

# TECNOLOGIAS, MÉTODOS E TEORIAS NA ENGENHARIA DE COMPUTAÇÃO 2



ERNANE ROSA MARTINS  
(ORGANIZADOR)

 **Atena**  
Editora

Ano 2020

# TECNOLOGIAS, MÉTODOS E TEORIAS NA ENGENHARIA DE COMPUTAÇÃO 2



ERNANE ROSA MARTINS  
(ORGANIZADOR)

 **Atena**  
Editora

Ano 2020

**Editora Chefe**

Profª Drª Antonella Carvalho de Oliveira

**Assistentes Editoriais**

Natalia Oliveira

Bruno Oliveira

Flávia Roberta Barão

**Bibliotecária**

Janaina Ramos

**Projeto Gráfico e Diagramação**

Natália Sandrini de Azevedo

Camila Alves de Cremo

Luiza Alves Batista

Maria Alice Pinheiro

**Imagens da Capa**

Shutterstock

**Edição de Arte**

Luiza Alves Batista

**Revisão**

Os Autores

2021 by Atena Editora

Copyright © Atena Editora

Copyright do Texto © 2021 Os autores

Copyright da Edição © 2021 Atena Editora

Direitos para esta edição cedidos à Atena Editora pelos autores.



Todo o conteúdo deste livro está licenciado sob uma Licença de Atribuição *Creative Commons*. Atribuição-Não-Comercial-NãoDerivativos 4.0 Internacional (CC BY-NC-ND 4.0).

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores, inclusive não representam necessariamente a posição oficial da Atena Editora. Permitido o *download* da obra e o compartilhamento desde que sejam atribuídos créditos aos autores, mas sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

Todos os manuscritos foram previamente submetidos à avaliação cega pelos pares, membros do Conselho Editorial desta Editora, tendo sido aprovados para a publicação com base em critérios de neutralidade e imparcialidade acadêmica.

A Atena Editora é comprometida em garantir a integridade editorial em todas as etapas do processo de publicação, evitando plágio, dados ou resultados fraudulentos e impedindo que interesses financeiros comprometam os padrões éticos da publicação. Situações suspeitas de má conduta científica serão investigadas sob o mais alto padrão de rigor acadêmico e ético.

**Conselho Editorial**

**Ciências Humanas e Sociais Aplicadas**

Prof. Dr. Alexandre Jose Schumacher – Instituto Federal de Educação, Ciência e Tecnologia do Paraná

Prof. Dr. Américo Junior Nunes da Silva – Universidade do Estado da Bahia

Prof. Dr. Antonio Carlos Frasson – Universidade Tecnológica Federal do Paraná

Prof. Dr. Antonio Gasparetto Júnior – Instituto Federal do Sudeste de Minas Gerais

Prof. Dr. Antonio Isidro-Filho – Universidade de Brasília  
Prof. Dr. Carlos Antonio de Souza Moraes – Universidade Federal Fluminense  
Prof. Dr. Crisóstomo Lima do Nascimento – Universidade Federal Fluminense  
Profª Drª Cristina Gaio – Universidade de Lisboa  
Prof. Dr. Daniel Richard Sant'Ana – Universidade de Brasília  
Prof. Dr. Deyvison de Lima Oliveira – Universidade Federal de Rondônia  
Profª Drª Dilma Antunes Silva – Universidade Federal de São Paulo  
Prof. Dr. Edvaldo Antunes de Farias – Universidade Estácio de Sá  
Prof. Dr. Elson Ferreira Costa – Universidade do Estado do Pará  
Prof. Dr. Eloi Martins Senhora – Universidade Federal de Roraima  
Prof. Dr. Gustavo Henrique Cepolini Ferreira – Universidade Estadual de Montes Claros  
Profª Drª Ivone Goulart Lopes – Istituto Internazionale delle Figlie de Maria Ausiliatrice  
Prof. Dr. Jadson Correia de Oliveira – Universidade Católica do Salvador  
Prof. Dr. Julio Candido de Meirelles Junior – Universidade Federal Fluminense  
Profª Drª Lina Maria Gonçalves – Universidade Federal do Tocantins  
Prof. Dr. Luis Ricardo Fernandes da Costa – Universidade Estadual de Montes Claros  
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte  
Prof. Dr. Marcelo Pereira da Silva – Pontifícia Universidade Católica de Campinas  
Profª Drª Maria Luzia da Silva Santana – Universidade Federal de Mato Grosso do Sul  
Prof. Dr. Pablo Ricardo de Lima Falcão – Universidade de Pernambuco  
Profª Drª Paola Andressa Scortegagna – Universidade Estadual de Ponta Grossa  
Profª Drª Rita de Cássia da Silva Oliveira – Universidade Estadual de Ponta Grossa  
Prof. Dr. Rui Maia Diamantino – Universidade Salvador  
Prof. Dr. Saulo Cerqueira de Aguiar Soares – Universidade Federal do Piauí  
Prof. Dr. Urandi João Rodrigues Junior – Universidade Federal do Oeste do Pará  
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande  
Profª Drª Vanessa Ribeiro Simon Cavalcanti – Universidade Católica do Salvador  
Prof. Dr. William Cleber Domingues Silva – Universidade Federal Rural do Rio de Janeiro  
Prof. Dr. Willian Douglas Guilherme – Universidade Federal do Tocantins

#### **Ciências Agrárias e Multidisciplinar**

Prof. Dr. Alexandre Igor Azevedo Pereira – Instituto Federal Goiano  
Prof. Dr. Arinaldo Pereira da Silva – Universidade Federal do Sul e Sudeste do Pará  
Prof. Dr. Antonio Pasqualetto – Pontifícia Universidade Católica de Goiás  
Profª Drª Carla Cristina Bauermann Brasil – Universidade Federal de Santa Maria  
Prof. Dr. Cleberton Correia Santos – Universidade Federal da Grande Dourados  
Profª Drª Diocléa Almeida Seabra Silva – Universidade Federal Rural da Amazônia  
Prof. Dr. Écio Souza Diniz – Universidade Federal de Viçosa  
Prof. Dr. Fábio Steiner – Universidade Estadual de Mato Grosso do Sul  
Prof. Dr. Fágner Cavalcante Patrocínio dos Santos – Universidade Federal do Ceará  
Profª Drª Girlene Santos de Souza – Universidade Federal do Recôncavo da Bahia  
Prof. Dr. Jael Soares Batista – Universidade Federal Rural do Semi-Árido  
Prof. Dr. Jayme Augusto Peres – Universidade Estadual do Centro-Oeste  
Prof. Dr. Júlio César Ribeiro – Universidade Federal Rural do Rio de Janeiro  
Profª Drª Lina Raquel Santos Araújo – Universidade Estadual do Ceará  
Prof. Dr. Pedro Manuel Villa – Universidade Federal de Viçosa  
Profª Drª Raissa Rachel Salustriano da Silva Matos – Universidade Federal do Maranhão  
Prof. Dr. Ronilson Freitas de Souza – Universidade do Estado do Pará  
Profª Drª Talita de Santos Matos – Universidade Federal Rural do Rio de Janeiro  
Prof. Dr. Tiago da Silva Teófilo – Universidade Federal Rural do Semi-Árido  
Prof. Dr. Valdemar Antonio Paffaro Junior – Universidade Federal de Alfenas

### **Ciências Biológicas e da Saúde**

Prof. Dr. André Ribeiro da Silva – Universidade de Brasília  
Profª Drª Anelise Levay Murari – Universidade Federal de Pelotas  
Prof. Dr. Benedito Rodrigues da Silva Neto – Universidade Federal de Goiás  
Profª Drª Daniela Reis Joaquim de Freitas – Universidade Federal do Piauí  
Profª Drª Débora Luana Ribeiro Pessoa – Universidade Federal do Maranhão  
Prof. Dr. Douglas Siqueira de Almeida Chaves – Universidade Federal Rural do Rio de Janeiro  
Prof. Dr. Edson da Silva – Universidade Federal dos Vales do Jequitinhonha e Mucuri  
Profª Drª Elizabeth Cordeiro Fernandes – Faculdade Integrada Medicina  
Profª Drª Eleuza Rodrigues Machado – Faculdade Anhanguera de Brasília  
Profª Drª Elane Schwinden Prudêncio – Universidade Federal de Santa Catarina  
Profª Drª Eysler Gonçalves Maia Brasil – Universidade da Integração Internacional da Lusofonia Afro-Brasileira  
Prof. Dr. Ferlando Lima Santos – Universidade Federal do Recôncavo da Bahia  
Profª Drª Fernanda Miguel de Andrade – Universidade Federal de Pernambuco  
Prof. Dr. Fernando Mendes – Instituto Politécnico de Coimbra – Escola Superior de Saúde de Coimbra  
Profª Drª Gabriela Vieira do Amaral – Universidade de Vassouras  
Prof. Dr. Gianfábio Pimentel Franco – Universidade Federal de Santa Maria  
Prof. Dr. Helio Franklin Rodrigues de Almeida – Universidade Federal de Rondônia  
Profª Drª Iara Lúcia Tescarollo – Universidade São Francisco  
Prof. Dr. Igor Luiz Vieira de Lima Santos – Universidade Federal de Campina Grande  
Prof. Dr. Jefferson Thiago Souza – Universidade Estadual do Ceará  
Prof. Dr. Jesus Rodrigues Lemos – Universidade Federal do Piauí  
Prof. Dr. Jônatas de França Barros – Universidade Federal do Rio Grande do Norte  
Prof. Dr. José Max Barbosa de Oliveira Junior – Universidade Federal do Oeste do Pará  
Prof. Dr. Luís Paulo Souza e Souza – Universidade Federal do Amazonas  
Profª Drª Magnólia de Araújo Campos – Universidade Federal de Campina Grande  
Prof. Dr. Marcus Fernando da Silva Praxedes – Universidade Federal do Recôncavo da Bahia  
Profª Drª Maria Tatiane Gonçalves Sá – Universidade do Estado do Pará  
Profª Drª Mylena Andréa Oliveira Torres – Universidade Ceuma  
Profª Drª Natiéli Piovesan – Instituto Federacl do Rio Grande do Norte  
Prof. Dr. Paulo Inada – Universidade Estadual de Maringá  
Prof. Dr. Rafael Henrique Silva – Hospital Universitário da Universidade Federal da Grande Dourados  
Profª Drª Regiane Luz Carvalho – Centro Universitário das Faculdades Associadas de Ensino  
Profª Drª Renata Mendes de Freitas – Universidade Federal de Juiz de Fora  
Profª Drª Vanessa Lima Gonçalves – Universidade Estadual de Ponta Grossa  
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande  
Profª Drª Welma Emidio da Silva – Universidade Federal Rural de Pernambuco

### **Ciências Exatas e da Terra e Engenharias**

Prof. Dr. Adélio Alcino Sampaio Castro Machado – Universidade do Porto  
Profª Drª Ana Grasielle Dionísio Corrêa – Universidade Presbiteriana Mackenzie  
Prof. Dr. Carlos Eduardo Sanches de Andrade – Universidade Federal de Goiás  
Profª Drª Carmen Lúcia Voigt – Universidade Norte do Paraná  
Prof. Dr. Cleiseano Emanuel da Silva Paniagua – Instituto Federal de Educação, Ciência e Tecnologia de Goiás  
Prof. Dr. Douglas Gonçalves da Silva – Universidade Estadual do Sudoeste da Bahia  
Prof. Dr. Eloi Rufato Junior – Universidade Tecnológica Federal do Paraná  
Profª Drª Érica de Melo Azevedo – Instituto Federal do Rio de Janeiro  
Prof. Dr. Fabrício Menezes Ramos – Instituto Federal do Pará  
Profª Dra. Jéssica Verger Nardeli – Universidade Estadual Paulista Júlio de Mesquita Filho  
Prof. Dr. Juliano Carlo Rufino de Freitas – Universidade Federal de Campina Grande

Profª Drª Luciana do Nascimento Mendes – Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte  
Prof. Dr. Marcelo Marques – Universidade Estadual de Maringá  
Prof. Dr. Marco Aurélio Kistemann Junior – Universidade Federal de Juiz de Fora  
Profª Drª Neiva Maria de Almeida – Universidade Federal da Paraíba  
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte  
Profª Drª Priscila Tessmer Scaglioni – Universidade Federal de Pelotas  
Prof. Dr. Sidney Gonçalves de Lima – Universidade Federal do Piauí  
Prof. Dr. Takeshy Tachizawa – Faculdade de Campo Limpo Paulista

#### **Linguística, Letras e Artes**

Profª Drª Adriana Demite Stephani – Universidade Federal do Tocantins  
Profª Drª Angeli Rose do Nascimento – Universidade Federal do Estado do Rio de Janeiro  
Profª Drª Carolina Fernandes da Silva Mandaji – Universidade Tecnológica Federal do Paraná  
Profª Drª Denise Rocha – Universidade Federal do Ceará  
Profª Drª Edna Alencar da Silva Rivera – Instituto Federal de São Paulo  
Profª Drª Fernanda Tonelli – Instituto Federal de São Paulo,  
Prof. Dr. Fabiano Tadeu Grazioli – Universidade Regional Integrada do Alto Uruguai e das Missões  
Prof. Dr. Gilmei Fleck – Universidade Estadual do Oeste do Paraná  
Profª Drª Keyla Christina Almeida Portela – Instituto Federal de Educação, Ciência e Tecnologia do Paraná  
Profª Drª Miraniide Oliveira Neves – Instituto de Educação, Ciência e Tecnologia do Pará  
Profª Drª Sandra Regina Gardacho Pietrobon – Universidade Estadual do Centro-Oeste  
Profª Drª Sheila Marta Carregosa Rocha – Universidade do Estado da Bahia

#### **Conselho Técnico Científico**

Prof. Me. Abrãao Carvalho Nogueira – Universidade Federal do Espírito Santo  
Prof. Me. Adalberto Zorzo – Centro Estadual de Educação Tecnológica Paula Souza  
Prof. Dr. Adaylson Wagner Sousa de Vasconcelos – Ordem dos Advogados do Brasil/Seccional Paraíba  
Prof. Dr. Adilson Tadeu Basquerote Silva – Universidade para o Desenvolvimento do Alto Vale do Itajaí  
Prof. Dr. Alex Luis dos Santos – Universidade Federal de Minas Gerais  
Prof. Me. Alessandro Teixeira Ribeiro – Centro Universitário Internacional  
Profª Ma. Aline Ferreira Antunes – Universidade Federal de Goiás  
Profª Drª Amanda Vasconcelos Guimarães – Universidade Federal de Lavras  
Prof. Me. André Flávio Gonçalves Silva – Universidade Federal do Maranhão  
Profª Ma. Andréa Cristina Marques de Araújo – Universidade Fernando Pessoa  
Profª Drª Andrezza Lopes – Instituto de Pesquisa e Desenvolvimento Acadêmico  
Profª Drª Andrezza Miguel da Silva – Faculdade da Amazônia  
Profª Ma. Anelisa Mota Gregoleti – Universidade Estadual de Maringá  
Profª Ma. Anne Karynne da Silva Barbosa – Universidade Federal do Maranhão  
Prof. Dr. Antonio Hot Pereira de Faria – Polícia Militar de Minas Gerais  
Prof. Me. Armando Dias Duarte – Universidade Federal de Pernambuco  
Profª Ma. Bianca Camargo Martins – UniCesumar  
Profª Ma. Carolina Shimomura Nanya – Universidade Federal de São Carlos  
Prof. Me. Carlos Antônio dos Santos – Universidade Federal Rural do Rio de Janeiro  
Prof. Me. Carlos Augusto Zilli – Instituto Federal de Santa Catarina  
Prof. Me. Christopher Smith Bignardi Neves – Universidade Federal do Paraná  
Profª Drª Cláudia de Araújo Marques – Faculdade de Música do Espírito Santo  
Profª Drª Cláudia Taís Siqueira Cagliari – Centro Universitário Dinâmica das Cataratas  
Prof. Me. Clécio Danilo Dias da Silva – Universidade Federal do Rio Grande do Norte  
Prof. Me. Daniel da Silva Miranda – Universidade Federal do Pará  
Profª Ma. Daniela da Silva Rodrigues – Universidade de Brasília  
Profª Ma. Daniela Remião de Macedo – Universidade de Lisboa  
Profª Ma. Dayane de Melo Barros – Universidade Federal de Pernambuco

Prof. Me. Douglas Santos Mezacas – Universidade Estadual de Goiás  
Prof. Me. Edevaldo de Castro Monteiro – Embrapa Agrobiologia  
Prof. Me. Edson Ribeiro de Britto de Almeida Junior – Universidade Estadual de Maringá  
Prof. Me. Eduardo Gomes de Oliveira – Faculdades Unificadas Doctum de Cataguases  
Prof. Me. Eduardo Henrique Ferreira – Faculdade Pitágoras de Londrina  
Prof. Dr. Edwaldo Costa – Marinha do Brasil  
Prof. Me. Eliel Constantino da Silva – Universidade Estadual Paulista Júlio de Mesquita  
Prof. Me. Ernane Rosa Martins – Instituto Federal de Educação, Ciência e Tecnologia de Goiás  
Prof. Me. Euvaldo de Sousa Costa Junior – Prefeitura Municipal de São João do Piauí  
Prof. Dr. Everaldo dos Santos Mendes – Instituto Edith Theresa Hedwing Stein  
Prof. Me. Ezequiel Martins Ferreira – Universidade Federal de Goiás  
Prof<sup>a</sup> Ma. Fabiana Coelho Couto Rocha Corrêa – Centro Universitário Estácio Juiz de Fora  
Prof. Me. Fabiano Eloy Atílio Batista – Universidade Federal de Viçosa  
Prof. Me. Felipe da Costa Negrão – Universidade Federal do Amazonas  
Prof. Me. Francisco Odécio Sales – Instituto Federal do Ceará  
Prof. Me. Francisco Sérgio Lopes Vasconcelos Filho – Universidade Federal do Cariri  
Prof<sup>a</sup> Dr<sup>a</sup> Germana Ponce de Leon Ramírez – Centro Universitário Adventista de São Paulo  
Prof. Me. Gevair Campos – Instituto Mineiro de Agropecuária  
Prof. Me. Givanildo de Oliveira Santos – Secretaria da Educação de Goiás  
Prof. Dr. Guilherme Renato Gomes – Universidade Norte do Paraná  
Prof. Me. Gustavo Krahl – Universidade do Oeste de Santa Catarina  
Prof. Me. Helton Rangel Coutinho Junior – Tribunal de Justiça do Estado do Rio de Janeiro  
Prof<sup>a</sup> Ma. Isabelle Cerqueira Sousa – Universidade de Fortaleza  
Prof<sup>a</sup> Ma. Jaqueline Oliveira Rezende – Universidade Federal de Uberlândia  
Prof. Me. Javier Antonio Albornoz – University of Miami and Miami Dade College  
Prof. Me. Jhonatan da Silva Lima – Universidade Federal do Pará  
Prof. Dr. José Carlos da Silva Mendes – Instituto de Psicologia Cognitiva, Desenvolvimento Humano e Social  
Prof. Me. Jose Elyton Batista dos Santos – Universidade Federal de Sergipe  
Prof. Me. José Luiz Leonardo de Araujo Pimenta – Instituto Nacional de Investigación Agropecuaria Uruguay  
Prof. Me. José Messias Ribeiro Júnior – Instituto Federal de Educação Tecnológica de Pernambuco  
Prof<sup>a</sup> Dr<sup>a</sup> Juliana Santana de Curcio – Universidade Federal de Goiás  
Prof<sup>a</sup> Ma. Juliana Thaisa Rodrigues Pacheco – Universidade Estadual de Ponta Grossa  
Prof<sup>a</sup> Dr<sup>a</sup> Kamilly Souza do Vale – Núcleo de Pesquisas Fenomenológicas/UFGA  
Prof. Dr. Kárpio Márcio de Siqueira – Universidade do Estado da Bahia  
Prof<sup>a</sup> Dr<sup>a</sup> Karina de Araújo Dias – Prefeitura Municipal de Florianópolis  
Prof. Dr. Lázaro Castro Silva Nascimento – Laboratório de Fenomenologia & Subjetividade/UFRP  
Prof. Me. Leonardo Tullio – Universidade Estadual de Ponta Grossa  
Prof<sup>a</sup> Ma. Lillian Coelho de Freitas – Instituto Federal do Pará  
Prof<sup>a</sup> Ma. Lilian de Souza – Faculdade de Tecnologia de Itu  
Prof<sup>a</sup> Ma. Liliani Aparecida Sereno Fontes de Medeiros – Consórcio CEDERJ  
Prof<sup>a</sup> Dr<sup>a</sup> Lívia do Carmo Silva – Universidade Federal de Goiás  
Prof. Dr. Lucio Marques Vieira Souza – Secretaria de Estado da Educação, do Esporte e da Cultura de Sergipe  
Prof. Dr. Luan Vinicius Bernardelli – Universidade Estadual do Paraná  
Prof<sup>a</sup> Ma. Luana Ferreira dos Santos – Universidade Estadual de Santa Cruz  
Prof<sup>a</sup> Ma. Luana Vieira Toledo – Universidade Federal de Viçosa  
Prof. Me. Luis Henrique Almeida Castro – Universidade Federal da Grande Dourados  
Prof. Me. Luiz Renato da Silva Rocha – Faculdade de Música do Espírito Santo  
Prof<sup>a</sup> Ma. Luma Sarai de Oliveira – Universidade Estadual de Campinas  
Prof. Dr. Michel da Costa – Universidade Metropolitana de Santos  
Prof. Me. Marcelo da Fonseca Ferreira da Silva – Governo do Estado do Espírito Santo  
Prof. Dr. Marcelo Máximo Purificação – Fundação Integrada Municipal de Ensino Superior  
Prof. Me. Marcos Aurelio Alves e Silva – Instituto Federal de Educação, Ciência e Tecnologia de São Paulo

Profª Ma. Maria Elanny Damasceno Silva – Universidade Federal do Ceará  
Profª Ma. Marileila Marques Toledo – Universidade Federal dos Vales do Jequitinhonha e Mucuri  
Prof. Dr. Pedro Henrique Abreu Moura – Empresa de Pesquisa Agropecuária de Minas Gerais  
Prof. Me. Pedro Panhoca da Silva – Universidade Presbiteriana Mackenzie  
Profª Drª Poliana Arruda Fajardo – Universidade Federal de São Carlos  
Prof. Me. Rafael Cunha Ferro – Universidade Anhembi Morumbi  
Prof. Me. Ricardo Sérgio da Silva – Universidade Federal de Pernambuco  
Prof. Me. Renan Monteiro do Nascimento – Universidade de Brasília  
Prof. Me. Renato Faria da Gama – Instituto Gama – Medicina Personalizada e Integrativa  
Profª Ma. Renata Luciane Polsaque Young Blood – UniSecal  
Prof. Me. Robson Lucas Soares da Silva – Universidade Federal da Paraíba  
Prof. Me. Sebastião André Barbosa Junior – Universidade Federal Rural de Pernambuco  
Profª Ma. Silene Ribeiro Miranda Barbosa – Consultoria Brasileira de Ensino, Pesquisa e Extensão  
Profª Ma. Solange Aparecida de Souza Monteiro – Instituto Federal de São Paulo  
Profª Ma. Taiane Aparecida Ribeiro Nepomoceno – Universidade Estadual do Oeste do Paraná  
Prof. Me. Tallys Newton Fernandes de Matos – Faculdade Regional Jaguaribana  
Profª Ma. Thatiany Jasmine Castro Martins de Carvalho – Universidade Federal do Piauí  
Prof. Me. Tiago Silvío Dedoné – Colégio ECEL Positivo  
Prof. Dr. Welleson Feitosa Gazel – Universidade Paulista



## Tecnologias, métodos e teorias na engenharia de computação 2

**Bibliotecária:** Janaina Ramos  
**Diagramação:** Maria Alice Pinheiro  
**Correção:** Maiara Ferreira  
**Edição de Arte:** Luiza Alves Batista  
**Revisão:** Os Autores  
**Organizador:** Ernane Rosa Martins

### Dados Internacionais de Catalogação na Publicação (CIP)

T255	Tecnologias, métodos e teorias na engenharia de computação 2 / Organizador Ernane Rosa Martins. – Ponta Grossa - PR: Atena, 2021.  Formato: PDF Requisitos de sistema: Adobe Acrobat Reader Modo de acesso: World Wide Web Inclui bibliografia ISBN 978-65-5706-945-5 DOI 10.22533/at.ed.455211604  1. Engenharia de Computação. I. Martins, Ernane Rosa (Organizador). II. Título.  CDD 621.39
<b>Elaborado por Bibliotecária Janaina Ramos – CRB-8/9166</b>	

**Atena Editora**  
Ponta Grossa – Paraná – Brasil  
Telefone: +55 (42) 3323-5493  
[www.atenaeditora.com.br](http://www.atenaeditora.com.br)  
contato@atenaeditora.com.br

## DECLARAÇÃO DOS AUTORES

Os autores desta obra: 1. Atestam não possuir qualquer interesse comercial que constitua um conflito de interesses em relação ao artigo científico publicado; 2. Declaram que participaram ativamente da construção dos respectivos manuscritos, preferencialmente na: a) Concepção do estudo, e/ou aquisição de dados, e/ou análise e interpretação de dados; b) Elaboração do artigo ou revisão com vistas a tornar o material intelectualmente relevante; c) Aprovação final do manuscrito para submissão.; 3. Certificam que os artigos científicos publicados estão completamente isentos de dados e/ou resultados fraudulentos; 4. Confirmam a citação e a referência correta de todos os dados e de interpretações de dados de outras pesquisas; 5. Reconhecem terem informado todas as fontes de financiamento recebidas para a consecução da pesquisa.

