressão equivalente ou semelhante com expressões usadas no português brasileiro, já que a tradução palavra por palavra pode não fazer sentido.

Durante o uso do aplicativo, notou-se a necessidade de estar conectado a internet para visualizar o conteúdo (expressões traduzidas), demonstrando assim que o conteúdo pode ser atualizado sem a necessidade de atualizar a instalação local. Não foi encontrado uma forma de contribuir com o conteúdo do aplicativo. O *Expressões Idiomáticas* está disponível para Android [Play 2019a] e iOS [Store 2019b].

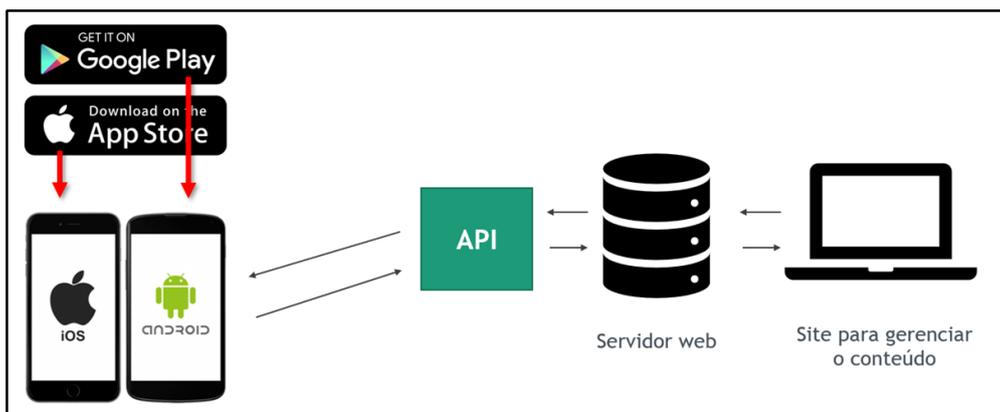
### **3. O Aprenda Talian**

O objetivo do aplicativo *Aprenda Talian* é proporcionar de uma forma simples o aprendizado do Talian para os seus usuários e permitir que eles contribuam com o conteúdo do app (categorias, frases, traduções e áudios) enquanto a aplicação web é destinada para usuários administradores, que têm acesso restrito e podem gerenciar o conteúdo do aplicativo e moderar as contribuições dos usuários.

Na Figura 1 são demonstrados os componentes que são parte do projeto e as suas interações. O site é usado para o gerenciamento do conteúdo e assim como o banco de dados MySQL [MySQL 2019], também fica hospedado no servidor Web. A aplicação gerada para as plataformas iOS [Store 2019a] e Android [Play 2019b] ficam disponíveis nas respectivas lojas e usam uma API (Application Programming Interface - Interface de Programação de Aplicativos) para acesso aos dados que ficam armazenados no servidor. A API contém um conjunto de funções que funcionam como uma intermediadora para acesso aos dados do banco de dados, fornecendo acesso por meio de requisições HTTP.

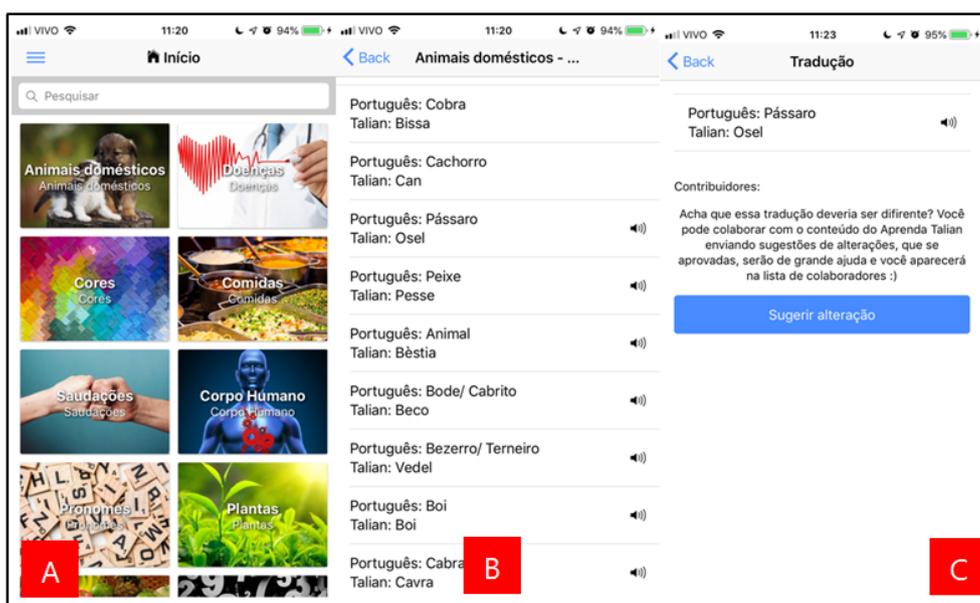
O aplicativo foi desenvolvido usando o Ionic [Framework 2019], um framework para desenvolvimento de aplicativos híbridos em HTML, CSS e Javascript, para dispositivos móveis. Nele há incorporado o Apache Cordova [Cordova 2019], que permite o acesso de aplicativos híbridos a recursos nativos dos smartphones (como câmera e vibração). Foi utilizado o tema menu lateral (sidemenu), que conta com os links para as páginas *Aprender*, *Contribuir* e *Sobre*.

A página *Aprender* possui um formulário para pesquisar traduções e uma lista de categorias de traduções disponíveis (Figura 2A). Quando clicado sobre uma categoria, é aberta uma página com uma lista de palavras/expressões em português e suas traduções



**Figura 1. Interação entre os componentes do projeto**

para o Talian (Figura 2B), com um ícone para ouvir a pronúncia em Talian caso disponível, e ao clicar sobre uma tradução é apresentada os seus contribuidores. No final da página encontra-se o botão “Sugerir alteração” (Figura 2C) que, caso autenticado, permite ao usuário enviar contribuições para editar o conteúdo que está sendo visualizado.



**Figura 2. Telas do aplicativo - categorias, traduções e detalhes da palavra**

Como o usuário precisa estar autenticado para poder enviar contribuições, ao acessar a página “Contribuir”, é redirecionado para a página “Entrar”, onde pode usar 3 métodos de autenticação diferentes, com e-mail e senha, com o Facebook Login [Facebook 2019] ou com o Google Sign-in [Developers 2019]. Caso o usuário ainda não possua um cadastro no aplicativo, pode criar acessando a página “cadastrar-se”. Com o usuário autenticado, a página “Contribuir” passa a exibir as contribuições do usuário, informando se elas foram ou não aceitas.

Pelo fato dos dados (categorias, traduções e áudios) serem dinâmicos, estes não são distribuídos junto com o instalador do APP e sim acessados via requisições HTTP

do aplicativo à API do site. O aplicativo foi desenvolvido para os sistemas operacionais Android e iOS e está disponível em suas respectivas lojas oficiais. Por ter sido utilizado o framework Ionic e suas técnicas de desenvolvimento híbrido, o aplicativo possui as mesmas funcionalidades em ambos sistemas operacionais, as únicas diferenças ao se comparar a versão do aplicativo para Android e para iOS são estéticas.

No final da página “aprender” sempre está presente um botão “Sugerir alteração”, cuja a ação desencadeada ao ser selecionado é contextual. Se for na tela onde são exibidas as categorias, é apresentado o formulário para sugerir a adição de uma nova categoria. Caso for selecionado enquanto se visualiza uma categoria, o usuário pode escolher entre adicionar uma tradução, uma categoria ou editar ou remover a categoria que está sendo visualizada. Caso clicado visualizando uma tradução, o usuário pode escolher entre editar ou excluir a tradução.

Para enviar contribuições é necessário que o usuário esteja autenticado. Um vez logado, a página Contribuições (Figura 3C) passa a exibir uma lista de contribuições do usuário, com o tipo de contribuição, data e se foi aceita, negada ou se está aguardando moderação. As Figuras 3A e 3B apresentam as telas de login e cadastro, respectivamente.

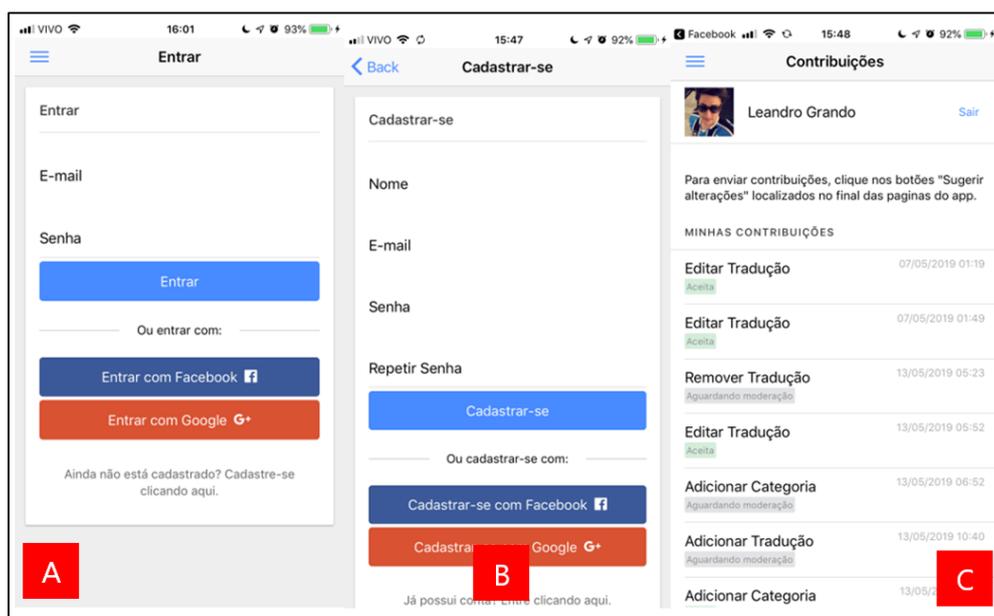


Figura 3. Telas do aplicativo: login, cadastro e contribuições

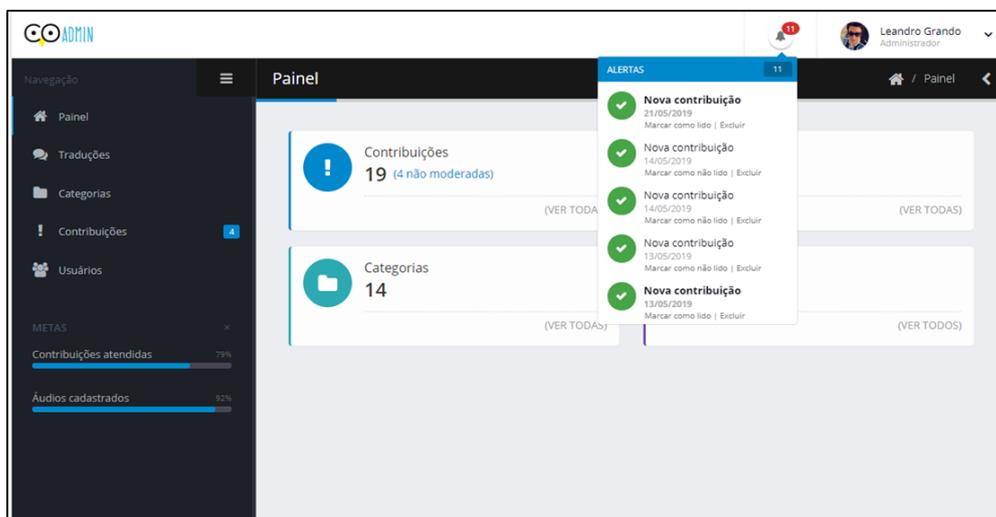
### 3.1. Site para gerenciamento do app

O site para gerenciamento dos dados do aplicativo foi desenvolvido utilizando os frameworks Laravel [Laravel 2019], Bootstrap [Bootstrap 2019] e jQuery [jQuery 2019]. O Laravel é um framework web em PHP, ou seja, um conjunto de classes e funções pré-implementadas para ajudar a “desenvolver aplicações seguras e performáticas de forma rápida, com código limpo e simples, já que ele incentiva o uso de boas práticas de programação” [Silva 2019]. O Bootstrap é um framework front-end direcionado para o desenvolvimento rápido e fácil de sites e aplicações web responsivos. jQuery é uma biblioteca javascript “desenvolvida para simplificar os scripts interpretados no navegador do usuário (client-side)” [Teixeira 2019].

No site os usuários administradores podem gerenciar o conteúdo a ser disponibilizado no aplicativo e também moderar as contribuições recebidas. É possível se autenticar com dois diferentes níveis de usuários: Comum e Administrador. O usuário comum é o usuário que se cadastra através do aplicativo e é registrado no site através da API. O usuário Administrador pode gerenciar as categorias e palavras/expressões com suas respectivas traduções e áudios de pronúncia, as contribuições dos usuários do aplicativo (usuários comuns), moderando-as para que façam parte ou não do conteúdo do APP, bem como os usuários do site. Ao acessar o site, o usuário é direcionado para a tela de autenticação, onde pode escolher entrar com e-mail e senha, Facebook Login ou Google Sign.

Após a autenticação, o usuário é redirecionado para a área administrativa que possui os seguintes itens de interface (Figura 4):

- menu lateral: que permite navegar entre as suas páginas;
- seção “metas”: mostra um percentual das contribuições atendidas e dos áudios cadastrados, se referindo ao percentual de traduções que já possuem um áudio da sua pronúncia;
- popup de alertas: aberto ao clicar sobre o ícone de sino na parte superior esquerda, que exibe as novas contribuições recebidas;
- menu do usuário: permite ao usuário sair ou atualizar os dados de sua conta;
- painel: exibe os totais de contribuições, traduções, categorias e usuários, com links para suas respectivas páginas.



**Figura 4. Página home do site**

A página Traduções permite ao administrador gerenciar as traduções, incluindo novas, editando ou excluindo as existentes. Ao adicionar ou editar uma tradução é possível informar a expressão em português, em Talian, escrever observações, gravar um áudio com a pronúncia da tradução em Talian e escolher em quais categorias essa tradução ficará visível.

A página Categorias possibilita gerenciar as categorias de traduções, incluindo novas, editando ou excluindo as existentes e ordenando a posição das categorias na hora de

listá-las. O formulário para adicionar ou remover uma categoria permite ao administrador inserir o nome da categoria em português e em Talian, além de uma imagem.

Na página Contribuições (Figura 5) são gerenciadas as contribuições enviadas via aplicativo pelos usuários. Nessa página é possível moderar as contribuições, para isso, basta clicar sobre uma das contribuições listadas e depois, clicar em “Aceitar” ou “Negar”. Caso aceita, de forma automática o site desempenha a ordem daquela contribuição, por exemplo, caso a contribuição seja do tipo “Adicionar categoria”, ao ser aceita, é automaticamente acrescentada uma nova categoria com as especificações presentes na contribuição. As contribuições listadas em amarelo estão aguardando moderação, as verdes foram aceitas e as vermelhas negadas.

#	Tipo	Usuário	Status	Criada	Ações
1	Adicionar categoria	Desconhecido	Aceita	22/04/2019 12:31	 
2	Adicionar categoria	Desconhecido	Aceita	22/04/2019 11:55	 
3	Adicionar categoria	Desconhecido	Aceita	23/04/2019 10:31	 
4	Editar tradução	Leandro Grando	Aceita	07/05/2019 01:19	 
5	Editar tradução	Leandro Grando	Aceita	07/05/2019 01:49	 
16	Remover tradução	Leandro Grando	Negada	13/05/2019 05:23	 
18	Editar tradução	Leandro Grando	Aceita	13/05/2019 05:52	 
19	Adicionar categoria	Leandro Grando	Aguardando moderação	13/05/2019 06:52	 
20	Adicionar tradução	Leandro Grando	Aguardando moderação	13/05/2019 10:40	 

Figura 5. Página contribuições do site

#### 4. Resultados Obtidos

O Aprenda Talian foi disponibilizado ao público através das lojas oficiais de aplicativos do Android [Play 2019b] e do iOS [Store 2019a]. Após usar o aplicativo, um grupo de 13 usuários, todos da cidade de Serafina Corrêa no Rio Grande do Sul, preencheram um questionário para avaliá-lo. As primeiras questões da pesquisa tinham o objetivo de verificar o perfil dos usuários. Constatou-se que a faixa etária varia entre 11 e 60 anos, com maior incidência entre 11 a 18 anos. Também constatou-se que 84,6% dos respondentes conhecem pessoas que falam Talian. Ao serem questionados sobre o quanto conhecem o dialeto Talian em uma escala de 1 a 5, sendo 1 “Desconheço totalmente” e 5 “Falo fluentemente”, a resposta média foi de 2,9.

Em outra questão foi solicitado para que o usuário escolhesse por quais meios ele já teve contato com o Talian, a alternativa mais escolhida foi “Pessoas conversando em Talian ao meu redor” com 76,9%, enquanto apenas 7,7% disseram nunca ter tido

contato com o Talian. Também foi questionado qual o interesse em aprender Talian em uma escala de 1 a 5, sendo 1 “nenhum interesse” e 5 “muito interessado”, a resposta média foi 3,6. O restante das questões diziam respeito sobre a experiência que o usuário teve ao usar o aplicativo e a opinião deles sobre a importância do uso de aplicativos para dispositivos móveis no ensino de línguas. Dos 13 participantes, 12 usaram o aplicativo em smartphones e 4 em tablets, 8 em sistema operacional Android e 7 em iOS e 53% disseram ter contribuído com o conteúdo do aplicativo. O restante das questões foram:

- O quanto o aplicativo te ajudou a aprender talian?
- O que você achou do aplicativo em geral?
- O que você achou da interface do aplicativo?
- O aplicativo foi intuitivo na questão de usabilidade?
- Você considera importante o uso de aplicativos como este?
- Na sua opinião, o uso de um dispositivo móvel pode auxiliar o processo de aprendizagem?

Para as questões acima, o usuário poderia classificar a resposta numa escala de 1 a 5, onde 1 representa a avaliação mais negativa e 5 a mais positiva. Após a análise das respostas recebidas, em relação ao aplicativo em geral, obteve-se que 92,3% avaliaram de maneira positiva, como mostrado na Figura 6. Considera-se uma avaliação positiva as respostas 4 e 5 da escala.

Em relação à interface, intuitividade e usabilidade, 100% avaliaram positivamente. 100% das avaliações foram positivas sobre a importância de aplicativos (Figura 6A) como esse, bem como para as questões a respeito do uso de dispositivos móveis no processo de aprendizagem (Figura 6B) e sobre o aplicativo ter se mostrado relevante para o aprendizado de Talian ao usuário (Figura 6C).

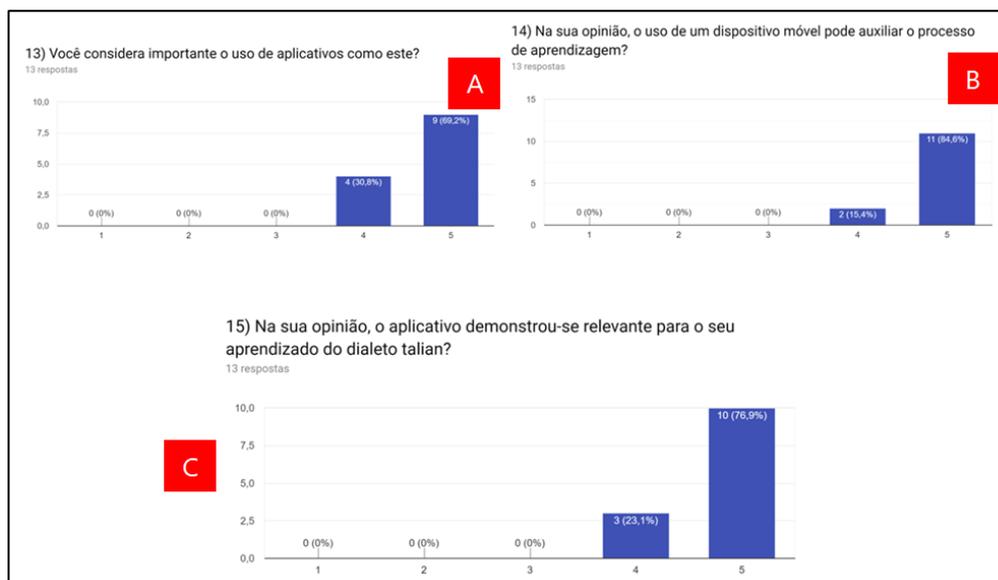


Figura 6. Gráfico representando as avaliações dos usuários

## 5. Considerações finais

Esse trabalho apresentou o aplicativo Aprenda Talian que tem como objetivo auxiliar no aprendizado do dialeto Talian, disponibilizando aos usuários um conteúdo colaborativo,

de expressões traduzidas do português para o Talian, em texto e áudio. Para o seu desenvolvimento, foi utilizado o framework Ionic, enquanto para o site foi utilizado framework Laravel junto ao banco MySQL. Com esse trabalho foi possível ter a experiência de unir várias áreas do conhecimento, tais como, desenvolvimento híbrido de aplicativos, API com autenticação, desenvolvimento de sites responsivos e o dialeto Talian.

A avaliação realizada com o usuários mostrou que é possível aprender um dialeto como o Talian utilizando dispositivos móveis e colaborações dos usuários para tornar o conteúdo de aprendizagem mais completo. Também percebeu-se que o método da criação do conteúdo de forma colaborativa, juntamente do uso de dispositivos móveis para o ensino, pode ser utilizado em outras áreas.

A fim de melhorar a experiência do usuário, formulou-se algumas ideias para futuramente serem implementadas no aplicativo, buscando dar mais incentivos para que os usuários colaborem criando conteúdo, tais como: notificações que lembrem o usuário para diariamente estudar a tradução de uma expressão diferente, incluir uma forma dos usuários gravarem áudio através do aplicativo para anexar as contribuições e acrescentar uma forma de estudar por lições, onde em cada lição o usuário terá uma lista de traduções para aprender.

## Referências

- Bootstrap (2019). Bootstrap - the most popular html, css, and js library in the world. <https://getbootstrap.com/>. Acessado em: 07 jun 2019.
- Cordova (2019). Apache cordova. <https://cordova.apache.org/>. Acessado em: 19 jun 2019.
- Corrêa, S. (2019). Talian - prefeitura municipal de serafina corrêa - rs. <http://www.serafinacorrea.rs.gov.br/talian/>. Acessado em: 19 jun 2019.
- Developers (2019). Google identity platform - google developers. <https://developers.google.com/identity>. Acessado em: 19 jun 2019.
- Facebook (2019). Login do facebook. <https://developers.facebook.com/docs/facebook-login/>. Acessado em: 07 jun 2019.
- Framework, I. (2019). Ionic - cross-platform mobile app development. <https://ionicframework.com/>. Acessado em: 07 jun 2019.
- jQuery (2019). jquery. <https://jquery.com/>. Acessado em: 19 jun 2019.
- Laravel (2019). Laravel - the php framework for web artisans. <https://laravel.com/>. Acessado em: 07 jun 2019.
- MySQL (2019). Mysql. <https://www.mysql.com/>. Acessado em: 07 jun 2019.
- Play, G. (2019a). Expressões idiomáticas - aplicações no google play. <https://play.google.com/store/apps/details?id=br.com.unicasolucoes.idiomaticexpressions>. Acessado em: 07 jun 2019.
- Play, G. (2019b). Google play. <https://play.google.com/store>. Acessado em: 07 jun 2019.
- Pure, A. (2019). Aprenda pomerano para android - apk baixar. <https://apkpure.com/br/aprenda-pomerano/br.com.aprendapomerano>. Acessado em: 28 mai 2019.

- Silva, W. A. L. (2019). Laravel tutorial: Introdução ao laravel framework php. <https://www.devmedia.com.br/laravel-tutorial/33173>. Acessado em: 25 mai 2019.
- Store, A. (2019a). App store - apple. <https://itunes.apple.com/br>. Acessado em: 19 jun 2019.
- Store, A. (2019b). Expressões idiomáticas na app store. <https://itunes.apple.com/br/app/express> Acessado em: 19 jun 2019.
- Teixeira, J. R. (2019). jquery tutorial: Introdução à biblioteca javascript jquery. <https://www.devmedia.com.br/jquery-tutorial/27299>. Acessado em: 14 jul 2019.

# Caracterizando as relações entre reputação e os desafios em atividades nas plataformas de *crowdtesting*

Fernando Luiz Costella, Alexandre Lazaretti Zanatta, Cristiano Cervi

Ciência da Computação – Universidade de Passo Fundo (UPF) – Campus Passo Fundo  
Av. Brasil Leste, 285 – São José – Passo Fundo/RS

nandocostella@gmail.com, [zanatta@upf.br](mailto:zanatta@upf.br), cervi@upf.br

**Abstract.** *Crowdsourced Software Testing model is a software testing outsourcing to an online community. The crowd play an important role in this model to become active members. However, previous research shows that many challenges, such as low reputation, poor time management, lack of opportunities to participate in projects, communication problems, and lack of experience, to contribute in this model. We conducted a qualitative study to identify relationships between low reputation and the others challenges. Thus, it was possible to observe and describe these relationships.*

**Resumo.** *Crowdsourced Software Testing é a terceirização de teste de software para uma comunidade on-line no modelo crowdsourcing. Testadores executam um importante papel neste modelo para se tornarem membros ativos. Entretanto, pesquisas anteriores mostram que enfrentam muitos desafios, como a baixa reputação, falta de tempo, dificuldade em receber novos convites, problemas de comunicação, falta de experiência para, enfim, contribuir em comunidades on-line. O objetivo do trabalho foi identificar se existem relações entre reputação e estes desafios. Realizou-se um estudo qualitativo para identificar estas relações com cinco plataformas de crowdtesting. Foi possível observar e descrever como ocorrem estas relações.*

## 1. Introdução

O modelo *Crowdsourcing* tem como objetivo transferir uma tarefa tradicionalmente feita por um empregado ou colaborador da organização, para um indefinido, desconhecido e geralmente grande grupo de pessoas, sob demanda, por meio de uma chamada pública, via uma plataforma de *crowdsourcing*. Este modelo possui quatro elementos: a multidão, a plataforma, o cliente e a tarefa.

O *Crowdsourcing* pode ser aplicado em diversas áreas, entre elas, o desenvolvimento de software, denominado de software *crowdsourcing*. Conforme Mao, Licia, Harman, e Yue (2016) *Software Crowdsourcing* (SWCS), é o ato de transferir, externamente, qualquer tarefa (requisitos, desenvolvimento, testes e evolução) no processo de desenvolvimento de software a um grande grupo indefinido de trabalhadores online em formato de chamada pública. Dentre estas tarefas, uma que tem chamado a atenção é a área de teste de software, denominada de *crowdtesting* (CST). O *crowdtesting* utiliza o conceito de *crowdsourcing*, pois os *crowdtesters* (testadores de software na multidão) realizam os testes remotamente em seu ambiente.

Porém, conforme Zanatta, Machado, e Steinmacher (2018), Zanatta, Steinmacher, Machado, Souza, e Prikladnicki (2017) os *crowdtesters* enfrentam muito desafios para

contribuírem, pela primeira vez, em tarefas de SWCS, dentre eles a falta de tempo para executar as tarefas, a concorrência com demais participantes mais experientes, a dificuldade de comunicação com os demais membros na plataforma, a falta de conhecimento sobre a tarefa ofertada pela plataforma, e a pouca experiência em tarefas SWCS, entre outros.

Kogler, Costela, e Zanatta (2018) identificaram que uma baixa reputação nas plataformas de CST pode ser também um desafio para o *crowdtester*. Por exemplo, quando um *crowdtester* envia uma tarefa, e em caso de aprovação desta tarefa ele melhora sua reputação, elevando, com isso, as chances de receber novos convites para testes. Por outro lado, em caso de rejeição da tarefa enviada, sua reputação pode baixar, impactando, de certa forma, negativamente no recebimento de novas tarefas. Ao não receber novas tarefas devido a baixa reputação, o *crowdtester* pode desistir de utilizar este modelo. Yang, Karim, Saremi, e Ruhe (2016) apontam um índice de 82,9% de taxa de desistência de submissão de tarefas de software *crowdsourcing*, considerando que a maioria dos participantes em comunidades on-line não contribuem, como apontam Hill, Hollan, Wroblewski, e Mccandless (1992).

Assim, o objetivo deste trabalho é iniciar um estudo exploratório para identificar se existe relação entre a reputação dos *crowdtesters* com os desafios que os trabalhadores na multidão enfrentam ao participarem em tarefas de software *crowdsourcing*. Entende-se, portanto, que existe oportunidade para uma investigação uma vez que o desinteresse em seguir contribuindo, participando e aceitando tarefas de testes pode ser um risco ao modelo de CST que precisa da força de trabalho online dos membros das plataformas para atender as demandas de teste de seus clientes. Assim, a questão de pesquisa do presente trabalho é: “*Quais são as relações entre a reputação e os desafios que um crowdtester enfrenta ao realizar tarefas de CST?*”.

Este trabalho está organizado da seguinte maneira: na seção 2 são apresentadas as plataformas de CST. A seção 3 descreve a metodologia. Na seção 4 são expostos os resultados preliminares das relações entre a reputação e os desafios. Na seção 5 são apresentadas as discussões. A seção 6 encerra com algumas considerações sobre a experiência obtida, limitações e trabalhos futuros.

## 2. Plataformas de CST

As plataformas que utilizam o modelo *crowdsourcing* são o elo entre o cliente e a multidão, tendo como função, basicamente, receber demandas dos clientes, organizar estas demandas em tarefas e realizar uma chamada pública para que a multidão possa executá-la. A seguir far-se-á uma breve descrição das plataformas utilizadas no trabalho.

A uTest<sup>1</sup>, considerada a pioneira de CST no Brasil, atua em mais de 150 países com aproximadamente 60 mil *crowdtesters*. Além do recebimento de uma compensação financeira (remuneração monetária), os *crowdtesters* participam de um sistema de reconhecimento e reputação e ranking. A plataforma utiliza apenas o idioma inglês e seus testes são do tipo funcionais. Na BugFinders<sup>2</sup>, os testes geralmente não possuem casos de teste ou especificação do escopo do projeto, são considerados testes a nível mais exploratório, o que torna o relato de defeitos um pouco complexo. Essa plataforma também possui o idioma inglês como padrão e seus testes são funcionais. No TestBirds<sup>3</sup>

---

1 <https://www.utest.com/>

2 <https://join.digivante.com/as-functional-tester>

3 <https://www.testbirds.com/>

o participante não necessita de nenhum conhecimento prévio em testes de software ou tecnologias em particular. O participante também realiza um projeto de teste não remunerado, para avaliação de suas capacidades de encontrar e reportar defeitos, porém, essa plataforma exige o uso de imagens detalhando esses defeitos. O TestBirds também utiliza o idioma inglês e segue nos testes funcionais. No PassBrains<sup>4</sup> o participante necessita aguardar um convite para integrar um time de testes de projeto remunerado, baseado nos seus conhecimentos e dispositivos informados no cadastro. Seus testes são do tipo funcional e o padrão do idioma é o inglês. O objetivo do 99Tests<sup>5</sup> é fornecer testes funcionais de segurança, de carga e teste de automação. Também trabalha com processos em ciclos de testes, seguindo exigências quanto aos componentes que devem ser validados em cada ciclo. Essa plataforma também utiliza o idioma inglês.

Todas as plataformas analisadas possibilitam a realização de testes funcionais, também conhecidos como teste da “caixa-preta”, cujo objetivo é entender o comportamento e as funcionalidades da aplicação.

As plataformas de CST podem oferecer flexibilidade e instantaneidade como diferenciais sobre os modelos empresariais tradicionais de validação de software, pois as conexões dos participantes ocorrem de forma integral diariamente. Outro diferencial se refere aos participantes, pois geralmente as validações não são realizadas por profissionais em teste de software, ou seja, a impressão deles sobre as aplicações podem ser as mesmas dos *crowdtesters* ao qual o sistema (ou software) é destinado.

### 3. Metodologia

Este trabalho caracteriza-se por uma pesquisa exploratória com abordagem qualitativa, buscando entender e interpretar através do ingresso dos autores, o funcionamento, os métodos, as regras e os mecanismos (sistemas) de reputações que determinam ou influenciam no perfil do indivíduo que está participando das plataformas de CST. Em um segundo momento, com os dados obtidos foram armazenados em diários individuais. Por meio deste armazenamento, foi possível, com a utilização da técnica da Análise de Conteúdo, como aborda Bardin (2006), explorar os dados, codificar as fontes e gerar quais foram os nós temáticos identificados, quais são usados com mais frequência, e como estes nós se relacionam. Ao fazer esta relação espera-se, também, como trabalho futuro oferecer ao *crowdtester* um guia ou um conjunto de instruções para a realização de tarefas em CST. A ferramenta Nvivo<sup>6</sup> auxiliou o processo de organização dos dados coletados e análise de relações.

A coleta de dados foi realizada por meio da submissão de tarefas de teste funcional de software nas plataformas de CST, observando-se as dificuldades encontradas em cada etapa durante a realização das ações de identificação e reporte dos casos de teste executados. Essas tarefas referem-se à participação de um trabalhador na multidão, desde o cadastro nas plataformas, realização do teste de qualificação, discussões em fóruns, interação com os demais participantes e, finalmente, submissão de um teste de software. É importante ressaltar que as plataformas de CST analisadas oferecem uma compensação financeira para o *crowdtester* vencedor da tarefa de teste funcional.

---

4 <https://www.passbrains.com/>

5 <https://99tests.com/>

6 <https://www.qsrinternational.com/>

## 4. Resultados Preliminares

A seguir far-se-á uma breve descrição das relações encontradas (nós temáticos) entre a reputação e os desafios que o *crowdtester* enfrenta ao participar de atividades em CST.

A Figura 1 mostra a relação entre reputação e os desafios.

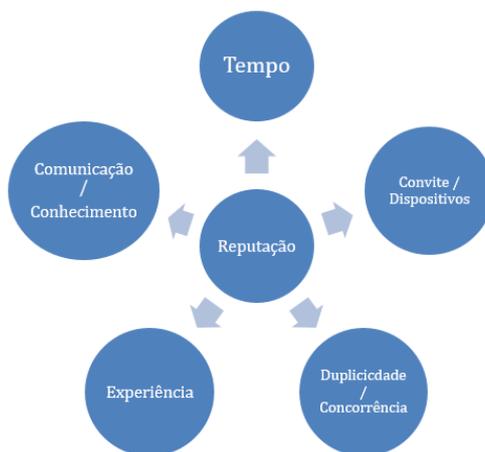


Figura 1. Relação da Reputação com os Desafios

### 4.1 Relação entre a Reputação e o Tempo

Quando a tarefa é enviada por e-mail, primeiramente é necessário que o *crowdtester* realize o aceite o quanto antes para que o convite não expire, pois isto pode ocorrer em poucas horas. Com isso, o mesmo pode ser recebido pelo *crowdtester* em um momento em que não está acompanhando os e-mails e dificultar o seu ingresso ao projeto. Quando o *crowdtester* aceita, precisa aguardar o início do ciclo de testes para a realização da tarefa, que em alguns casos se inicia no momento do aceite do convite. Caso realize corretamente o aceite e ao iniciar o ciclo de teste, não o realiza, baixará sua reputação na ferramenta. Com isso, é necessário que o *crowdtester* se programe para não perder o início dos testes, pois se demorar a iniciar estes testes, muitos defeitos já estarão abertos. Isso requer que seja mais criterioso na sua análise, pois os defeitos mais evidentes provavelmente já serão relatados pelos demais.

Como o período estipulado para o ciclo de testes normalmente se encerra em poucos dias (entre 5 e 7 dias), o *crowdtester* terá que organizar seu tempo disponível para que possa realizar os testes, pois os convites não são enviados em um período/dia padrão. Essas questões influenciam na regularidade de acesso e realização de tarefas nas plataformas. A não organização do tempo pode acarretar na perda de reputação e por consequência, o recebimento de menos tarefas. O recebimento de menos tarefas, pode acarretar na diminuição do entusiasmo. Esses fatores podem ser determinantes para a desistência do *crowdtester* em realizar tarefas na plataforma. Assim, a organização do tempo é um fator preponderante para o sucesso na execução de tarefas em CST.

### 4.2 Relação Entre Reputação e Duplicidade/Concorrência

Os *crowdtesters* devem ficar preparados para receber os convites. Aqueles que possuem uma “boa” reputação tem a possibilidade de receberem mais projetos pagos. Para esses

projetos, o nível de concorrência entre os participantes é alto e exige um nível alto de atenção, pois é muito importante a forma de reportar um defeito encontrado.

Os requisitos apresentados para o projeto devem ser lidos e entendidos de uma forma clara. O entendimento incorreto de um requisito pode acarretar na abertura errônea de um defeito, ocasionando a perda de reputação. Com os requisitos entendidos, o *crowdtester* pode explorar a aplicação na busca de defeitos.

Ao encontrar um defeito, o *crowdtester* deve reportá-lo na plataforma respeitando o padrão exigido pela plataforma. Algumas plataformas exigem apenas uma descrição e uma imagem da tela do defeito encontrado, outras, por exemplo, solicitam a gravação de um vídeo do defeito. É importante primeiramente entender a forma correta para reportar um defeito na plataforma que está sendo utilizada e se certificar de que o defeito não foi aberto por outro *crowdtester*, pois a abertura de um defeito duplicado acarreta na rejeição do mesmo e na perda de reputação na plataforma. Assim, devido a possível concorrência do projeto, o *crowdtester* tem dois caminhos: a) abre o defeito o mais rápido possível, sujeitando-se que o defeito já esteja aberto; b) dedica-se em verificar se o defeito já está aberto, correndo o risco de o defeito ser reportado por outro *crowdtester*.

### **4.3 Relação Entre Reputação e Comunicação/Conhecimento**

A comunicação é um dos principais elementos em tarefas de desenvolvimento de software e fundamental quando realizadas no modelo *crowdsourcing*. Uma forma errônea de reportar defeitos ocasiona a dificuldade de entendimento pelo avaliador. Quando o avaliador não interpreta corretamente o erro relatado, ele pode solicitar uma explicação ao *crowdtester* ou até mesmo rejeitar o defeito aberto. Essa reavaliação depende da plataforma utilizada, sendo que algumas permitem a realização de algumas revisões antes de realizar a rejeição do defeito.

Para reportar um defeito, o *crowdtester* deve estar atento a todos os detalhes necessários para que o avaliador possa reproduzir fielmente o defeito. Rejeições de defeitos podem ser ocasionadas em virtude da falta de passos para reprodução do mesmo. Por isso é fundamental que ele seja reportado de uma forma que sejam apresentados todos os passos necessários, desde o acesso a página até o ponto onde o defeito é apresentado, sendo importante que o *crowdtester* descreva de forma clara o comportamento errôneo da aplicação.

A comunicação correta pelo *crowdtester* é de fundamental importância para reportar o defeito, pois deve ser considerado que o avaliador utilizará somente as informações descritas para reproduzir o que foi relatado. Quando um defeito é rejeitado, acarreta na perda de reputação na plataforma e por consequência podendo desmotivar o *crowdtester* a participar de novos testes.

### **4.4 Relação entre Reputação e Convite**

O *crowdtester* pode se sentir desmotivado quando não receber convites para realização dos testes. Esse ponto está relacionado diretamente com a reputação conquistada durante a realização das tarefas de testes. Por isso é fundamental ter atenção e entender como funcionam os mecanismos de reputação. Quando a reputação do *crowdtester* é baixa, os convites para projetos pagos na mesma diminuem.

O “volume baixo” de convites recebidos, pode estar relacionado também aos dispositivos necessários para realização dos testes. Quando o *crowdtester* realiza o

cadastro, algumas plataformas, por exemplo, exigem que informe quais dispositivos possui para realização dos testes. Essas informações no cadastro são consideradas quando a plataforma disponibiliza novos projetos, sendo que serão enviados apenas aqueles que condizem com o cadastro. Caso o *crowdtester* possua um número pequeno de dispositivos cadastrados, poderá restringir o recebimento de convites para projeto.

Esses pontos são decisivos na realização frequente de testes na plataforma, pois o fato de receber o convite para um projeto não significa que isso poderá ocorrer diariamente ou semanalmente.