## APRESENTAÇÃO

A Engenharia de Computação é a área que estuda as técnicas, métodos e ferramentas matemáticas, físicas e computacionais para o desenvolvimento de circuitos, dispositivos e sistemas. Assim, este segundo volume busca apresentar a matemática e a computação com foco no desenvolvimento de soluções de software e na solução de problemas de Engenharia.

Dentro deste contexto, esta obra apresenta diversos aspectos tecnológicos computacionais, tais como: um software que reúna informações científicas sobre vacinas e doenças imunopreveníveis de forma lúdica; um modelo preditivo com objetivo de identificar a correlação entre o valor predito e o preço de fechamento das ações listadas na bolsa de valores brasileira; ensino de programação para crianças; o algoritmo genético e o método da evolução diferencial; uma modelagem matemática para o cenário de um ciclo de desenvolvimento do Scrum; simulações computacionais; um sistema háptico sonoro para auxiliar a navegação e locomoção de deficientes visuais em ambientes fechados; uma solução ótima de despacho de geração de energia elétrica para 4 usinas térmicas, através de simulação no software MATLAB; uma rede neural perceptron multicamadas para previsão de séries temporais de nível de água de uma bacia hidrográfica; uma rede neural artificial (Multilayer Perceptron) para a classificação de perfis de passageiros no setor aéreo brasileiro; um modelo de aprendizado de máquina que combina diferentes técnicas de regressão; a complexidade na inteligência artificial dos mascotes virtuais.

Sendo assim, esta obra é composta por trabalhos pertinentes da área, que permitem aos leitores, analisar e discutir assuntos importantes. Por fim, agradecemos aos autores pelas significativas contribuições, e desejamos aos nossos leitores uma excelente leitura, repleta de reflexões significativas.

Ernane Rosa Martins

## SUMÁRIO

### **CAPÍTULO 1..... 1**

#### **VACINA.COM: A SOFTWARE FOR TEACHING AND PROFESSIONAL UPDATING ABOUT VACCINES AND IMMUNO-PREVENTABLE DISEASES**

Paôla de Oliveira Souza  
José Maria Parente de Oliveira  
Letícia Helena Januário  
Daniel Moraes dos Reis  
Paula Luciana Gonçalves Pereira  
André Almeida Gonçalves

**DOI 10.22533/at.ed.4552116041**

### **CAPÍTULO 2..... 13**

#### **UMA ANÁLISE DE VANTAJOSIDADE EM MODELOS DE PREVISÃO EM SÉRIES TEMPORAIS**

Rafael Diniz Toscano de Lima  
Sérgio Murilo Maciel Fernandes  
Sidney Marlon Lopes de Lima  
Ricardo Paranhos Pinheiro  
Sthéfano Henrique Mendes Tavares Silva

**DOI 10.22533/at.ed.4552116042**

### **CAPÍTULO 3..... 24**

#### **SENTECH: UM COMBINADOR DE ANÁLISE TÉCNICA E DE SENTIMENTO PARA O MERCADO DE AÇÕES**

Isabela Nunes Caetano  
Érica Ferreira de Souza  
Giovani Volnei Meinerz

**DOI 10.22533/at.ed.4552116043**

### **CAPÍTULO 4..... 34**

#### **PROGRAMAÇÃO DE JOGOS COM SCRATCH PARA AUXÍLIO À ALFABETIZAÇÃO DE CRIANÇAS**

Rute Vitorino Oliveira  
Jemima Vitorino de Oliveira  
Luciene Cavalcanti Rodrigues  
Ana Paula Garrido de Queiroga

**DOI 10.22533/at.ed.4552116044**

### **CAPÍTULO 5..... 46**

#### **OTIMIZAÇÃO GEOMÉTRICA DAS PÁS DE UMA TURBINA EÓLICA DE EIXO HORIZONTAL**

Rafael Romão da Silva Melo

**DOI 10.22533/at.ed.4552116045**

**CAPÍTULO 6..... 59**

**OTIMIZAÇÃO DO SPRINT BACKLOG COM O PROBLEMA DA MOCHILA 0/1**

Michel Willian Alves  
Elisa de Fátima Andrade Soares  
Thalia Katiane Sampaio Gurgel  
José Weliton de Vasconcelos Filho  
Dario José Aloise

**DOI 10.22533/at.ed.4552116046**

**CAPÍTULO 7..... 68**

**MODELOS EPIDÊMICOS: PROJETO DE INICIAÇÃO CIENTÍFICA NO CONTEXTO DA COVID-19**

Vinícius R. da Silva  
Felipe Y. Hatanaka  
Olavo H. Menin

**DOI 10.22533/at.ed.4552116047**

**CAPÍTULO 8..... 78**

**GUIDE2BLIND: SISTEMA HÁPTICO-SONORO DE ORIENTAÇÃO PARA DEFICIENTES VISUAIS EM AMBIENTES FECHADOS - FASE 2**

Lucas Rafael da Silva Martins  
Mikael Tolotti da Silva  
Bernardo Moreira  
Diego Afonso da Silva Lima  
Carlos Francisco Soares de Souza  
Luis Gustavo Fernandes dos Santos  
Carlos Arthur Carvalho Sarmanho Junior

**DOI 10.22533/at.ed.4552116048**

**CAPÍTULO 9..... 96**

**DESPACHO DE GERAÇÃO ÓTIMA ATRAVÉS DO MÉTODO DOS PONTOS INTERIORES VERSÃO PRIMAL-DUAL**

Jean Ferguson Pimentel  
João Vitor Gerevini Kasper  
Juliana Almansa Malagoli  
Thelma Solange Piazza Fernandes

**DOI 10.22533/at.ed.4552116049**

**CAPÍTULO 10..... 105**

**COMBINING RAINFALL AND WATER LEVEL DATA FOR MULTISTEP HIGH TEMPORAL RESOLUTION EMPIRICAL HYDROLOGICAL FORECASTING**

Cintia Pereira de Freitas  
Michael Macedo Diniz  
Glauston Roberto Teixeira de Lima  
Marcos Gonçalves Quiles  
Stephan Stephany  
Leonardo Bacelar Lima Santos

**DOI 10.22533/at.ed.45521160410**

<b>CAPÍTULO 11</b> .....	<b>118</b>
<b>CLASSIFICAÇÃO DE PASSAGEIROS DOMÉSTICOS DE LINHAS AÉREAS UTILIZANDO REDES NEURAIS ARTIFICIAIS DO TIPO MLP</b>	
Sidnei Gouveia Junior	
Narciso Ferreira dos Santos Neto	
Nilton Alves Maia	
<b>DOI 10.22533/at.ed.45521160411</b>	
<b>CAPÍTULO 12</b> .....	<b>129</b>
<b>APRENDIZADO CONJUNTO APLICADO NA PREDIÇÃO DO MERCADO DE AÇÕES BRASILEIRO</b>	
Alvaro Pedroso Queiroz	
Giovani Volnei Meinerz	
Érica Ferreira de Souza	
<b>DOI 10.22533/at.ed.45521160412</b>	
<b>CAPÍTULO 13</b> .....	<b>138</b>
<b>INFORMATIZAÇÃO DE PROCESSOS GERENCIAIS EM UM SETOR DE ASSISTÊNCIA ESTUDANTIL: ESTUDO DE CASO NO IFMG – CAMPUS BAMBUÍ</b>	
Eduardo Cardoso Melo	
Gabriel da Silva	
<b>DOI 10.22533/at.ed.45521160413</b>	
<b>CAPÍTULO 14</b> .....	<b>151</b>
<b>A SIMULAÇÃO DE EMOÇÕES EM JOGOS DIGITAIS</b>	
Pedro Henrique Senkio Cardoso	
<b>DOI 10.22533/at.ed.45521160414</b>	
<b>SOBRE O ORGANIZADOR</b> .....	<b>158</b>
<b>ÍNDICE REMISSIVO</b> .....	<b>159</b>

# CAPÍTULO 1

## VACINA.COM: A SOFTWARE FOR TEACHING AND PROFESSIONAL UPDATING ABOUT VACCINES AND IMMUNO-PREVENTABLE DISEASES

*Data de aceite: 01/04/2021*

*Data de submissão: 09/02/2021*

### **Paôla de Oliveira Souza**

Instituto Tecnológico de Aeronáutica  
São José dos Campos - São Paulo  
<http://lattes.cnpq.br/2714652578820025>

### **José Maria Parente de Oliveira**

Instituto Tecnológico de Aeronáutica  
São José dos Campos - São Paulo  
<http://lattes.cnpq.br/9413052240755786>

### **Letícia Helena Januário**

Universidade Federal de São João del Rei  
Divinópolis - Minas Gerais  
<http://lattes.cnpq.br/2284103972327626>

### **Daniel Morais dos Reis**

Centro Federal de Educação Tecnológica de  
Minas Gerais  
Divinópolis - Minas Gerais  
<http://lattes.cnpq.br/2020021419382172>

### **Paula Luciana Gonçalves Pereira**

Universidade Federal de São João del Rei  
Divinópolis - Minas Gerais  
<http://lattes.cnpq.br/1867644608303091>

### **André Almeida Gonçalves**

Centro Federal de Educação Tecnológica de  
Minas Gerais  
Divinópolis - Minas Gerais  
<http://lattes.cnpq.br/4939552189201118>

**ABSTRACT:** Introduction: vaccination is one of the main procedures accomplished in public healthcare to promote health and disease prevention. If on one hand vaccination seems simple enough, still it houses a sequence of information and procedures that are essential to the success of the task. Diverse situations may complicate the process and expose professionals and healthcare users to risk of error and damages, demanding greater agility from services and health professionals in the update of procedures. Usually, consultation documents from students and professionals are technical reports and manuals that fragment themselves as new procedures are introduced. Studies show the necessity of strategies that allow the education of students and professionals in a fast and effective manner. Objective: develop a software that gathers scientific information about vaccines and immuno-preventable diseases in a ludic manner. Method: The utilization of a web platform with Java, HTML, CSS, JavaScript, the graphical framework Bootstrap, and MySQL database. Results: the “Vacina.com” portal was developed to interface with teaching, integrating knowledge concerning the National Immunization Program’s vaccines and respective preventable diseases. The portal was implemented in three modules: Game, Wiki and Mirror Card. Conclusions: Vacina.com allows working with complex information in a simple way and fills a gap in the education of professionals responsible for vaccination. With dynamic design, it values ease of access, knowledge update, information completeness and quick answering. As such, it contributes to health promotion, minimizing

mistakes in the vaccination process and reducing preventable risks and aggravations.

**KEYWORDS:** Health, Vaccination, Educational Software.

## VACINA.COM: UM SOFTWARE PARA O ENSINO E ATUALIZAÇÃO PROFISSIONAL SOBRE VACINAS E DOENÇAS IMUNO-PREVENÍVEIS

**RESUMO:** Introdução: a vacinação é um dos principais procedimentos realizados na rede pública de saúde para promoção da saúde e prevenção de doenças. A vacinação, se por um lado parece simples, por outro, abriga uma cadeia de processos e informações essenciais para o sucesso da ação. Situações diversas podem tornar o procedimento mais complexo e expor profissionais e usuários ao risco de erros e danos, exigindo maior agilidade dos serviços e profissionais de saúde na atualização dos procedimentos. Normalmente os documentos de consulta de alunos e profissionais são manuais e relatórios técnicos, que se fragmentam à medida que novos procedimentos são introduzidos. Estudos mostram a necessidade de estratégias que possibilitem a formação de alunos e profissionais de forma mais rápida e eficaz. Objetivo: desenvolver um software que reúna informações científicas sobre vacinas e doenças imunopreveníveis de forma lúdica. Método: Utilizamos a plataforma web com Java, HTML, CSS, JavaScript, o framework gráfico Bootstrap e o banco de dados MySQL. Resultado: o portal “Vacina.com” foi desenvolvido para fazer interface com o ensino, integrando conhecimentos sobre vacinas PNI e respectivas doenças evitáveis. O portal foi implementado em três módulos: Game, Wiki e Mirror Card. Conclusões: O Vacina.com permite trabalhar com informações complexas de forma simples e preenche uma lacuna na formação dos profissionais responsáveis pela vacinação. Com design dinâmico, valoriza a facilidade de acesso, a atualização do conhecimento, a completude das informações e a rapidez de resposta. Dessa forma, contribui para a promoção da saúde, minimizando erros no processo de vacinação e reduzindo riscos e agravos evitáveis.

**PALAVRAS - CHAVE:** Saúde, Vacinação, Software Educacional.

### 1 | INTRODUCTION

Large epidemics have devastated nations across the timeline of human evolution. Despite important scientific and technological advances in the field of healthcare, humanity is still threatened by pandemics such as Covid-19. Nonetheless, diseases that are eradicable by immunization, like mumps, measles, and rubella, also prevail, mainly in economically less-privileged populations. In this sense, the National Immunization Program (or *Programa Nacional de Imunização* – PNI), performs an important role, implementing actions of reduction and eradication of immuno-preventable diseases. Since its creation in 1973, the actions of the PNI have been characterized by assisting all Brazilian citizens, with no distinction, through the assessment of epidemiological situations, risks and social particularities, such as those of indigenous peoples [1]. Today, more than 300 million vaccines are applied in Brazil each year[2]. The basic health units administer specific vaccines to Brazilian population from birth to old age. However, the constant increase in the



number of vaccines entails the larger complexity of its handling, especially concerning the person responsible for administering the vaccine to the individual [3].

If vaccination may seem simple enough, on the other hand, this process houses a framework and essential information for a successful action. The regular vaccination calendar from Brazil's Ministry of Health (*Ministério da Saúde* – MS), in 2020, advocates the administering of 27 doses of vaccines in the first 10 years of life [4]. These 27 doses contain 14 different immunobiologicals for the prevention of 17 specific diseases, notwithstanding extra dosages from eventual vaccination campaigns. For the administering of such doses in the regular calendar, it is necessary to consider not only the correct vaccine, the administering volume, the means of application (be it oral, intramuscular, subcutaneous or intradermal); the anatomical region (intramuscular in the right or left thigh's vastus lateralis muscle, or the deltoid); but also specific guidelines to parents or guardians concerning side effects related to each vaccine. Many of these occurrences are common and may have light, moderate or serious consequences.

The possibilities of mistake increase in situations with migrating users [5], delays that may be related to the user (such as forgetting or a sickness) or healthcare issues (be it a lack of the immunobiological product or because the professionals chose to schedule specific dates for specific vaccines) [3]. In such circumstances, it is needed that the worker take into account not only the previously mentioned technical recommendations, but also data related to a safe vaccination procedure, such as: the minimal interval between doses of a same vaccine, the minimum and maximum age limit for a vaccine, and the feasibility of simultaneous application of two or more vaccines. There could also be guidelines, contraindications (and false contraindications) in the moment of administering these substances [6]. The conservation of immunobiological products is also a complex theme. It is up to cold chain logistics ensure quality of storage, conservation, handling, distribution, and transportation, from producing laboratory to the moment of an individual's vaccination. According to specific characteristics of immunobiologicals, these products can or cannot be frozen, must be stored at the first or second shelf of a freezer, and can or cannot be administered 4 hours after dilution[7].

The PNI establishes guidelines and technical-scientific procedures handed to health secretaries by means of strategic systemic actions of populace vaccination. The combination of differentiated vaccination strategies such as the case of vaccination campaigns and routine calendar – actions of epidemiological surveillance implemented as means of lowering incidence of preventable diseases. The administering of vaccines in healthcare units is responsibility of the nursing team, usually composed of a nurse and a responsible auxiliary or nursing technician. However, all maintenance chains are responsible for keeping the needed quality of immunobiological administering processes that can be disrupted if, in the moment of application, during classes, the professional do not receive appropriate training.

## 2 | BACKGROUND

Studies generally recommend the creation of strategies that allows updating of professionals responsible for quick and effective vaccination. But generally, real circumstances do not contribute to the quality of the procedure. Oftentimes, both guidelines manuals and recommended specifications makes difficult the access of the professional, be it by the volume of data that is higher than needed for quick consultations, or the low reproduction quality of photocopied materials. Moreover, due to the insufficiency of human resources, or work overload, professionals are taken to clarify their doubts in faster sources, like Google. Besides other situations, such as conduct related to inadequate techniques of professionals [8], distraction, and unfamiliarity of adverse effects [9,10]. Another concerning factor are fake news related to anti-vaccination movements.

The use of informatized resources related to vaccination is widespread. The MS [11], presents a digital system to coverage assessment. The Center for Disease Control and Prevention (CDC) of Atlanta, [12], besides systems for monitoring vaccination coverage, publishes digitized programs to orient parents about the recommended vaccination calendar for children. The World Health Organization, WHO, [13], also presents software applications directed to parents. In that sense, literature is consensual in pointing towards the necessity of developing more effective educational techniques [14,15], as well as the rising importance of mobile applications usage [16,17,18]. Many studies present applications and mobile applications' reviews related to vaccination [19,20,21,22], and most describe results concerning acceptance of vaccination[18,23], information about vaccines[24,25], vaccination records [26,27], vaccination coverage[20,22], vaccination alerts [28], knowledge for the decision making of parents and/or professionals [23,29], communication about vaccine safety [30,31].

The use of digitized resources in teaching environment is widespread, however, no papers were found specifically related to vaccination. That is, no apps were found to help the process of teaching/learning about vaccination. The application “Vacinação em Dia” (Up to date Vaccination) from the Ministry of Health (MS) [11], available for Android operating system, has as its target-audience the citizens and their families.

As such, the objective of this research was to develop an interactive software, accessible and with ludic language, that provides information and practical solutions to the most usual questions and problems from the daily routines of professionals responsible for vaccination.

## 3 | METHODS

This study concerns an exploratory research of technological development. To guarantee the reciprocal influence of the distinct areas that are involved, a focal group was utilized. Nursing and computer-engineering graduation students and teachers, and

primary attention healthcare nurses composed the group. The technical content related to vaccination and immuno-preventable diseases to subsidize the modules of Vacina.com were extracted from manuals and directives from Ministry of Health[4,6,7,32,33] and the Brazilian Society of Immunizations (SBIIm) [34].

The implementation of code from the Vacina.com (Vaccine.com) portal was made in JavaServer Pages (JSP), a technology that aids the creation of webpages using HTML and Java. The model of used software architecture was MVC: Model, View and Controller, where: Model concerns logic and functions that act as the “cogs” of software; View consists in the creation of interface pages and the direct interaction with the user; and Controller connects and organizes both previous modules. For the implementation of graphical interface, it was opted to use Bootstrap, a framework of HTML5 and CSS3. Its function is to help the creation of websites and web applications, and one of its greatest advantages is being compatible with current versions of the most used web browsers, besides supporting responsive design, that allows adaptations to different kinds of devices. The platforms used for implementation were NetBeans to the development of JSP, and MySQL Workbench to work directly with the database.

The Vacina.com portal was designed with three modules: a game, a virtual encyclopedia and a vaccination mirror card. The game approaches knowledge concerning immunobiologicals and immuno-preventable diseases. The encyclopedia (wiki) must gather information about vaccines and diseases that allow for treatment by means of vaccines. Lastly, the vaccination mirror card is a digital version of the current vaccination record found in basic healthcare units.

## **4 | RESULTS AND DISCUSSION**

The portal was implemented to house three modules: game, wiki and vaccination card. In the Game, the player is provided with tips that bring specific information about previously registered themes. At the end of each match, score is stored in a database to allow a player ranking. Reports for the teacher’s evaluation are also provided. On the Wiki, from the insertion of a term that refers to the intended subject matter, the software provides information about vaccines and the diseases related to the research, that the user could browse, obtaining more item-specific information. The Vaccination Card is a digital version of the vaccination record found in healthcare units. The user has access to a record of the administered vaccinations, and is informed if he is or is not up to date with deadlines established by the National Program of Immunizations.

When the user accesses the system, he is directed to a screen that presents the three modules proposed by the project: the game, the encyclopedia, and the vaccination mirror card (see Fig. 1).



Sua forma interativa de aprendizado...



Fig. 1. Vacina.com homepage

## 4.1 The Game

The first module of the platform is the game concerning immunobiologicals and immuno-preventable diseases. This application integrates knowledge of vaccinations presented by the National Vaccination Calendar [4]. The purpose of this core is to provide an interactive and singular form of learning that could be aggregated into the institution's pedagogical project, helping the effectiveness of the teaching/learning process. The provided information in this software was extracted from the Manual of Norms and Procedures for Vaccination [6], the Manual of Reference Centers for Special Immunobiologicals [32], and the Manual of Epidemiologic Surveillance of Post-Vaccination Adverse Events [33], from technical guidelines, and from the Brazilian Society of Immunization [34].

In this game's model, only one player can participate at a time. A player may choose hints that provide specific information about the theme, to guess to which immunobiological or immuno-preventable disease this information concerns. The player starts the round with 1000 points and each hint is valued 100 points.

If the player chooses a hint, the box that refers to the number of a hint will be presented in the white color, and the guideline is presented in the "Hint" box, as can be seen in the Figure 2 example. According to the game's rules, each new information the user solicits causes him to lose 100 points.

When the player knows what the hints refer to, he can submit an answer to the system, losing 10 points with each wrong guess. In case of all the hints being solicited, and the student giving up on trying to answer what is the vaccine or disease in question, he can press the "View Answer" button.

An important fact is that the target audience of Vacina.com are students of health-related courses, such as Nursing. These students will be professionals that coexist daily with the possibility of mistakes, and as such, the game aims helping them test their necessary

knowledge. By the end of a match, the player's score will be given by means of the number of hints that were not used.



Fig. 2. Game and the tip box

The use of ludic approaches to the work and qualification of health professionals can increase motivation. Studies show the relevancy of serious games and gamification when promoting health and prevention, especially in the case of specific groups such as healthcare professionals, or teenagers, possibly disinterested [35,36]. The use of serious games for vaccination is an innovative tool for public health [37].

Although there is a relative scarcity of studies concerning the use of digital games when promoting health, studies suggest that digital gamification has a strong potential to increase knowledge about vaccination, and eventually, the vaccination coverage itself [20].

## 4.2 Immunopedia

The other implemented module is the Immunopedia. The student or health professional types in a vaccine or disease and the portal provides a list of results for that search, and is up to the visitor to choose one an option to access the vaccine or disease. The topics of vaccine are: Immunity against, Indication, Contraindication, Before vaccination (including guidelines to the person responsible and adverse events), Calendar, Vaccination preparation, Dosage, Route of Administration, Place of application, Conservation, Production laboratory, Composition, Culture, Vaccine efficiency, and Titration. The topics referring to immuno-preventable diseases are: Disease description, etiological agent, reservoir, period and mode of transmission, period of incubation, clinical manifestations, prevention, and treatment.

The mobile devices enhanced the health care for individuals that need it as much as for the providers of the service. Smartphones offer the opportunity to enhance the

immunization practice and facilitate personalized communication among individuals and public health authorities [7,9]. With the increased use of mobile devices and applications, there is the potential of capturing, transmitting and monitoring information of post-immunization experience in real time using personal mobile devices [27].

### 4.3 Vaccination Card

The Vaccination Card module is a digital version of the mirror card used in basic health units to control and update individual vaccines. In this module, when used as a teaching/learning tool, allows the teacher to create hypothetical situations of already administered vaccines, and the student must design the next doses according to the national vaccination calendar. When used in basic healthcare units, the user reports the registration of vaccines already applied to an individual and the application will indicate the vaccines still to be administered. The application is also able to alert both professionals and individuals about the need to apply upcoming vaccines. And the user can also contribute to updating the mirror card database of basic healthcare units.

In the Vaccination Card module, there is the display page of the vaccines already applied to the patient and the pending vaccines, both arranged in squares marked by the blue and red colors respectively, as shown in Figure 4.

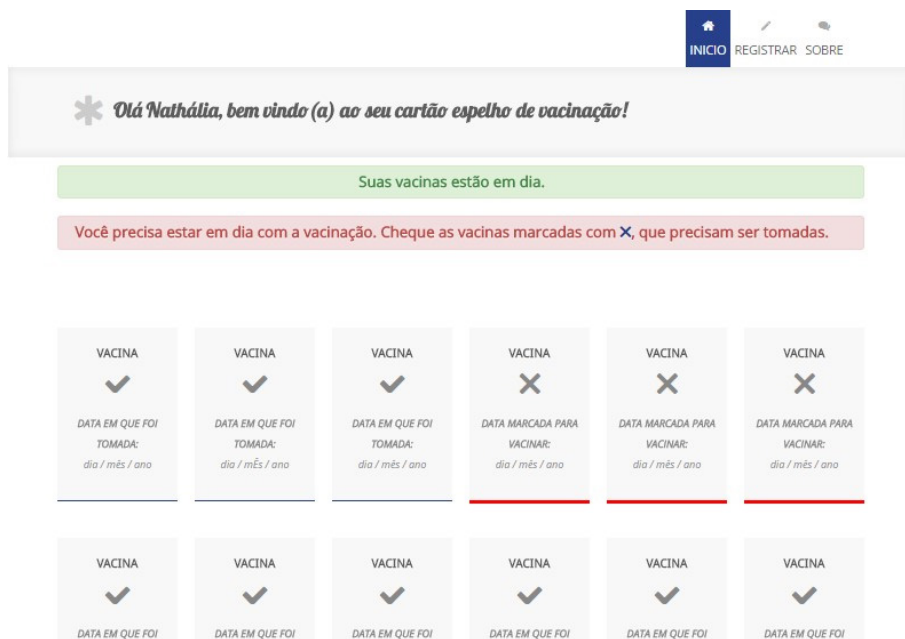


Fig. 3. Vaccination Cards

The portal can be used in healthcare units, by which the professional updates the database of vaccination records. By means of Vacina.com complex information can be worked with in a simple way. With a dynamic design, it values ease of access, the updating of knowledge, information wholeness, and answering speed, facilitating the access of necessary information for safe vaccination. As such, it contributes to promoting healthcare by means of minimizing errors in the vaccination process, reducing risks and preventable aggravations.

## 5 | CONCLUSION

Vaccination is complex in method, and it is necessary to have deep knowledge about each vaccine so that one reaches a successful outcome. It is also known that health professionals are not always the bearers of such crucial specificities during the vaccination process.

The produced software presents itself as a tool with the ability to help remedy this demand of professionalization, as it contains essential information concerning immunobiologicals and immuno-preventable diseases in a dynamic way, and by means of a highly popularized tool: the Internet.

However, it is known that the work does not end there. The research group intends to make improvements in the already produced core with, for example, an increase in the number of players, and the possibility of player interaction.

## REFERENCES

1. Ministério da Saúde do Brasil. Departamento de Vigilância de Doenças Transmissíveis. Coordenação Geral do Programa Nacional de Imunizações. Brasília, 2016 Homepage, <https://portal.arquivos2.saude.gov.br/images/pdf/2016/maio/05/relatorio-01-criterios-orientar-decisao-vacina-dengue.pdf>
2. Cruz A. **A queda da imunização no Brasil**. Revista Consensus, 2017; 20-29. Homepage [https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/revistaconsensus\\_25\\_a\\_queda\\_da\\_imunizacao.pdf](https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/revistaconsensus_25_a_queda_da_imunizacao.pdf)
3. Souza IBJ, Lago EC, Araújo, TME, Almeida, APL, Tapety FI, Carvalho ML. **Conhecimento de enfermeiros e médicos de uma microrregião do Nordeste brasileiro sobre a vacinação infantil**. Nursing Dezembro 2018; 21 (247): 2498-2505. ID: biblio-970888
4. Ministério da Saúde do Brasil. **Calendário Nacional de Vacinação 2020**. Homepage, <https://www.saude.gov.br/files/imunizacao/calendario/Calendario.Nacional.Vacinacao.2020.atualizado.pdf>
5. McCallum AD, Kanagarajah S, Lisa F, Patel D. **To err is human: Clinical incident calls to a national travel health advice line**. Vaccine. December 2019; 37 (52): 7535-7538. <https://doi.org/10.1016/j.vaccine.2019.10.007>.

6. Ministério da Saúde do Brasil. **Manual de Normas e Procedimentos para Vacinação**. Brasília, 2014. Homepage, [http://bvsmis.saude.gov.br/bvs/publicacoes/manual\\_procedimentos\\_vacinacao.pdf](http://bvsmis.saude.gov.br/bvs/publicacoes/manual_procedimentos_vacinacao.pdf), ISBN 978-85-334-2164-6.
7. Ministério da Saúde do Brasil. **Manual de Rede de Frios do Programa Nacional de Imunizações**. Brasília, 2017. Homepage, [https://portalarquivos2.saude.gov.br/images/pdf/2017/dezembro/15/rede\\_frio\\_2017\\_web\\_VF.pdf](https://portalarquivos2.saude.gov.br/images/pdf/2017/dezembro/15/rede_frio_2017_web_VF.pdf), ISBN 978-85-334-2534-7.
8. Dutra FC, Vasconcelos PF, Monteiro FM, Freire VE, Souza Neto PH. **Falhas na administração de imunobiológicos: análise de causa raiz**. Rev enferm UFPE. 2019; 13: <https://doi.org/10.5205/1981-8963.2019.239254>
9. Aragão RF, Albuquerque IMN, Ribeiro MA, Barreto RM, Sousa JA. **Percepções e conhecimentos da equipe de enfermagem sobre o processo de imunização**. Rev. bras. promoç. Saúde. 2019; 32: 01-08. <https://doi.org/10.5020/18061230.2019.8809>
10. Silva MRB, Oliveira RB, Armada e Silva HCD, Medeiros CS, Cunha AL, Messias CM. **Imunização: o conhecimento e práticas da equipe de enfermagem na sala de vacina**. Nursing. Janeiro 2020; 23 (260): 3533-3536. ID: biblio-1095551
11. Ministério da Saúde do Brasil. **Aplicativo Vacinação em Dia**. Homepage <http://www.blog.saude.gov.br/promocao-da-saude/32865-como-baixar-o-aplicativo-vacinacao-em-dia>
12. Centers for Disease Control and Prevention. **Recommended immunization schedules for persons aged 0--18 years**---United States, 2011. MMWR, 60(5), (2011) Homepage, <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6005a6.htm>,
13. World Health Organization. Health topics. **New app will help parents keep track of their children's vaccinations**. Homepage, [http://www.who.int/mediacentre/news/releases/2013/world\\_health\\_assembly\\_20130527/en/](http://www.who.int/mediacentre/news/releases/2013/world_health_assembly_20130527/en/)
14. Bechini A, Moscadelli A, Sartor G, Shtylla J, Guelfi MR, Bonanni P, Boccacini S. **Impact assessment of an educational course on vaccinations in a population of medical Students**. J Prev Med Hyg. September 2019; 60 (3): 171-177. <https://doi.org/10.15167/2421-4248/jpmh2019.60.3.1201>
15. Costantino C, Amodio E, Calamusa G, Vitale F, Mazzucco W. **Could university training and a proactive attitude of coworkers be associated with influenza vaccination compliance? A multicentre survey among Italian medical residents**. BMC Med Educ. January 2016; 16 (38): <https://doi.org/10.1186/s12909-016-0558-8>.
16. Kumanan W, Atkinson KM, Westeinde J. **Apps for immunization: Leveraging mobile devices to place the individual at the center of care**. Hum Vaccin Immunother. May 2015, 11 (10): 2395-2399. <https://doi.org/10.1080/21645515.2015.1057362>
17. Amicizia D, Domnich A, Gasparini R, Bragazzi NL, Lai PL, Panatto D. **An overview of current and potential use of information and communication technologies for immunization promotion among adolescents**. Hum Vaccin Immunother. December 2013; 9 (12): 2634-2642. <https://doi.org/10.4161/hv.26010>



18. Kumanan W, Atkinson K, Crowcroft N. **Teaching children about immunization in a digital age.** Hum Vaccin Immunother. May 2017; 13 (5): 1155-1157. <https://doi.org/10.1080/21645515.2016.1271519>
19. Kumanan W, Atkinson KM, Westeind J, Bell C, Marty K, Fergusson D, Deeks SL, Crowcroft N, Bettinger JA. **An evaluation of the feasibility and usability of a proof of concept mobile app for adverse event reporting post influenza vaccination.** Hum Vaccin Immunother. July 2016; 12 (7): 1738-1748. <https://doi.org/10.1080/21645515.2016.1152434>
20. Montagni I, Mabchou I, Tzourio C. **Digital Gamification to Enhance Vaccine Knowledge and Uptake: Scoping Review.** JMIR Serious Games. May 2020; 8(2): <https://doi.org/10.2196/16983>.
21. Burgess, Atkinson KM, Westeinde J, Crowcroft N, Deeks SL, **Kumanan W. Barriers and facilitators to the use of an immunization application: a qualitative study supplemented with Google Analytics data.** J Public Health (Oxf). September 2017; 39 (3): 118-126. <https://doi.org/10.1093/pubmed/fdw032>.
22. Schelling Jörg , Isabella Thorvaldsson, Linda Sanftenberg. **[Digital vaccination management systems may improve immunization rates in primary healthcare].** Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. April 2019; 62 (4): 433-439. <https://doi.org/10.1007/s00103-019-02912-2>
23. Handy LK, Maroudi S, Powel M, Bakanuki N, Moser C, Japa I, Monyatsi N, Tzortzi E, Kouzeli I, Luberti A, Theodoridou M, Offit P, Steenhoff A, Shea JA, Feemster KA. **The impact of access to immunization information on vaccine acceptance in three countries.** PLoS One. August 2017; 12 (8): <https://doi.org/10.1371/journal.pone.0180759>
24. Bednarczyk RA, Frew PM, Salmon DA, Whitney E, and Omer SB. ReadyVax: **A new mobile vaccine information app.** Hum Vaccin Immunother. May 2017; 13 (5): 1149–1154. <https://doi.org/10.1080/21645515.2016.1263779>
25. Minoletti S, Rapisarda R, Giraldo L, Grande M, Sommer J, Plazzotta F, Luna D. **User-Centered Design of a Pediatric Vaccination Module for Patients.** Stud Health Technol Inform. August 2019; (264): 1096-1100. <https://doi.org/10.3233/SHT1190395>.
26. Atkinson KM, El-Khatib Z, Barnum G, Bell C, Turcotte MC, Murphy MSQ, Teitelbaum M, Chakraborty P, Laflamme L, Kumanan W. **Using Mobile Apps to Communicate Vaccination Records: A City-wide Evaluation with a National Immunization App.** Maternal Child Registry and Public Health Authorities. Healthc Q. 2017; 20 (3):41-46. <https://doi.org/10.12927/hcq.2017.25289>.
27. Kumanan W, Atkinson KM, Deeks SL, Crowcroft NS. **Improving vaccine registries through mobile technologies: a vision for mobile enhanced immunization information systems.** J Am Med Inform Assoc. January 2016; 23 (1): 207-211. <https://doi.org/10.1093/jamia/ocv055>
28. Jacobson JCV, Jacobson RM, Coyne-Beasley T, Asafu-Adjei JK, Szilagyi PG. **Patient reminder and recall interventions to improve immunization rates.** Cochrane Database Syst Ver. January 2018; 1 (1): <https://doi.org/10.1002/14651858.CD003941.pub3>.
29. Corriero R, Redmon G. **Planning and Implementing Immunization Billing Programs at State and Local Health Departments: Barriers and Possible Solutions.** J Public Health Manag Pract. May/June 2018; 24 (3): 248-254. <https://doi.org/10.1097/PHH.0000000000000676>

30. Rath B. Vienna Vaccine Safety Initiative. **Hum Vaccin Immunother**. January 2018; 14 (5):1038-1041. <https://doi.org/10.1080/21645515.2017.1412666>
31. Kumanan W, Atkinson K, Deeks S. **Opportunities for utilizing new technologies to increase vaccine confidence**. *Expert Rev Vaccines*. June 2014; 13 (8): 969-977. <https://doi.org/10.1586/14760584.2014.928208>
32. Ministério da Saúde do Brasil. **Manual dos Centros de Referência para Imunobiológicos Especiais**. Brasília, 2019. Homepage, <https://portalarquivos2.saude.gov.br/images/pdf/2019/dezembro/11/manual-centros-referencia-imunobiologicos-especiais-5ed.pdf> ISBN 978-85-334-2746-4,
33. Ministério da Saúde do Brasil. Manual de Vigilância Epidemiológica de Eventos Adversos Pós-Vacinação. Brasília, 2014. Homepage, <http://vigilancia.saude.mg.gov.br/index.php/download/manual-de-vigilancia-epidemiologica-de-eventos-adversos-pos-vacinacao-2014/?wpdmdl=5343>, ISBN 978-85-334-2175-253
34. Sociedade Brasileira de Imunizações (SBIm). Homepage, <https://sbim.org.br/informes-e-notas-tecnicas/sbim>
35. Tolks D, Lampert C, Dadaczynski K, Maslon E, Paulus P, Sailer M. **[Game-based approaches to prevention and health promotion: serious games and gamification]**. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. June 2020; 63 (6): 698-707. <https://doi.org/10.1007/s00103-020-03156-1>
36. Cates JR, Fuemmeler BF, Diehl SJ, Stockton LL, Porter J, Ihekweazu C, Gurbani AS, Coyne-Beasley T. **Developing a Serious Videogame for Preteens to Motivate HPV Vaccination Decision Making: Land of Secret Gardens**. *Games Health J*. February 2018; 7 (1): 51-56. <https://doi.org/10.1089/g4h.2017.0002>
37. Ohannessian R, Yaghobian S, Verger P, Vanhems P. **A systematic review of serious video games used for vaccination**. *Vaccine*. August 2016; 34 (38): 4478-4483. <https://doi.org/10.1016/j.vaccine.2016.07.048>

## UMA ANÁLISE DE VANTAJOSIDADE EM MODELOS DE PREVISÃO EM SÉRIES TEMPORAIS

Data de aceite: 01/04/2021

### Rafael Diniz Toscano de Lima

Universidade de Pernambuco, Escola  
Politécnica - POLI  
Recife - Pernambuco  
<http://lattes.cnpq.br/7679503812101994>

### Sérgio Murilo Maciel Fernandes

Universidade de Pernambuco, Escola  
Politécnica - POLI  
Recife - Pernambuco  
<http://lattes.cnpq.br/4520293519781462>

### Sidney Marlon Lopes de Lima

Universidade Federal de Pernambuco,  
Departamento de Eletrônica e Sistemas  
Recife - Pernambuco  
<http://lattes.cnpq.br/0323190806293435>

### Ricardo Paranhos Pinheiro

Universidade de Pernambuco, Escola  
Politécnica - POLI  
Recife - Pernambuco  
<http://lattes.cnpq.br/9889280450483185>

### Sthéfano Henrique Mendes Tavares Silva

Universidade de Pernambuco, Escola  
Politécnica - POLI  
Recife - Pernambuco  
<http://lattes.cnpq.br/8682401562704469>

**RESUMO:** A modelagem e previsão em séries temporais nem sempre é um processo direto. Existem inúmeras técnicas para construir modelos que podem modelar, estimar e prever momentos

futuros. Posteriormente, esses modelos podem ajudar a tomar decisões calculadas que, por sua vez, podem mitigar os riscos e maximizar os retornos. Além disso, a construção de modelos de previsão confiáveis e robustos é essencial para prever o comportamento dos dados, como por exemplo, movimentos do mercado financeiro, geração de energia e precipitação de chuvas. Dessa forma, o presente artigo tem por objetivo a análise das vantagens e desvantagens dos modelos mais utilizados na literatura, suportando pesquisas futuras no tema.

**PALAVRAS - CHAVE:** Séries temporais, previsão, modelos, forecasting.

**ABSTRACT:** Modeling and forecasting time series is not always a straightforward process. There are numerous techniques for building models that can model, estimate, and predict future moments. Subsequently, these models can help you make calculated decisions that, in turn, can mitigate risks and maximize returns. In addition, the construction of reliable and robust forecasting models is essential to predict the behavior of the data, such as financial market movements, power generation, and rainfall. Thus, this article aims to analyze the advantages and disadvantages of the most used models in the literature, supporting future research on the topic.

**KEYWORDS:** Time series, prediction, models, forecasting.

## INTRODUÇÃO

Uma série temporal é um conjunto de observações ordenadas que apresentam

dependência entre instantes de tempo. Os métodos para predição de séries temporais são baseados essencialmente na ideia de que dados históricos apresentam padrões intrínsecos, que geralmente são de difícil identificação.

Uma vez que os padrões são descritos e interpretados, estes podem auxiliar na previsão de valores futuros de fenômenos investigados. Essa descrição constitui um dos principais objetivos do processamento de séries temporais, pois visa responder em que circunstâncias os padrões encontrados irão se repetir e quais os principais tipos de variação que os mesmos poderão sofrer no decorrer do tempo .

De acordo com o intervalo de tempo em que os dados são adquiridos, as séries temporais podem ser categorizadas em contínua e discreta. As contínuas são aquelas onde as observações dos dados são realizadas de maneira contínua sobre um intervalo de tempo específico. Já as discretas, ocorrem quando as observações dos dados são convertidas em intervalos fixos de tempo e geralmente são espaçadas igualmente . Os principais elementos das séries temporais são:

- **Tendência:** A tendência pode ser definida como o movimento regular desenvolvido ao longo da série. Esse elemento é referente ao comportamento durante toda a duração, podendo ser crescente ou decrescente, além de assumir variação de padrões de crescimento: linear, exponencial, amortecido ;
- **Sazonalidade:** É referente a repetição de um determinado comportamento durante diferentes períodos de tempo dentro da série temporal. As variações sazonais podem ser observadas pela oscilação ao longo do tempo da tendência segundo um determinado fator ;
- **Resíduo:** São representados por movimentos aleatórios causados por fatores pontuais e inesperados. Um grande exemplo de resíduo foram as interferências causadas pela pandemia decorrente da COVID-19, impactando em várias observações ao longo do tempo, tais quais: nos número de mortes por doenças respiratórias, nos resultados financeiros das empresas, nas lotações de centros hospitalares, etc. Tais externalidades são ao mesmo tempo incontroláveis e inesperadas nas séries temporais em comento. Uma vez que esses tipos de acontecimento não são regulares e não se repetem em um padrão, eles podem comprometer o resultado de alguns estudos das séries temporais.

Sendo assim, a identificação do componente residual é fundamental tanto para sua remoção quando oportuno for, ou então para verificação das variações cíclicas que podem ocorrer durante esse evento .

A análise de séries temporais podem ser realizadas com base em diversos objetivos, que podem ser segmentados em quatro grupos :

- **Descrição:** Consiste em descrever o comportamento da série temporal, como a existência ou não de tendências, sazonalidade e observações dos resíduos;
- **Explicação:** É referente a identificação de duas ou mais variáveis para expli-

cação da variação de uma série com base em outra;

- **Previsão:** Tem como objetivo prever valores futuros a partir dos valores passados, caracterizando o comportamento da série, para esse grupo é realizada uma análise de toda a série é utilizado uma janela temporal de dados passados para prever o valor futuro. Por exemplo, é analisado os últimos 20 dias para prever os próximos 2 dias;
- **Controle:** Seu principal objetivo é mensurar a qualidade dos dados do processo.

## PREVISÃO DE SÉRIES TEMPORAIS

*Forecasting* é a previsão de ocorrência de um evento futuro, dadas as observações de uma série de eventos numa escala temporal . O principal objetivo na previsão da ocorrência desse evento vindouro é para que os tomadores de decisão possam fazer melhores escolhas. Os autores em Wheelwright et al. (1998), elucidam que existem duas abordagens principais para a previsão, quais sejam: as séries explicativas (causais) e temporais .

A previsão explicativa assume uma relação de causa e efeito entre as entradas e saídas. De acordo com a previsão explicativa, alterar as entradas afetará a saída do sistema de uma maneira previsível, assumindo que a relação de causa e efeito é constante.

Ao contrário da previsão explicativa, a previsão de séries temporais trata o sistema como uma “caixa preta” e procura descobrir os fatores que afetam seu comportamento. Há duas importantes motivações para querer tratar um sistema como uma caixa preta. Primeiramente, o sistema pode não ser compreendido e, mesmo que fosse entendido, pode ser extremamente difícil medir as relações que governam seu comportamento. Em segundo lugar, a principal preocupação pode ser apenas a previsão em si, e não o porquê do fenômeno observado.

Por ter ampla aplicabilidade, a previsão de séries temporais é utilizada em diversas áreas de conhecimento, por exemplo: aviação, para prever a demanda por viagens ; em medicina, para prever a quantidade de óbitos por determinada doença , ; e em economia, para estimar os valores de ações no mercado financeiro.

## MODELAGEM E PREVISÃO DE SÉRIES TEMPORAIS

A modelagem e previsão em séries temporais nem sempre é um processo direto. Existem inúmeras técnicas para construir modelos que podem modelar, estimar e prever momentos futuros.

Posteriormente, esses modelos podem ajudar a tomar decisões calculadas que, por sua vez, podem mitigar os riscos e maximizar os retornos. Além disso, a construção

de modelos de previsão confiáveis e robustos é essencial para prever o comportamento dos dados, como por exemplo, movimentos do mercado financeiro, geração de energia e precipitação de chuvas.

## MODELOS DE REGRESSÃO

Há muitos problemas que exigem a investigação de relacionamentos entre duas ou mais variáveis. A análise de regressão é um método estatístico típico, que há muito é usado para a modelagem e solução desta classe de problema. O objetivo da análise de regressão é estimar as dependências entre a variável principal e um conjunto de fatores externos aos regressores .

O modelo de regressão linear é um modelo simples e largamente utilizado. Assume-se que existe um conjunto de fatores externos  $X_1(t), X_2(t), \dots, X_p(t)$ , que têm um impacto no processo e a relação entre eles é linear .

Os autores em Halimi et al. (2011), afirmam que os modelos de regressão não-linear são baseados em suposições , que são dadas a uma função matemática, que descreve a relação entre o processo e o fator externo. No entanto, na prática, não é muito comum esse tipo de dependência funcional entre o processo e o fator externo. Portanto, os modelos de regressão não-linear são usados com menos frequência do que os lineares.

Esses modelos têm como principais vantagens a simplicidade, flexibilidade e uniformidade de cálculos, simplicidade de construção de modelos, transparência dos cálculos intermediários. E como desvantagens a ineficiência em situações com externalidades ao modelo e baixa adaptabilidade de modelos de regressão linear para processos não lineares.

## MODELOS DE SUAVIZAÇÃO EXPONENCIAL

Ainda no início da década de 90, Barakat et al. (1990) provou a dificuldade em analisar e prever os dados aplicando unicamente o método de Holt-Winters, pelo fato dos conjuntos de dados geralmente conterem valores esparsos . Para superar o problema da esparsidade, nesta mesma década, El-Keib et al. (1995), propôs uma estratégia combinada onde o método de suavização exponencial foi combinado com outros métodos, como o auto-regressivo e modelos de análise de espectro .

Apesar do fato de que os métodos de suavização exponencial foram inventados em meados do século 20, eles ainda são frequentemente usados no enfrentamento de diversos problemas até a atualidade, como pode-se atestar nos trabalhos de Goodwin et al. (2010), nas variações e adaptações do modelo clássico de Holt-Winters , ou ainda como os modelos de suavização exponencial de Sasongko et al. (2017), específicos para a criação e modelagem financeiras e de processos econômicos .

A base da suavização exponencial é uma ideia de revisão repetitiva da função de previsão, sendo comparada ao valor recentemente observado. O modelo de suavização exponencial atribui pesos exponencialmente decrescentes a valores anteriores, de acordo com a antiguidade .

A utilização da técnica de suavização exponencial trás de vantagens a transparência dos cálculos intermediários, simplicidade e eficácia relativa, modelos de fácil construção. Porém, assim como os modelos de regressão, sofrem com a baixa adaptabilidade.