#### **4.5 Relação Entre Reputação e Experiência**

Para realização correta dos testes na plataforma, é necessária uma clara compreensão do escopo do teste a ser realizado e a descrição do projeto. Escopo pode ser entendido como o objetivo do teste ou a abrangência que o teste deve ser realizado, como por exemplo, realização apenas de testes funcionais. Caso o *crowdtester* relate um defeito que não faz parte dos testes funcionais, mesmo que seja um defeito válido, será rejeitado pelo avaliador porque está fora do escopo dos testes. Nesse caso a reputação será afetada, pois o defeito foi rejeitado.

O entendimento do teste a ser realizado impacta diretamente no sucesso ou fracasso do mesmo. Isso porque se o *crowdtester* entender de forma incorreta o teste a ser realizado, poderá abrir defeitos inválidos e com isso decrementar a sua reputação. Caso não possua experiência, o aprendizado na compreensão dos testes pode vir com a experiência adquirida a cada participação ativa nos projetos. É importante que esse conhecimento seja adquirido para que a participação nos demais projetos seja cada vez mais vantajosa.

Outra forma de adquirir experiência e conhecimento é a comunicação, participação e interação com os demais participantes da plataforma. Por mais que seja um ambiente de competição, também é um ambiente colaborativo, sendo possível absorver informações que servirão como auxílio nos testes dos projetos que são remunerados.

### **5. Discussão**

Motivação e personalidade, são exemplos de atributos que impulsionam um participante a ingressar e contribuir em uma comunidade on-line de acordo com Stol e Fitzgerald (2014). Também são considerados fatores como identificação com a comunidade, satisfação, envolvimento e confiança, que demonstram as intenções e comportamentos do participante de acordo com Kittur, Chi, e Suh (2008). A identificação e o envolvimento são demonstrados por meio de uma participação ativa nas atividades da comunidade de cada plataforma como destacam Straub, Gimpel, Teschner, e Weinhardt (2015).

Neste contexto, um dos desafios relaciona-se ao posicionamento do participante no ranking geral de cada plataforma. Esse ranking denota experiência, participações em testes remunerados, conexões com os demais participantes, o quanto o participante está contribuindo de forma positiva na comunidade. Esse desafio também se conecta com a participação ativa do usuário, pois gerar conteúdos de interesse sobre testes em CST e suas ferramentas é um exemplo de contribuição que agrega valor ao participante, tanto diante da plataforma quanto dos demais participantes da comunidade. Rohrmeier (2012) comenta que mesmo com a diversidade de participação em uma comunidade, é crucial

que os participantes forneçam conteúdo suficiente e mantenham uma interação com os demais participantes para preservação da comunidade.

Porém, a participação na plataforma também se demonstrou como um desafio para os trabalhadores na multidão em CST. Estar conectado à plataforma não certifica que o usuário conseguirá convites para projetos remunerados, e se faz necessária uma participação ativa de forma diária. Identificar-se com as ferramentas disponíveis na plataforma, assim como com os demais participantes, colabora para que haja uma participação constante. Essa identificação social está relacionada com o tipo de consciência de cada participante. É definido como um sentimento de conexão intrínseca entre os membros e um senso de diferença coletivo de outros que não estão na mesma comunidade.

Outro desafio identificado foi quanto ao “grau” de conhecimento que o participante consegue obter ao fazer uso das ferramentas dentro de cada comunidade de CST. Nas comunidades sociais online, o valor é inerente à conexão entre os pares, a comunicação entre eles e a troca e combinação de conhecimento, informação e bens digitais como aponta Rohrmeier (2012). A participação do usuário em ferramentas de comunicação como fóruns e chats, deve ocorrer sempre de forma a discutir assuntos que agreguem valor às experiências com as tarefas que vem sendo realizadas, neste caso, tarefas de teste de software. Comunicar-se com a plataforma, utilizando canais síncronos como o chat ou assíncronos como o e-mail, com o intuito de solucionar dúvidas quanto o escopo ou processo de teste, também gera um agregado de conhecimentos e ações objetivas para dúvidas futuras. Essa comunicação ocorre diretamente com os gestores de cada plataforma, sem envolver os demais participantes.

Além dessas ferramentas de comunicação, interagir com a comunidade da plataforma, pode resultar em troca de experiências e crescimento coletivo. Para Rohrmeier (2012), estar em contato com os demais participantes, em uma comunidade online, pode levar a mais interações, pois isso significa uma diversidade de fontes de informações disponíveis para o usuário focal. Ocorre, então, uma identificação social, que expressará um tipo de força de empatia para todos os contatos do participante. Ou seja, a identificação com outros participantes potencializa relacionamentos estáveis.

A realização da avaliação do novo participante por parte da plataforma foi outra barreira identificada. Essa etapa do processo consiste em comprovar as qualificações informadas no cadastro do participante na plataforma. Também está interligada ao posicionamento no ranking geral, pois cada defeito relatado no teste de entrada, conta pontos, negativos ou positivos. Este teste, serve como um primeiro contato do trabalhador na multidão com o processo de validação dos testes que cada plataforma exerce. A partir desta avaliação inicial, o participante já é inserido no ranking geral da plataforma, porém, não garante ao participante uma seleção para um convite de projeto de teste remunerado. Ao receber um convite para projeto remunerado, o participante deve comprometer-se com as tarefas propostas, buscando sempre entender seus escopos, pois as pontuações no ranking geralmente são acumulativas, e defeitos invalidados prejudicam o destaque positivo do participante.

Por meio da identificação dos desafios e relacionando-as com o impacto gerado em relação à reputação do *crowdtester* nas ferramentas. Na literatura foram verificados princípios e práticas de desenvolvimento e projeto que influenciam ou são influenciados por aspectos sociais, humanos e econômicos.

## 6. Considerações Finais

O *Crowdsourced Software Testing* (CST) auxilia no processo de validação da qualidade de um software por meio de testes realizados pela multidão de trabalhadores online. Porém, os trabalhadores deste modelo enfrentam algumas barreiras para executarem tais atividades. Foi possível identificar, preliminarmente, algumas relações entre reputação e desafios. Para um trabalhador na multidão em atividades de CST receber um convite para atividades remuneradas é fundamental para manter-se ativo na plataforma, comunicar-se com seus pares e com a plataforma e isso, pode auxiliar e guiar o trabalhador durante sua participação nos projetos de teste. Experiências e conhecimento coletivo agregam valor na qualidade dos testes realizados.

Considerando o estágio exploratório e preliminar do trabalho, os resultados não podem ser generalizados pois as atividades foram realizadas pelos autores em cinco plataformas de CST. Como trabalho futuro realizar-se-á um experimento com alunos de graduação de cursos da área da tecnologia da informação em disciplinas relacionadas ao teste de software.

### Referências

- Bardin, L. (2006). *Análise de conteúdo* (L. de A. Rego & A. Pinheiro, Trads.). Lisboa: Edições 70.
- Hill, W. C., Hollan, J. D., Wroblewski, D., and Mccandless, T. (1992). Edit wear and read wear. *Conference on Human Factors in Computing Systems*, (pp. 3–9). New York.
- Kittur, A., Chi, E., and Suh, B. (2008). Crowdsourcing user studies with Mechanical Turk. *Human factors in computing systems*, (pp. 453– 456).
- Kogler, R., Costela, F., e Zanatta, A. (2018). *Sistemas de Reputação: Uma Análise em Plataformas de Crowdttesting*. Escola Regional de Engenharia de Software, (pp. 97-104). Dois Vizinhos.
- Mao, K., Licia, C., Harman, M., and Yue, J. (2016). A Survey of the Use of Crowdsourcing in Software Engineering. *Journal of Systems and Software*, 126(1), pp. 57-84.
- Rohrmeier, P. (2012). *Social networks and online communities—managing user acquisition, activation and retention*. Technischen Universität München. München: Tese de Doutorado. Universität München.
- Stol, K., and Fitzgerald, B. (2014). *Researching crowdsourcing software development: perspectives and concerns*. 1st International Workshop on CrowdSourcing in Software Engineering, (pp. 7-10). Hyderabad.
- Straub, T., Gimpel, H., Teschner, F., and Weinhardt, C. (2015). How (not) to Incent Crowd Workers - Payment Schemes and Feedback in Crowdsourcing. *Business e Information Systems Engineering*, 57(3), pp. 167-179.
- Yang, Y., Karim, M. R., Saremi, R., and Ruhe, G. (2016). Who Should Take This Task?: Dynamic Decision Support for Crowd Workers. *10th International Symposium on Empirical Software Engineering and Measurement*, (p. 8).
- Zanatta, A. L., Machado, L. S., and Steinmacher, I. (2018). *Competence, Collaboration and Time Management: Barriers and Recommendations for crowdworkers*. 5th International Workshop on Crowd Sourcing in Software Engineering. Gotemburgo.
- Zanatta, A. L., Steinmacher, I., Machado, L., Souza, C. R., and Prikladnicki, R. (2017). *Barriers Faced by Newcomers in Software Crowdsourcing Projects*. *IEEE Software*, 34(2), pp. 37-43.

# Comparação de Performance entre o PostgreSQL e o MongoDB

Emanuele A. Solagna<sup>1</sup>, Alexandre T. Lazzaretti<sup>2</sup>

<sup>1</sup>Instituto de Ciências Exatas e Geociências – Universidade de Passo Fundo (UPF)  
99.052-900 – Passo Fundo – RS – Brasil

<sup>2</sup>Instituto Federal Sul-Rio-Grandense (IFSUL) – Passo Fundo, RS – Brazil

175478@upf.br, alexandre.lazzaretti@passofundo.ifsul.edu.br

**Abstract.** *In recent years there was a increase use of scalable applications, making possible to use non-relational or NoSQL databases. This is due to the big quantity of data generated and the need to meet the demand for scalable web applications, as well as the inability of traditional models to suply new demands for this large amount of data. This paper presents a performance (response time) comparison of one of these NoSQL databases: MongoDB. Comparing it to a relational database PostgreSQL, applying several data manipulation commands, using a weather dataset provided by the AgroDB database. As a result, it can be concluded that the MongoDB database has been shown to perform better when subjected to the proposed tests.*

**Resumo.** *Nos últimos anos houve um aumento significativo do uso de aplicações escaláveis de grande porte, fazendo necessária a utilização principalmente de sistemas de banco de dados que utilizam o modelo não-relacional e NoSQL. Isso se deve ao aumento do volume de dados gerado e a necessidade de suprir a demanda de aplicativos web escaláveis, bem como a falta de capacidade de modelos tradicionais atenderem novas demandas dessas grandes quantidades de dados. Este trabalho apresenta uma comparação de desempenho (tempo de resposta), de um desses bancos de dados NoSQL: o MongoDB. Comparando o mesmo ao banco relacional PostgreSQL, através de diversos comandos de manipulação de dados, utilizando um dataset meteorológico fornecido pelo banco de dados AgroDB. Como resultado, pode-se concluir que o banco de dados MongoDB demonstrou ter uma melhor performance ao ser submetido aos testes propostos.*

## 1. Introdução

Com a globalização, a expansão virtual se tornou efetiva e após o ano 2000 a geração de dados vem crescendo de forma exponencial. Segundo a IBM [IBM ], em 2008, 90% dos dados no mundo foram gerados nos dois anos anteriores. Este fato é decorrente da adesão de grandes empresas à Internet, como redes sociais, companhias de telefonia móvel, dentre outras. E esse crescimento gerou novos conflitos quanto ao armazenamento e a disponibilidade desses dados. O paradigma sobre o armazenamento de dados traz à tona a necessidade por estudos sobre as características necessárias em modelos de banco de dados para que esse crescente volume de informações seja gerenciado de forma adequada e fazem com que requisitos, até então indispensáveis, como consistência de dados, sejam revistos [VOLGELS 2009].

Segundo [MADISON et al. 2015], um modelo de banco de dados é uma descrição dos tipos de informações que estão armazenadas em um banco de dados. Um dos modelos de bancos de dados amplamente usados até os dias de hoje é o modelo relacional. Porém, constatou-se que os modelos de bancos de dados relacionais apresentam limitações ao trabalhar com grandes volumes de dados. Como opção para a solução deste e outros problemas, tecnologias emergentes recentemente chamadas de movimento NoSQL (Not Only SQL), surgiram como uma nova classe de SGBDs que utilizam outros modelos, diferentes do relacional e que trazem outros paradigmas de armazenamento de dados, como por exemplo, armazenamento de documentos sem esquema, onde os dados são armazenados sem necessidade de fazer normalização.

Uma das motivações para o movimento NoSQL ter surgido foi para contornar a dificuldade de escalabilidade presente nos bancos tradicionais, tendo em vista que pode ser muito caro e/ou complexo escalar um banco de dados relacional. Sendo assim, o objetivo deste trabalho é realizar uma comparação de desempenho de consultas entre dois bancos de dados gratuitos que vêm sendo amplamente utilizados, um relacional e outro NoSQL. Os bancos escolhidas para a realização dos experimentos foram o PostgreSQL [PostgreSQL 2016] e o MongoDB [MONGODB 2016]. Os testes foram executados através da implementação de um ambiente, onde foi possível realizar a comparação, analisar os resultados e obter conclusões sobre o desempenho das duas tecnologias no cenário utilizado.

## **2. Referencial Teórico**

Nesta seção são detalhados os conceitos e definições necessárias para o entendimento do trabalho realizado.

### **2.1. Banco de Dados**

Para [ELMASRI and NAVATHE 2011], “na expressão Banco de Dados, subentende-se que, um banco de dados representa algum aspecto do mundo real”, às vezes chamado de minimundo ou de universo de discurso. As mudanças no minimundo são refletidas no Banco de Dados. Um banco de dados é uma coleção logicamente coerente de dados com algum significado inerente a eles. Uma variedade aleatória de dados não pode ser corretamente chamada de banco de dados, pois ele deve ser projetado, construído e populado com dados para uma finalidade específica. Deve possuir um grupo definido de usuários e algumas aplicações previamente concebidas nas quais esses usuários possuem interesse. Os componentes de um banco de dados são: o próprio banco de dados e o Software Gerenciador do Banco de Dados (SGBD), um conjunto de programas que fazem o intermédio entre o banco de dados armazenado e os usuários e aplicações que têm acesso aos seus componentes [DATE 2004].

“O principal objetivo de um SGDB é proporcionar um ambiente tanto conveniente quanto eficiente para a recuperação e armazenamento das informações do banco de dados”, [SILBERSCHATZ et al. 2012]. Para tanto, o SGBD disponibiliza recursos para definir, construir, manipular, compartilhar, proteger e manter bancos de dados [ELMASRI and NAVATHE 2011], garantindo que todas as instâncias satisfaçam ao esquema ou modelo do banco de dados, respeitando sua estrutura e suas restrições. Portanto, podemos entender o SGBD como o conjunto de programas responsável por gerenciar o

banco de dados, pois fornecem suporte aos métodos de acesso, ou seja, maneiras de armazenar e recuperar informações de forma apropriada.

## 2.2. Modelo de Dados

Um modelo de dados é “uma coleção de conceitos que podem ser usados para descrever a estrutura de um banco de dados” [ELMASRI and NAVATHE 2011]. Portanto, o modelo de um banco de dados pode ser entendido como a forma de descrever os seus dados, relacionamentos, semântica e restrições de consistência. Entre os modelos existentes cita-se o relacional e o não-relacional, que foram utilizados neste trabalho, os quais são apresentados a seguir.

### 2.2.1. Modelo Relacional

Um banco de dados que segue o Modelo Relacional (MR) representa seus dados como um conjunto de relações. Considerando que uma relação é, de certo modo, similar a uma tabela de valores e aplicando a terminologia do MR diz-se que as linhas denominam-se tuplas; as colunas, atributos; e a tabela em si denomina-se relação [ELMASRI and NAVATHE 2011].

id [PK] integer	description character varying(200)	data_type character(1)	unit integer
1	Temperatura	3	1
2	Umidade Relativa	3	2
3	Fracao de Cobertura de Nuvens Baixas	3	1
4	Fracao de Cobertura de Nuvens Medias	3	1
5	Fracao de Cobertura de Nuvens Altas	3	1
6	Precipitacao	3	1
7	Radiacao de Onda Curta	3	1
8	Pressao Atmosferica	3	1
9	Velocidade Rajada do Vento	3	1
10	Direcao Vento	3	1

Figura 1. Exemplo de armazenamento em formato de tabela no MR.

As Relações ou tabelas, são constituídas de uma série de elementos (atributos), que possuem características comuns. No exemplo apresentado na Figura 1 tem-se uma tabela denominada *variable*. Nessa figura, pode-se observar também o conceito de atributo ou campo. Cada atributo indica uma determinada característica associada a tabela. Nesse caso, tem-se os atributos: *id*, *description*, *data\_type*, *unit*. O domínio é caracterizado como sendo os valores possíveis do atributo. As tuplas ou registros, indicam as linhas da tabela.

### 2.2.2. Modelo NoSQL

NoSQL é um termo genérico para uma classe definida de banco de dados que rompe o paradigma dos bancos de dados baseados nos Modelos Relacionais. O termo NoSQL foi primeiramente utilizado em 1998 como o nome de um banco de dados não-relacional de código aberto, criado por Carlo Strozzi. Segundo ele “o movimento NoSQL é completamente distinto do modelo relacional e, portanto, deveria ser mais apropriadamente chamado NoREL” [STROZZI 2016]. Porém, somente em 2009 o termo obteve maior

visibilidade quando abordado pela Google no artigo “BigTable: A Distributed Storage System for Structured Data”, que resgatou o NoSQL como um conceito de gerenciamento de Big Data. No artigo, o BigTable foi descrito como um banco extremamente escalável e tolerante a falhas onde os dados inseridos já seriam indexados, tornando assim mais rápida a consulta aos dados armazenados.

O NoSQL foi projetado para promover uma alternativa de armazenamento com velocidade e disponibilidade elevada. E começaram a ser desenvolvido por grandes empresas para suprir a necessidade de alto armazenamento e escalabilidade. Atualmente são tratados como o futuro de grande armazenamento de informações, conseguindo assim trabalhar com esse novo cenário que os bancos de dados Relacionais não atendem como desejado.

Os bancos de dados NoSQL são classificados de acordo com a forma de armazenar os dados. Existem 4 tipos de categorias: chave/valor, orientado a colunas, orientado a documentos e baseados em grafos [Shertil 2016]. O principal deles é o Armazenamento orientado a Documentos. Esse tipo de banco, armazena, recupera e gerencia dados semiestruturados. O elemento dos dados é chamado documento. Os documentos são endereçados no banco de dados via uma chave única que representa o documento. Uma das características de um banco de dados orientado a documentos é que, além da simples chave/valor de pesquisa, é possível recuperar um documento através de uma API, ou linguagem de consulta disponibilizada pelo banco [LEAVITT 2010]. Os documentos possuem uma estrutura comparável a um documento JSON. O qual é um padrão aberto de estruturação de dados baseado em texto e legível por humanos, utilizado para a serialização de dados. Foi projetado com o objetivo de ser simples, portátil e textual[JSON 2016].

### **2.3. MongoDB**

O MongoDB (do inglês humongous, “gigantesco”) é uma aplicação de código aberto, de alta performance, sem esquemas, orientado a documentos. Foi escrito na linguagem de programação C++, inicialmente sendo desenvolvido como um componente de serviço pela empresa 10gen em outubro de 2007 [MONGODB 2016].

O MongoDB é considerado um dos projetos mais notáveis NoSQL devido sua linguagem de consulta de alta performance, baseada em documentos e facilidade de migração de um banco de dados relacional, pelo fato de suas consultas serem convertidas facilmente.

### **2.4. Postgresql**

O PostgreSQL é um sistema de gerenciamento de banco de dados objeto-relacional (SGB-DOR). Foi desenvolvido pelo Departamento de Ciência da Computação da Universidade da Califórnia em Berkeley em 1998 [PostgreSQL 2016].

É compatível com a linguagem SQL ANSI, uma linguagem de consulta declarativa, que é a padrão para SGBD relacionais. Apresenta suporte a chave estrangeira, operações join, views, triggers, procedures, entre outros. Além disso, suporta a maioria dos tipos de dados da especificação SQL:2008, dentre eles estão: INTEGER, VARCHAR, BOOLEAN, NUMERIC, CHAR, DATE, INTERVAL e TIMESTAMP.