## MODELOS AUTOREGRESSIVOS DE MÉDIAS MÓVEIS

Os modelos autorregressivos baseiam-se na ideia de que os valores do processo são linearmente dependentes de algum número de valores passados do mesmo processo. Segundo Dimitrios e Hall (2011), nesse modelo, o valor real do processo é expresso como uma soma da combinação linear finita de valores anteriores e dos impulsos, chamada ruído branco . Este modelo é frequentemente denotado como  $AR(p)$ . Os parâmetros  $c$  e  $\varphi_i$  são geralmente estimados pelos métodos de mínimos quadrados ou máxima verossimilhança. O modelo de média móvel, desempenha um papel muito importante na descrição das séries temporais e é frequentemente utilizado em relação aos modelos auto-regressivos.

Na literatura, o modelo de média móvel de ordem  $q$  é geralmente denotado como  $MA(q)$ . Na verdade, o modelo de média móvel é um filtro de resposta de impulso finito aplicado ao ruído branco. Como observado nos trabalhos de previsão de índices de Wang et al. (2012), para a obtenção de uma melhor qualidade de previsão, os dois modelos anteriores são frequentemente mesclados em um modelo híbrido, autorregressivo e modelo de média móvel .

O modelo comum é denotado como  $ARMA(p, q)$  e une um filtro de média móvel de ordem  $q$  e autorregressão de valores filtrados de ordem  $p$ . Se os dados da série temporal mostrarem evidência de não-estacionariedade, a etapa de diferenciação inicial poderá ser aplicada para reduzir a não-estacionariedade. Este modelo é geralmente denotado como  $ARIMA(p, d, q)$ . O parâmetro  $d$  representa o grau de diferenciação, corresponde à parte integrada do modelo.

Médias Móveis Integradas Autoregressivas (do inglês, *Autoregressive Integrated Moving Average* - ARIMA) são classes de algoritmos aplicáveis para cenários onde o tempo é uma variante imprevisível. O modelo usa os dados atuais combinados com valores coletados anteriormente, juntamente com os erros enfrentados nas previsões passadas para prever valores autênticos .

O modelo ARIMA é capaz de corrigir a inconsistência e os preconceitos nas estimativas de regressão. Inconsistências geralmente surgem quando as variáveis dependentes atrasadas são selecionadas como regressoras. Este modelo é considerado superior a outras regressões multivariadas comuns e análises de séries temporais, isto

porque leva em consideração o erro residual e prevê o erro residual atual com a ajuda da computação dos erros prévios .

## MODELOS ESTOCÁSTICOS

Alguns dados possuem características ruidosas, bem como múltiplas dimensionalidades. Assim, a maioria dos métodos acima mencionados são comprovadamente inconsistentes e imprevisíveis em termos de desempenho da previsão. Isto é devido à presença de processos não-lineares dentro dos modelos de previsão de séries temporais, um exemplo desses dados são as séries financeiras.

Para superar esses problemas, Vapnik (1999), desenvolveu um modelo de previsão de séries temporais financeiras utilizando um algoritmo de aprendizado de máquina de vetor de suporte ( do inglês, *Support Vector Machines* - SVM) que utiliza soluções lineares para resolver esses problemas .

As pesquisas do SVM mostraram que esses modelos superam as questões existentes no plano não-linear de métodos baseado em dados. O desempenho do SVM depende dos vários recursos de entrada dos dados que são levados para o processamento dentro dos modelos. Combinando os modelos de regressão SVM com outros parâmetros e técnicas de otimização, como o Algoritmo Genético (AG), a previsão pode ser melhorada.

Em An et al. (2007), os autores apresentam um modelo combinatório entre o SVM com AG para previsão de séries temporais financeiras . Um método muito semelhante foi proposto por Huang e Wang (2006), que envolve processos de otimização simultânea, onde comprovou a eficiência de classificação em SVM .

Em sequência, Huang propôs um tipo semelhante de modelo híbrido para seleção e análise utilizando o AG e regressão vetorial de suporte (do inglês, *Support Vector Regression* - SVR).

O SVR é usado para gerar as previsões dos resultados reais do estoque com base no ranking. As ações altamente cotadas são finalmente tomadas por construção de portfólio. A paralelização do parâmetro otimização e modelagem são feitas usando AG.

Já Kumar et al. (2014), desenvolveu modelos híbridos para prever retornos de índices de ações, combinando modelos lineares ARIMA e não-lineares, SVM, redes neurais artificiais e *Random Forest* (RF). A análise demonstrou que o modelo híbrido ARIMA-SVM foi o melhor modelo de previsão para alcançar alta precisão de previsão e melhores retornos .

## Modelos de Redes Neurais Artificiais

Redes Neurais Artificiais (RNA) são ferramentas que, atualmente, estão sendo usadas para resolver uma enorme quantidade de tarefas de diferentes áreas. Segundo Sze et al. (2017), a maioria dos exemplos frequentes são a previsão de séries temporais, reconhecimento de padrões, *clustering* de dados e classificação .



Os autores em Zhang et al. (2015), forneceram uma investigação sobre como modelar a predição de séries temporais de dados de tendências e sazonais com vários tipos de RNAs . Eles usaram técnicas de redes neurais para modelar previsões trimestrais de séries temporais para padrões de tendência e sazonalidade. Em Hamzacebi (2008), o autor propôs um novo tipo de modelo RNA para melhorar o desempenho da predição de séries temporais sazonais .

Uma coletânea das técnicas híbridas de RNA usadas para predição de séries temporais foi organizada por Taskaya-Temizel and Casey (2005), para a otimização dos modelos de RNA, para construção desta coletânea foram utilizados algoritmos de busca aleatória e algoritmos de busca gradiente, de acordo com as especificidades dos dados das séries temporais estudadas .

Kim e Han (2000), idealizaram um método baseado em AG para discretizar as características do estoque e descobrir os pesos de conexão no modelo de RNA para prever o índice do mercado de ações . Muitas outras pesquisas e modificações foram feitas nos modelos baseados no AG que são usados para classificar os ativos com base nas funções de adequação pré-definidas. Os autores em Egriglu et al. (2013), propuseram um método híbrido combinando a RNA ao Fuzzy C-Means (FCM), para previsão de séries temporais difusas .

A principal vantagem desses modelos é a não linearidade. As redes neurais podem lidar facilmente com as dependências não lineares entre valores futuros e passados dos processos, tem-se uma grande adaptabilidade, escalabilidade e paralelização dos modelos. Além da possibilidade de treinamento sem o conhecimento do problema em si, necessitando apenas de uma boa base de dados. Porém, como desvantagens têm-se os altos requisitos de desempenho de *hardware* ou ainda de tempo durante o processo de treinamento da rede. Complexidade da arquitetura e ausência de transparência dos cálculos intermediários.

## MODELOS BASEADOS EM CADEIA DE MARKOV

Modelos de previsão baseados nas cadeias de Markov assumem que o estado futuro do processo depende apenas do seu estado atual e não depende de seus estados mais antigos. Os modelos de cadeia de Markov são aplicáveis nas séries temporais de memória curta (*Long-short Memory*).

Como explicam os autores em Hochreiter and Schmidhuber (1997), que propôs a arquitetura LSTM que em sua arquitetura leva em consideração os períodos da série temporal e utiliza como parâmetro em sua forma de ligação nos neurônios em sua topologia da rede neural .

Ao construir o modelo da cadeia de Markov, o conjunto de estados e as probabilidades de transições correspondentes são definidos. Se o estado atual do processo for definido, o

estado futuro é selecionado como o estado com probabilidade de transição máxima.

Em seus estudos, Liu (2010), demonstrou que se as probabilidades de transição forem adequadamente armazenadas na matriz, os valores futuros subsequentes podem ser determinados pela multiplicação da matriz de probabilidade e pela seleção máxima de probabilidade. Esse tipo de técnica tem como principal desvantagem a impossibilidade da obtenção de uma modelagem e previsão de longo prazo.

## ANÁLISE DOS MODELOS

Todos esses trabalhos motivam a investigação de métodos eficazes na modelagem e predição de séries temporais, bem como fornecem informações que mostram a relevância do principal objetivo da pesquisa. Assim, é natural que existam vantagens e desvantagens associadas a cada modelo e técnica em particular. A Tabela 1 traça um resumo analítico sobre estes fatores.

Modelo de Previsão	Vantagens	Desvantagens
Regressão	Simplicidade, flexibilidade e uniformidade de cálculos. Simplicidade de construção de modelos. Transparência dos cálculos intermediários	Ineficiência em situações com externalidades ao modelo e baixa adaptabilidade de modelos de regressão linear para processos não lineares
Suavização Exponencial	Transparência dos cálculos intermediários, simplicidade e eficácia relativa. Modelos de fácil construção	Inflexibilidade semelhante aos modelos de regressão
ARIMA e Estocásticos	Transparência e uniformidade de cálculos e construção de modelos. Construção de modelo relativamente simples. Grande capacidade de hibridização com os demais modelos	Possibilidade de grande número de parâmetros requeridos para serem configurados. Limitações de adaptação na resolução de processos não lineares
RNA	Grande adaptabilidade e escalabilidade. Habilidade de cálculos paralelos. Treinamento sem o conhecimento do problema em si, necessitando apenas de uma boa base de dados	Grande número de parâmetros. Altos requisitos de desempenho de hardware ou ainda de tempo durante o processo de treinamento de rede. Complexidade da arquitetura e ausência de transparência dos cálculos intermediários
Cadeia de Markov	Transparência dos cálculos e passos realizados para a solução do problema	Impossibilidade da obtenção de uma modelagem e previsão de longo prazo

Tabela 1. Análise das vantagens e desvantagens dos modelos estudados.

## REFERÊNCIAS

- [Aladag et al. 2012] Aladag, C. H., Yolcu, U., Egrioglu, E., and Dalar, A. Z. (2012). A new time invariant fuzzy time series forecasting method based on particle swarm optimization. *Applied Soft Computing*, 12(10):3291–3299.
- [An et al. 2007] An, S., Liu, W., and Venkatesh, S. (2007). Fast cross-validation algorithms for least squares support vector machine and kernel ridge regression. *Pattern Recognition*, 40(8):2154–2162.
- [Asteriou and Hall 2011] Asteriou, D. and Hall, S. G. (2011). Arima models and the box–jenkins methodology. *Applied Econometrics*, 2(2):265–286.
- [Barakat et al. 1990] Barakat, E., Qayyum, M., Hamed, M., and Al Rashed, S. (1990). Short-term peak demand forecasting in fast developing utility with inherit dynamic load characteristics. i. application of classical time-series methods. ii. improved modelling of system dynamic load characteristics. *IEEE Transactions on Power Systems*, 5(3):813–824.
- [Botchkarev 2018] Botchkarev, A. (2018). Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. *arXiv preprint arXiv:1809.03006*.
- [Brocklebank and Dickey 2003] Brocklebank, J. C. and Dickey, D. A. (2003). *SAS for forecasting time series*. John Wiley & Sons.
- [Brockwell et al. 2002] Brockwell, P. J., Davis, R. A., and Calder, M. V. (2002). *Introduction to time series and forecasting, volume 2*. Springer.
- [Chatfield 2013] Chatfield, C. (2013). *The analysis of time series: theory and practice*. Springer.
- [Chen et al. 2009] Chen, C.-F., Chang, Y.-H., and Chang, Y.-W. (2009). Seasonal arima forecasting of inbound air travel arrivals to taiwan. *Transportmetrica*, 5(2):125–140.
- [de Alencar 2020] de Alencar, V. C. (2020). *Uso de Técnicas de Data Science na Previsão: de Febre Amarela Utilizando o Twitter*. Editora Appris.
- [de Lima et al. 2019] de Lima, R. D. T., Fernandes, S. M. M., and Melo, I. P. L. (2019). Fourier genetic series: An evolutionary time series modeling technique. In *2019 IEEE Latin American Conference on Computational Intelligence (LA-CCI)*, pages 1–6. IEEE.
- [Dolacio et al. 2017] Dolacio, T. A. et al. (2017). Modelagem na predição de risco de infestação de *aedes aegypti* (diptera, culicidae) e subsídios governamentais na tomada de decisão para vigilância e monitoramento.
- [Egrioglu et al. 2013] Egrioglu, E., Aladag, C. H., and Yolcu, U. (2013). Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks. *Expert Systems with Applications*, 40(3):854–857.
- [Ehlers 2007] Ehlers, R. S. (2007). *Análise de séries temporais*. Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná .

- [El-Keib et al. 1995] El-Keib, A., Ma, X., and Ma, H. (1995). Advancement of statistical based modeling techniques for short-term load forecasting. *Electric Power Systems Research*, 35(1):51–58.
- [Goodwin et al. 2010] Goodwin, P. et al. (2010). The holt-winters approach to exponential smoothing: 50 years old and going strong. *Foresight*, 19(19):30–33.
- [Halimi et al. 2011] Halimi, A., Altmann, Y., Dobigeon, N., and Tourneret, J.-Y. (2011). Nonlinear unmixing of hyperspectral images using a generalized bilinear model. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4153–4162.
- [Hamzac,ebi 2008] Hamzacebi, C. (2008). Improving artificial neural networks performance in seasonal time series forecasting. *Information Sciences*, 178(23):4550–4559.
- [Hochreiter and Schmidhuber 1997] Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.
- [Huang and Wang 2006] Huang, C.-L. and Wang, C.-J. (2006). A ga-based feature selection and parameters optimization for support vector machines. *Expert Systems with applications*, 31(2):231–240.
- [Hyndman and Koehler 2006] Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4):679–688.
- [Ishaque and Ziblim 2013] Ishaque, M. and Ziblim, S. (2013). Use of some exponential smoothing models in forecasting some food crop prices in the upper east region of ghana. *Mathematical Theory and Modeling*, 3(7):16–27.
- [Kim and Han 2000] Kim, K.-j. and Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert systems with Applications*, 19(2):125–132.
- [Kirchg' assner and Wolters 2007] Kirchgassner, G. and Wolters, J. (2007). *Introduction to modern time series analysis*. Springer Science & Business Media.
- [Kumar and Thenmozhi 2014] Kumar, M. and Thenmozhi, M. (2014). Forecasting stock index returns using arima-svm, arima-ann, and arima-random forest hybrid models. *International Journal of Banking, Accounting and Finance*, 5(3):284–308.
- [Liu 2010] Liu, T. (2010). Application of markov chains to analyze and predict the time series. *Modern Applied Science*, 4(5):162.
- [Malska and Wachta 2015] Malska, W. and Wachta, H. (2015). Wykorzystanie modelu arima do analizy szeregu czasowego. *Zeszyty Naukowe Politechniki Rzeszowskiej. Elektrotechnika*, (34 [292], nr 3):23–30.
- [Montgomery et al. 2012] Montgomery, D. C., Peck, E. A., and Vining, G. G. (2012). *Introduction to linear regression analysis*, volume 821. John Wiley & Sons.
- [Peng et al. 2003] Peng, H., Ozaki, T., Haggan-Ozaki, V., and Toyoda, Y. (2003). A parameter optimization method for radial basis function type models. *IEEE Transactions on neural networks*, 14(2):432–438.

[Ravichandran et al. 2005] Ravichandran, K., Thirunavukarasu, P., Nallaswamy, R., and Babu, R. (2005). Estimation of return on investment in share market through ann. *Journal of Theoretical and Applied Information Technology*, 3.

[Sasongko et al. 2017] Sasongko, R. S., Prasetyo, E., and Purbaningtyas, R. (2017). System prediction production pt. vico indonesia using method holt winters. *JOURNAL OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCES*, VOL 2 NUMBER 1, JUNE 2017, 2(1).

[Sorjamaa et al. 2007] Sorjamaa, A., Hao, J., Reyhani, N., Ji, Y., and Lendasse, A. (2007). Methodology for long-term prediction of time series. *Neurocomputing*, 70(16-18):2861–2869.

[Sze et al. 2017] Sze, V., Chen, Y.-H., Yang, T.-J., and Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE*, 105(12):2295–2329.

[Taskaya-Temizel and Casey 2005] Taskaya-Temizel, T. and Casey, M. C. (2005). A comparative study of autoregressive neural network hybrids. *Neural Networks*, 18(5-6):781–789.

[Vapnik 1999] Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE transactions on neural networks*, 10(5):988–999.

[Wang et al. 2012] Wang, J.-J., Wang, J.-Z., Zhang, Z.-G., and Guo, S.-P. (2012). Stock index forecasting based on a hybrid model. *Omega*, 40(6):758–766.

[Wheelwright et al. 1998] Wheelwright, S., Makridakis, S., and Hyndman, R. J. (1998). *Forecasting: methods and applications*. John Wiley & Sons.

[Zhang and Qi 2005] Zhang, G. P. and Qi, M. (2005). Neural network forecasting for seasonal and trend time series. *European journal of operational research*, 160(2):501–514.

# CAPÍTULO 3

## SENTECH: UM COMBINADOR DE ANÁLISE TÉCNICA E DE SENTIMENTO PARA O MERCADO DE AÇÕES

*Data de aceite: 01/04/2021*

*Data de submissão: 14/01/2021*

### **Isabela Nunes Caetano**

Universidade Tecnológica Federal do Paraná  
Cornélio Procópio – Paraná  
<http://lattes.cnpq.br/6454612302361407>

### **Érica Ferreira de Souza**

Universidade Tecnológica Federal do Paraná  
Cornélio Procópio – Paraná  
<http://lattes.cnpq.br/8904855809524041>

### **Giovani Volnei Meinerz**

Universidade Tecnológica Federal do Paraná  
Cornélio Procópio – Paraná  
<http://lattes.cnpq.br/0746723713646857>

**RESUMO:** O mercado financeiro apresenta extrema importância para as relações econômicas modernas. Além de se consolidar com o passar do tempo, cresce o número de pessoas que se tornam investidores. Para tanto, é necessário estratégias que auxiliem o processo de tomada de decisões. Para resolver este problema da previsão do comportamento das ações, encontram-se propostas que se separam por séculos. A primeira, baseada em indicadores matemáticos e a segunda baseada nos sentimentos identificados em texto. A fusão das técnicas já foi explorada em panorama internacional, entretanto, o mesmo não acontece em âmbito nacional. Visto isso, este trabalho apresenta o SENTECH, um modelo preditivo que combina as duas abordagens técnicas

citadas, desenvolvido com objetivo de identificar a correlação entre o valor predito e o preço de fechamento das ações listadas na bolsa de valores brasileira. Como resultado da tarefa de classificação de dados, tem-se resultados de correlação moderados, com alta significância (95%), para duas empresas da carteira teórica do índice IBOVESPA.

**PALAVRAS - CHAVE:** Bolsa de valores. Twitter. Inteligência Artificial. Aprendizado do Computador. Correlação.

### SENTECH: A STOCK MARKET SENTIMENT AND TECHNICAL ANALYSIS COMBINER

**ABSTRACT:** The financial market is extremely important for modern economic relations. In addition to consolidating itself over time, the number of people who become investors grows. For this purpose, strategies are needed to assist the decision-making process. To solve this problem of forecasting the behavior of actions, there are proposals that are separated for centuries. The first based on mathematical indicators and the second based on the feelings identified in text. The fusion of techniques has already been explored on the international scene, however, the same does not happen at the national level. Having said that, this paper presents the SENTECH, a predictive model combining the two techniques mentioned, developed with the objective of identifying the correlation between the predicted value and the closing price of the stocks listed on the brazilian stock exchange. As a result of the data classification task, there are moderate correlation results, with high

significance (95%), for two companies listed in the IBOVESPA index.

**KEYWORDS:** Stock markets. Twitter. Artificial Intelligence. Machine Learning. Correlation.

## 1 | INTRODUÇÃO

A expansão da digitalização acelerou a crescente esfera de dados global atingindo a produção de Zettabytes (ZB) a partir de 2010, já estimada para 50 ZB no ano corrente (REINSEL, GANTZ e RYDNING, 2018).

Consequência de tamanho volume será devido a uma maior interação dos usuários, com foco para as redes sociais, cenário dos dados que descrevem os comportamentos do público, já estimado em 140 milhões (STATISTA, 2020).

Nesse passo, um dos métodos para extrair informação é a **Análise de Sentimentos (AS)**, que tem como objetivo classificar a emoção e o humor denotados em texto. A referida técnica tem sido amplamente utilizada como estratégia para descrever o comportamento do mercado de ações. Isso, com maior assertividade no curto prazo e, possivelmente, menor assertividade no longo prazo (SHAH, ISAH e ZULKERNINE, 2019). Assim, BOLLEN, MAO e ZENG (2011) foram os precursores na utilização do Twitter como fonte de dados para a AS.

Em outro viés, para a mesma finalidade de acompanhamento do mercado financeiro, tem-se uma abordagem atemporal, a **Análise Técnica (AT)**, comumente adotada por investidores e que trata de dados quantitativos, a fim de buscar padrões recorrentes e previsíveis nos preços das ações. Ademais, a filosofia técnica, pressupõe que o preço é influenciado pela psicologia do investidor tanto quanto pelas variáveis financeiras (DAMODARAN, 2010).

Na literatura, ATTANASIO et al. (2019) se destacou ao explorar, em conjunto, as estratégias de AT e AS para identificação da reversão de tendência, que obtiveram maior sucesso do que em relação as estratégias analisadas individualmente. O seu trabalho se assimila com o executado por PICASSO et al. (2019), que se diferencia pelo intuito de prever o preço de fechamento das ações. No entanto, apesar dos trabalhos citados explorarem a combinação das técnicas, não há registros desta aplicação no âmbito nacional.

Assim sendo, este trabalho de pesquisa tem como objetivo desenvolver um modelo preditivo que combina as análises técnica e de sentimento, visando **identificar a correlação** entre o valor predito com o preço de fechamento das ações listadas na **bolsa de valores brasileira**.

## 2 | MATERIAIS E MÉTODOS

Definido o objetivo, levanta-se a hipótese da relação existente entre as variáveis, objeto de estudo deste trabalho. Isto posto, optou-se por realizar a validação da hipótese do ponto de vista da metodologia de um projeto fundamentado na ciência de dados, ilustrado

na Figura 1, que tem como principais etapas para o desenvolvimento (*pipeline*) de um modelo: coleta, processamento e modelagem de dados e, por fim, a validação do modelo.

As etapas que constituem o *pipeline* do modelo, requerem a realização de tarefas específicas. Para tanto, escolheu-se a linguagem Python, que possui bibliotecas que dão subsídio para que os requisitos necessários fossem alcançáveis, tais como: Pandas, para manipulação e análise de dados; TA-Lib, para cálculo dos indicadores técnicos; SciPy, para análises estatísticas e a Matplotlib, para visualização dos dados por meio de gráficos.

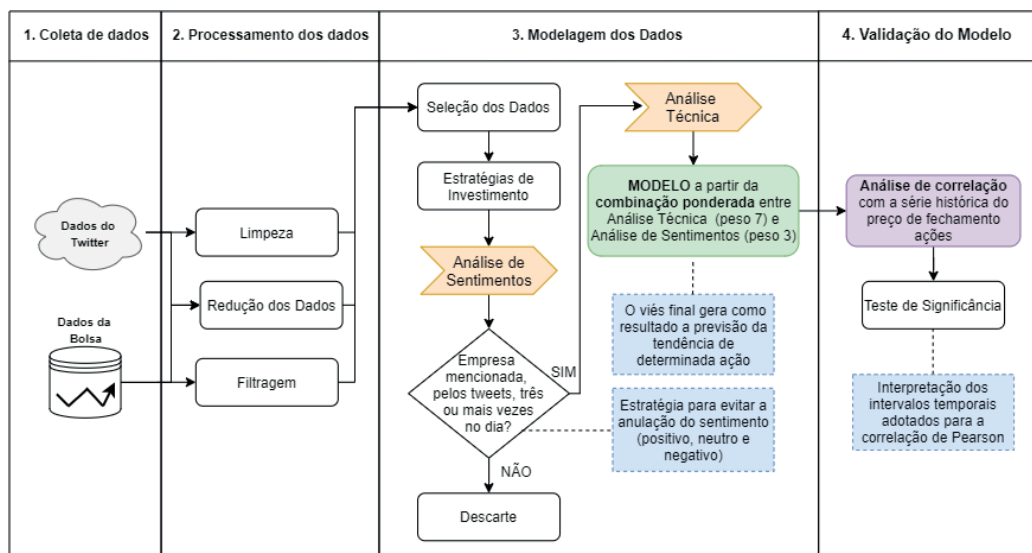


Figura 1 - Fluxograma de desenvolvimento do pipeline do modelo.

Fonte: Próprio autor (2020).

A etapa inicial do *pipeline*, a **coleta de dados**, consiste em mapear o escopo do trabalho. Para tanto, foi utilizado um conjunto refinado de *tweets* emitidos por veículos de comunicação especializados no mercado de ações brasileiro, em um intervalo temporal que vai de 12/11/2015 à 16/05/2019. O processo para a identificação dos perfis válidos foi realizado manualmente por MELO e MEINERZ (2019), bem como a rotulação da polaridade do sentimento presente em uma mensagem de texto, classificado em: positivo (1), neutro (0) ou negativo (-1).

Para a segunda base de dados, referente a **série histórica diária das ações** das empresas que fazem parte da Carteira Teórica do IBOVESPA, utilizou-se como fonte de dados a Brasil, Bolsa, Balcão (B3), no intervalo temporal correspondente ao da base de dados do Twitter.

Em seguida, para o refino dos dados, cerne do desenvolvimento do modelo, são



necessárias tarefas de **filtragem e limpeza**, típicas da etapa de **processamento de dados**. Assim, os dados do Twitter foram adaptados para análise diária, enquanto que do *dataset* da B3 foram filtrados os registros do segmento ações, a vista e fracionário, equivalente a 36,4% dos registros da base original.

Finalizada a etapa de processamento, inicia-se a da **modelagem de dados**, que tem como objetivo principal a construção do modelo preditivo. A primeira tarefa empregada foi a **seleção dos dados**, com a decisão de escolha pelas empresas que seriam exploradas. Para isso, criaram-se variáveis explanatórias, de agrupamento dos sentimentos, para responder à seguinte questão: quais empresas são mencionadas, nos *tweets*, três ou mais vezes em um dia?

A limitação do volume válido (mínimo de três *tweets*), mencionado na questão, se deu a fim de evitar a anulação entre os sentimentos positivo e negativo, o que também garante que haja uma maior convicção do sentimento, por parte do mercado, para o próximo pregão.

A resposta manteve, então, somente 35 empresas, correspondente a 56,45% do total de 62 empresas que compunham a carteira teórica do índice Ibovespa considerado em maio de 2019. Ainda assim, as tarefas seguintes prosseguiram somente com duas empresas que, em algum dia, apresentaram um notório volume de dados, sendo este o caso da PETROBRAS e VALE.

O volume dos *tweets* delimitam o intervalo temporal de análise, onde para todas as empresas observa-se um volume notório e válido a partir de 2019, conforme ilustra a Figura 2. Entretanto, a frente, são analisadas as ações das empresas no mês em que há maior quantidade de postagens, com destaque para o mês de abril, para a PETROBRAS, e janeiro, para a VALE.

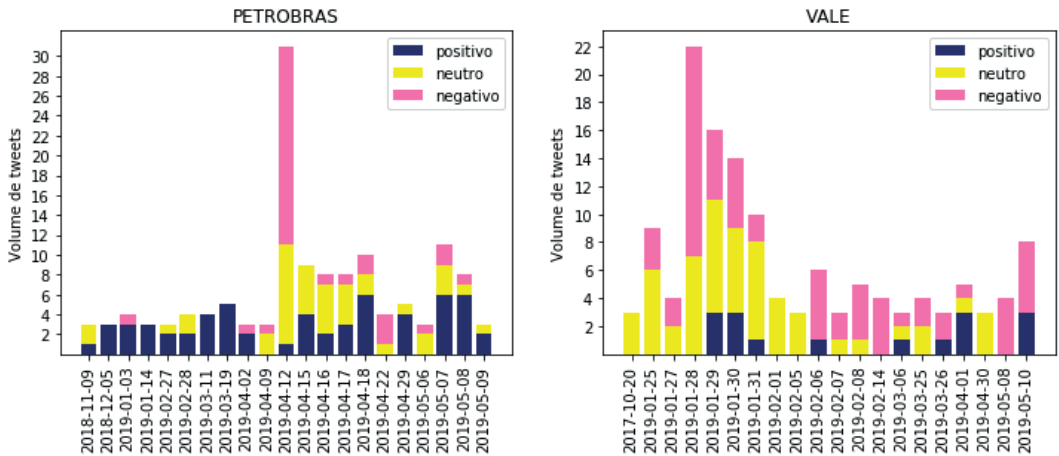


Figura 2 - Sentimentos rotulados para as empresas analisadas (PETROBRAS E VALE).

Fonte: Próprio autor (2020).

Sabendo que uma empresa listada em uma bolsa de valores pode ter mais do que um tipo de ação sendo negociada, decidiu-se selecionar aquelas com maior liquidez, ou seja, as mais negociadas. São elas:

- a) PETR4 - Petróleo Brasileiro S.A;
- b) VALE3 - Vale S.A.

Para a tarefa de **análise de sentimento**, sabe-se que o sentimento coletado pelo conjunto de *tweets* em um dia  $t$  representará a tendência, o que chamamos de viés, para o próximo pregão ( $t + 1$ ) – dia em que ocorre negociação das ações. Para este fim, calculou-se a força do sentimento diário ( $F$ ) a partir da divisão entre a soma ( $S$ ) dos sentimentos classificados em positivo (1), neutro (0) e negativo (-1) e a quantidade total de *tweets* ( $T$ ), conforme a Equação 1:

$$F = \frac{S}{T} \tag{1}$$

Os valores emitidos para os vieses da AS foram mapeados, em uma escala de -1 a 1, de forma análoga a variação de tendência nos níveis de venda forte, venda, neutro, compra e compra forte, conforme a Tabela 1.

Força do sentimento (F)	Viés de tendência
-1,0 à -0,6	VENDA FORTE
-0,59 à -0,2	VENDA
-0,19 à +0,19	NEUTRO
+0,2 à +0,59	COMPRA
+0,6 à +1,0	COMPRA FORTE

Tabela 1 - Escala de tendência para a polaridade dos sentimentos.

Fonte: Próprio autor (2020).

Para a modelagem do ponto de vista da AT, tem-se **uma estratégia de investimento** que combina grupos de indicadores técnicos. Sendo assim, o primeiro indicador escolhido foi o das **Medias Móveis Exponenciais (MME)**, que levam em consideração a comparação entre duas médias de períodos (n) diferentes, sendo uma de curto prazo (5 dias) e outra de prazo maior (20 dias), atribuindo maior peso aos dados mais recentes (MURPHY, 1999), demonstrado na Equação 2:

$$MME_{(atual)} = Preço_{(atual)} - MME_{(anterior)} * \frac{2}{n+1} + MME_{(anterior)} \quad (2)$$

Uma aplicação mais ampla das médias móveis é o indicador **Convergência/ Divergência das Médias Móveis (MACD)**, reconhecido por MURPHY (1999) como um híbrido ao combinar indicadores de oscilação com o de médias móveis. Este, compõe-se de duas linhas que sinalizam o início ou fim de uma tendência: a linha MACD, onde o valor da MME mais curta (12 dias) é subtraído pelo valor da MME mais longa (26 dias), e a linha de Sinal, composta pela linha MACD ajustada por uma MME de 9 dias.

Por último, buscou-se um indicador capaz de realizar a confirmação das tendências encontradas. A decisão foi pelo **Volume**. Ele se refere a quantidade de dinheiro sendo negociado, representando a força do mercado e, ainda, retratando a espinha dorsal da AT - oferta/demanda. E para o viés final da AT, a combinação dos indicadores técnicos ficou definida conforme a Equação 3:

$$Viés_{AT} = (0.5 * MME) + (0.3 * MACD) + (0.2 * Volume) \quad (3)$$

Assim como realizado na tarefa de identificação do viés da AS, o conjunto dos indicadores técnicos foram mapeados para a escala de -1 a 1, como pode ser observado na Tabela 2.

MME	MACD	VOLUME	ESCALA	F
COMPRA	COMPRA	SIM	+1	COMPRA FORTE
COMPRA	COMPRA	NÃO	+0,5	COMPRA
COMPRA	VENDA	SIM	+0,5	COMPRA
COMPRA	VENDA	NÃO	+0,5	COMPRA
VENDA	COMPRA	SIM	-0,5	VENDA
VENDA	COMPRA	NÃO	-0,5	VENDA
VENDA	VENDA	NÃO	-0,5	VENDA
VENDA	VENDA	SIM	-1	VENDA FORTE

Tabela 2 - Escala de tendência para a AT.

Fonte: Autoria própria (2020).

Cumpridas as técnicas, a fase de Modelagem dos Dados chega à última tarefa de implementação: a **combinação ponderada entre o viés de ambas técnicas** para a consolidação do modelo preditivo, nomeado como SENTECH (Combinador de Análise Técnica e Sentimento), que atribui maior peso para a AT (0,7), visto que a AS (0,3) apresenta maior sensibilidade em seu viés.

Com o modelo desenvolvido, inicia-se a etapa de **Validação do Modelo**, que consiste da aplicação e análise de testes estatísticos, com vistas a apoiar o processo de inferência. Portanto, aplicou-se a **análise por correlação**, exploração descritiva que mede se há, e em que grau, dependência entre duas variáveis, embora não implique em causalidade (MOORE e KIRKLAND, 2007). Empregou-se o **Coefficiente de Correlação de Pearson ( $\rho$ )**, alcançável pela divisão entre a covariância (cov) e o desvio padrão ( $\sigma$ ), conforme a Equação 4. Logo, quanto mais próximo de 1 ou -1 está o valor ( $\rho$ ), mais forte é a relação entre as variáveis.

$$\rho_{x,y} = \frac{cov_{x,y}}{\sigma_x \sigma_y} \quad (4)$$

### 3 | RESULTADOS E DISCUSSÃO

A Figura 3 apresenta a aplicação dos indicadores técnicos em comparação com o Gráfico de Candles, uma das principais ferramentas da AT, com informações de preço de fechamento, abertura, mínimo e máximo.

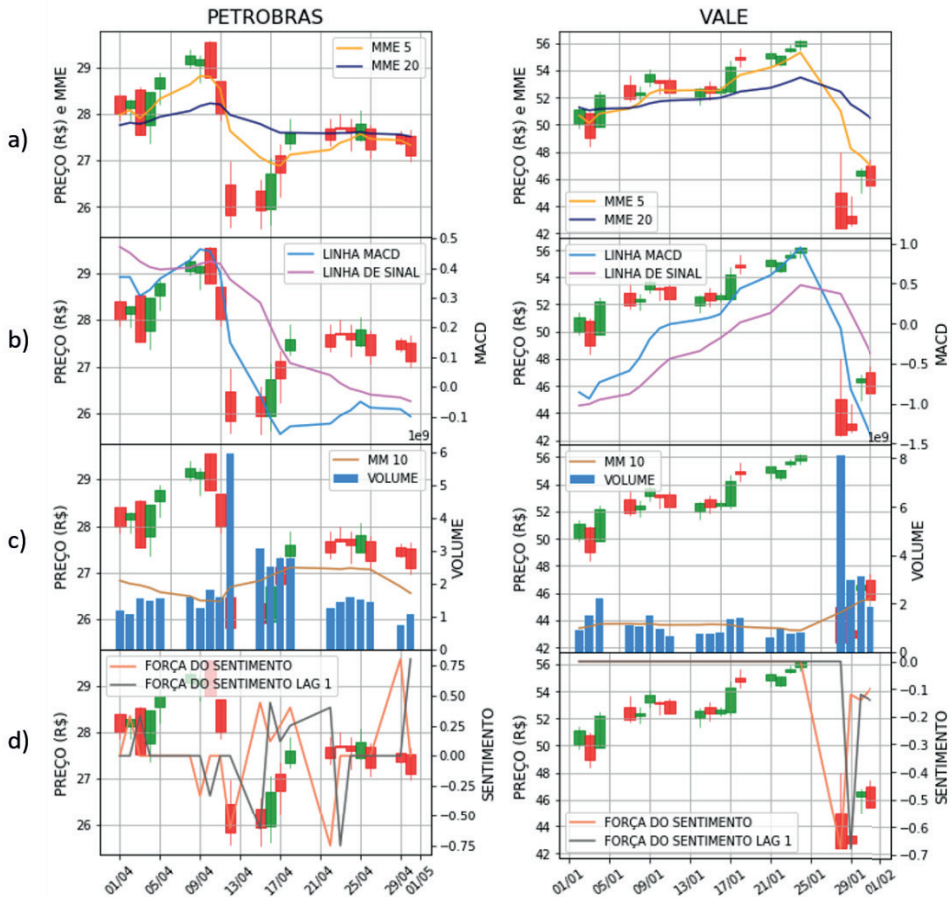


Figura 3 - AT e AS aplicadas para PETROBRAS e VALE.

Fonte: Autoria própria (2020).

Sendo assim, é possível notar o comportamento dos indicadores técnicos, que acompanham a movimentação real do preço das ações analisadas. Além do mais, nos momentos em que há inversão de tendência, nota-se o cruzamento das linhas que compõe a MME (a) e o MACD (b), além do Volume (c) acima da média. Já para o indicador que representa a força do sentimento (d), nota-se um atraso de um dia (*lag 1*) em relação ao comportamento traçado pelos *candles*.

Já na etapa **Validação do Modelo**, a hipótese foi validada para a correlação entre o viés da combinação dos indicadores e o preço de fechamento das ações. Assim, pelo **teste de significância**, obteve-se valores de correlação válidos com um **nível de confiança de 95%** ao adotar a probabilidade de erro de 5%, conforme elenca a Tabela 3, com os valores de *p*-valor menores que o  $\alpha$  adotado (0,05), sendo, portanto, significativos.

Empresa	Intervalo	Correlação ( $\rho$ )	p-valor
PETROBRAS	1 mês 01/04/2019 à 31/04/2019	0,57 MODERADA	0,0073
	1 ano 01/05/2018 à 31/05/2019	0,24 DESPREZÍVEL	5,29e-05
VALE	1 mês 01/01/2019 à 31/01/2019	0,65 MODERADA	0, 0013
	1 ano 01/05/2018 à 31/05/2019	0,34 FRACA	9,76e-09

Tabela 3 - Resultado da correlação para a PETROBRAS e VALE

Fonte: Autoria própria (2020).

Ainda, para ambas ações foi encontrado um baixo nível de correlação ao considerar a janela de análise de um ano, o que ainda pode-se justificar pelo baixo volume de *tweets* encontrados no longo prazo. A justificativa foi analisada, na Tabela 4, para a empresa VALE, aumentando a janela de análise para dois meses, intervalo no qual houve uma maior constância de postagens referentes a empresa.

Empresa	Intervalo	Correlação ( $\rho$ )	p-valor
VALE	2 meses 01/01/2019 à 31/02/2019	0,7 FORTE	4.30e-07

Tabela 4 – Correlação para a VALE em janela de dois meses.

Fonte: Autoria própria (2020).

## 4 | CONCLUSÃO

Voltando à questão colocada no início deste estudo, agora é possível afirmar, com 95% de confiança, que o resultado predito pelo SENTECH, modelo desenvolvido a partir da combinação do viés das abordagens de análise técnica e análise de sentimentos, está correlacionado, para o curto prazo, com o preço de fechamento das ações.

A pesquisa apresenta como principais benefícios propiciados pelo modelo a quantificação da correlação entre o resultado do modelo e a série histórica dos preços de fechamento das ações e a alta significância dos testes de correlação.

Em consequência, a aplicação do modelo reforça a importância do mercado financeiro para as relações modernas, unindo o elo do conservadorismo da análise técnica

e o contemporâneo da análise de sentimentos.

Para a continuidade deste trabalho, destaca-se como principal recomendação, caracterizada como uma extensão natural do SENTECH, a validação do modelo por meio do *backtesting*, a fim de avaliar o percentual de lucro obtido a partir das decisões tomadas baseadas na previsão realizada, produto deste trabalho.

## REFERÊNCIAS

ATTANASIO, G. et al. **Combining News Sentiment and Technical Analysis to Predict Stock Trend Reversal**. In: 2019 International Conference on Data Mining Workshops (ICDMW), Beijing, 2019. Disponível em: <https://ieeexplore.ieee.org/document/8955650>. Acesso em: 23 abr. 2020.

BOLLEN, J.; MAO, H.; ZENG, X. **Twitter mood predicts the stock market**. Journal of Computational Science, v. 2, n. 1, p. 1-8, 2011. Disponível em: <http://www.sciencedirect.com/science/article/pii/S187775031100007X>. Acesso em: 23 mai. 2020.

DAMODARAN, A. **Avaliação de investimentos: Ferramentas e Técnicas para Determinação do Valor de Qualquer Ativo**. 2. ed. Rio de Janeiro: Qualitymark, 2010.

MELO, J. G.; MEINERZ, G. V. **Análise de sentimentos de textos voltados ao mercado de ações**. In: XXIV Seminário de Iniciação Científica e Tecnológica, Pato Branco, 2019. Disponível em: <https://eventos.utfpr.edu.br/sicite/sicite2019/paper/viewFile/4739/1331>. Acesso em: 20 mar. 2020.

MOORE, D. S.; KIRKLAND, S. **The basic practice of statistics**. New York: WH Freeman, 2007.

MURPHY, J. **Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications**. New York Institute of Finance, 1999.

PICASSO, A. et al. **Technical analysis and sentiment embeddings for market trend**. Expert Systems with Applications, v. 135, p. 60-70, 2019. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0957417419304142?via%3Dihub>. Acesso em: 23 abr. 2020.

REINSEL, D.; GANTZ, J.; RYDNING, J. **The digitization of the world from edge to core**. Framingham: International Data Corporation, 2018. Disponível em: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>. Acesso em: 01 jun. 2020.

SHAH, D.; ISAH, H.; ZULKERNINE, F. **Stock Market Analysis: A Review and Taxonomy of Prediction Techniques**. International Journal of Financial Studies, Kingston, v. 7, n. 2, p. 26, 2019. Disponível em: <https://www.mdpi.com/2227-7072/7/2/26>. Acesso em: 23 mar. 2020.

STATISTA. **Number of social network users in Brazil from 2017 to 2023 (in millions)**. 2020. Disponível em: <https://www.statista.com/statistics/278408/number-of-social-network-users-in-brazil/>. Acesso em: 01 abr. 2020.

## PROGRAMAÇÃO DE JOGOS COM SCRATCH PARA AUXÍLIO À ALFABETIZAÇÃO DE CRIANÇAS

*Data de aceite: 01/04/2021*

*Data de submissão: 20/01/2021*

### **Rute Vitorino Oliveira**

Graduanda em Licenciatura em Física  
Instituto Federal de São Paulo, campus  
Votuporanga  
Votuporanga/SP  
<https://orcid.org/0000-0001-5602-9265>

### **Jemima Vitorino de Oliveira**

Graduanda em Engenharia Elétrica  
Instituto Federal de São Paulo, campus  
Votuporanga  
Votuporanga/SP  
<http://lattes.cnpq.br/4583262643190371>

### **Luciene Cavalcanti Rodrigues**

Tecnóloga em Processamento de Dados,  
Mestre em Engenharia Elétrica e Doutora em  
Ciências com foco em Física Computacional  
pela USP-SP, docente do Instituto Federal de  
São Paulo, campus Votuporanga e da FATEC  
Rio Preto  
São José do Rio Preto/SP  
<http://lattes.cnpq.br/9346178640224901>

### **Ana Paula Garrido de Queiroga**

Bacharel em Engenharia de Produção,  
Mestranda em Ciências Ambientais UEMG,  
docente da UNILAGO  
São José do Rio Preto/SP  
<http://lattes.cnpq.br/7071401572450281>

**RESUMO:** A linguagem, a comunicação e elementos comunicacionais são eixos fundamentais de uma proposta educativa que

visa ajudar no processo do ensino-aprendizagem. Paulo Freire (1960), propõe que se trabalhe em favor do alfabetismo, método de ferramenta alternativa face ao momento em que estamos vivendo. Nessa era tecnológica, devemos trazer dados, fatos, conhecimento científico e tecnológico de maneira que o conhecimento não se torne banal. O objetivo da criação de jogos é de o professor ensinar a criança a fazer o jogo, porque a criança vai interagir tanto com o mundo dos jogos, quanto com o mundo da aprendizagem (alfabetização). Introduzir o uso de jogos no processo de aprendizagem torna-se um método inovador e estimula a solução de problemas de forma prática. Para tanto estão disponíveis diversas ferramentas computacionais, sendo o Scratch uma ferramenta fácil e intuitiva, utilizada por diversas crianças e que ajudará tanto o professor a seguir a norma da BNCC quanto ao aluno a conhecer as letras do alfabeto e aprender mais sobre o pensamento computacional. O Scratch tem mecanismos de envolver a criança em programar por meio de blocos que se encaixam para que a ação desejada seja compreendida e executada, sendo mais visual e intuitiva para o processo de aprendizagem, incentivando a diversão entre crianças de todas as idades. Os jogos que são desenvolvidos ajudarão as crianças a refletir de forma sistemática e, além de treinar suas habilidades essenciais para que possa aprender o som de cada palavra, desenvolverá os processos cognitivos, como a aprendizagem, atenção, memória, linguagem e o raciocínio. Esse trabalho está sendo desenvolvido em conjunto com o projeto CODE CLUBE do IFSP



Campus Votuporanga, com a finalidade de ensinar programação para crianças e auxiliar professores a criar jogos para uso em sala de aula.

**PALAVRAS - CHAVE:** Tecnologia; jogo de alfabetização; linguagem de programação.

**ABSTRACT:** Language, communication and communicational elements are fundamental axes of an educational proposal that aims to help in the teaching-learning process. Paulo Freire (1960), proposes to work in favor of literacy, a method of alternative tool in face of the moment in which we are living. In this technological age, we must bring data, facts, scientific and technological knowledge in a way that knowledge is not trivial. The goal of creating games is for the teacher to teach the child to make the game, because the child will interact with both the world of games and the world of learning (literacy). Introducing the use of games in the learning process becomes an innovative method and encourages the solution of problems in a practical way. For that, several computational tools are available, being Scratch an easy and intuitive tool, used by several children and that will help both the teacher to follow the BNCC norm and the student to know the letters of the alphabet and learn more about computational thinking. Scratch has mechanisms to involve the child in programming through blocks that fit together so that the desired action is understood and performed, being more visual and intuitive for the learning process, encouraging fun among children of all ages. The games that are developed will help children to reflect systematically and, in addition to training their essential skills so that they can learn the sound of each word, it will develop cognitive processes, such as learning, attention, memory, language and reasoning. This work is being developed in conjunction with the CODE CLUBE project of the IFSP Campus Votuporanga, with the purpose of teaching programming to children and helping teachers to create games for use in the classroom.

**KEYWORDS:** Technology; literacy game; programming language.

## 1 | INTRODUÇÃO

Este trabalho tem como fundamento demonstrar a criação de jogos para a alfabetização desenvolvidos com uso da ferramenta Scratch, bem como descrever os fundamentos teóricos que o embasam. Com a evolução da tecnologia e a informação a humanidade necessita dos métodos de aquisição de leitura e escrita, ainda mais tendo em vista que a geração dos chamados nativos digitais interage facilmente com games e crescem mexendo em celulares e internet, então essas ferramentas são de suma importância tanto para o ensino, quanto para aprendizagem.

Esse fenômeno vem se espalhando pela educação, aplicando como estratégia de ensino e aprendizagem, dirigida a um público-alvo inserido na chamada geração gamer, e existem resultados positivos sendo obtidos através dessas experiências (Sheldon, 2012)<sup>1</sup>.

---

<sup>1</sup> SHELDON, Lee. The Multiplayer Classroom: Designing Coursework as a Game. Boston, MA: Cengage Learning, 2012.

Paulo Freire (1960), educador e filósofo brasileiro, buscava a educação por meio do diálogo, concebendo a apropriação do conhecimento como fonte vital para a libertação das pessoas.

A alfabetização não pode se fazer de cima para baixo, nem de fora para dentro, como uma doação ou uma exposição, mas de dentro para fora pelo próprio analfabeto, somente ajustado pelo educador. Esta é a razão pela qual procuramos um método que fosse capaz de fazer instrumento também do educando e não só do educador. (Freire, 1979, p. 72)<sup>2</sup>.

Para ele, diversos fatores chegam à autonomia; o papel do educador e aluno, a consciência do inacabamento, o papel da pesquisa, todos esses fatos resultam na autonomia tanto do educador, e ambos partem de uma autonomia para chegada a um autêntico conhecimento.