### 3. Metodologia

A implementação dos testes a seguir foi realizada a partir de teste práticos em ambos os bancos de dados fazendo uso de métricas previamente definidas, com base nas principais métrica utilizadas para avaliar o desempenho desses bancos de dados.

#### 3.1. Ambiente de Testes

Os testes foram realizados em um computador com processador Intel Core i7-3517U CPU de 2.4 GHz, 8 GB de RAM, com o sistema operacional Windows 10 Pro de 64 bits. A versão do PostgreSQL usada foi a 9.3 e a do MongoDB foi a 3.2.3. Para implementar os testes, a linguagem Java [JAVA 2016] foi utilizada juntamente com os respectivos drivers necessários para a conexão com os bancos. Para o PostgreSQL foi utilizado o driver PostgreSQL JDBC e para o MongoDB foi utilizado o driver Java Driver MongoDB. E a medição do tempo de execução das consultas, foi realizada utilizando o método *System.currentTimeMillis()* da linguagem JAVA.

Como sistema gerenciador do MongoDB, foi utilizada a interface gráfica Robomongo 0.9.0-RC6, uma aplicação nativa do MongoDB, criada pela empresa de desenvolvimento Paralect, a qual publicou a primeira versão do Robomongo em 28 de maio de 2012. Já no gerenciamento do PostgreSQL, foi utilizada a interface pgAdmin, uma ferramenta gráfica open source, desenvolvida pela equipe de desenvolvimento do PostgreSQL. Ela pode ser utilizada em sistemas como o Linux, Mac OSX, Windows, entre outros e gerencia as versões do PostgreSQL acima da 7.3.

Os testes foram divididos em cinco categorias: busca com a utilização de índices, busca sem a utilização de índices, inserção, atualização e exclusão. Os comandos de busca foram executados em loops de 1, 1.000, 10.000 repetições, bem como os comandos de inserção, ambos foram efetuados três vezes cada, sendo considerada a média aritmética dos mesmos. As demais operações (atualização e exclusão), foram realizadas em apenas uma execução, sem loops.

#### 3.2. Modelo do Estudo de Caso

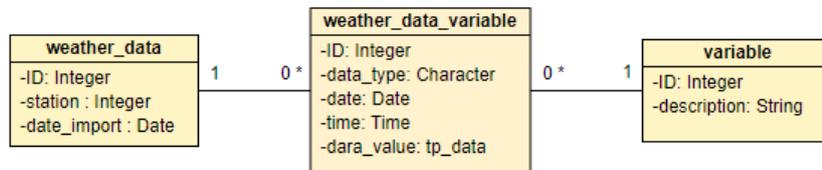
O banco de dados AgroDB [LAZZARETTI 2013] foi utilizado para realização dos testes, em específico, o conjunto de dados meteorológicos já existente. A base de dados utilizada no PostgreSQL, possui aproximadamente 89.000.000 (oitenta e nove milhões) de registros sendo que 45.000.000 (quarenta e cinco milhões) desses dados foram copiados para a base de dados MongoDB, para que os testes pudessem ser realizados.

A estrutura do banco no PostgreSQL foi estabelecida através do diagrama de classe, apresentado na Figura 2.

A tabela *variable* possui um atributo identificador (*id*) e um atributo para descrever o valor do registro (*description*), neste campo encontram-se valores como por exemplo: temperatura, radiação solar, latitude, entre outros. Já na tabela *weather\_data*, estão presentes, o identificador, os atributos *station*, onde é armazenado o número da estação de onde o registro foi coletado e o atributo *date\_import* que representa a data de importação do registro.

A tabela *weather\_data\_variable* possui os seguintes atributos:

1. Id – atributo identificador do registro no banco;



**Figura 2. Diagrama de classes.**

2. Data\_type – atributo que faz referência ao tipo de dado armazenado no atributo data\_value;
3. Date – data em que o registro foi armazenado no banco;
4. Time – hora em que o registro foi armazenado no banco;
5. Data\_value – campo de tipo abstrato "tp\_data", "com ele, é possível armazenar em um mesmo campo dados com formatos diferentes" [LAZZARETTI 2013];

A estrutura do documento no MongoDB foi construída utilizando as 3 tabelas do banco presentes no PostgreSQL como demonstrado na Figura 3.

```

{
  ID: 1,
  DATA TYPE: "2",
  DATE: "2016-05-02",
  TIME: "12:32",
  DATA_VALUE: 27,
  VARIABLE: {
    ID: 7,
    DESCRIPTION: "TMAX",
  }
  WHEATHER_DATA: {
    ID: 754,
    STATION: 8,
    DATE_IMPORT: "2016-01-30",
  }
}
  
```

**Figura 3. Estrutura do documento MongoDB [O AUTOR].**

Na modelagem orientada a documentos, a tabela weather\_data\_variable passou a ser a collection e seus registros, os documentos. Já as tabelas variable e weather\_data estão inseridas em cada documento como subdocumentos.

## 4. Resultados e Discussões

Nesta seção são apresentados os resultados obtidos e as discussões levantadas a partir da aplicação dos testes.

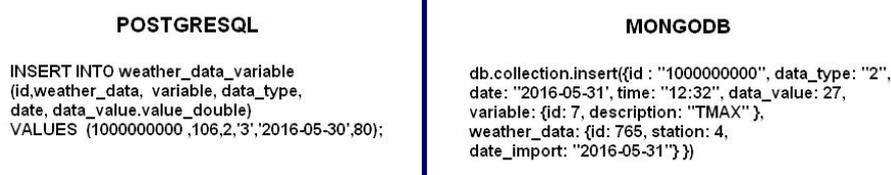
### 4.1. Comandos de Teste

Os testes foram divididos em cinco categorias, descritas a seguir:

1. Inserção;

2. Busca com a utilização de índices;
3. Busca sem a utilização de índices;
4. Atualização;
5. Exclusão.

Os testes se iniciaram pelos comandos de inserção, apresentados na Figura 4.



**Figura 4. Comandos de inserção utilizados.**

As inserções foram feitas englobando todas as tabelas no PostgreSQL e um documento completo e seus subdocumentos com todos os dados preenchidos no MongoDB. Os dados foram inseridos em duas execuções, na primeira execução foram inseridos 1.000 (mil) registros, já na segunda foram inseridos 10.000 (dez mil) registros, sendo no total 11.000 (onze mil) registros em cada um dos bancos. A seguir, na Tabela 1, são apresentados os resultados obtidos utilizando esses comandos.

**Tabela 1. - Resultados referentes ao tempo de resposta dos comandos de inserção em ambos os bancos de dados representados em segundos.**

Nº Registros	Banco de Dados	
	MongoDB	PostgreSQL
<b>1.000</b>	1,608s	2,015s
<b>10.000</b>	5,663s	15,993s

Nos resultados das operações de inserção, ambos os bancos de dados obtiveram bons resultados. No PostgreSQL, em aproximadamente 2 segundos foi realizada a inserção de 1.000 (mil) registros e com cerca de 16 segundos pode-se efetuar a inserção de 10.000 (dez mil) registros. O banco MongoDB apresentou tempos de inserção ainda menores. Sendo assim, em aproximadamente 1 segundo e meio foram inseridos 1.000 (mil) registros e em 5 segundos e meio 10.000 (dez mil) registros.

A busca utilizando índices consiste na busca de todos os registros ou documentos da tabela ou coleção weather\_data.variable que possuem 106 como o valor do campo weather\_data (weather\_data.id no caso do MongoDB). Na Figura 5, estão os comandos utilizados para a busca.



**Figura 5. Comandos de busca com índices.**

Os resultados da utilização dos comandos de busca utilizando índices, estão dispostos na Tabela 2.

**Tabela 2. Resultados das execuções de comandos de busca utilizando índices.**

Nº Execuções	Banco de Dados	
	MongoDB	PostgreSQL
<b>1</b>	0,350s	0,410s
<b>1.000</b>	0,613s	7,695s
<b>10.000</b>	4,020s	54191,975s

Na busca por registros utilizando índices, pode-se notar que o banco MongoDB obteve um melhor resultado em relação ao PostgreSQL quando a consulta foi executada apenas uma vez. O PostgreSQL executou a consulta 1.000 (mil) vezes em 7 segundos e meio, e o MongoDB em meio segundo. A mesma operação, com 10.000 (dez mil) execuções, foi finalizada após 15 horas no PostgreSQL, já no MongoDB, foram necessários apenas 4 segundos.

Na busca sem índices o resultado deve retornar todos os registros/documentos que possuem o campo date com valor entre 1980-01-01 e 1999-12-31. Os parâmetros \$gte e \$lte presentes no comando do MongoDB são operadores de comparação, esses parâmetros são utilizados para filtrar documentos onde o valor do campo é maior do que ou igual a, e menor ou igual a determinado valor.



**Figura 6. Comandos de busca sem índices.**

Na Figura 6, estão os dispostos os comandos utilizados para busca sem a utilização de índices, os quais foram utilizados em ambos os bancos. Os resultados obtidos a partir desses comandos são apresentados Tabela 3:

**Tabela 3. - Resultados dos comandos de busca sem a utilização de índices.**

Nº Execuções	Banco de Dados	
	MongoDB	PostgreSQL
<b>1</b>	0,466s	3,125s
<b>1.000</b>	0,926s	8,791s
<b>10.000</b>	3,500s	66555,960s

As consultas sem utilização de índices, retornaram aproximadamente 1.100.000 (um milhão e cem mil) registros/documentos por execução. E como demonstrado, o MongoDB obteve melhores resultados que o PostgreSQL quando índices não foram utilizados. No PostgreSQL foram necessárias 18 horas para finalizar as 10.000 (dez mil) consultas, sendo que no MongoDB as mesmas consultas foram executadas em 3,5 segundos.

Nos comandos de atualização de registros, foram modificados 1.200.000 (um milhão e duzentos mil) registros/documentos. A operação atualizou o campo data\_value para 0.95 em todos os registros/documentos da tabela/coleção weather\_data\_variable onde

o valor do campo variable fosse igual a 2 e o valor do campo weather\_data estivesse entre 98 e 109. Na Figura 7 estão os comandos utilizados para atualização.



**Figura 7. Comandos de atualização.**

Os resultados obtidos através da execução de comandos de atualização, estão presentes na Tabela 4. Onde é possível notar que os registros foram atualizados em 12 minutos no PostgreSQL e em 5 minutos no MongoDB.

**Tabela 4. Resultados obtidos ao executar os comandos de atualização de registros em ambos os bancos.**

Nº Execuções	Banco de Dados	
	MongoDB	PostgreSQL
1	300,540s	725,340s

Nos testes de exclusão foram apagados cerca de 1.100.000 (um milhão e cem mil) registros/documentos. A operação excluiu todos os registros/documentos que possuíam 106 como o valor do campo weather\_data (weather\_data.id no caso do MongoDB). Na Figura 8 estão os comandos utilizados para exclusão de registros.



**Figura 8. Resultados da execução de comandos de exclusão de registros em ambos os bancos [O AUTOR]**

**Tabela 5. Resultados referentes aos comandos de exclusão de registros.**

Nº Execuções	Banco de Dados	
	MongoDB	PostgreSQL
1	61,212s	163,932s

A tabela acima (Tabela 5), apresenta os resultados dos comandos de exclusão, onde é possível perceber que em 2 (dois) minutos a operação foi concluída no banco PostgreSQL e em 1 (um) minuto a mesma operação foi efetuada pelo MongoDB.

## 5. Conclusão

O presente trabalho apresentou uma comparação de desempenho do tempo de resposta de instruções que incluem, removem, alteram e consultam dados em bancos de dados relacionais e NoSQL.

Utilizando os dois bancos de dados na sua configuração padrão, sem nenhum tipo de ajuste de otimização e baseados no ambiente de testes proposto, pode-se concluir que o MongoDB obteve melhores resultados. Isso se dá pelo fato de que os bancos NoSQL terem sido desenvolvidos para suprir exatamente a demanda por performance, sua estrutura, forma de busca, de armazenamento e distribuição dos dados facilita a execução dos comandos, os tornando mais rápidos e eficientes nas respostas. Com isso, vale salientar, que os bancos NoSQL não vieram para substituir bancos tradicionais, e sim para servir de uma opção em potencial para situações onde estes ainda possuírem maiores dificuldades. Pois ambos os bancos utilizam paradigmas diferentes e, por sua vez, possuem diferentes finalidades.

Acredita-se que pode-se atribuir a melhor performance do SGBD MongoDB, mostrada neste trabalho, pelo fato de sua modelagem estar de acordo com o propósito e os padrões das aplicações NoSQL, evitando a existência de junções e tornando os procedimentos mais rápidos.

## References

- DATE, C. J. (2004). *Introdução a Sistemas de Bancos de Dados*. Elsevier, 6 edition.
- ELMASRI, R. and NAVATHE, S. B. (2011). *Sistemas de Bancos de dados*. Addison-Wesley, 6 edition.
- IBM. Big Data : at the speed of business. <http://www01.ibm.com/software/data/bigdata/what-is-big-data.html>. Acesso em 27 de Outubro 2019.
- JAVA (2016). *JAVA*. Accessed: 2016-06-17.
- JSON (2016). *JSON Schema*. Accessed: 2016-06-15.
- LAZZARETTI, T. A. (2013). *Integração de banco de dados e modelos de simulação de culturas para estimular o impacto de mudanças do clima no rendimento de grãos e na severidade da giberela em trigo*. PhD thesis. Faculdade de Agronomia e Medicina Veterinária – FAMV.
- LEAVITT, N. (2010). Will nosql databases live up to their promise? *Computer*, 43(2):12–14.
- MADISON, M., BARNHILL, M., NAPIER, C., and GODIN, J. (2015). Nosql database technologies. *Journal of International Technology and Information Management*, 24(1).
- MONGODB (2016). *MongoDB Documentation*. Accessed: 2016-03-10.
- PostgreSQL (2016). *PostgreSQL Documentation*. Accessed: 2016-03-10.
- Shertil, M. (2016). Traditional rdbms to nosql database: New era of databases for big data. *Journal of Humanities and Applied Science (JHAS)*, 1(29).
- SILBERSCHATZ, A., Korth, H. F., and Sudarshan, S. (2012). *Sistema de Banco de Dados*. Elsevier, 6 edition.
- STROZZI, C. (2016). *NOsql Documentation*. Accessed: 2016-04-18.
- VOLGELS, W. (2009). Eventually consistent. *Commun. ACM*, 52(1):40–44.

# Caracterizando os Desenvolvedores de Ecosystemas de Software

Marina Moreira<sup>1</sup>, Caio Steglich<sup>1</sup>, Sabrina Marczak<sup>1</sup>, Fernando Figueira Filho<sup>2</sup>,  
Marcelo Perin<sup>3</sup>, Cleidson de Souza<sup>4</sup>,

<sup>1</sup>Escola Politécnica – PUCRS  
Porto Alegre - RS - Brasil

<sup>2</sup>Departamento de Informática e Matemática Aplicada – UFRN  
Natal - RN - Brasil

<sup>3</sup>Fundação Getúlio Vargas - FGV  
São Paulo - SP - Brasil

<sup>4</sup>Faculdade de Computação – UFPA  
Belém - PA - Brasil

{marina.moreira, caio.borges}@acad.pucrs.br, sabrina.marczak@pucrs.br  
fernando@dimap.ufrn.br, cleidson.desouza@acm.org, mperin25@gmail.com

**Abstract.** *Software Ecosystems (SECO) are a set of communities composed by internal and external elements in which applications of various kinds have been developed, such as mobile applications. However, it is critical to understand who are the actors who develop these applications, what are the common and divergent characteristics between themselves, and what can influence them to participate in these ecosystems. This research aimed to identify the profile of mobile technology developers, as well as which dimensions (Technical, Business or Social) of a Software Ecosystem can influence these developers. In order to accomplish it, we conducted a survey, which receives 106 answers through an online questionnaire. Thus, this research confirms studies by other researchers who find Android and iOS as market leaders, and also we found out distinct perceptions of each dimension among the developers of these ecosystems. By understanding the profile of these developers, it collaborates to filling gaps that still present in the Software Ecosystem literature.*

**Resumo.** *Os Ecosystemas de Software (ECOS) são um conjunto de comunidades compostas por elementos internos e externos, nas quais se têm desenvolvido aplicativos de diversas naturezas, por exemplo, os aplicativos para dispositivos móveis. Contudo, é fundamental entender quem são os atores que desenvolvem esses aplicativos, quais são suas características em comum e divergentes entre si e o que os influencia a participar desses ecosystemas. Nesta pesquisa, desejou-se identificar o perfil dos desenvolvedores de tecnologias móvel, bem como quais dimensões (Técnica, Negócio ou Social) de um Ecosystema de Software podem influenciar estes desenvolvedores. Para isso, realizou-se uma survey, aonde coletou-se 106 respostas através de um questionário online. Assim sendo, confirmou-se estudos de outros pesquisadores que constatam Android e*

*iOS como os líderes de mercado, bem como percepções distintas de cada dimensão entre os desenvolvedores destes ecossistemas. Ao entender o perfil desses desenvolvedores, colabora-se assim preenchendo lacunas ainda presentes na literatura de Ecossistemas de Software.*

## **1. Introdução**

Ecossistemas de software (ECOS) são ecossistemas em torno das plataformas digitais nas quais têm se produzido soluções em software nos tempos atuais (exemplo: aplicativos). Um ECOS é composto por uma plataforma tecnológica, um conjunto interno e externo de desenvolvedores, os quais em parceria com uma comunidade que entende do domínio produzem aplicativos para atender as demandas de usuários [Bosch and Bosch-Sijtsema 2010].

Em alguns estudos da área de ECOS, como, por exemplo, De Souza et al. [de Souza et al. 2016] apontam que os elementos sociais (ou seja, estudos que abordem desenvolvedores ou usuários de um ecossistema móvel) são menos abordados do que os de outras naturezas dentro da literatura de ECOS, e isso demonstra a necessidade de mais estudos que explorem os atores que compõem um ecossistema, como os desenvolvedores. Por sua vez, [Dhungana et al. 2010] definem um ecossistema sustentável como aquele que consegue manter e melhorar sua capacidade de atrair e manter colaboradores, bem como seu potencial de inovação. Logo, os autores explicam que um ECOS apresenta dois grandes pilares para a sustentabilidade, sendo eles: 1) adaptar-se à novas tecnologias ou tendências; 2) atrair e manter pessoas ao entorno deste ECOS (por exemplo, usuários ou desenvolvedores). Em outras palavras, a sustentabilidade e sobrevivência de um ECOS estão relacionados à sua capacidade de atrair e manter os colaboradores engajados na comunidade.

Visando compreender mais a respeito do perfil dos desenvolvedores que colaboram, durante a condução desta pesquisa, convidou-se 1000 desenvolvedores que estavam em projetos de aplicativos móveis no Github. Além disso, buscou-se entender um pouco sobre a motivação destes desenvolvedores, tentando entender quais dimensões os têm levado a participar de ECOS.

O restante deste trabalho está estruturado da seguinte maneira: a Seção 2 contextualiza os principais conceitos do ECOS. A Seção 3 apresenta a metodologia utilizada e a condução da coleta de dados. Na Seção de 4 são apresentados os resultados da *survey*. A Seção 5 contém a discussão sobre a contribuição desse estudo. Adiante, a Seção 6 explicará as limitações do presente estudo. Por fim, a Seção 7 contém conclusões identificadas durante a condução deste estudo.

## **2. Fundamentação Teórica**

De acordo com [Bosch and Bosch-Sijtsema 2010], um ECOS é composto por desenvolvedores, uma comunidade de domínio, uma plataforma tecnológica e aplicativos móveis como resultado. Contudo, [de Lima Fontão et al. 2015] explicam que um ECOS é composto basicamente por sete elementos, sendo eles: Plataforma tecnológica, usuários, desenvolvedores, aplicativos, loja de aplicativos, comunidade e evangelistas (usuários ou desenvolvedores especiais que buscam inovação).

[Campbell and Ahmed 2010] explicam que existem três dimensões que compõem um ECOS, sendo elas:

**Técnica:** Compreende todos os elementos técnicos, como a plataforma, os aplicativos, as linguagens de programação, as técnicas de desenvolvimento, as ferramentas para desenvolvimento, entre outros elementos.

**Negócio:** Compreende todos os elementos relacionados ao negócio, como a monetização de aplicações, o salário dos desenvolvedores, as formas de distribuição das aplicações, entre outros elementos.

**Social:** Compreende todos os elementos sociais, como os desenvolvedores, os usuários, seus sentimentos, emoções, interações ou decisões quanto aos demais elementos de um ECOS ou quanto a si mesmos.

Quanto aos ECOS que se dedicam ao desenvolvimento de aplicativos móveis, [Mallinson 2015] pontua que atualmente Android e iOS são os líderes de mercado, sendo muito pequena a participação dos demais ECOS, praticamente vestigiais. Dessa forma, tem-se a necessidade de investigar qual o perfil dos desenvolvedores pertencentes aos maiores ECOS móveis e identificar suas características em comum e divergentes para entender possíveis motivos de sua adoção a tal ECOS.

Por fim, segundo [de Souza et al. 2016], o qual aponta que os estudos voltados para a dimensão Social são mais escassos quando comparado às dimensões de Negócio e Técnica, vê-se a necessidade de entender o fator social dentro do perfil dos desenvolvedores e qual o seu peso na adoção e permanência para um ECOS móvel.

## **2.1. Trabalhos Correlatos**

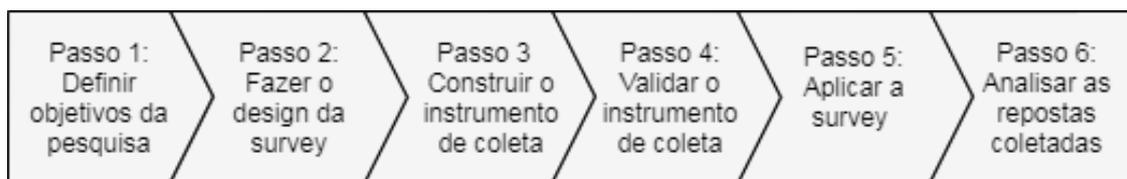
Existem alguns outros estudos que estudam os desenvolvedores e suas motivações para participarem de um ECOS, como, por exemplo, [Lim et al. 2014], que investigaram, através de uma *survey* (com o uso de um questionário), as diferenças de comportamento dos usuários em diferentes países e os desafios que isto trouxe à Engenharia de Software.

Outro estudo correlato seria o de [Steglich et al. 2019], que através de um estudo empírico identificaram 11 fatores referentes a dimensão social e exploraram como eles são vistos pelos desenvolvedores de ECOS.

Por fim, o estudo de [Ryu et al. 2014] investigaram a lealdade dos desenvolvedores de ECOS para com suas plataformas. Utilizando-se de uma *survey*, obtiveram 133 respondentes através de questionário e assim, elaboraram um constructo que envolve fatores técnicos, de negócio e sociais, indicando que através desses o desenvolvedor torna-se fiel a plataforma.

## **3. Metodologia de Pesquisa**

A metodologia de pesquisa adotada para execução desse estudo foi uma *survey*. Segundo [Kitchenham and Pfleeger 2001] uma *survey* não é apenas um instrumento de coleta (*e.g.* um questionário ou uma *checklist*), mas sim um sistema compreensivo de coleta de informações para descrever, comparar ou explicar conhecimentos, atitudes ou comportamentos, do qual faz parte de um processo maior. O restante dessa seção, apresenta os passos definidos



**Figura 1. Passos da coleta de dados adotada**

pelos autores [Kitchenham and Pfleeger 2001, Kitchenham and Pfleeger 2002a, Kitchenham and Pfleeger 2002b, Kitchenham and Pfleeger 2002c, Kitchenham and Pfleeger 2002d, Kitchenham and Pfleeger 2003]. Suas etapas são apresentadas na Figura 1.

### 3.1. Definição de Objetivos

Como o primeiro passo descrito por [Kitchenham and Pfleeger 2001], definiu-se como objetivo deste estudo a caracterização do perfil dos desenvolvedores de Ecosistemas de Software que têm lidado com tecnologias móveis (*e.g.*, Android ou iOS). Também, desejou-se identificar a importância que o desenvolvedor atribui a cada uma das dimensões apresentadas por [Campbell and Ahmed 2010], sendo essas: Técnica, Negócio ou Social.

Este objetivo torna-se importante pois, como citado na introdução, [Dhungana et al. 2010] explicam que um dos fatores principais para a sustentabilidade e sobrevivência de um ecossistema ao longo do tempo é sua capacidade de atrair ou manter pessoas engajadas. Dessa forma, os desenvolvedores representam uma força vital para o bom funcionamento de um ecossistema, por serem esses que elaboram as soluções em software para os usuários finais.

Para tanto, assumiu-se como questões de pesquisa:

(QP1) Qual o perfil profissional do desenvolvedor de um ECOS?

(QP2) Qual a importância de cada uma das dimensões (Técnica, Negócio e Social) de um ECOS percebida pelo desenvolvedor para começar e para continuar participando de um ecossistema?

### 3.2. Design da Survey

Baseado nas questões de pesquisa, o *design* desta *survey* foi elaborado e validado para tentar ampliar a qualidade das questões realizadas aos desenvolvedores. O perfil inicial desejado para o contexto desta pesquisa foram desenvolvedores de tecnologias móveis, que conseqüentemente estejam inseridos nos Ecosistemas de Software.

Para a elaboração do questionário, foram utilizadas um conjunto de questões fechadas com poucas questões abertas desejando identificar detalhes sobre suas experiências ou alguma peculiaridade. Desta forma, a *survey* identificou dados em formato quantitativo, mas a análise realizada foi descritiva.

### 3.3. Instrumento de Coleta

Definiu-se as seguintes questões para a *Survey*, presentes na Tabela 1, visando entender desde dados demográficos até o impacto que cada dimensão tem para que o desenvolvedor

**Tabela 1. Questões Realizadas aos Desenvolvedores durante a Survey**

Questão de Pesquisa	Questões da Survey
QP1	(A) Você já desenvolveu aplicativos para dispositivos móveis? Opções: múltipla escolha - Sim, Profissionalmente - Sim, mas não profissionalmente - Sim, eu já contribuí com um ou mais projetos open-source.
QP1	(B) Você normalmente desenvolve aplicativos móveis em uma equipe? Opções: escolha única - Sim, eu geralmente desenvolvo em equipe (com outros) - Não, eu geralmente desenvolvo sozinho
QP1	(C) Para qual plataforma você mais desenvolveu recentemente? Opções: escolha única - Android - iOS - Windows Phone
QP1	(D) Quantos aplicativos você desenvolveu para essa plataforma móvel? Opções: de 0 a 50+
QP1	(E) Há quanto tempo (em anos) você desenvolve aplicativos para essa plataforma móvel? Opções: de 0 a 50+
QP1	(F) Antes de desenvolver para essa plataforma, você já teve alguma experiência no desenvolvimento de aplicativos para outras plataformas móveis? Se sim, qual e como? Opção: Questão dissertativa (aberta)
QP2	(G) Até que ponto as dimensões (Técnica, Negócio ou Social) abaixo influenciaram sua decisão de <b>começar</b> a desenvolver para essa plataforma? Opções - Para cada fator: - Me influenciou muito - Me influenciou um pouco - Nunca me influenciou
QP2	(H) Até que ponto as dimensões (Técnica, Negócio ou Social) abaixo influenciaram sua decisão de <b>continuar</b> a desenvolver para essa plataforma? Opções - Para cada fator: - Me influenciou muito - Me influenciou um pouco - Nunca me influenciou

comece a desenvolver em uma plataforma ou continue desenvolvendo em uma determinada plataforma.

### 3.4. Validação da Survey

Após a elaboração do instrumento de coleta, dois pesquisadores com mais de 20 anos de atuação em Engenharia de Software revisaram o protocolo. Todas as divergências foram debatidas entre os pesquisadores quanto ao questionário e devidamente alinhadas de forma que o questionário aplicado na Tabela 1 foi o resultado final.

Além disso, realizou-se 2 testes piloto, nos quais concluiu-se que os respondentes

levariam em média de 5 a 10 minutos para responderem ao questionário, de tal maneira que não fosse preciso reorganizar o questionário em decorrência de excesso de perguntas ou de tempo de respostas.

### 3.5. Aplicação da *Survey*

O questionário foi enviado para uma população de 1000 desenvolvedores de tecnologias móveis no Github, tendo como resposta 112 desenvolvedores. Contudo, 6 destas respostas estavam incompletas, resultando assim no conjunto de 106 respondentes.

Os perfis selecionados através do Github foram de desenvolvedores que trabalham em projetos de tecnologias móveis, tendo esses participado de um ou mais projetos hospedados no Github.

### 3.6. Análise dos Dados Coletados

A análise dos dados foi realizada através da estatística descritiva, a fim de identificar padrões nos dados ou possíveis tendências. Desta forma, o resultado dessa análise está presente na Seção 4, na qual os dados auxiliam na resposta das questões de pesquisa propostas para este estudo.

## 4. Resultados

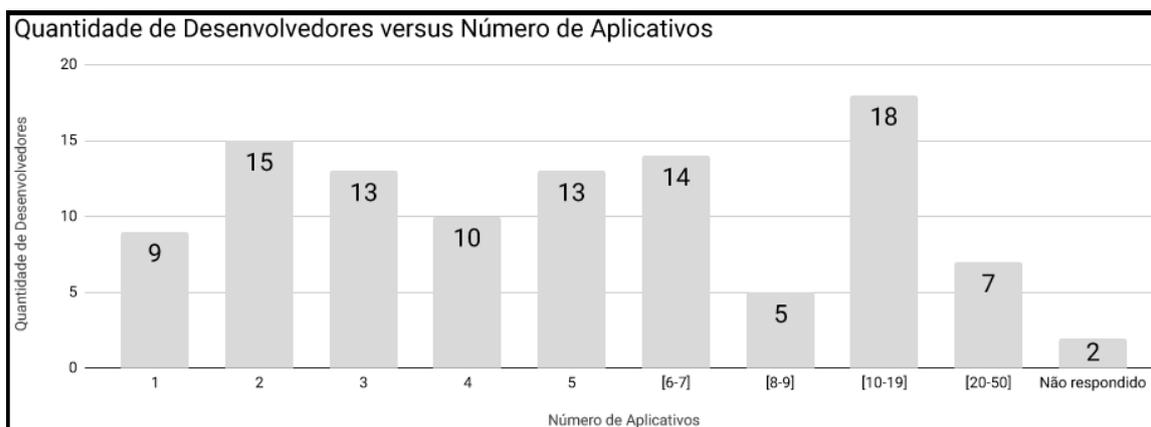
Nesta seção serão apresentados os resultados obtidos da *survey*, aonde as subseções são focadas em responder as questões de pesquisa. Os resultados das questões realizadas na *survey* são demonstrados textualmente com a respectiva letra da questão entre parênteses (Tabela 1).

### 4.1. O Perfil Atual dos Desenvolvedores de Ecossistemas de Software (QP1)

Os desenvolvedores convidados são desenvolvedores que podem ter diferentes relacionamentos com o ecossistema no qual estão inseridos, pois quando questionados se desenvolveram aplicativos móveis Questão (A), 56,6% dos respondentes informam que desenvolveram ou desenvolvem profissionalmente. Além disso, 77,4% dos respondentes informam que atuaram de forma não profissional (por exemplo, em projetos de curso/graduação ou pessoais) em algum momento. Por fim, 34,9% dos desenvolvedores entrevistados informam que atuaram em minimamente um projeto *open-source*.

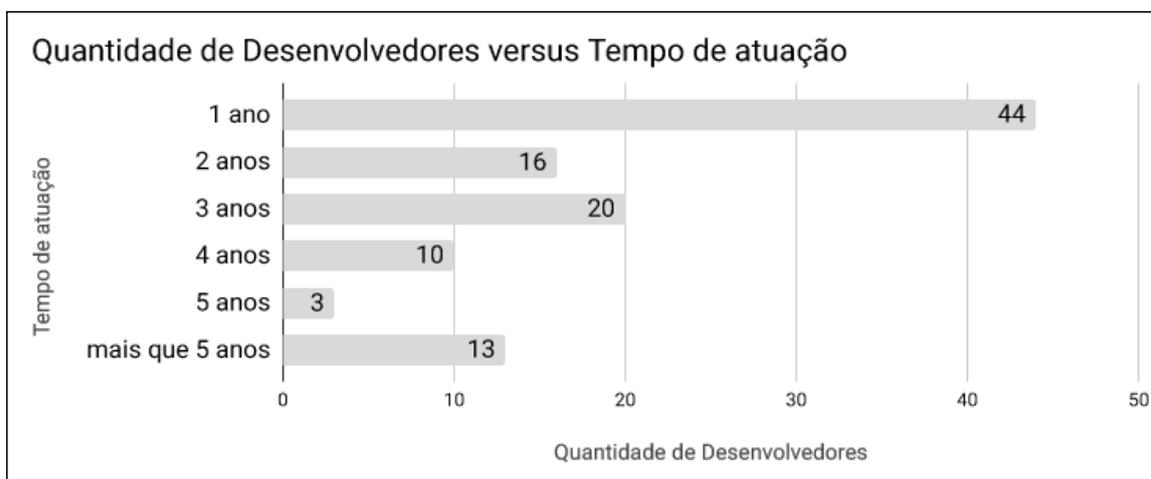
Quando questionados sobre desenvolverem aplicativos em equipe (B), 52,8% dos desenvolvedores informam que costumam desenvolver aplicativos sozinhos, enquanto 47,2% informam que geralmente desenvolvem em equipe. Nenhum desenvolvedor alegou nunca ter desenvolvido aplicativos antes. As plataformas que os desenvolvedores tem colaborado (C), são majoritariamente as apontadas por [Mallinson 2015], aonde 63,2% dos respondentes informam desenvolver para Android e 34,9% para iOS. Também, 1,9% dos respondentes informaram desenvolver para Windows Phone, mas nenhum dos entrevistados afirmou desenvolver para outra plataforma além destas três.

Estes desenvolvedores possuem diferentes níveis de experiência quanto a elaboração de soluções móveis (D). A Figura 2 representa quantos aplicativos foram desenvolvidos por X número de desenvolvedores. Assim, o eixo Y representa a quantidade de desenvolvedores e o eixo X diz respeito ao número de aplicativos. Por exemplo, 9 desenvolvedores afirmam ter desenvolvido apenas 1 aplicativo, enquanto 7 afirmam terem



**Figura 2. Quantidade de Aplicativos por Desenvolvedores**

desenvolvido mais de 20 aplicativos, sendo o máximo que um respondente afirmou ter desenvolvido foram aproximadamente 50 aplicativos. Além disso, o tempo de atuação com tecnologias móveis também se altera (E), sendo esse tempo disposto na Figura 3, aonde aproximadamente 42% dos respondentes estão atuando em plataformas de tecnologias móveis a apenas 1 ano, mas 12.5% estão atuando com tecnologias móveis a mais de 5 anos.



**Figura 3. Tempo de Experiência dos Desenvolvedores**

Apesar de alguns desenvolvedores terem informado que colaboraram com outras plataformas anteriormente (F), a maioria destes informou basicamente os mesmos. Por exemplo, desenvolvedores que atuam com iOS apontam que tentaram anteriormente Android ou Windows Phone, assim como alguns de Android também tentaram iOS ou Windows Phone. Alguns desenvolvedores mais antigos tentaram Blackberry entre outras que atualmente, de acordo com [Mallinson 2015], estão cada vez em maior desuso.

#### **4.2. Influência das Dimensões de ECOS na Perspectiva dos Desenvolvedores (QP2)**

Em um panorama geral, tanto para começar (G) quanto para continuar (H), ambas questões tiveram três opções, sendo essas: 1) Influenciou-me muito, 2) Influenciou-me

um pouco, ou 3) Não me influenciou em nada. Para tanto, realizou-se um cálculo proporcional, aonde todas as respostas "muito importante" valeriam 1 ponto, todas "um pouco importante" valeriam 0.5 e as "não importante" valeriam 0. Após, dividiu-se pelo número total de respondentes, para demonstrar a importância percebida ao começar e ao continuar.

Para começar suas participações (G), o desenvolvedor considera os fatores técnicos extremamente importantes (77,4%), os de negócio pouco menos importantes (60,4%) e a dimensão social a menos importante (56,1%).

Para continuarem suas participações (H), o desenvolvedor considera todas dimensões um pouco mais importante, sendo a dimensão técnica a mais considerada (81,3%), seguida pela dimensão de negócio (64,6%) e pela social (60,8%). Desta forma, segue um cenário similar aonde as tecnologias e adaptações continuam sendo os fatores mais desejados pelos desenvolvedores. Contudo, a dimensão social foi a que apresentou o maior crescimento, apesar de singelo (4,7%) se comparado com a técnica (3,9%) e de negócio (4,2%).

Analisando apenas os respondentes que desenvolvem iOS, a dimensão técnica é a que mais cresceu, sendo de aproximadamente 9,5% esse crescimento, considerando o momento que os desenvolvedores iniciaram sua participação e para continuarem participando. Também, a dimensão social cresceu mais que a de negócio, sendo respectivamente 8,2% e 5,3%.

Na perspectiva dos desenvolvedores Android, apesar da dimensão técnica ser considerada a mais importante, foi a única que perdeu importância (1,5%) de quando começam para se manterem participando. Além disso, a dimensão social (58,9% para começar e 61,9% para continuar) foi considerada mais importante que a de negócio (55,9% para começar e 59,7% para continuar) em ambos cenários, contudo, a dimensão de negócio é a que mais tem crescido em importância (3,8%).

## **5. Discussão**

O perfil dos desenvolvedores é variado, aonde muitos têm atuado em projetos pessoais ou profissionais, com aproximadamente um terço dos desenvolvedores interessados em projetos open source, sendo que esses têm trabalhado sozinhos ou em equipe. Confirma-se o descoberto por [Mallinson 2015], uma vez que a maior parte dos desenvolvedores atuam na plataforma Android, seguidos pelos desenvolvedores iOS. Ao todo, boa parte dos desenvolvedores de tecnologias móveis desenvolveram até 10 aplicativos. Também, mais da metade dos desenvolvedores entrevistados possuíam menos de 3 anos de experiência, ou seja, ainda são principiantes nos ecossistemas a que fazem parte.

Das dimensões apontadas por [Campbell and Ahmed 2010], em todos os cenários a dimensão técnica foi a melhor avaliada pelos desenvolvedores, quanto ao impacto em suas decisões de começar ou se manter participando de um ecossistema. Para desenvolvedores Android, a dimensão social aparece como mais importante que a de negócio, mas para desenvolvedores iOS, a dimensão de negócio segue sendo mais importante que a social.

Além disso, ao iniciar, o desenvolvedor busca principalmente se inteirar das tecnologias do ecossistema, não se preocupando tanto com os elementos da dimensão social. Identificou-se assim a necessidade de estudos futuros para o entendimento em de-

talhe de como a dimensão social é vista por desenvolvedores de ecossistemas diferentes, contrastando como os desenvolvedores Android ou iOS percebem estas dimensões ou a importância destas na sua escolha de começar ou permanecer participando de um ecossistema de software.

## **6. Limitações**

Os desenvolvedores convidados foram todos identificados no Github, o que poderia gerar alguma distinção se comparado com desenvolvedores de outras plataformas digitais. Contudo, os cenários apontados por demais autores (como [Mallinson 2015] e [de Souza et al. 2016]) se confirmam com os dados coletados, pela escassez de pessoas atuando em outros ecossistemas diferentes do Android e iOS. Também, a dimensão técnica é a mais considerada em panorama geral.

Baseado nos dados identificados, esses podem ser confirmados com estudos qualitativos e empíricos visando entender em detalhe as respostas dos entrevistados, por exemplo, identificar e entrevistar especialistas (desenvolvedores experientes em um ecossistema) visando entender o real impacto de cada dimensão na influência aos desenvolvedores. Por fim, outra limitação é a amostragem, que foi coletada por conveniência (não probabilística), verificando quais membros estavam ativos dentro dos projetos no GitHub, sem filtrar tempo de participação ou demais critérios de aceite para os participantes da *survey*.

## **7. Conclusão**

A realização deste estudo auxiliou no entendimento dos perfis dos desenvolvedores de tecnologias móveis no contexto atual, bem como confirmar estudos de outros pesquisadores, que demonstram cenários reais sobre os ecossistemas de software da atualidade.