Uma das tarefas mais importantes da prática educativa crítica é propiciar as condições em que os educandos em suas relações uns com os outros e todos com o professor ou a professora ensaiam a experiência profunda de assumir-se. Assumir-se como ser social e histórico, como ser pensante, comunicante, transformador, criador, realizador de sonhos, capaz de ter raiva porque capaz de amar. (Freire, 2000, p.46)<sup>3</sup>.

Paulo Freire (1960) defendia o método de alfabetização para além da decodificação dos códigos linguísticos, ou seja, não basta ler e escrever, mas fazer uso de ferramentas que possam ajudar na aprendizagem. A resolução de problemas dentro de jogos voltados à educação visa obter a atenção da criança, principalmente no período de letramento e alfabetização, que é um desafio para todos os educadores tendo em vista que hoje em dia existem inúmeras distrações que competem por um pouco de atenção e concentração. O que os professores precisam ter em mente é que interagir com algumas destas distrações é uma das formas de se tornar mais eficaz.

O uso de jogos e da gamificação é uma boa estratégia para vencer este desafio, pois é uma metodologia de ensino pautada na dinâmica e no uso de jogos, sejam eles eletrônicos ou aqueles tradicionais, para aplicar os conteúdos das disciplinas escolares. Como cita ADAIL (2013), “No imerso mar de informações que é a internet, há duas bússolas que podem orientar a pesquisa: os sites de busca e os diretórios pulmões de um portal”.

Uma das formas de aliar o uso de jogos e a aprendizagem é com o uso de ferramentas para a criação de jogos, tais como Scratch, Construct, Unity entre outros, desta forma este trabalho apresenta a ferramenta Scratch como uma nova linguagem de programação desenvolvida pelo Lifelong Kindergarten Group no MIT Media Lab. O Scratch permite criar as próprias histórias interativas, animações, jogos, músicas e arte e compartilhá-las através de websites.

2 FREIRE, Paulo; MACEDO, D. Alfabetização: Leitura do mundo, leitura da palavra. Tradução OLIVEIRA, Lólio Lourenço. Educar em Revista, 2015.

3. FREIRE, Paulo.; A importância do ato de ler: em três artigos que se completam. São Paulo, Autores Associados & Cortez. Coleções polêmicas do nosso tempo, 2012.

Os projetos Scratch são baseadas em objetos gráficos chamados Sprites. Pode-se mudar aparência de um Sprite, dando-lhe um novo traje, ou fazê-lo parecer-se com uma pessoa, um objeto ou até mesmo um animal. Ainda, pode usar qualquer imagem como traje: pode desenhar no “Editor de Pintura”, importar uma imagem do disco rígidos, ou arrastar uma imagem a partir de um site (SCRATCH, 2013)<sup>4</sup>.

Os jogos do Scratch podem ser criados a partir de comandos de blocos lógicos que devem ser agrupados com peças de Lego, é uma ferramenta gratuita e pode ser utilizada nos principais sistemas operacionais (Windows, Linux e Mac) ou diretamente pelo site. Foi criada com o propósito de introduzir a programação e a conceitos matemáticos, ao mesmo tempo em que incentiva o pensamento criativo, o raciocínio sistemático e o trabalho colaborativo (SCRATCH, apud PEREIRA, 2012). Relacionar ideias de linguagens de programação como estas, não são de fácil compressão para crianças que não demonstram familiaridade com o ambiente de programação.

Segundo Oliveira (2000)<sup>5</sup> ao brincar a criança pode desenvolver capacidades importantes como atenção, memória, imitação, imaginação, ainda propiciando à criança o desenvolvimento de áreas da personalidade como afetividade, motricidade, inteligência, sociabilidade e criatividade.

## 2 | ALFABETIZAÇÃO

Segundo Soares (2007, p.16)<sup>6</sup> – destaca a etimologia da palavra Alfabetização: “[...] O termo alfabetização não ultrapassa o significado de ‘levar à aquisição do alfabeto’, ou seja, ensinar o código da língua escrita, ensinar as habilidades de ler e escrever”. Para entender o processo de alfabetização, há um crescente nível de exigência sobre as competências da professora alfabetizadora. É preciso ter domínio da disciplina, capacidade de organizar, para enfatizar os principais objetivos, capacidade de comunicação com os alunos. O uso dos jogos pode despertar nas crianças a motivação, a expressividade, imaginação, linguagem comunicativa, atenção, concentração, raciocínio lógica, como de acordo com Soares.

“[...] Jogos para o desenvolvimento da consciência fonológica, se realizados sistematicamente na educação infantil, criam condições propícias, até mesmos necessários, para apropriação do sistema alfabético.” (SOARES, 2017, p.142)<sup>7</sup>.

---

4. SCRATCH. ABOUT Scratch (Scratch Documentation Site). Disponível em: <https://scratch.mit.edu/>. Acesso em: 29 dezembro 2020.

5. OLIVEIRA, Marta Kohl de. Vygotsky: aprendizado e desenvolvimento um processo sócio-histórico. São Paulo: Scipione, 1997

6. SOARES, Magda. Alfabetização e letramento. 7. ed. São Paulo. Contexto, 2017. E-book.

7. SOARES, Magda. Alfabetização: a questão dos métodos. São Paulo. Contexto, 2016. E-book.

No Brasil a ideia de ler é sinônimo de compreender. Esse conceito equivocado reside na confusão entre o objetivo de ler, que é compreender, e o processo de aprender a ler, que envolve a decodificação com processo no qual o leitor transforma as retas e letras em sons oralizando ou numa imagem mental do som.

## 3 | OS JOGOS

### 3.1 Jogo com sons

O primeiro jogo a ser apresentado chama-se “*Alfabetização*”. Qual o objetivo de criar o jogo? O objetivo é o professor ensinar a criança a fazer o jogo, porque a criança vai interagir tanto com o mundo da gamificação, quanto com o mundo da aprendizagem (alfabetização).

Abaixo são apresentados os passos para a criação do primeiro jogo, junto com os comandos de blocos lógicos que representam a programação.

Os sons das letras foram gravados com uso do microfone diretamente na plataforma Scratch, sem a necessidade de uso de softwares auxiliares. Tais sons não estão disponíveis diretamente no Scratch, sendo necessária a gravação da voz para que fosse possível ouvir os sons dentro do jogo, o uso dos sons das letras aprimora o funcionamento do jogo e trabalha a fonética.



Figura 1. Tela inicial do jogo onde o jogador coloca seu nome e inicia o jogo

Fonte: Tela do Scratch captura pelos autores.

**Tela inicial:** Quando o jogador inicia o jogo, aparecerá a personagem chamada Amora. Ela dará as “Boas-vindas” e perguntará o nome do jogador, desta forma a criança poderá digitar seu nome com o teclado e iniciar o jogo.

Serão apresentados os botões INSTRUÇÕES E JOGAR, ao clicar em jogar o jogador deverá utilizar o mouse, desta forma a criança poderá trabalhar a coordenação motora.

No Scratch todas estas ações devem ser previamente programadas com os respectivos blocos, então os seguintes comandos devem ser inseridos no cenário do jogo (Figura 2):

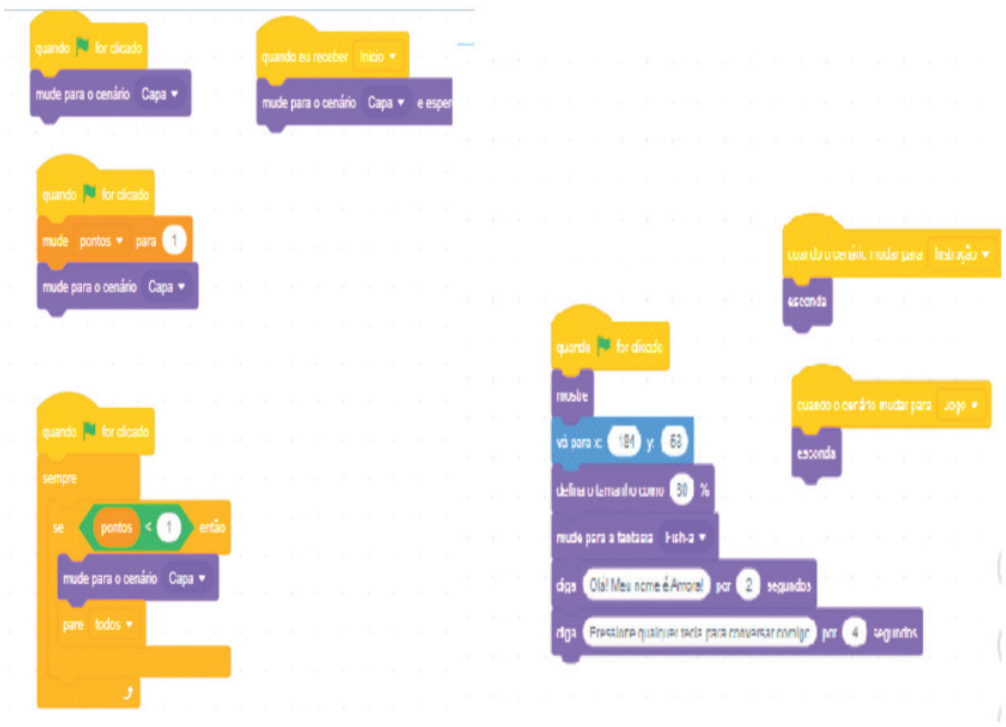


Figura 2. Programação em bloco utilizada para mostrar o cenário de início e a pontuação.

Fonte: Tela do Scratch captura pelos autores.

**Instrução:** Após a abertura da Figura 1, o usuário deve clicar no botão abrir “instrução”, em seguida será apresentado o roteiro de como iniciaria o jogo (Figura 3).

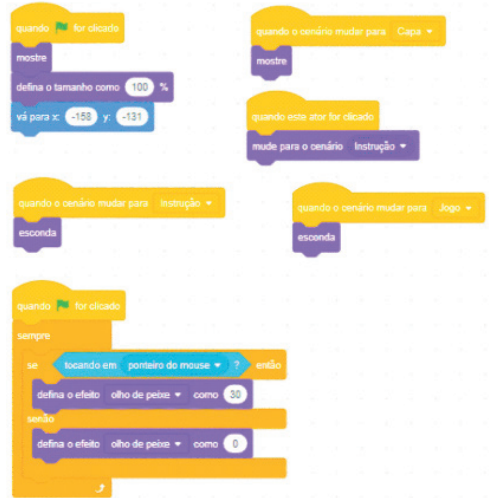
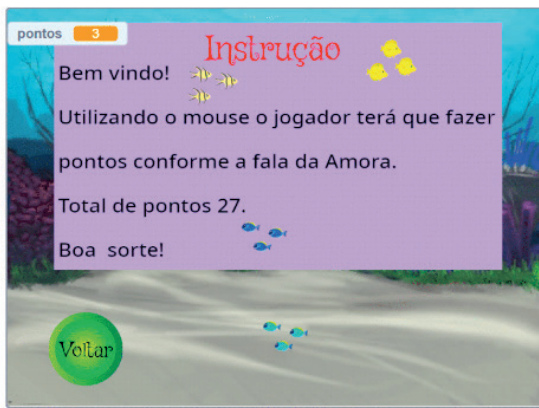


Figura 3. Janela que será mostrada depois do clique no botão instrução, ao lado os blocos necessários para a programação.

Fonte: Tela do Scratch captura pelos autores.

**Jogar:** Ao iniciar o jogo, o jogador deverá clicar na letra do alfabeto de acordo com a fala da personagem Amora. A cada letra em que o jogador acertar receberá 01 (um) ponto, se errar não perderá pontos (Figura 4).



Figura 4. Tela quando clicar o iniciar já com a primeira letra indicada para dar início

Fonte: Tela do Scratch captura pelos autores.

A programação para a apresentação das letras e emissão dos sons estão representadas na Figura 5.

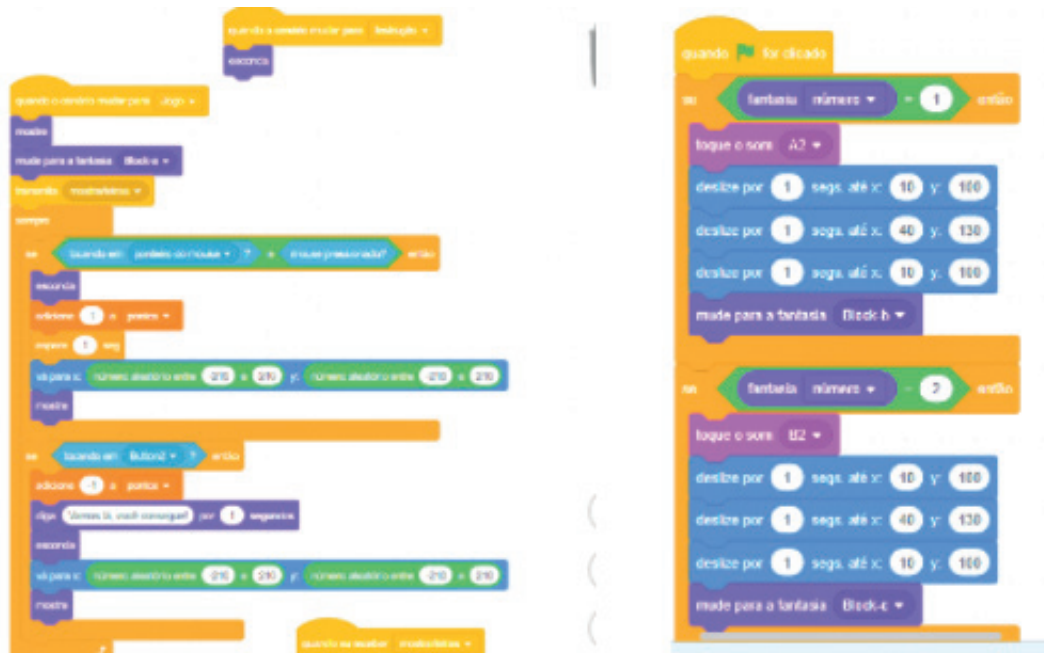


Figura 5. A programação de troca de fantasia no jogo e a reprodução do som.

Fonte: Tela do Scratch captura pelos autores.

### 3.2 Jogo das vogais

Neste jogo o personagem Sapo inicia a conversa com o jogador, depois disso serão apresentadas várias letras e o jogador ganha pontos ao clicar nas vogais e perde pontos ao clicar nas consoantes.

A primeira etapa é colocar o cenário, os atores e criar a variável Pontos. Depois inicia-se a programação, na Figura 6 é mostrada a programação do cenário.

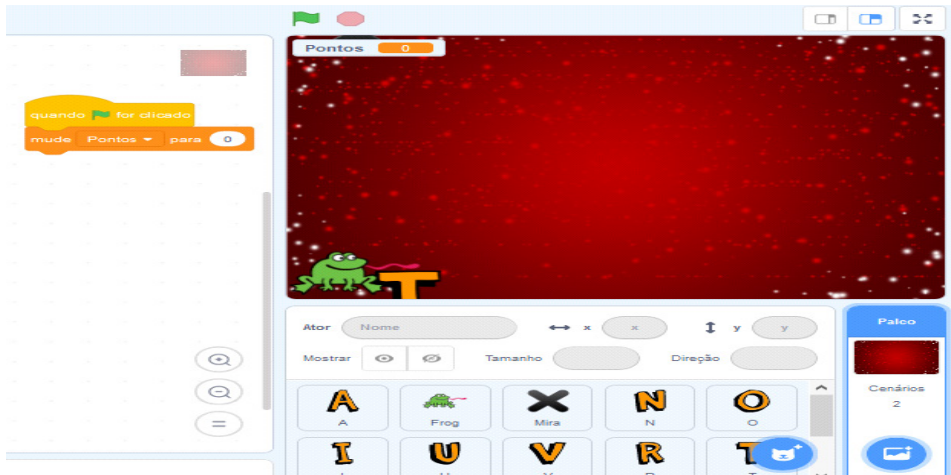


Figura 6. Janela do início do jogo com a programação

Fonte: Tela do Scratch captura pelos autores.

É possível notar que foram inseridos vários atores, as letras que representam vogais, as que representam consoantes, o sapo e a mira. Tais atores podem ser modificados de acordo com a vontade do professor. O ator do sapo vai conversar com o jogador, então a programação da Figura 7 deve ser colocada no respectivo ator.

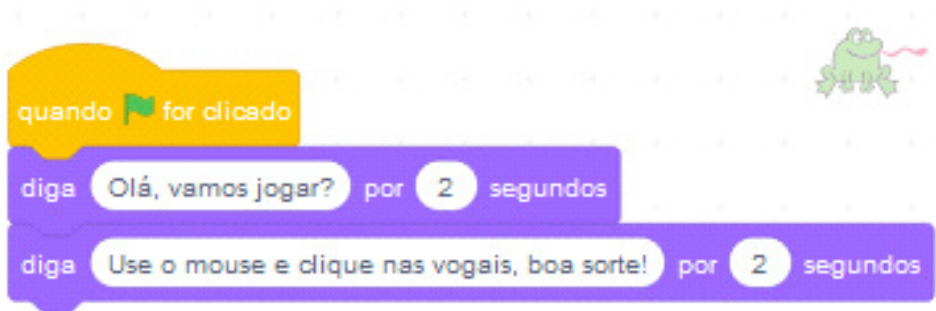


Figura 7. Programação em bloco chamando a variável de início do jogo

Fonte: Tela do Scratch captura pelos autores.

Para facilitar a jogabilidade o mouse terá uma mira, então é necessário inserir o ator da imagem X e a programação da Figura 8, desta forma ao mexer o mouse a figura se movimentará junto.



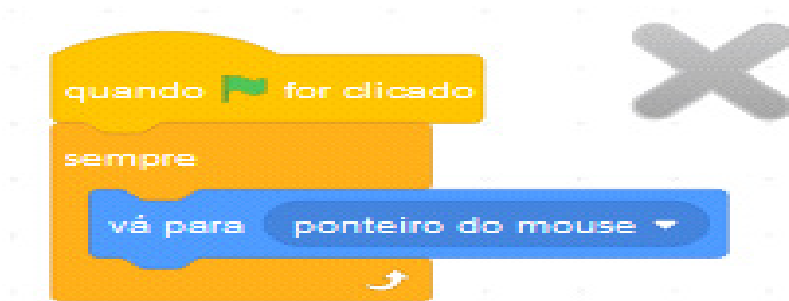


Figura 8. Programação utilizada para indicar um x na ponteira do mouse para facilitar na jogabilidade

Fonte: Tela do Scratch captura pelos autores.

Para as vogais a programação será colocada para que o jogador ganhe pontos ao clicar nelas, então para cada letra deve ser inserida a programação da Figura 9.

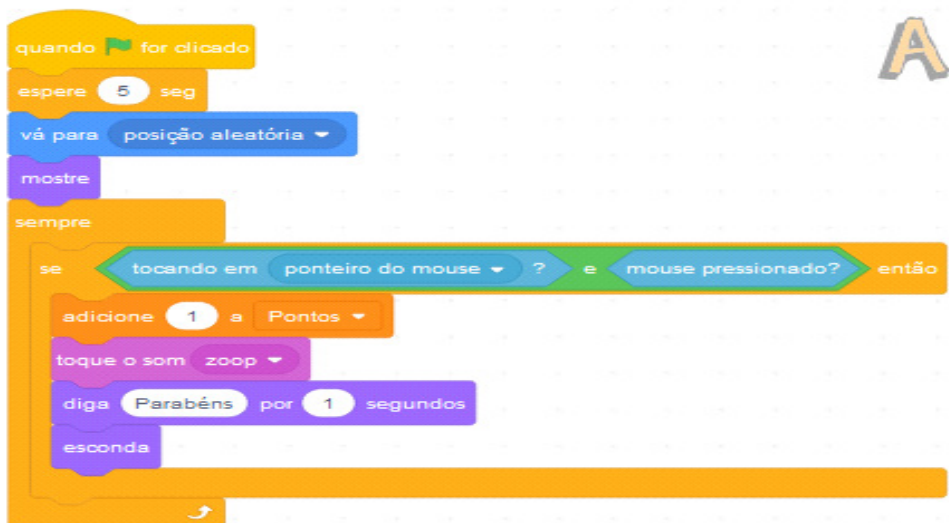


Figura 9. Programação com bloco lógico para as vogais (letras corretas)

Fonte: Tela do Scratch captura pelos autores.

Para as consoantes a programação é similar, mas o jogador perderá pontos ao clicar (Figura 10).

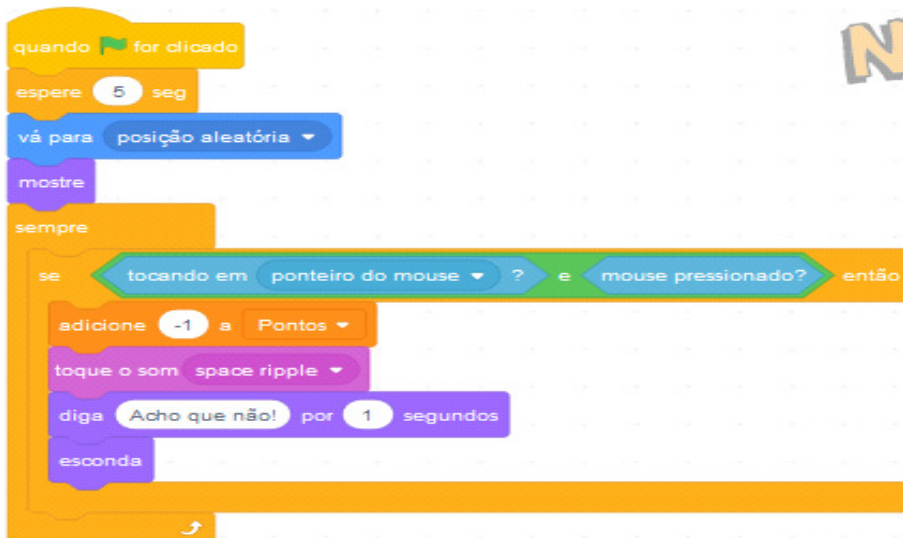


Figura 10. Bloco lógico para as consoantes (letra errada)

Fonte: Tela do Scratch captura pelos autores.

A plataforma de acesso ao jogo Scratch na alfabetização – Jogos Vogais está disponível no site <https://scratch.mit.edu/projects/215199041/>, bem como os vídeos interativos para criação deste jogo está disponível no site <https://www.youtube.com/watch?app=desktop&v=jm7jKwGmsnc>.

Desta forma, objetivo desse jogo é encontrar as vogais misturadas entre as consoantes, desta forma os alunos poderão tentar diferenciar umas das outras. Como atividade para o professor trabalhar com seus alunos, é possível colocar diferentes letras, adicionar sons e animações, bem como alterar o conteúdo, trabalhando por exemplo com números, letras maiúsculas e minúsculas, cores, animais, objetos e diversos outros conteúdos.

Existem diversos jogos disponíveis na plataforma do Scratch que trabalham com Alfabetização, alguns estão listados no link: <https://scratch.mit.edu/search/projects?q=Alfabetização>, todos criados pela comunidade e disponíveis gratuitamente no site do Scratch.

#### 4 | CONSIDERAÇÕES FINAIS

É importante salientar que a plataforma Scratch é intuitiva tanto para professores quanto para alunos, utilizada mundialmente por crianças e adolescentes para trabalhar o raciocínio lógico, criatividade, leitura, escrita, matemática, coordenação motora fina e direção. Neste trabalho são enfatizados o trabalho com o tema alfabetização, cujos jogos

podem ser desenvolvidos por professores e posteriormente apresentados aos alunos, ou o professor pode trabalhar a concepção e desenvolvimento dos jogos junto aos alunos e posterior apresentação dos trabalhos, deixando em evidência que os alunos aprendem tanto sobre o tema do jogo quanto sobre programação, transformando as crianças em produtores de conteúdo e não apenas consumidores.

## REFERÊNCIAS

LOUREIRO, Gustavo. **Curso Superior de Formação Específica em Gestão de Ambientes Internet Webmaster/Webdesigner** [manual]. [Capturado em 2008 jun 20]. Disponível em: <http://www.scratchbrasil.net.br/index.php/materiais/tutoriais.html>. Acesso em: 31 dezembro 2020.

LOPES, V. C. Melo.; **Projeto e-jovem- modulo II: Lógica de Computadores. Lógica de programação Scratch**. ScratBrasil.v. 1, n. 1, p. 1-70, dez. 2009. Disponível em: [http://www.scratchbrasil.net.br/images/download-materiais/Apostila\\_L%C2%A2gica\\_Scratch\\_Python\\_E-Jovem.pdf](http://www.scratchbrasil.net.br/images/download-materiais/Apostila_L%C2%A2gica_Scratch_Python_E-Jovem.pdf). Acesso em: 28 set. 2020.