Constatou-se um comportamento geral dos desenvolvedores, mas se observou que desenvolvedores de ecossistemas distintos possuem necessidades diferentes, por exemplo, os desenvolvedores Android considerarem a dimensão social mais importante que a de negócio. Também, os desenvolvedores iOS têm se preocupado ainda mais com a dimensão social após obterem determinado tempo de experiência.

Os dois ecossistemas que atualmente são líderes de mercado (Android e iOS) possuem propostas diferentes, permanecendo ativos por possuírem maior número de vendas de dispositivos aos usuários e usuários que investem nestes ecossistemas. Android é o líder de mercado, alcançando maior número de usuários e de desenvolvedores porém, iOS segue consolidado no mercado apostando na qualidade dos aplicativos e serviço a seus usuários. Fato que se confirma pela preocupação com o negócio por parte de seus desenvolvedores.

## **Agradecimentos**

Os autores agradecem ao CNPq (processos 420801/2016-2 e 311256/2018-0) pelo apoio financeiro.

## **Referências**

[Bosch and Bosch-Sijtsema 2010] Bosch, J. and Bosch-Sijtsema, P. (2010). From integration to composition: On the impact of software product lines, global development and ecosystems. *Journal of Systems and Software*, 83(1):67–76.

- [Campbell and Ahmed 2010] Campbell, P. R. and Ahmed, F. (2010). A three-dimensional view of software ecosystems. In *Proceedings of the European Conference on Software Architecture: Companion Volume*, pages 81–84. ACM.
- [de Lima Fontão et al. 2015] de Lima Fontão, A., dos Santos, R. P., and Dias-Neto, A. C. (2015). Mobile software ecosystem (mseco): a systematic mapping study. In *IEEE Annual Computer Software and Applications Conference*, volume 2, pages 653–658. IEEE.
- [de Souza et al. 2016] de Souza, C. R., Figueira Filho, F., Miranda, M., Ferreira, R. P., Treude, C., and Singer, L. (2016). The social side of software platform ecosystems. In *Proceedings of the Conference on Human Factors in Computing Systems*, pages 3204–3214. ACM.
- [Dhungana et al. 2010] Dhungana, D., Groher, I., Schludermann, E., and Biffel, S. (2010). Software ecosystems vs. natural ecosystems: learning from the ingenious mind of nature. In *Proceedings of the European Conference on Software Architecture: Companion Volume*, pages 96–102. ACM.
- [Kitchenham and Pfleeger 2001] Kitchenham, B. and Pfleeger, S. (2001). Principles of survey research: part 1: turning lemons into lemonade. *ACM SIGSOFT Software Engineering Notes*, 26:16–18.
- [Kitchenham and Pfleeger 2002a] Kitchenham, B. and Pfleeger, S. (2002a). Principles of survey research part 2: designing a survey. *ACM SIGSOFT Software Engineering Notes*, 27:18–20.
- [Kitchenham and Pfleeger 2002b] Kitchenham, B. and Pfleeger, S. (2002b). Principles of survey research: part 3: constructing a survey instrument. *ACM SIGSOFT Software Engineering Notes*, 27:20–24.
- [Kitchenham and Pfleeger 2002c] Kitchenham, B. and Pfleeger, S. (2002c). Principles of survey research part 4: questionnaire evaluation. *ACM SIGSOFT Software Engineering Notes*, 27:20–23.
- [Kitchenham and Pfleeger 2002d] Kitchenham, B. and Pfleeger, S. (2002d). Principles of survey research: part 5: populations and samples. *ACM SIGSOFT Software Engineering Notes*, 27:17–20.
- [Kitchenham and Pfleeger 2003] Kitchenham, B. and Pfleeger, S. (2003). Principles of survey research part 6: data analysis. *ACM SIGSOFT Software Engineering Notes*, 28:24–27.
- [Lim et al. 2014] Lim, S. L., Bentley, P. J., Kanakam, N., Ishikawa, F., and Honiden, S. (2014). Investigating country differences in mobile app user behavior and challenges for software engineering. *IEEE Transactions on Software Engineering*, 41(1):40–64.
- [Mallinson 2015] Mallinson, K. (2015). Smartphone revolution: Technology patenting and licensing fosters innovation, market entry, and exceptional growth. *IEEE Consumer Electronics Magazine*, 4(2):60–66.
- [Ryu et al. 2014] Ryu, M. H., Kim, J., and Kim, S. (2014). Factors affecting application developers’ loyalty to mobile platforms. *Computers in Human Behavior*, 40:78–85.
- [Steglich et al. 2019] Steglich, C., Marczak, S., de Souza, C. R., Guerra, L. P., Mosmann, L. H., Perin, M., et al. (2019). Social aspects and how they influence mseco developers. In *Proceedings of the International Workshop on Cooperative and Human Aspects of Software Engineering*, pages 99–106. IEEE Press.

# Uso de Redes Neurais Convolucionais para a identificação e classificação de Insetos

Bruno Bearsi Paixão, João Mário Lopes Brezolin<sup>1</sup>

<sup>1</sup>Instituto Federal Sul-Riograndense - Campus Passo Fundo  
Estrada Perimetral Leste, 135

bruno.bearsi@gmail.com

joao.brezolin@passofundo.ifsul.edu.br

**Abstract.** *The winter crops are affected by these insects that feed on their sap causing damage. Currently, the monitoring of these populations is made manually by counting species of insects caught in traps. The process is manual and error-prone. To assist in this process, a system was developed to perform the counting of elements extracted from images and classified by using a convolutional neural network. The present article presents the results of this study.*

**Resumo.** *As culturas de inverno são afetadas por insetos que alimentam da sua seiva ocasionando danos. Atualmente o monitoramento dessas populações ocorre de forma manual a partir da contagem de espécies capturadas em armadilhas. Atualmente esse processo, é manual e passível de erros. Para auxiliar nessa tarefa desenvolveu-se um sistema que realiza a contagem de elementos extraídos de imagens e classificadas por meio de uma rede neural convolucional. O presente artigo descreve o sistema implementado e apresenta os resultados desse estudo.*

## 1. Introdução

Os insetos exercem papel fundamental para o equilíbrio ecológico do planeta. Entretanto, para garantir a sua subsistência, os mesmos se alimentam de culturas de grãos como trigo e milho. Desse fato, decorrem perdas significativas o que demanda estabelecer o controle das populações dessas espécies visando minimizar e prevenir danos causados por insetos [Fabretti 2013]. Atualmente o controle dessas populações é realizado a partir de amostras de capturadas em armadilhas. As armadilhas compreendem bacias que são nas lavouras durante um determinado espaço de tempo e que armazenam que tem contato com as plantas. Ao final desse período as armadilhas são retiradas das plantações e encaminhadas a um especialista de domínio que realiza a contagem manual das espécies capturadas. O processo manual, ocasiona demora na contabilização da população de insetos o que acaba por atrasar a tomada de medidas que poderiam minimizar os danos nas lavouras. Além disso, o processo é suscetível a erros, uma vez, que é cansativo e demanda muita atenção do técnico responsável.

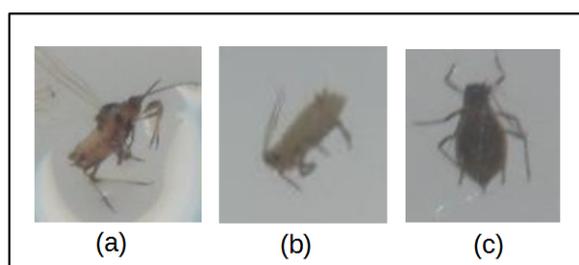
Observou-se que esse processo poderia ser automatizado com o uso de técnicas de visão computacional e aprendizado de máquina. Dessa forma, na primeira etapa desse projeto, optou-se pelo desenvolvimento de uma ferramenta computacional que realiza a identificação de afídeos com auxílio de uma Rede Neural Artificial (RNA). Realiza-se

o recorte dos potenciais contornos que permitirão estabelecer o banco de imagens que serão utilizadas para o treinamento da rede neural [Lins 2018]. Nesse contexto, a RNA era treinada a partir das características (forma, tamanho e área) extraídas dos elementos identificados nas imagens [Belloni 2018]. Apesar dos resultados satisfatórios, observou-se, que as Redes Neurais Convolucionais (RNC) poderiam permitir obter resultados melhores aos obtidos na estratégia utilizada anteriormente. Com RNC, torna-se possível realizar a classificação das espécies a partir das amostras de imagens dos insetos capturados. Isto permite identificar um número maior de características o que torna o classificador mais preciso. O uso de imagens para realizar a classificação de insetos, porém, demanda que o treinamento da rede neural seja realizado com imagens qualificadas. Nesse sentido observou-se que seria necessário uma metodologia para melhorar a segmentação dos insetos identificados nas imagens dado ao fato de que é possível identificar nessas imagens insetos sobrepostos o que acabaria por induzir o classificador a erros no processo de identificação de espécies de insetos.

O presente artigo descreve o desenvolvimento de um sistema de classificação de insetos baseado em Redes Neurais Convolucionais assim como o processo de validação do mesmo. Para realizar a implementação da RNC utilizou-se o *framework* Keras/Tensorflow. O tratamento e segmentação das imagens foi realizado com auxílio da biblioteca OpenCV. Este artigo está estruturado como segue: a seção 2 apresenta o referencial teórico. Nessa apresenta-se a problemática sobre o controle de insetos e a revisão acerca das tecnologias utilizadas. A seção 3 apresenta a arquitetura do sistema desenvolvido. A seção 4 traz os resultados da validação do sistema e, por fim, na seção 5 encontram as considerações finais sobre o sistema desenvolvido.

## 2. Referencial teórico

Afídeos são insetos pequenos que se alimentam dos nutrientes produzidos pela seiva da planta [Stern 2008]. Estes insetos causam diferentes tipos de danos. O dano direto ocorre quando o afídeo se alimenta da seiva da planta causando o enfraquecimento da mesma.



**Figura 1. Imagens de afídeos**

O dano indireto ocorre quando o afídeo instalado na planta elimina uma substância conhecida como *honeydew*, que promove o surgimento de fungos comprometendo a respiração da planta. Por fim, os insetos podem transmitir doenças como o Vírus do Nanismo Amarelo da Cevada (VNAC), que adoce a planta, podendo causar o amarelamento de suas folhas e diminuição no tamanho da planta [Gallo 2002]. Também conhecidos como pulgões, os afídeos são classificados em três categorias durante sua vida: a fase inicial do afídeo intitula-se ninfa; ao entrar na fase adulta, podem variar entre duas

classes, desta forma, são classificados como ninfas, ápteros (adulto sem asas) ou alado (adulto com asas). As três classes podem ser observadas na Figura 1.

## 2.1. Redes Neurais Convolucionais

Uma rede neural pode ser classificada como um processador paralelo que contém unidades de processamento (neurônios) que tem como objetivo armazenar elementos que serão utilizados no processo de classificação. O conhecimento é adquirido a partir de um processo de aprendizagem que utilizam pesos sinápticos para armazenar, e refinar o conhecimento produzido. O algoritmo de aprendizagem, que é usado para o treinamento da rede neural, tem a função modificar os pesos sinápticos da rede, com o intuito de conduzir o classificador a identificação dos objetos desejados. As RNAs são compostas por uma r uma camada de entrada, uma ou mais camadas ocultas de neurônios na qual o sinal de entrada se propaga em um único sentido, da entrada da rede até a saída (*feed-forward*) [Haykin 2003].

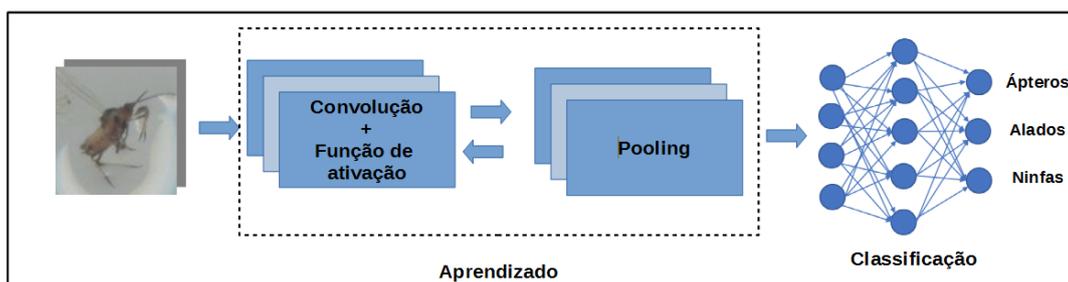
As Redes Neurais Convolucionais (RNC) são uma variação das RNAs de múltiplas camadas. As RNC caracterizam-se por aplicar filtros em dados visuais, mantendo a relação de vizinhança entre os pixels da imagem durante o processamento da rede [WU 2017]. O processo de classificação inicia-se nas camadas de convolução, os neurônios dessas camadas aplicam filtros e adiciona-se pesos as suas conexões. A combinação dessas entradas de neurônios produz saídas que são processada pelas camadas seguintes. Nas RNCs apenas um subconjunto de entradas e conectado a cada neurônio. O tamanho do filtro define o tamanho da vizinhança que cada neurônio da camada irá processar. Além disso, o passo, representa a quantidade de pixels que serão pulados e qual o tamanho da camada que será processada no neurônio seguinte. Após cada convolução aplica-se uma função de ativação não-linear. Para realizar a tarefa de classificação deve-se acrescentar ao menos uma camada totalmente conectada que é responsável por traçar um caminho de decisão a partir das resposta dos filtros vindos das camadas anteriores. Finalmente, depois da camada completamente conectada o ultimo passo é a função de classificação, que é a camada responsável pelo treinamento da rede. que são processada pelas camadas seguinte.

Nas RNCs apenas um subconjunto de entradas e conectado a cada neurônio. O tamanho do filtro define o tamanho da vizinhança que cada neuronio da camada irá processar Além disso, o passo, representa a quantidade de pixels que serão pulados e qual o tamanho da camada que será processada no neurônio seguinte. Após cada convolução aplica-se uma função de ativação não-linear. Para realizar a tarefa de classificação deve-se acrescentar ao menos uma camada totalmente conectada que é responsável por traçar um caminho de decisão a partir das resposta dos filtros vindos das camadas anteriores. Finalmente, depois da camada completamente conectada, o ultimo passo é a função de classificação, que é a camada responsável pelo treinamento da rede [Haykin 2003].

## 2.2. Framework Keras/Tensorflow

TensorFlow é uma *framework* de aprendizado de máquina aplicável que permite implementar e treinar redes neurais artificiais [Martín Abadi 2016]. Keras é um freamework que permite implementar redes neurais e realizar a interface com o TensorFlow.

O *framework* Keras/Tensorflow é atualmente a biblioteca mais completa na implementação de diversos tipos de redes neurais e funções auxiliares para a prática do



**Figura 2. Processo de classificação RNC**

aprendizado de máquina. A RNC utilizada para a classificação de insetos foi desenvolvida através deste *framework*.

### 2.3. Numpy

Numpy é uma biblioteca para computação científica que implementa um array multidimensional muito eficiente e flexível para aplicações com transformações e álgebra linear.

### 2.4. OpenCV: Open Computer Vision Library

A biblioteca OpenCV (*Open Source Computer Vision Library*) compreende um conjunto de ferramentas de processamento de imagem e vídeo. Na biblioteca encontram-se diversas funcionalidades de visão computacional tais como como: Filtros de imagem, calibração de câmera, reconhecimento de objetos, análise estrutural e entre outros [Huaman et al. 2018]. A biblioteca foi utilizada para refinar a segmentação das imagens de insetos como será descrito a seguir.

### 2.5. Python Flask

O Flask é um microframework que permite criar webservices em Python [Grimberg 2018]. O microframework é baseado na biblioteca WSGI (*Web Server Gateway Interface*). Na aplicação desenvolvida o Flask foi utilizado para estabelecer um mecanismo para automatizar o recebimento de imagens de insetos e o classificador.

## 3. Arquitetura do sistema desenvolvido

No sistema desenvolvido, a entrada de uma imagem é realizada por uma interface desenvolvida utilizando o Python Flask. A imagem então é carregada no sistema em escala de cinza. Uma cópia dessa imagem é criada são identificados a partir do tamanho os contornos que podem ser considerados potenciais insetos. O inseto é recortado do resto da imagem e inserido em uma imagem de cor sólida semelhante ao fundo da imagem original. A imagem então é convertida para um array, que passa pelas camadas convolucionais que possuem 9, 18 e 36 filtros. O tamanho do quadro utilizado nas convoluções é 3x3 e a função de ativação utilizada nas camadas é a ReLU (*Rectified Linear Unit*).

A ReLU permite facilitar o treinamento de redes profundas, pois age como função linear ao receber valores positivos, permitindo aprendizado eficiente através do *backpropagation* de erros, e age como não-linear caso receba valores negativos. Nesse último caso, a saída da função será o valor 0, o que permite o aprendizado de relações complexas entre os dados ao mesmo tempo que não satura ou suprime a sensibilidade de seus

nós o que dificulta a aprendizagem. Na camada densa da RNC foi utilizada a função Leaky Relu, ela possui características semelhantes a ReLU porém foi escolhida por ter uma chance menor de ocorrer a situação de morte dos nós, caracterizada quando a saída de um nó é sempre 0 independente das entradas fornecidas. Na última camada foi usada a função Softmax, ela é muito utilizada em redes neurais com o objetivo de classificação, os valores recebidos por ela são normalizados em uma escala de 0 a 1, representando uma probabilidade de cada um de seus nós ser a classe da qual a entrada representa.

Durante o treinamento do modelo foi usado o otimizador Adam, atualmente o algoritmo de otimização mais recomendado pela sua eficiência, definindo as mudanças a serem feitas nos pesos dos nós durante a busca pela redução da perda do modelo durante o treinamento, em consideração com seu baixo custo computacional. A perda é calculada pela função categorical *crossentropy*, ela deve ser usada junto de uma camada Softmax na saída do modelo. Essa função é utilizada no modelos em que apenas uma saída é correta para cada tipo de entrada, como o é o caso do modelo desenvolvido para a classificação de afídeos.

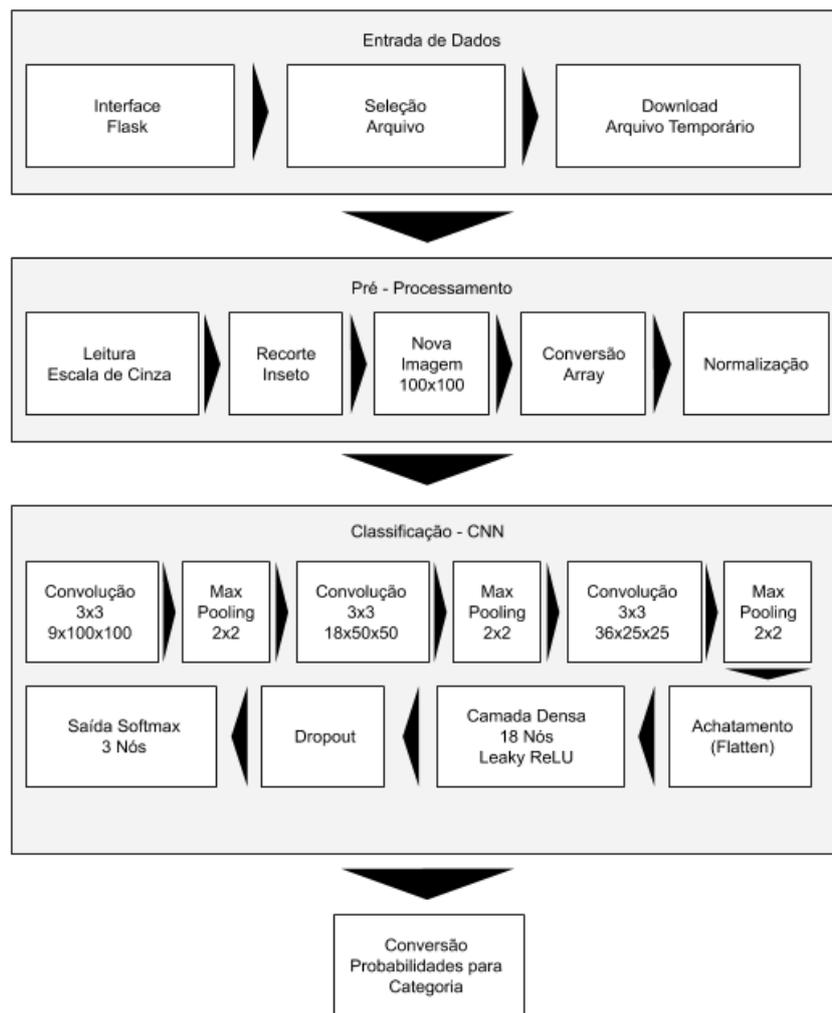
Após cada camada convolucional ocorre um processo de max pooling de tamanho 2x2 e por último uma camada de flatten. Essa camada transforma a matriz de saída das camadas convolucionais para um array para ser usada em camadas densas. A saída desses processos alimenta uma camada densa de 18 nós, com função de ativação *leaky\_ReLu* e pesos iniciados aleatoriamente. Em seguida os dados passam por uma camada Dropout e, por último, uma camada Softmax. A saída do modelo será um array com valores que permite classificar a classe ao qual o inseto deverá ser associado. A sequência de passos desse processo esta descrita na Figura 3.

O tratamento das imagens com o OpenCV tem como objetivo remover partículas ou insetos sobrepostos que dificultavam a retirada de características do objeto na imagem que deveria ser classificado, atrapalhando tanto o treinamento quanto o uso da rede neural na aplicação (Figura 4). Outro objetivo é a restrição das redes neurais que só podem receber como entrada um formato padrão, no caso foi definido que a entrada como uma imagem 100x100 pixels convertida em array. Através desse tratamento, não importa como o crop dos insetos foi feito, o inseto será encontrado na imagem e recortado para uma imagem aceita pelo sistema.

#### **4. Validação do sistema**

O primeiro modelo foi criado para a base de testes Iris Dataset através da API Estimators do TensorFlow. Esse modelo permitiu avaliar o uso de funções do *framework* e parâmetros para testes e análise de dados . Após esse processo partiu-se para o desenvolvimento do primeiro modelo usando dados extraídos das imagens (área, perímetro e momentos invariantes de Hu). Após testes o modelo obteve 86% de precisão em suas predições. Optou-se então pelo desenvolvimento de uma rede neural convolucional. Diversos testes foram realizados variando em parâmetros:Batch size, tamanho dos filtros de convolução, quantidade de convoluções, quantidade de camadas densas, chance de Dropout. Verificou-se que os dados obtidos dos testes que o modelo chegou a 99% de precisão.

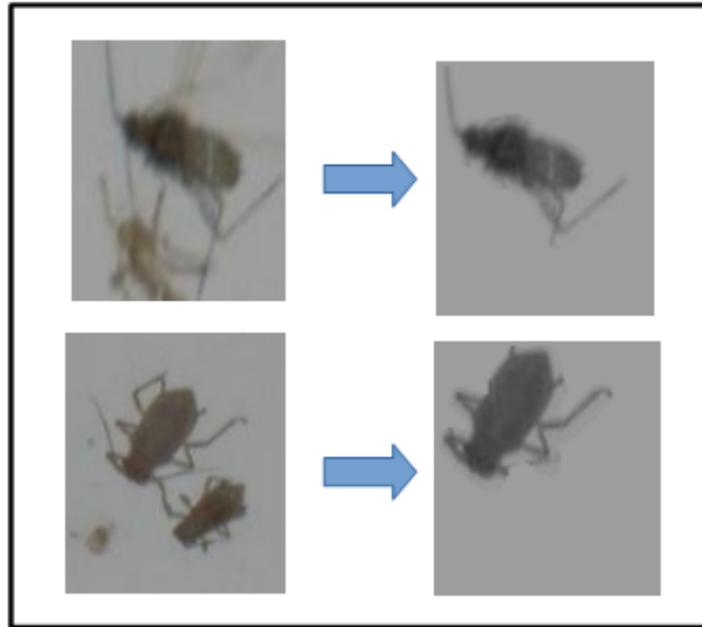
Após o desenvolvimento do modelo, passou-se para a fase de testes com a base de dados de insetos. O primeiro modelo criado durante os primeiros 6 meses de pesquisa era uma Rede Neural Convolucional(RNC) que possuía 3 camadas convolucionais: a primeira



**Figura 3. Arquitetura do sistema desenvolvido**

com 32 filtros e um kernel de tamanho  $7 \times 7$  e as demais possuíam 16 filtros com kernel  $5 \times 5$ , junto de cada convolução havia uma camada de Pooling de tamanho  $2 \times 2$ . Após as camadas os resultados eram analisados por uma camada densa de 8 nós com função de ativação ReLU, mesma das camadas convolucionais. Por fim os dados eram analisados por uma camada Softmax de 3 saídas.

Os primeiros testes da RNC foram realizados com o uso do dataset MNIST e os resultados obtidos foram bons com precisão chegando aos 92%, porém ao iniciar os testes com o dataset disponibilizado pela Embrapa Trigo, composto de imagens de pulgões em 3 estágios de vida (ninfas, alados e ápteros), a precisão caiu drasticamente chegando aos 50%. Observou-se que este fato deveria ocorrer porque o modelo poderia estar com problemas e também porque não se havia dado o tratamento necessário as imagens. Devido aos resultados iniciais com a base Minist serem positivos, inferiu-se que o problema poderia estar na base de imagens. No modelo original o tratamento dado para as imagens consistia e redimensionar as imagens para  $128 \times 128$  pixels e normalizar as mesmas em um array e dividindo seus elementos por 255 (número máximo que os elementos poderiam ter levando em conta a escala de cinza). Redimensionar era necessário para usar



**Figura 4. Imagem dos afideos após tratamento**

as imagens como entrada na RNC pois todas precisam ter as mesmas dimensões.

Observou-se entretanto que o recorte das imagens era irregular. Quando as imagens, que tinham tamanhos variados entre 40x40 e 220x220 pixels, eram redimensionadas seu conteúdo ficava disforme horizontal ou verticalmente, perdendo características essenciais dos insetos para o aprendizado de máquina. Além disso, muitas imagens eram compostas de concentrações de insetos impedindo a análise das características corretas. Para resolver o problema foi codificado um filtro utilizando a biblioteca para Python OpenCV, através dela foram removidos os contornos dos objetos na imagem. Utilizando os contornos era possível calcular o tamanho das áreas que eles representavam e através disso pegar o maior que seja válido, pois levando em consideração o tamanho de cada tipo de inseto, percebe-se que se uma área for muito grande ela representa uma concentração de insetos e deve ser descartada. Após definir um contorno que de fato representa um único inseto ele era recortado da imagem original e colado sobre um template 200x200 de cor semelhante ao fundo da imagem onde os insetos foram recortados. Isso resolve dois problemas: o descarte das “sujeiras” que antes estavam ao redor do inseto e as concentrações de insetos em uma imagem. Outro problema encontrado posteriormente foi a grande quantidade de imagens de alados na pasta onde deveria haver somente ápteros. Separar manualmente levaria muito tempo logo para tratar esse problema treinou-se uma rede neural para ser capaz de reconhecer alados e assim separar em duas pastas as imagens misturadas.

Depois desse processo de limpeza no dataset a RNC foi treinada novamente, levando em consideração que mesmo mais ‘limpo’ que as imagens de insetos já estavam melhores classificadas nas devidas pastas. Os resultados não melhoraram muito, os problemas na rede indicavam o caso de unidades ReLU estarem “morrendo”, termo que indica que não importa quais sejam as entradas no nó o resultado entregue pela função de ativação será sempre 0 (zero).

A RNC foi remodelada de várias formas alternando quantidade de convoluções, quantidade e tamanho dos filtros, quantidade de nós na camada densa, inicializadores dos pesos e vies(bias) foram alterados, porém nenhuma das alterações foi capaz de retirar a rede deste estado incapaz de generalizar as entradas. Observou-se que o problema estava ocorrendo porque havia semelhança de insetos considerados ninfas com as outras duas classes. Testou-se então utilizar o modelo já treinado, capaz de classificar alados e ápteros, junto de técnica de ajuste fino para classificar as ninfas. Após remover a última camada, Softmax de 2 saídas, e adicionar uma de 3 saídas, retreinou-se o modelo. Mas esse novamente não demonstrou resultados e continuou a ter como saída o valor 0.

A solução para o problema foi recortar as imagens de 200x200 para 100x100, como os insetos estavam sempre no canto superior esquerdo da imagem, a rede considerava toda a área vazia como um característica, atrapalhando o treinamento. Outro implemento ao dataset foi um aumento na quantidade de imagens, havia em torno de 1900 imagens de cada inseto, foi aplicado rotações aleatórias para transformar essas imagens em 10000 de cada tipo. A modelagem da rede neural e o processamento das imagens do dataset permitiram um desempenho de 93%.

## 5. Conclusão

A modelagem da rede neural e o processamento das imagens do dataset permitiram um desempenho de 93% podendo chegar a valores maiores. No início do projeto foi modelada uma rede neural que treinada utilizando momentos invariantes, perímetro e área de insetos, ela demonstrou bons resultados em torno de 86%. Mesmo com bons resultados preferiu-se utilizar redes convolucionais com o treinamento através de imagens pois apresentava um potencial maior.

A primeira abordagem poderia ter melhores resultados considerando o pequeno escopo de três classes do projeto, através de aprimoramentos na rede seria possível ter uma precisão maior que 90%. É possível distinguir cada um dos 3 tipos de inseto através de forma e tamanho representados no dataset pelos momentos invariantes e área. Além de que uma rede neural processa os dados muito mais rápido que uma rede neural convolucional processaria. Porém essa rede não possui muito potencial caso a quantidade de classes aumente. Em uma situação onde fosse necessário classificar 7 tipos de insetos mais “sujeiras” de diferentes formas e tamanhos dificilmente, não importa como fosse modelada, teria mais de 80% de precisão. As RNCs podem retirar das imagens muitas características, o suficiente para classificar milhares conforme o desenvolvimento, outro motivo é o tamanho das imagens dos insetos não serem pequenas possibilitando um bom desempenho da rede.

uso das redes neurais convolucionais mostrou-se adequada para os objetivos do projeto dada a característica do tipo de objeto a ser classificado (insetos). Observou-se que as redes neurais além de obterem uma alta precisão no processo de classificação abrem caminho para a classificação de outras espécies de insetos. Como trabalho futuro objetiva-se pesquisar outras modalidades de RNC como as Redes Convolucionais recorrentes que permite ampliar a precisão do classificador.

## Referências

- Belloni, Bruno, e. a. (2018). Ferramenta para identificação e classificação de afídeos utilizando momentos invariantes de hu e redes neurais artificiais. *I Congresso de Tecnologia da Informação do IFSUL - Campus Passo Fundo*, pages 19–28.
- Fabretti, P. K. M. J. O. (2013). Seres indesejáveis. *New Trends in Revista Cultivar*, n. 167,, pages 03–10.
- Gallo, D. (2002). *Entomologia Agrícola*. FEALQ.
- Grimberg, M. (2018). O'Reilly Media Inc.
- Haykin, Simon, E. P. M. (2003). *Redes Neurais: Princípios e Prática*. Bookman.
- Huaman, A. et al. (2018). *OpenCV Documentation*. Addison-Wesley.
- Lins, E. A. (2018). *Uma metodologia de contagem e classificação de afídeos utilizando visão computacional*. Dissertação de mestrado. UPF.
- Martín Abadi, e. a. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *CoRR*, abs/1603.04467.
- Stern, D. L. (2008). Aphids. *Current Biology*, 18(12):R504 – R505.
- WU, J. (2017). *Introduction to convolutional neural networks*. . National Key Lab for Novel Software Technology Nanjing University.

# Predição da Nota do ENEM 2018 usando Técnicas de Regressão

Pedro Marques Lessa<sup>1</sup>, Karin Satie Komati<sup>2</sup>

<sup>1</sup> Programa de Pós-graduação em Ciência de Dados com Big Data

<sup>2</sup> Programa de Pós-graduação em Computação Aplicada (PPComp)  
Campus Serra do Instituto Federal do Espírito Santo (Ifes)

pedromlessa@gmail.com, kkomati@ifes.edu.br

**Abstract.** *This study aims to generate a predictive model, where the dependent variable is the average of the objective tests and the essay of the ENEM 2018. The participant's socioeconomic data and the Censo Escolar 2018 data were used, especially the school infrastructure data. With the data processed properly, three regression techniques were chosen: Linear Regression, k-NN and Decision Tree. The models were trained and tested using Weka software, and the best result was Linear Regression, with an MAE of 51.70. The study showed that there is a correlation between the participant's socioeconomic data, the structure of their school and their performance.*

**Resumo.** *Esse estudo tem como objetivo gerar um modelo preditivo, onde a variável dependente é a média das provas objetivas mais a prova de redação do ENEM de 2018. Foram utilizados, principalmente os dados socioeconômicos do participante, e os dados do Censo Escolar 2018, principalmente os dados relativos à infraestrutura da escola. Com os dados devidamente tratados, foram escolhidas três técnicas de regressão: Regressão Linear, k-NN e Árvore de Decisão. Os modelos foram treinados e testados no software Weka, e o de melhor resultado foi o de Regressão Linear, com MAE de 51,70. O estudo apontou que existe uma correlação entre os dados socioeconômicos do participante, estrutura de sua escola e o seu desempenho.*

## 1. Introdução

O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) [INEP 2019] é uma autarquia federal vinculada ao Ministério da Educação (MEC) e tem por objetivo subsidiar a formulação de políticas educacionais dos diferentes níveis e, com isso, contribuir para o desenvolvimento socioeconômico do Brasil. O INEP é o responsável pela elaboração do Exame Nacional do Ensino Médio (ENEM) e pelo Censo Escolar.

O ENEM permite o acesso à Educação Superior e ao financiamento estudantil, além de promover o desenvolvimento pessoal e melhorias na educação, pois, além de avaliar o nível de conhecimento em diversas áreas do conhecimento, apresenta dados socioeconômicos do inscrito. Já o Censo Escolar é o maior levantamento de dados estatísticos da educação brasileira. É realizado todos os anos com a colaboração de todas

as instituições públicas e privadas do país de educação básica (ensino infantil, ensino fundamental e ensino médio). O INEP, através de seu Plano de Dados Abertos (PDA), disponibiliza em seu site informações estatísticas que ajudam a construir e a interpretar a educação brasileira. Dentre os diversos dados disponíveis, há os relativos ao ENEM e ao Censo Escolar<sup>1</sup>.

A proposta desse trabalho é, através dos dados do ENEM 2018 e do Censo Escolar 2018, prever as notas dos inscritos do ENEM com base, principalmente, nos seus dados socioeconômicos e os de sua escola. A partir desse estudo espera-se responder à seguinte pergunta: “Será que a estrutura escolar e os dados socioeconômicos do participante refletem no desempenho (nota) do participante do ENEM?”. Para isso, serão utilizadas técnicas de aprendizado de máquina na tarefa de regressão, a regressão linear, o k-NN e a árvore de decisão. Entende-se que, um dos principais fatores para a definição da nota no ENEM é o esforço pessoal do inscrito e, não há dados disponíveis com este tipo de informação.

Esse artigo está organizado da seguinte forma: a seção 1 faz uma breve introdução sobre o tema; a seção 2 apresenta o Referencial Teórico, as tarefas de Mineração de Dados utilizadas e as métricas de avaliação dos modelos de Mineração de Dados criados; a seção 3 descreve trabalhos relacionados; a seção 4 informa como foi realizada a etapa de pré-processamento de dados; a seção 5 traz uma discussão sobre os resultados obtidos; a seção 6 apresenta a conclusão do estudo e sugestões para trabalhos futuros.

## 2. Referencial Teórico

Da Silva, Peres e Boscaroli [da Silva et al. 2017] definem Mineração de Dados como um processo automático ou semiautomático de explorar analiticamente uma grande quantidade de dados com a finalidade de descobrir padrões relevantes que ocorrem nos dados, apresentando-os de forma a facilitar a sua assimilação e conhecimento. Tal processo de descoberta está associado a um processo de KDD (do inglês, “Knowledge Discovery in Databases”).

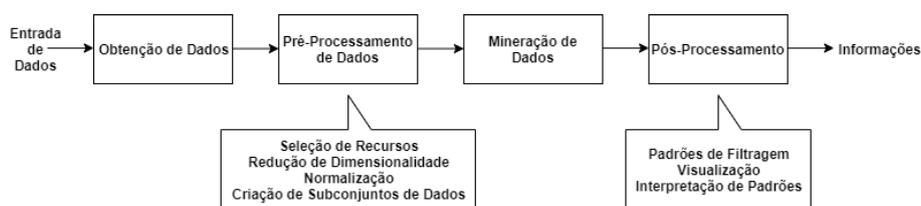


Figura 1. Etapas do KDD. Adaptado de [Tan et. al 2009].

Na Figura 1, as etapas que constituem o processo KDD são ilustradas. Esse processo consiste em uma série de etapas e de transformações, da obtenção dos dados até o pós-processamento dos resultados da Mineração de Dados. Cabe ressaltar que todo o processo do KDD pode ser iterativo e interativo, isto é, cada etapa pode ser executada mais de uma vez. A seguir descrevem-se cada um desses passos:

- Obtenção de Dados: nessa etapa é onde o processo KDD se inicia, nela há a organização das bases de dados que contém os dados da área de interesse.

<sup>1</sup><http://portal.inep.gov.br/web/guest/microdados>

- **Pré-Processamento dos Dados:** o propósito dessa etapa é transformar os dados de entrada brutos em um formato apropriado para posteriores análises. Envolve as ações de fusão de dados de múltiplas fontes, limpeza dos dados para remoção de dados duplicados e, seleção de registros e características que sejam relevantes à tarefa de Mineração de Dados.
- **Mineração de Dados:** envolve resolução de tarefas como predição, agrupamento e associação. Segundo Goldshmidt e Passos [Goldschmidt and Passos 2005], é o processo de busca de conhecimento através de algoritmos inteligentes.
- **Pós-Processamento:** os resultados obtidos da etapa anterior são validados, avaliados e formatados em gráficos, tabelas e relatórios estruturados.

## 2.1. Tarefas de Mineração de Dados

De forma mais abrangente e simplificada, as tarefas de mineração de dados são geralmente divididas em duas categorias principais: tarefas de previsão e tarefas descritivas [da Silva et al. 2017]. Nas tarefas de previsão, o objetivo é prever o valor de um determinado atributo baseado nos valores de outros atributos. O atributo a ser previsto é comumente conhecido com variável dependente ou alvo, enquanto que os atributos que foram usados para fazer a previsão são conhecidos como variáveis independentes. Já nas tarefas descritivas o objetivo é derivar padrões que resumam os relacionamentos dos dados.

Tarefas de predição são, em geral, técnicas supervisionadas, que exigem uma fase de treinamento para se descobrir o modelo de predição. Assim, no conjunto de treinamento já deve existir exemplos de respostas corretas. Classificação e regressão são exemplos de tarefas supervisionadas. A classificação consiste em definir um modelo capaz de mapear uma das instâncias (registros) aos seus rótulos (classes). Já a Regressão é a tarefa de aproximar uma função de mapeamento das variáveis de entrada para uma variável de saída contínua. Isto é, na classificação os resultados são discretos enquanto na regressão é um valor contínuo.

Nesse trabalho, através do software Weka [Hall et al. 2009], os seguintes algoritmos de regressão foram selecionados:

- Regressão Linear (RL) [Rodrigues et al. 2013], no WEKA sob o nome LinearRegression;
- k-nearest neighbors (k-NN), no WEKA sob o nome IBk;
- Árvore de Decisão (AD), no WEKA sob o nome REPTree.

A Regressão Linear é uma técnica para, tendo dois conjuntos de dados com uma correlação significativa, prever o valor de uma variável, tendo o valor de outra. Imagine dois conjuntos de dados desenhados em um gráfico de dispersão com eixos X e Y. A Regressão Linear constrói uma linha de melhor ajuste entre os dados das variáveis. Assim, X pode prever Y e vice-versa. O gráfico é utilizado para fins didáticos. As previsões são calculadas utilizando expressões matemáticas [Amaral 2016].

No k-NN, não é criado um modelo. Em tempo de execução, o algoritmo busca nos dados de treinamento qual instância mais se parece com a nova instância e a classifica como a mais semelhante, usando um algoritmo de vizinho mais próximo (como o IBk por exemplo). Embora esse tipo de algoritmo possa ser bastante eficiente, tem um custo computacional altíssimo, pois além de carregar e manter os dados em memória, a cada

nova instância a classificar terá que comparar com todos os dados históricos até encontrar o de melhor semelhança [Amaral 2016].