MALUF, M. R.; CARDOSO-MARTINS, C. **Alfabetização no século XXI: como se aprende a ler e a escrever**. Porto Alegre: Penso Editora, 2013.

NEGRINE, Aírton. **Aprendizagem e desenvolvimento infantil**. Conteúdo: v. 1. Simbolismo e jogo. Porto Alegre: Prodil, 1994.

OLIVEIRA, Elaine Cecília Lima. **O uso do software Scratch no Ensino Fundamental: possibilidades de incorporação curricular segundo professoras dos anos iniciais**. Belo Horizonte, 2009, p. 106. Dissertação (Mestrado Acadêmico em Educação). Pontifícia Universidade Católica de Minas Gerais.

SOUZA, C.; **Círculos de Cultura Infantil: O método Paulo Freire na Alfabetização de Crianças: Um Estudo Aproximativo Socioconstruivista**. Biblioteca Digital de Teses e Dissertações, p. 34, 2006.

## OTIMIZAÇÃO GEOMÉTRICA DAS PÁS DE UMA TURBINA EÓLICA DE EIXO HORIZONTAL

Data de aceite: 01/04/2021

Data de submissão: 05/01/2021

**Rafael Romão da Silva Melo**

Universidade Federal de São João del-Rei  
Departamento de Ciências Térmicas e dos  
Fluidos (DCTEF)  
São João del-Rei – MG  
<http://lattes.cnpq.br/3435986407796162>

**RESUMO:** Uma das etapas em busca da melhor eficiência energética da turbina é a análise do escoamento em torno de suas pás, onde a mínima variação na geometria das mesmas ocasiona uma mudança na potência gerada. Este trabalho apresenta a otimização de uma turbina eólica de eixo horizontal, utilizando modelo integral através da Teoria do Elemento de Pá e Teoria do Balanço da Quantidade de Movimento, abordando o problema em um volume de controle no qual a turbina eólica está imersa, fazendo sobre este volume um balanço global da energia, calculando a potência gerada a partir de dados como a velocidade do vento, velocidade angular e o perfil da pá do rotor. As técnicas de otimização utilizadas foram o Algoritmo Genético o método da Evolução Diferencial, ambos métodos heurísticos. Foi realizada a comparação dos resultados utilizando os dois métodos, e o método da Evolução diferencial apresentou melhor eficiência, atingindo uma melhor solução para a função objetivo, ou seja, uma turbina eólica com maior eficiência energética.

**PALAVRAS - CHAVE:** Otimização, Turbinas Eólicas, Evolução Diferencial, Algoritmo Genético, Formulação Integral.

### GEOMETRIC OPTIMIZATION OF THE BLADES OF A HORIZONTAL AXIS WIND TURBINE

**ABSTRACT:** One of the steps in search of the turbine's best energy efficiency is the analysis of the flow around its blades, where the minimum variation in their geometry causes a change in the power generated. This work presents a horizontal axis wind turbine optimization, using an integral formulation through the Blade Element Theory and the Balance of Linear Momentum, solving mathematically the problem in a control volume in which the wind turbine is immersed, doing a global energy balance, calculating the power generated from data such as wind speed, angular speed, and the rotor blade profile. The optimization techniques used were the Genetic Algorithm and the Differential Evolution, both heuristic methods. The results were compared using the two methods, and the Differential Evolution method showed better efficiency, achieving a better solution for the objective function, that is, a wind turbine with greater energy efficiency.

**KEYWORDS:** Optimization, Wind Turbines, Differential Evolution, Genetic Algorithm, Integral Formulation.

## 1 | INTRODUÇÃO

Uma das etapas em busca da melhor eficiência energética da turbina é a análise do escoamento em torno de suas pás, onde a

mínima variação na geometria das mesmas ocasiona uma mudança na potência gerada. Para resolver este problema de mecânica dos fluidos pode-se decidir por dois caminhos, o modelo teórico e o modelo experimental.

A experimentação em laboratório tem a vantagem de tratar com a configuração aproximada do real ou um modelo idêntico ao real, porém, geralmente, um experimento possui altíssimo custo e apresenta grande dificuldade de reprodução das condições reais, como por exemplo, simulações a grandes altitudes e velocidades variadas. Este elevado custo existe pela necessidade de se investir em um laboratório adequado que atenda às necessidades mínimas para testes, e também a necessidade de se produzir um novo protótipo a cada modelo projetado. Na ausência de modelos matemáticos estabelecidos e em geometrias extremamente complexas, modelo experimental é, muitas vezes, a única alternativa disponível ao projetista (Maliska, 1995).

Para análise teórica de tal problema pode-se escolher entre dois métodos: o método integral e o método diferencial. O método integral aborda o problema em um volume de controle no qual a turbina eólica está imersa, fazendo sobre este volume um balanço global da energia, calculando a potência gerada a partir de dados como a velocidade do vento, velocidade angular e o perfil da pá do rotor (Melo e Silveira-Neto, 2012). A partir de não se conhecer com detalhe o escoamento em cada ponto espacial em torno do gerador, o modelo integral pode fornecer resultados satisfatórios de maneira rápida.

O método diferencial é uma ferramenta poderosa, a qual nos permite simular, visualizando de fato o que ocorre em torno de um corpo imerso em um meio fluido, neste caso uma turbina eólica em meio a uma corrente de ar. Para isto utiliza-se a dinâmica dos fluidos computacional, do inglês Computational Fluid Dynamics (CFD), a qual é a área da computação científica que estuda métodos computacionais para simulação de fenômenos que envolvem fluidos em movimento com ou sem troca de calor (Fortuna, 2000).

Na busca da otimização de uma turbina eólica o método diferencial pode ser desvantajoso, uma vez que a cada nova geometria é necessária uma nova simulação, onde dependendo do refinamento da malha pode acarretar em um elevado custo computacional. Neste ponto o método integral pode ser uma excelente alternativa, gerando resultados relativamente precisos de maneira rápida e sem muito custo.

Este trabalho apresenta a otimização de uma turbina eólica de eixo horizontal, utilizando modelo integral através da Teoria do Elemento de Pá e Teoria do Balanço da Quantidade de Movimento. As técnicas de otimização utilizadas foram o clássico método heurístico Algoritmo Genético, proposto por Holland (1975), e o método da Evolução Diferencial, outro método heurístico proposto por Storn e Price (1996).

## 2.1 FORMULAÇÃO INTEGRAL DE UMA HAWT

Nesta seção será apresentada uma formulação integral para uma HAWT. O objetivo final é determinar a potência gerada pelo rotor, e para isto é necessário determinar as forças que agem nas pás. Nesta formulação são utilizadas duas metodologias, a Teoria do Elemento de Pá (Blade Element Theory, ou BEM (Morcos, 1994; Kishinami et al., 2005; Lanzafame e Messina, 2010) e o Balanço da Quantidade de Movimento (Çengel e Cimbala, 2007). Para cada uma destas formulações é apresentado o modelo físico e o modelo matemático, e em seguida o acoplamento entre as duas teorias, fechando a formulação.

### 2.1 Teoria do elemento de pá (BEM)

A teoria do elemento de pá, BEM, é uma forma de determinar as forças que agem nas pás da turbina, sendo elas a força de sustentação  $F_l$  e de arrasto  $F_d$ . Para isto a pá é dividida em um número discreto de elementos, ou planos, nos quais as forças são determinadas. Nessa teoria admite-se que o escoamento seja bidimensional sobre cada elemento, ou seja, existe apenas as componentes de velocidade normal (direção do escoamento) e tangencial (direção de rotação da turbina), sendo a componente de velocidade na direção radial nula.

A Figura 1 apresenta a pá de uma HAWT de raio  $R$ , cujo elemento de pá genérico possui raio  $r$ . Ao se aproximar deste elemento de pá a velocidade do vento livre  $U_\infty$  reduz a uma velocidade menor, denominada  $U_n$ . Ao passar pela pá a velocidade da corrente passa a ser  $U_2$ . Este elemento ainda é atingido por uma parcela de velocidade tangencial  $U_t$  devida a rotação da turbina, que gira com velocidade angular  $\omega$ . Esta figura ainda mostra que cada elemento de pá possui comprimento de corda  $c$  e ângulo de passo  $\lambda$  diferentes, sendo ângulo de passo o ângulo formado pela corda do elemento e a direção tangencial.

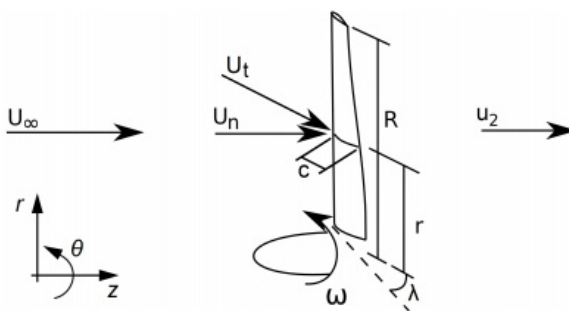


Figura 1: Características geométricas de uma pá de uma HAWT e um plano genérico para análise utilizando a BEM.

A **Figura 2** apresenta um diagrama representativo de um plano de um elemento de pá. Esta figura mostra que este elemento de comprimento de corda  $c$  e um ângulo de

passo  $\lambda$  é atingido por duas parcelas de velocidade, normal  $U_n$  e tangencial  $U_t$ . A velocidade resultante  $V$  sobre o elemento é a soma vetorial das duas velocidades. O ângulo de ataque  $\alpha$  é formado entre o vetor da velocidade resultante  $V$  e a linha da corda  $c$ . Outra variável importante apresentada neste diagrama é o ângulo de fluxo  $\phi$ , formado entre o vetor velocidade resultante  $V$  e a direção tangencial. Por fim esta figura mostra a força de arrasto  $dF_d$  na direção da velocidade  $V$ , e de sustentação  $dF_l$  perpendicular à força de arrasto.

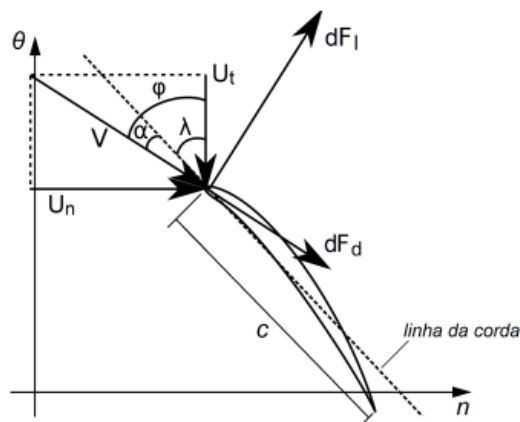


Figura 2: Diagrama de velocidades e forças que agem em um elemento de pá numa posição genérica.

Desconsiderando a presença da geometria pode-se considerar que a velocidade normal  $U_n$  é igual a velocidade de corrente livre  $U_\infty$ , e a velocidade tangencial  $U_t$  é igual a velocidade  $\omega r$  proveniente da rotação da turbina. Porém a presença da turbina interfere nas velocidades, onde  $U_n$  diminui devido a desaceleração da corrente livre, e  $U_t$  aumenta devida a transferência de momento angular da turbina para o fluido. Define-se então dois fatores de interferência: fator de interferência axial  $a$ , proporcional a  $U_\infty$ ; e fator de interferência tangencial  $a_t$ , proporcional a  $\omega r$ . Quanto maior for  $a$ , menor será  $U_n$ , e quanto maior  $a_t$  maior será  $U_t$ . Conhecendo as velocidades corretas,  $U_n$  e  $U_t$ , calcula-se então as forças  $dF_l$  e  $dF_d$  exatas. A formulação matemática para a BEM pode ser encontrada em Melo (2018).

Com a teoria do elemento de pá conseguimos calcular as forças em todos os elementos de pá, ou seja, todas as forças que agem na turbina. Porém os fatores de interferência não são conhecidos, então se faz necessário buscar uma nova alternativa, a qual será apresentada na próxima subseção.

## 2.2 Balanço da quantidade de movimento

Outra forma de calcular as forças normal e tangencial em cada ponto das pás, é através da teoria do balanço de quantidade de movimento em um volume de controle determinado. A Figura 3 apresenta um volume de controle no formato de um anel, onde o raio do rotor é dividido em um número finito de elementos de comprimento  $dr$ . Este comprimento percorre uma seção circular, como visto na Figura 3.a, englobando os elementos de pá da teoria BEM de todas pás no raio genérico  $r$ , onde o diferencial de força normal  $dF_n$  e tangencial  $dF_t$  que atua nesse volume de controle possui o mesmo valor dos diferenciais de força da teoria anterior (BEM). Este volume de controle na direção normal não é cilíndrico, mas segue uma curva delimitada por duas superfícies de corrente, fechando este volume de controle, como visto na Figura 3.b. A entrada de fluido neste volume de controle se inicia no plano 1, onde a velocidade do escoamento é a velocidade da corrente livre  $U_\infty$ , e a pressão é a pressão atmosférica. No plano 2 a velocidade do escoamento é uniforme e igual a  $u_2$ , e a pressão também é igual a pressão atmosférica. Como a pressão na entrada e na saída são iguais, não existe uma força de pressão resultante sobre este volume. Para determinar os diferenciais de força normal  $dF_n$  e tangencial  $dF_t$ , aplica-se o balanço da quantidade de movimento nas duas direções. A formulação matemática para o balanço da quantidade de movimento linear e o acoplamento desta teoria com a Teoria do elemento de Pá também pode ser encontrada em Melo (2018).

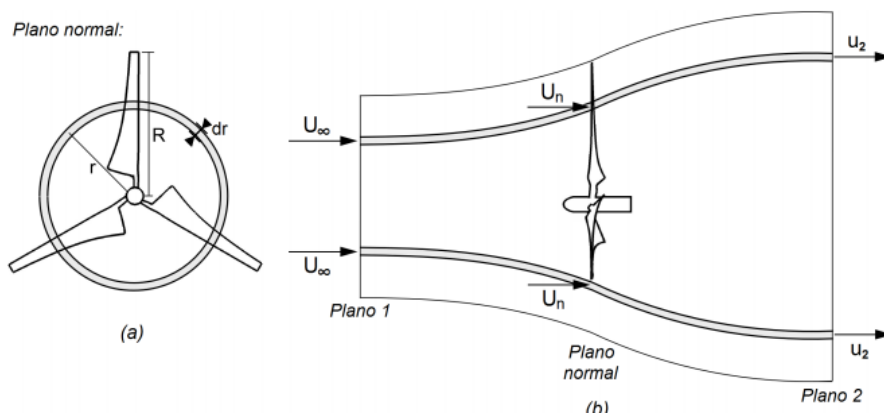


Figura 3: Volume de controle em um anel anular

Nas duas subseções seguintes são apresentados os métodos de otimização utilizados, Algoritmo Genético e Evolução Diferencial, respectivamente.



## 2.3 Algoritmo Genético

O método de otimização Algoritmo Genético, proposto por Hollan (1975), baseiam-se no mecanismo de evolução natural das espécies: sobrevivência do mais apto. Toda a informação sobre um indivíduo (todas suas características) está contida em seus genes. Em uma população, é mais provável que: o mais apto sobreviva, os mais fortes passam os seus genes para a próxima geração. Cada indivíduo corresponde a uma possível solução para o problema.

Os operadores genéticos básicos são:

- **Seleção:** Nesta etapa ocorre a escolha dos pais que vão gerar novos descendentes. Existem vários métodos, dentre eles os métodos do Truncamento, Ranking, Roleta, Amostragem Universal Estocástica, Torneio, entre outros.
- **Crossover:** Nesta etapa ocorre a mistura de dados genéticos dos dois pais selecionados na etapa de seleção. Estas informações dependem dos dados de cada indivíduo, sendo que pode haver a troca de vetores de binários, vetores de reais, vetor de inteiros (permutação) ou também árvores.
- **Mutação:** Nesta etapa ocorre a modificação de um dado gene de um indivíduo, dentro de uma certa probabilidade. Pode ser uma modificação aleatória, ou ainda a troca de informações entre dois genes de duas posições diferentes no mesmo indivíduo.

A seguir é apresentado a idéia básica de execução de um algoritmo genético:

- **ETAPA 1:** Gera-se a População Inicial com TP indivíduos;
- **ETAPA 2:** Avaliam-se todos os indivíduos da População;
- **ETAPA 3:** Selecionam-se os indivíduos que formarão os pares do crossover;
- **ETAPA 4:** Para cada par, sorteia-se um Ponto de Crossover e 2 novos indivíduos são gerados por crossover;
- **ETAPA 5:** Submetem-se os novos indivíduos a uma mutação com probabilidade de mutação;
- **ETAPA 6:** Avalia-se o desempenho dos novos indivíduos formados após Crossover e Mutação;
- **ETAPA 7:** Selecionam-se os TP melhores indivíduos dentre a população original e os novos indivíduos;
- **População atingiu Critério de Término?** Se SIM, FIM. Se NÃO, eetornar a ETAPA 3.

## 2.4 Evolução Diferencial

A Evolução Diferencial (Storn e Price, 1996) foi apresentada como uma versão melhorada do Algoritmo Genético, resolvendo de forma mais eficiente problemas de otimização. A ideia do método é simples, gera-se um novo através da adição da diferença ponderada entre dois indivíduos a um terceiro indivíduos. A ponderação previne a criação de indivíduos idênticos. Acriação deste novo individuo é considerado como uma mutação. Este individuo mutado, donor, é combinado com um individuo da população, target, gerando um individuo teste, trial, através de um cruzamento (crossover). Este cruzamento aumenta a diversidade das soluções mutadas. Caso o individuo teste apresente melhor resultado que o individuo target, este é substituído pelo individuo recém criado. Um detalhe importante é que em cada iteração do método todos os indivíduos da população são indivíduos target uma única vez, obrigatoriamente. A Figura 4 apresenta esquematicamente a evolução diferencial aplicado a um individuo da população.

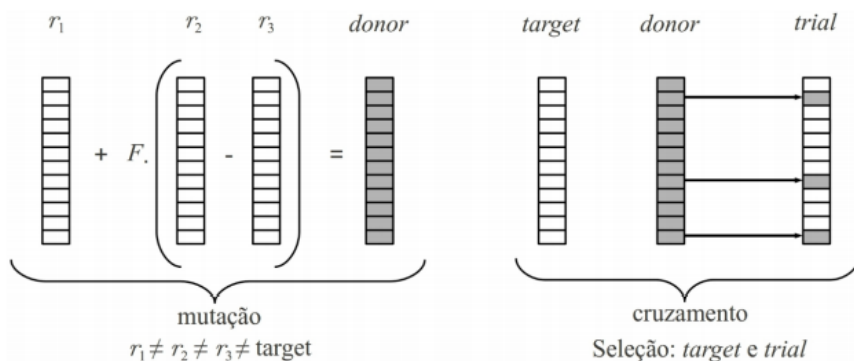


Figura 4: Mutação, cruzamento e seleção da evolução diferencial (Lopes et al., 2013).

Uma das vantagens da evolução diferencial é a necessidade de poucos parâmetros de controle para os operadores, os quais têm seus valores ajustados de maneira relativamente simples (Lopes et al., 2013).

A seção seguinte serão mostrados os resultados aplicando este método a formulação integral do escoamento em torno de uma turbina eólica.

## 3 | RESULTADOS E DISCUSSÃO

Será considerado uma turbina com raio constante  $R = 5\text{m}$ , onde as pás serão divididos em  $N$  partes,  $N$  elementos de pá. Cada elemento de pá possui um comprimento de corda  $c$  e ângulo de passo  $\lambda$ . Busca-se o melhor valor para estas duas variáveis de cada elemento para gerar a maior potência possível. Assim as variáveis de projeto são um

vetor de tamanho  $N$  para comprimento de corda e outro valor de tamanho  $N$  para ângulo de passo de passo.

A Figura 5 apresenta um indivíduo inicial simples, onde o comprimento de corda e o ângulo de passo são funções lineares. Toda a população de tamanho  $Z$  foi criada a partir deste indivíduo, variando o comprimento de corda em  $\pm 80\%$  e o ângulo de passo em  $\pm 20\%$  aleatoriamente. Foram utilizados 16 elementos de pá, resultando em 32 variáveis de projeto.

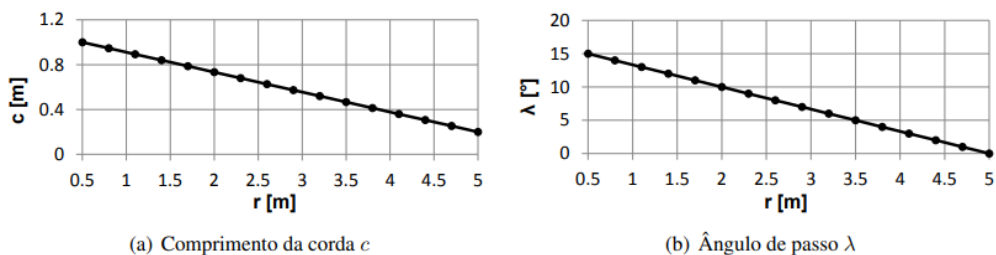


Figura 5: Dados geométrico do indivíduo inicial.

A Figura 6 apresenta a curva do coeficiente de potência em função da razão de velocidade  $t_{sr}$  para o indivíduo inicial. Esta figura auxiliara também a compreender como avaliar a função objetivo. Se uma dada turbina eólica possuir um controle ótimo de rotação a função objetivo poderia ser o ponto máximo da curva  $C_p \times t_{sr}$ , uma vez que o sistema de controle pode ajustar a rotação do rotor até a  $t_{sr}$  atingir este ponto máximo. Porém se uma turbina possuir um sistema de controle não eficaz, ou que possua certa tolerância, ou ainda nem possua sistema de controle pode-se trabalhar com uma faixa de operação de  $t_{sr}$ . Neste caso a função objetivo pode ser determinada pela área do gráfico abaixo da curva limitado lateralmente pela faixa de operação, como apresentado nesta figura.

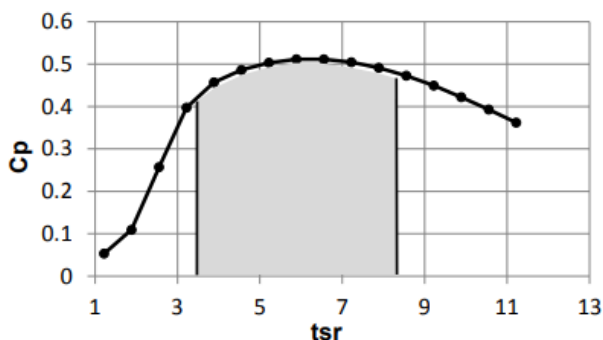


Figura 6: Coeficiente de potência gerado pelo indivíduo inicial.

Neste trabalho foi definida uma faixa de operação de  $t_{sr}$  igual a 3, onde o ponto central desta faixa está justamente no ponto da curva de maior coeficiente de potência. Para encontrar este ponto de eficiência máxima, e então determinar de onde até onde varia a  $t_{sr}$  foi utilizado o método da seção áurea.

Os parâmetros de entrada dos dois métodos de otimização apresentados neste trabalho foram: população de diferentes tamanhos, sendo  $Z = 5, 10, 15, 20, 30, 40, 50$  e  $100$ , sendo que a quantidade de filhos é a mesma quantidade de pais no algoritmo genético; probabilidade de cruzamento de 50% para ambos os métodos; a constante de ponderação aplicado na subtração de dois indivíduos na evolução diferencial de  $F = 0.8$ .

O perfil aerodinâmico utilizado foi o NACA 23018, e os coeficientes de arrasto e sustentação foram obtidos através do programa de domínio livre XFOIL, desenvolvido por Drela (1989).

Algumas restrições foram impostas: O comprimento de corda de cada elemento de pá deve ser maiores que 0 e menores que 2.0m, e o ângulo de passo deve ser inferior a  $45^\circ$ . Caso alguma variável ultrapasse estes limites ela é automaticamente atualizada com o valor limitante. Outra restrição é pela geometria, por exemplo, a corda na ponta da pá não pode ser maior que a corda da base, se assim fosse a turbina seria irrealista. Para contornar este problema o vetor de comprimentos de corda e o vetor de ângulo de passo são ordenados em ordem crescente após a criação da população inicial e após cada cruzamento.