Já na Árvore de Decisão, o modelo cria uma estrutura de decisão em forma de árvore, a cada nova instância ela é percorrida até chegar em um nodo terminal onde está a classe (nos casos de classificação) ou a variável dependente (nos casos de regressão). A árvore sempre começa por um único nodo raiz e vai sendo dividida até levar na classe ou no valor previsto. Um algoritmo determina como cada nodo será dividido [Amaral 2016].

## 2.2. Métricas de Avaliação

A avaliação de modelos preditivos para a resolução de tarefas de regressão, se dá por meio de funções de perda contínua, que são capazes de medir o erro entre a resposta obtida, pelo modelo preditivo, e a resposta esperada [da Silva et al. 2017]. Algumas funções de perda contínua comumente usadas são:

- Erro absoluto médio (MAE):

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\hat{\theta}_i - \theta_i|$$

- Erro quadrado médio (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta_i)^2}$$

- Erro absoluto (RAE):

$$\text{RAE} = \frac{\sum_{i=1}^N |\hat{\theta}_i - \theta_i|}{\sum_{i=1}^N |\bar{\theta} - \theta_i|}$$

- Erro quadrado (RRSE):

$$\text{RRSE} = \sqrt{\frac{\sum_{i=1}^N (\hat{\theta}_i - \theta_i)^2}{\sum_{i=1}^N (\bar{\theta} - \theta_i)^2}}$$

onde  $\theta$  representa o valor real já  $\hat{\theta}$  representa o valor previsto.

Todas as funções comparam os valores verdadeiros com os valores previstos, mas o fazem de forma diferente. De qualquer maneira, todas dizem o quão distante um valor estimado está do valor verdadeiro.

Além das funções de perda contínua, a correlação entre os valores reais e os previstos é calculada. A correlação informa o quanto os valores reais e previstos estão relacionados. O valor varia na faixa de -1 a 1, sendo que a correlação 1 significa que x e y possuem um relacionamento linear positivo perfeito, enquanto -1 significa um relacionamento linear negativo perfeito. O valor 0 (zero) significa que não há correlação alguma entre as variáveis.

Em um ambiente de aprendizado de máquina, os dados são separados em treino e teste. Dados de treino são submetidos ao classificador para produzirem o modelo. Posteriormente, os dados de teste são submetidos para que seja feita a previsão [Amaral 2016]. A divisão foi feita da seguinte forma: 70% dos dados para treinamento e 30% para testes.

## 3. Trabalhos Relacionados

A utilização dos microdados do ENEM para a criação de modelos preditivos não é exclusividade desse estudo. Nessa seção apresentam-se trabalhos relacionados à Mineração de

Dados aplicados aos dados do ENEM, em ordem cronológica.

No estudo de Adeodato, Santos Filho e Rodrigues [Adeodato et al. 2014], foram utilizados os dados do ENEM e do Censo Escolar (ambos de 2011). A análise foi feita apenas nos dados das escolas privadas, através de técnicas de Regressão Logística e Árvore de Decisão. O escore de propensão teve seu desempenho avaliado pelos testes Kolmogorov-Smirnov (KS) e a Curva ROC, onde se alcançou AUC\_ROC no valor de 0,897 e Máx\_KS em 0,632. Os resultados mostraram que o fator econômico é relevante e aparece indiretamente na região da escola. A árvore de decisão corroborou com o fato de que a boa qualidade das escolas está nos aspectos econômico-financeiros, seja direta (renda familiar) ou indiretamente (número de banheiros da casa ou opção do ProUni) ou aspectos culturais (nível de educação da mãe ou do pai) da família.

Já no artigo de Stearns e colegas [Stearns et al. 2017], para prever a performance dos estudantes no ENEM 2014, foram utilizadas apenas as suas informações socioeconômicas. Foi escolhida Árvore de Decisão, Gradient Boosting e AdaBoost. Chegou à conclusão que o algoritmo Gradiente Boosting apresentou um melhor resultado com uma taxa de Erro Absoluto Médio (MAE) de 65,90 pontos.

O trabalho de Simon e Cazella [Simon and Cazella 2017] utilizou o conjunto de dados do ENEM 2015, agrupados por escola, e a partir dele, criou um modelo preditivo onde a variável dependente é a nota em Ciências da Natureza e suas Tecnologias. Através do software WEKA, escolheu Árvore de Decisão (com o algoritmo j48) para a tarefa de Mineração de Dados. Apresentou um resultado relevante, com uma acurácia de 77,02%.

Bravin, Lee e Rissino [Bravin et al. 2019] utilizaram os dados do ENEM 2015, agrupados por escola, criaram um modelo preditivo onde a variável dependente é a nota média (das escolas) na língua portuguesa. Encontrou forte relação entre os dados socioeconômicos da escola e a nota de língua portuguesa. Esse estudo teve influência do artigo de [Simon and Cazella 2017].

#### **4. Processamento dos Dados**

Nessa etapa, o primeiro passo foi fazer o *download* dos microdados do ENEM 2018 e do Censo Escolar 2018 no site do INEP. Os arquivos estão no formato csv e possuem um tamanho de 3.4 GB e 99,1 MB, respectivamente.

A relação (tabela) do ENEM possui 5.513.748 instâncias (linhas), isto é, representa mais de 5 milhões de inscritos e possui 137 atributos (colunas), divididos da seguinte forma (conforme o dicionário de dados):

- 20 atributos com os dados do participante (inscrito);
- 8 atributos com os dados da escola;
- 13 atributos com dados referentes a pedidos de atendimento especializado (deficiência visual, auditiva, mental, dentre outros);
- 4 atributos referentes a pedidos de atendimento específico (gestante, lactante, idoso e atendimento hospitalar);
- 33 atributos referentes a pedidos de recursos especializados e específicos para a realização da prova (tais como, prova em braile, libras e cadeira para canhoto);
- 4 atributos referentes ao local da aplicação da prova;
- 21 atributos referentes à prova objetiva;

- 7 atributos referentes à prova de redação;
- 27 atributos referentes ao questionário socioeconômico do participante.

A relação do Censo Escolar possui 286.015 instâncias (escolas) e possui 167 atributos, divididos da seguinte forma:

- 1 atributo que representa o ano do Censo Escolar;
- 137 atributos com os dados da escola (tais como localização e infraestrutura);
- 29 atributos com dados de oferta de matrícula.

#### 4.1. Pré-Processamento

O pré-processamento de dados iniciou com a seleção de atributos. Através dos dicionários de dados (do ENEM e do Censo Escolar), verificou-se que diversos atributos seriam irrelevantes para a construção do nosso modelo de Mineração de Dados. Ao final, foram mantidos apenas 41 atributos da relação do ENEM e 134 atributos da relação do Censo Escolar. Essa primeira seleção foi feita totalmente de forma manual.

Em seguida, foi realizada a seleção de instâncias (linhas) com a ajuda da biblioteca Pandas<sup>2</sup>. Foram realizadas as seguintes tarefas:

- foram selecionados apenas os participantes do ensino médio regular e que o concluiriam no ano de 2018;
- foram removidas as instâncias que possuíam pelo menos um dado faltante;
- para evitar divisões por zero, foram selecionados apenas os participantes que possuíam notas diferentes de zero.

Após esses procedimentos, o número de instâncias na relação do ENEM 2018 foi reduzido de 5.513.748 para 1.015.172 instâncias.

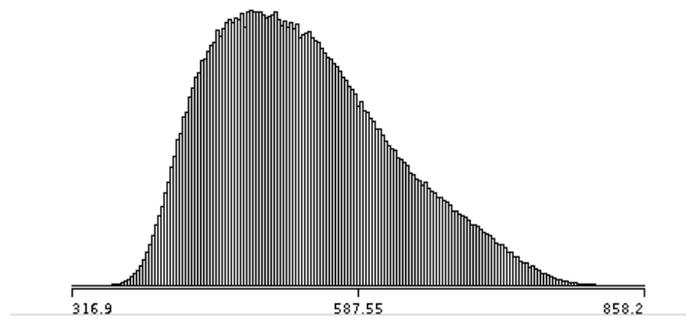
Para a relação do Censo Escolar foram removidas as instâncias que possuíam pelo menos um dado faltante. Depois, foi feita a junção da relação do ENEM e do Censo Escolar, através dos atributos **CO\_ESCOLA** (código da escola - relação ENEM) e **CO\_ENTIDADE** (código da entidade - relação Censo Escolar). Com isso, passou-se a ter apenas uma relação.

Foi criado um atributo que é a média das 4 provas objetivas mais a prova de redação, denominou-se esse atributo de **MD\_NOTA\_TOTAL** (média nota total). Este atributo é que contém os valores que se quer prever. Com este novo atributo, foram retiradas as 4 notas das provas objetivas e a nota da prova de redação, e gerou-se um arquivo csv. Para entender melhor esse atributo criado, e conforme observado na Figura 2, verifica-se que a nossa variável dependente (**MD\_NOTA\_TOTAL**) tem um valor mínimo de 316,9 pontos, um valor máximo de 858,2 pontos e um valor médio (da soma do valor máximo e mínimo) de 587,55 pontos.

Importou-se esse arquivo csv para o software Weka. O Weka possui algoritmos para seleção de atributos via estratégia de seleção de subconjuntos de características (sob a classe de nome **CfsSubsetEval**) e o melhor espaço de subconjuntos de atributos (sob a classe de nome **BestFirst**). Ou seja, através desses dois algoritmos o Weka selecionou os atributos mais relevantes para a construção do nosso modelo.

---

<sup>2</sup><https://pandas.pydata.org/>



**Figura 2. Histograma das notas dos participantes do ENEM 2018.**

A classe **CfsSubsetEval** avalia um conjunto de atributos considerando a capacidade preditiva individual de cada um, juntamente com o grau de redundância entre eles. Os conjuntos de atributos que são altamente correlacionados com a classe (ou a variável dependente) enquanto apresentam baixa correlação entre eles são os escolhidos [Hall 1998]. **BestFirst** é uma classe de busca genérica onde o nó de menor custo aparente (no nosso caso atributo) é avaliado primeiro. A ideia básica é prosseguir com a busca sempre a partir do nó mais promissor, considerando que ele ainda não foi expandido [Kohavi and John 1997].

Após a execução do algoritmo, a relação ficou com 46 atributos. A seguir, descreve-se, de forma geral, os atributos utilizados:

- 3 atributos individuais representando idade, sexo e cor/raça;
- 18 atributos com os dados socioeconômicos do participante;
  1. Até que série seu pai, ou o homem responsável por você, estudou?
  2. Até que série sua mãe, ou a mulher responsável por você, estudou?
  3. Ocupação do seu pai ou do homem responsável por você.
  4. Ocupação da sua mãe ou da mulher responsável por você.
  5. Incluindo você, quantas pessoas moram atualmente em sua residência?
  6. Qual é a renda mensal de sua família?
  7. Em sua residência trabalha empregado(a) doméstico(a)?
  8. Na sua residência tem banheiro?
  9. Na sua residência tem carro?
  10. Na sua residência tem freezer (independente ou segunda porta da geladeira)?
  11. Na sua residência tem máquina de lavar louça?
  12. Na sua residência tem aspirador de pó?
  13. Na sua residência tem televisão em cores?
  14. Na sua residência tem TV por assinatura?
  15. Na sua residência tem telefone fixo?
  16. Na sua residência tem computador?
  17. Na sua residência tem acesso à Internet?
  18. Em que tipo de escola você frequentou o Ensino Médio?
- 25 atributos com os dados da escola.
  1. Dependência Adm. (1. Federal; 2. Estadual; 3. Municipal; 4. Privada)
  2. Abastecimento de energia elétrica - Gerador?

3. Abastecimento de energia elétrica - Outros (Energia alternativa)?
4. Esgoto sanitário - Rede pública?
5. Destinação do lixo - Recicla?
6. Dependências existentes na escola - Laboratório de ciências?
7. Dependências existentes na escola - Cozinha?
8. Dependências existentes na escola - Parque infantil?
9. Dependências existentes na escola - Banheiro fora do prédio?
10. Banheiro adequado ao uso dos alunos com deficiência ou mobilidade reduzida?
11. Dependências existentes na escola - Banheiro com chuveiro?
12. Dependências existentes na escola - Auditório?
13. Dependências existentes na escola - Alojamento de professor?
14. Equipamentos existentes na escola - Fax?
15. Acesso à Internet?
16. Internet Banda Larga?
17. Alimentação escolar para os alunos?
18. Atendimento Educacional Especializado (AEE)?
19. Atividade Complementar?
20. Escola abre aos finais de semana para a comunidade?
21. Mediação didático-pedagógica oferecida pela escola - Educação a Distância - EAD?
22. Modo, maneira ou metodologia de ensino correspondente às turmas destinadas a pessoas que não cursaram o ensino fundamental e/ou médio em idade própria?
23. Escola oferece Creche em classes comuns do ensino regular?
24. Oferece Ensino Médio (Médio Propedêutico) em classes comuns do ensino regular?
25. Educação Profissional - Curso Técnico Concomitante e/ou Subsequente e/ou Curso FIC Concomitante em classes comuns?

## 5. Resultados e Discussão

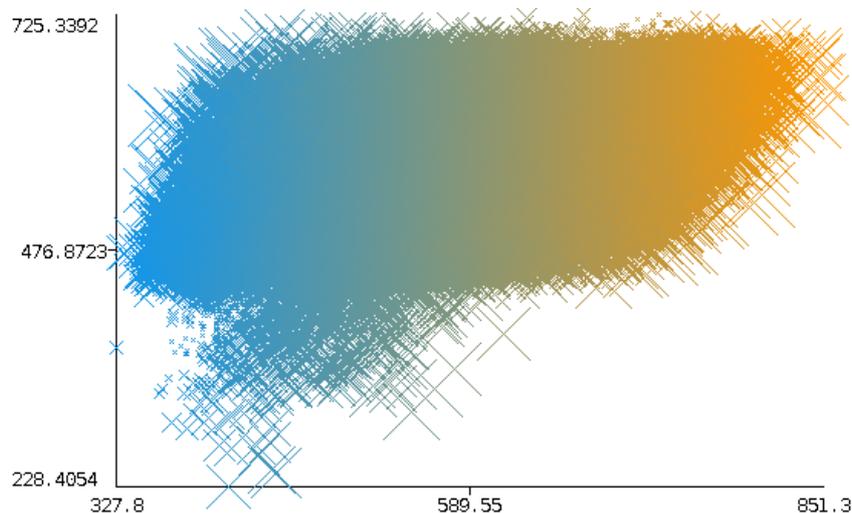
Os resultados são apresentados na Tabela 1. O número de instâncias na base de dados de testes foi de 304.552. A técnica da Árvore de Decisão (AD) apresentou um desempenho bem próximo da Regressão Linear (RL), com um coeficiente de correlação (entre o valor real e o previsto da variável dependente) de 0,63 e MAE de 51,70. Já o k-NN apresentou o pior resultado, coeficiente de correlação de 0,45 e MAE de 68,71.

**Tabela 1. Resultado da avaliação do modelo de Regressão Linear**

Algoritmo	RL	k-NN	AD
Coeficiente de correlação	0,6459	0,4563	0,6339
MAE	51,16	68,7119	51,70
RMSE	63,554	86,7004	64,4354
RAE	72,2655%	101,0712%	76,0503%
RRSE	763405%	104,1438%	77,3993%

Um MAE de 51,16 pontos representa 9,45% da diferença entre o valor máximo e mínimo dos valores preditos. O MAE nada mais é o quanto o modelo, em média e em

valores absolutos, pode errar a previsão. Suponha que se tenha apenas um participante do ENEM e que sua nota foi de 700 pontos e a previsão foi de 750 pontos. Nessa situação, o MAE seria de 50 pontos.



**Figura 3. Gráfico de Classificação de Erros. Média das notas reais (eixo X) vs. média das notas previstas (eixo Y).**

Uma vez que a tarefa de Regressão Linear apresentou o melhor resultado, com um coeficiente de correlação de 0,64 e MAE de 51,16, gerou-se um gráfico de Classificação de Erros (Figura 3), isto é, um gráfico que no eixo X estão os valores reais (**MD\_NOTA\_TOTAL**), e, no eixo Y, os valores previstos. A partir do gráfico da Figura 3, observa-se que há muitas notas reais baixas com uma previsão alta (região superior com a cor azul). Em contrapartida, há muitas notas reais altas e que foram previstas corretamente (região superior com a cor laranja). Observa-se também que há poucas notas reais altas e que a previsão seja de nota baixa (espaço a direita vazio no gráfico). Diante dessa última observação, surgiu mais uma indagação: "Quantos participantes superaram a previsão de suas notas (considerando apenas previsões acima do valor médio de 587,55)?"

Com a ajuda da biblioteca Pandas, verificou-se que, em um universo de 304.551 participantes (base de testes), 69.194 deles possuem uma nota real além da prevista e superior a 587,55 pontos, ou seja, 22,7% dos participantes (contidos na base de testes), mesmo em uma possível situação adversa considerada pelo modelo, superaram as expectativas.

## 6. Conclusão

Diante do estudo apresentado, conclui-se que os fatores socioeconômicos dos participantes do ENEM 2018 e a respectiva estrutura de sua escola tem sim influência no resultado do exame, tal como mostrado em trabalhos anteriores que usaram dados de 2011, 2014 e 2015. Não de maneira preponderante, mas, com os dados utilizados, consegue-se prever o resultado do participante de maneira razoável. E mesmo com um cenário possivelmente considerado adverso pelo modelo (prevendo uma nota baixa para o participante), muitos inscritos superaram a nota prevista. Dentre as técnicas de regressão utilizadas, a

Regressão Linear apresentou um resultado com menor erro do que as técnicas k-NN e Árvore de Decisão.

Para trabalhos futuros, pode-se indicar o uso de outros algoritmos e técnicas como, por exemplo, Redes Neurais Artificiais para a previsão das notas dos participantes e o uso de técnicas de detecção de *outliers* para verificação dos casos de alunos que tiraram notas muito acima do esperado (levando em consideração a sua característica socioeconômica e a estrutura de sua escola), tentando identificar as mentes brilhantes do Brasil. Também pode-se indicar um estudo mais aprofundado sobre quantos participantes superaram de forma significativa as previsões dos modelos e de que maneira obteve tal resultado.

## Referências

- Adeodato, P. J., Santos Filho, M. M., and Rodrigues, R. L. (2014). Predição de desempenho de escolas privadas usando o enem como indicador de qualidade escolar. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 25, page 891.
- Amaral, F. (2016). *Aprenda mineração de dados: teoria e prática*, volume 1. Alta Books Editora.
- Bravin, G. F., Lee, L., and das Dores Rissino, S. (2019). Mineração de dados educacionais na base de dados do enem 2015. *Brazilian Journal of Production Engineering-BJPE*, 5(4):186–201.
- da Silva, L. A., Peres, S. M., and Boscarioli, C. (2017). *Introdução à mineração de dados: com aplicações em R*. Elsevier Brasil.
- Goldschmidt, R. and Passos, E. (2005). *Data mining: um guia prático*. Gulf Professional Publishing.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1):10–18.
- Hall, M. A. (1998). *Correlation-based Feature Subset Selection for Machine Learning*. PhD thesis, University of Waikato, Hamilton, New Zealand.
- INEP (2019). Portal inep. <http://www.inep.gov.br/>. Acesso em: 11 Out. 2019.
- Kohavi, R. and John, G. H. (1997). Wrappers for feature subset selection. *Artificial intelligence*, 97(1):273–324.
- Rodrigues, R. L., de Medeiros, F. P., and Gomes, A. S. (2013). Modelo de regressão linear aplicado à previsão de desempenho de estudantes em ambiente de aprendizagem. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 24, page 607.
- Simon, A. and Cazella, S. (2017). Mineração de dados educacionais nos resultados do enem de 2015. In *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, volume 6, page 754.
- Stearns, B., Rangel, F., Firmino, F., Rangel, F., and Oliveira, J. (2017). Prevendo desempenho dos candidatos do enem através de dados socioeconômicos. In *36º Concurso de Trabalhos de Iniciação Científica (CTIC 2017)*, volume 36. SBC.