A Tabela 1 apresenta os resultados aplicando os métodos de otimização para os vários números de população. O critério de parada utilizado foi o erro entre a média da função objetivo de toda a população entre uma iteração e outra, com um resíduo de  $10^{-10}$ . Nota-se que a melhor solução encontrada com algoritmo genético foi utilizando o maior número de população, igual a 100. Observa-se também que apesar da evolução diferencial levar mais iterações até atingir o critério de parada, a função objetivo final é sempre superior à encontrada utilizando A. G. A explicação para o maior número de iterações da evolução diferencial é que a população no algoritmo genético se torna uniforme com mais rapidez, atingindo a função objetivo mais rápido, porém em um valor mais baixo. Já as combinações algébricas realizadas na evolução diferencial resultam em indivíduos novos, alcançando com mais facilidade uma melhor função objetivo.

População	Função Objetivo (melhor indivíduo)		Num. Iterações		Tempo Computacional (s)	
	A. G.	E. D.	A. G.	E. D.	A. G.	E. D.
5	1,4645	1,5133	3	15	0,1701	0,4081
10	1,5547	1,5789	25	69	1,3404	3,5664
15	1,5650	1,5785	27	73	2,3041	5,6930
20	1,5603	1,5798	35	99	3,7086	10,2893
30	1,5683	1,5809	31	85	5,4280	12,8516
40	1,5676	1,5816	36	128	7,3492	25,2577
50	1,5699	1,5815	42	128	10,7619	31,8600
100	1,5706	1,5821	49	222	24,4563	106,9315

Tabela 1: Comparação dos resultados obtidos pelos metodos de otimização utilizando critério de parada.

A Tabela 2 também apresenta o comparativo dos dois métodos de otimização, porém agora com um número fixo de iterações,  $N = 150$ . Para cada teste com diferentes números de população o tempo computacional para ambos os métodos foram próximos. Contudo o valor da função objetivo final encontrado utilizando evolução diferencial sempre foi superior ao encontrado utilizando algoritmo genético.

População	Função Objetivo (melhor indivíduo)		Tempo Computacional (s)	
	A. G.	E. D.	A. G.	E. D.
5	1,4645	1,5697	4,3455	3,8520
10	1,5547	1,5807	7,1076	7,7580
15	1,5650	1,5804	11,5730	11,7831
20	1,5603	1,5807	15,0398	15,5690
30	1,5683	1,5818	24,4366	22,3343
40	1,5676	1,5818	29,4074	29,3016
50	1,5699	1,5816	36,5354	37,2196
100	1,5706	1,5819	73,0528	73,5668

Tabela 2: Comparação dos resultados obtidos pelos métodos de otimização utilizando número de iterações  $N = 150$

A Figura 7 apresenta a evolução da média função objetivo entre toda população utilizando os dois métodos de otimização, com população de tamanho  $N = 20$ . É possível notar que o método algoritmo genético atinge a função objetivo máxima mais rapidamente, constatando que sua população se torna uniforme também mais rápido, enquanto que utilizando evolução diferencial a evolução vai caminhando para uma melhor solução de uma maneira mais simples, somente com operações algébricas como mostrado em teoria, e com a população sempre diversificada.

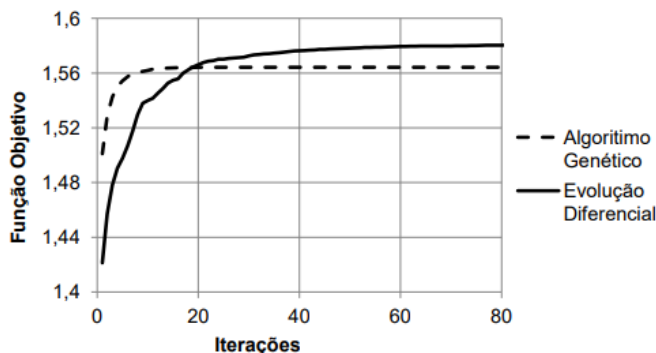


Figura 7: Evolução da média função objetivo entre toda população de tamanho  $N = 20$

Por fim, a Figura 8 apresenta a comparação do comprimento de corda  $c$  e do ângulo de passo  $\lambda$  do indivíduo inicial e do indivíduo que gera a melhor solução, utilizando evolução diferencial. Nota-se que a solução otimizada respeita as restrições impostas ao problema, as características geométricas estão em conformidade com as das turbinas fabricadas até então, e que o método se mostrou eficiente, buscando o ponto ótimo de cada elemento de pá.

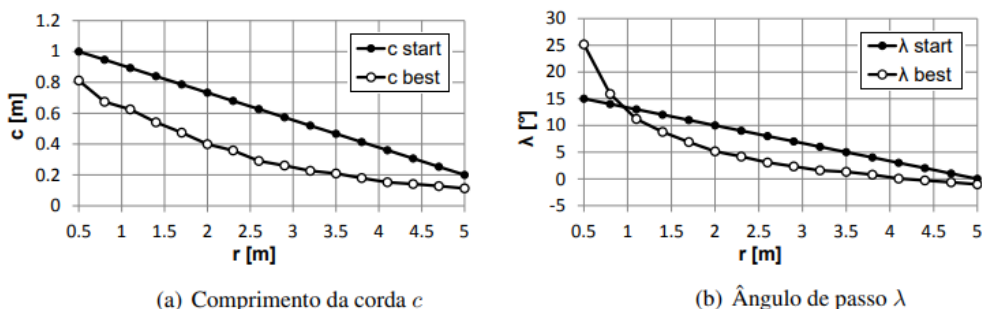


Figura 8: Dados geométrico do melhor indivíduo e do indivíduo inicial.

A seguir serão apresentados as conclusões e trabalhos futuros

## 4 | CONCLUSÕES E TRABALHOS FUTUROS

A energia eólica é parte da solução para a demanda de energia do mundo. A análise computacional aplicada em turbinas eólicas é notavelmente versátil, uma vez que não existem gastos com laboratórios e protótipos como em análises experimentais, e ainda é possível facilmente alterar as variáveis de projeto em busca da melhor solução para um problema específico.

Com a formulação integral sobre rotores eólicos é possível determinar com facilidade a eficiência da turbina, uma vez que é possível calcular as forças que agem nas pás do rotor em função das variáveis de entrada do modelo, as quais são velocidade do vento, velocidade de rotação da turbina, comprimento da pá e tipo de perfil aerodinâmico.

O método da evolução diferencial, aplicado a modelagem integral utilizada neste trabalho, se mostrou mais eficiente que o método algoritmo genético na busca da melhor solução para condições preestabelecidas, respeitando as restrições de projeto encontrando o melhor indivíduo de maneira rápida e eficiente.

Como trabalhos futuros pode-se melhorar o modelo integral, incluído, por exemplo, efeitos de perda de ponta da pá, e ainda uma maneira eficiente de obter os coeficientes de arrasto e sustentação para um perfil genérico. Com relação ao método de otimização pode-se incluir mais objetivos, como projetar uma turbina com raio de tamanho variado, com restrições como tensão máxima permitida, e ainda minimizar a quantidade de material utilizado na fabricação.

## REFERÊNCIAS

Betz, A., 1928. Windmills in the Light of Modern Research, Technical Report, National Advisory Committee for Aeronautics, Washington, EUA

Camporeale, S. M., Magi, V., 1999. Streamtube model for analysis of vertical axis variable pitch turbine for marine currents energy conversion. Energy Conversion and Management.

Degraire, P., 2010. Analytical Aerodynamic Simulation Tool for Vertical Axis Wind Turbines, Tese de Doutorado, Uppsala Universitet.

Drela, M., 1989. Xfoil: An analysis and design system for low reynolds number airfoils, Lecture Notes in Engineering: Low Reynolds Number Aerodynamics.

Fortuna, A. O., 2000. Técnicas Computacionais Para Dinâmica dos Flúidos: Conceitos Básicos e Aplicações, EDUSP.

Freitas, A. W., 2008. Controlo de Passo de Aero geradores de Baixa Potência, Dissertação de Mestrado, Faculdade de Engenharia da Universidade do Porto.

- Holland, J. H., 1975. *Adaptation in natural and artificial systems*. The University of Michigan Press, Ann Arbor, MI.
- Janajreh, I., Qudaih, R., Talab, I., Ghenai, C., 2010. Aerodynamic flow simulation of wind turbine: Downwind versus upwind configuration, *Energy Conversion and Management*.
- Kishinami, K., Taniguchi, H., Suzuki, J., Ibano, H., Kazunou, T., Turuhami, M., 2005. Theoretical and experimental study on the aerodynamic characteristics of a horizontal axis wind turbine, *Energy*, v. 30 pp. 2089 - 2100.
- Kjellin, J., Bulow, F., Eriksson, S., Deglaire, P., Leijon, M., Bernhoff, H., 2010. Power coefficient measurement on a 12 kw straight bladed vertical axis wind turbine, *Renewable Energy*.
- Lanzafame, R., Messina, M., 2010. Horizontal axis wind turbine working at maximum power coefficient continuously, *Renewable Energy*, v. 35 pp. 301 - 306.
- Lopes, H. S., Rodrigues, L. C. A., Steiner, M. T. A., 2013. *Meta-Heurísticas em Pesquisa Operacional*, Omnipax, Curitiba, PR, 1 edition.
- Maliska, C., 1995. *Transferência de Calor e Mecânica dos Fluidos Computacional*, LTC.
- Melo, R. R. S., Silveira-Neto A., 2012. Integral analysis of rotors of a wind generator, *Renewable and Sustainable Energy Reviews*, v. 16, n 7, pp. 4809-4817.
- Melo, R. R. S. 2018. *Determinação da eficiência de turbinas eólicas: teoria do elemento de pá e balanço de quantidade de movimento*. 1. ed. Uberlândia: Edibrás. v. 1. 62p .
- Morcos, V., 1994. Aerodynamic performance analysis of horizontal axis wind turbines, *Renewable Energy*, v. 4 pp. 505 - 518.
- Sharma, R. N. e Madawala, U. K., 2011. The concept of a smart wind turbine system, *Renewable Energy*.
- Sheldahl, R. E., Klimas, P. C., Feltz, L. V., 1980. *Aerodynamic Performance of a 5 Metre Diameter Darrieus Turbine With Extruded Aluminium NACA 0015 Blades*, Technical Report SAND 80 0179, EUA, Sandia National Laboratories, Albuquerque, NM, EUA.
- Storn, R., Price, K., 1996. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces, *Journal of Global Optimization*, v. 11 pp. 341 - 359.
- Çengel, Y. A., Cimbala, J. M., 2007. *Mecânica dos Fluidos Fundamentos e Aplicações*, Mc-Graw Hill Interamericana do Brasil Ltda.



## OTIMIZAÇÃO DO SPRINT BACKLOG COM O PROBLEMA DA MOCHILA 0/1

Data de aceite: 01/04/2021

Data de submissão: 05/01/2021

### Michel Willian Alves

Universidade do Estado do Rio Grande do  
Norte (UERN)  
Mossoró – RN  
<http://lattes.cnpq.br/7291561381054505>

### Elisa de Fátima Andrade Soares

Universidade Federal de Pernambuco (UFPE)  
Recife – PE  
<http://lattes.cnpq.br/6163215496349667>

### Thalia Katiane Sampaio Gurgel

Universidade do Estado do Rio Grande do  
Norte (UERN)  
Mossoró – RN  
<http://lattes.cnpq.br/8112466614884045>

### José Weliton de Vasconcelos Filho

PATHFIND Soluções de Tecnologia em  
Logística  
Fortaleza - CE  
<http://lattes.cnpq.br/3783191885590629>

### Dario José Aloise

Universidade do Estado do Rio Grande do  
Norte (UERN)  
Mossoró – RN  
<http://lattes.cnpq.br/7266011798625538>

**RESUMO:** O *Scrum* é um framework de gerenciamento ágil, que possui, entre seus propósitos, a divisão do desenvolvimento de sistemas organizada em ciclos de tempo, não

superior a 1 mês, chamados de *Sprints*. O *Sprint Backlog* é uma lista de itens (requisitos), cada um com um determinado tempo de desenvolvimento, que precisam ser selecionados da melhor forma possível para ser desenvolvida dentro de uma *Sprint*. Neste contexto, este trabalho apresenta uma modelagem matemática para o cenário de um ciclo de desenvolvimento do *Scrum* como um Problema da Mochila 0/1 para alocar, de maneira ótima, os itens do *Sprint Backlog*, utilizando um fator de priorização para os requisitos. A partir desta modelagem o problema foi implementado em C++ e solucionado através do *solver* Gurobi Optimizer. Para tanto, o modelo foi aplicado às instâncias construídas de forma aleatória, simulando situações reais, de tamanho variando entre 30 a 150 itens, o qual retornou para cada instância uma solução em tempo de execução satisfatório.

**PALAVRAS - CHAVE:** Otimização Combinatória. *Sprint Backlog*. Problema da Mochila. NP-Difícil.

### SPRINT BACKLOG OPTIMIZATION WITH THE 0/1 KNAPSACK PROBLEM

**ABSTRACT:** Scrum is an agile management framework, which has, among its purposes, the division of systems development organized in time cycles, not exceeding 1 month, called Sprints. The Sprint Backlog is a list of items (requirements), each with a specific development time, that need to be selected in the best possible way to be developed within a Sprint. In this context, this work presents a mathematical modeling for the scenario of a development cycle of Scrum as a 0/1 Knapsack Problem to allocate, in an optimal way, the items of the Sprint Backlog,

using a prioritization factor for the requirements. Based on this modeling, the problem was implemented in C ++ and solved using the Gurobi Optimizer solver. For this purpose, the model was applied to the instances constructed at random, simulating real situations, with sizes ranging from 30 to 150 items, which returned a satisfactory execution time solution for each instance.

**KEYWORDS:** Combinatorial Optimization. Sprint Backlog. Backpack problem. Np-hard.

## 1 | INTRODUÇÃO

Assumindo a atual tendência do mundo competitivo em que vivemos, a habilidade de possuir um planejamento eficaz nas empresas se torna cada vez mais importante. Visto que, um planejamento e um método de gestão escolhido de forma incompetente pode prejudicar todo segmento de uma empresa ou *startup*. Assim pode acarretar na maximização dos custos avaliados, bem como atrasos na entrega do produto final a um cliente. Desta forma ocasiona pontos negativos para reputação da empresa.

O gerenciamento de projeto exige rapidez e demanda das equipes de desenvolvimento, uma boa organização dos projetos com finalidade de atingir os objetivos de produção com menor tempo e com a melhor qualidade. Além do mais, necessita de agilidade, qualidade e eficiência dos softwares (VITIELLO, 2001). Portanto, o fato de não percebermos que o planejamento é um item básico para a obtenção de resultados satisfatórios, resulta em processos indesejados.

De acordo com Meredith e Mantel (2000), a área de engenharia de software traz diversas discussões a respeito dos fracassos resultantes da omissão ou má aplicação de metodologias em torno do desenvolvimento de software.

Desta forma, este trabalho utiliza o método de gestão ágil *Scrum*, juntamente com o Problema da Mochila 0/1 para otimizar um *Sprint Backlog*, de modo que ocorra um impacto no processo de desenvolvimento de um projeto, visando melhorias e obtendo uma escolha confiável na hora de desenvolver os requisitos demandados pelos clientes a fim de obter-se um produto final.

Este trabalho está organizado da seguinte maneira: a Seção 2 apresenta conceitos sobre o *framework Scrum*. A Seção 3 descreve o problema. **A Seção 4** apresenta a formulação matemática do problema como um Problema da Mochila 0/1. A Seção 5 mostra como os testes computacionais foram realizados e descreve como foram construídas as instâncias para validação do modelo matemático proposto. Por fim, resultados e conclusões são expostos nas seções 6 e 7, respectivamente.

## 2 | SCRUM

Segundo Schwaber e Beedle (2004), o *Scrum* **não é um processo previsível, ele não define o que fazer em toda circunstância. Ele é usado em trabalhos complexos**

nos quais não é possível prever tudo o que irá ocorrer e oferece um *framework* e um conjunto de práticas que torna tudo visível. Isso permite aos praticantes do *Scrum* saber exatamente o que está acontecendo ao longo do projeto e fazer os devidos ajustes para manter o projeto se movendo ao longo do tempo visando alcançar os seus objetivos.

*Sprint* representa um ciclo de trabalho no *Scrum*. Esse ciclo pode ser de qualquer tamanho, mas normalmente é de 2 ou 4 semanas, que é o *Time box* das *Sprints*. As *Sprints* devem ter sempre a mesma duração. A cada *Sprint* um conjunto de requisitos (histórias de usuários) é implementado, tendo como resultado um incremento do produto que está sendo desenvolvido.

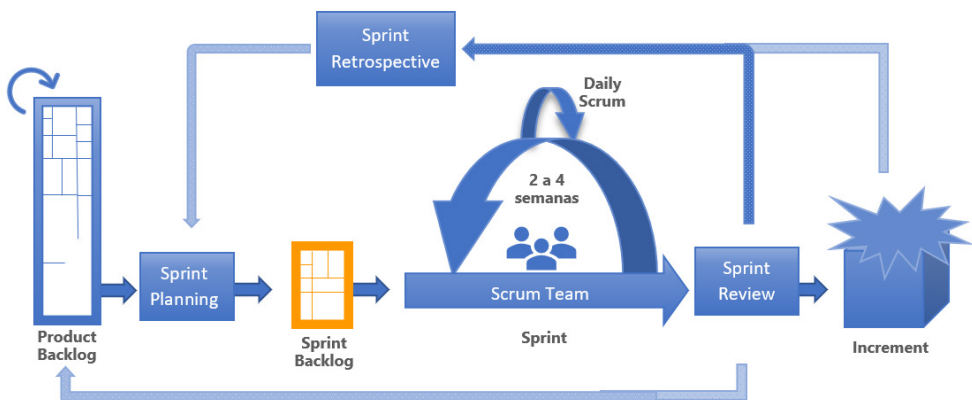


Figura 1 - Ciclo de desenvolvimento do Scrum

Fonte: Adaptada de Scrum.org (2020).

O *Sprint Planning Meeting* é uma reunião na qual estão presentes o Product Owner, o Scrum Master e todo o Scrum Team, bem como qualquer pessoa interessada que esteja representando a gerência ou o cliente.

Durante o *Sprint Planning Meeting*, o Product Owner descreve as funcionalidades de maior prioridade para a equipe. A equipe faz perguntas durante a reunião de modo que seja capaz de quebrar as funcionalidades em tarefas técnicas, após a reunião. Essas tarefas irão dar origem ao *Sprint Backlog*.

O *Sprint Backlog* é uma lista de tarefas que o Scrum Team se compromete a fazer em um *Sprint*. Os itens do *Sprint Backlog* **são extraídos do Product Backlog**, pela equipe, com base nas prioridades definidas pelo Product Owner e a percepção da equipe sobre o tempo que será necessário para completar as várias funcionalidades.

O termo *Product Backlog* diz respeito a todas as necessidades que o projeto demanda, as quais devem ser postas em ordem de acordo com a sua prioridade. Durante a reunião, determina-se quais dessas instâncias ou histórias serão incluídas no ciclo, essas

são chamadas de *Sprint Backlog*.

No *Planning Poker*, segundo Carvalho (2019) para estimar é necessário do *Product Backlog* e de um baralho de cartas, onde as cartas devem ter os seguintes números 0, ½, 1, 2, 3, 5, 8, 13, 20, 40 e 100, ou seja, semelhante à sequência de Fibonacci.

Cada membro da equipe deve possuir esse baralho de cartas, pois cada item do *Product Backlog* corresponde a um valor definido através de rodadas entre os membros.

A cada funcionalidade do *Product Backlog*, os membros relacionam uma carta com o valor que acham ser o ideal. Após todos os membros jogarem, é discutido qual o valor ideal para aquele item, caso a equipe não chegue a um consenso, e feito mais rodadas até se chegar a um consenso, e assim, continua com os outros itens.

### 3 | DESCRIÇÃO DO PROBLEMA

De acordo com Krishnan (2019), os grandes sistemas de *software* apresentam diversos requisitos. Assim como, várias restrições, recursos limitados, restrições orçamentárias, restrição de tempo e viabilidade, dentre outras restrições. Assim as empresas precisam priorizar os requisitos.

Para priorizar a demanda de requisitos, pretende-se fazer a escolha dos requisitos que serão desenvolvidos na próxima *Sprint* de forma otimizada, utilizando uma escala de priorização com 5 níveis. Considerando o nível 5 o mais crítico e o nível 1 menos crítico.

De acordo com Brennan (2015), a priorização de requisitos é um processo de decisão usado para determinar a importância dos requisitos que pode ser de acordo com o valor relativo, ao risco, a dificuldade de implementação ou qualquer outro critério. Assim, existem 8 fatores que devem ser considerados ao organizar os requisitos através de prioridades para o desenvolvimento de sistemas:

**Benefício:** é a vantagem que os negócios acumulam como resultado da implementação de requisitos. O benefício obtido pode se referir à funcionalidade, qualidade ou objetivos estratégicos e/ou de negócios.

**Penalidade:** é a consequência de não implementar um requisito. Pode se referir à perda de multas regulatórias, baixa satisfação do cliente ou usabilidade do produto.

**Custo:** é o esforço e os recursos necessários para implementar um requisito. Um recurso pode se referir a finanças, mão de obra ou até tecnologia.

**Risco:** é a probabilidade de o requisito não fornecer o valor esperado. Isso pode ser devido a vários motivos, desde dificuldade em entender o requisito até sua implementação.

**Dependências:** é o relacionamento entre os requisitos. Como tal, um requisito exigirá a conclusão de outro requisito para sua implementação.

**Sensibilidade ao tempo:** tudo vem com um prazo de validade. É necessário mencionar a que horas o requisito expirará ou também se o requisito é sazonal.

**Estabilidade:** refere-se à probabilidade de o requisito permanecer estático.

**Conformidade com regulamentos / políticas:** os requisitos que devem ser implementados para atender aos requisitos regulamentares.

Dentre os fatores existentes, consideramos utilizar neste trabalho um fator que é apresentado na Tabela 1, com a finalidade de selecionar os requisitos a serem desenvolvidos, utilizando o critério de prioridade.

Objetivo	Prioridade	Tipo	Descrição
<b>Benefício (B)</b>	5	Muito Alto	Quanto maior o valor da prioridade mais vantajosa é a implementação do requisito.
	4	Alto	
	3	Moderado	
	2	Baixo	
	1	Muito Baixo	

Tabela 1 - Prioridade dos requisitos de acordo com o objetivo

Fonte: elaborado pelos autores

#### 4 I FORMULAÇÃO MATEMÁTICA COMO UM PROBLEMA DA MOCHILA 0/1

A partir do que foi exposto sobre o problema, é possível criar modelagens matemáticas para serem resolvidas por softwares especiais chamados de otimizadores (*solvers*). Os otimizadores fornecem uma ou mais soluções ótimas, ou seja, soluções que possuem o melhor valor possível dentre todas as soluções factíveis para o problema.

O problema como descrito acima pode ser formulado como um Problema da Mochila, mais especificamente da Mochila 0/1, para escolha dos requisitos que irão fazer parte da próxima *Sprint* a partir da tabela de prioridade informada anteriormente.

O Problema da Mochila 0/1, ou mochila binária, consiste em, dada uma mochila de capacidade  $W$  e  $n$  itens, cujos peso  $w_i$  e valor de utilidade  $c_i$  são dados, escolher quais itens serão alocados na mochila. A escolha é feita pela variável de decisão  $x_i$ , de maneira que, se o item  $i$  é alocado na mochila, então  $x_i = 1$ , caso contrário, então  $x_i = 0$ . O problema consiste em encontrar um conjunto de itens que maximize o valor de utilidade (benefício), sem ultrapassar a capacidade da mochila. O Problema da Mochila 0/1 é um problema de programação linear inteira e é classificado como NP-Difícil (GAREY; JOHNSON, 2002). O modelo matemático do Problema da Mochila 0/1 apresentado por (MARTELLO; TOTH, 1990) é como segue:

$$\text{Max } Z = \sum_{i=1}^n c_i x_i \quad (1)$$

Sujeito à:

$$\sum_{i=1}^n w_i x_i \leq W \quad (2)$$

$$x_i \in \{0,1\}; \quad i = 1, \dots, n \quad (3)$$

Na otimização do *sprint backlog* esta formulação pode ser usada com as seguintes considerações:

$Z$  : Benefício de desenvolver cada requisito;

$W$  : *Time box* da *Sprint*;

$w_i$  : Tempo de desenvolver cada requisito  $i$ , com  $i = 1, \dots, n$ ;

$n$  : Números de requisitos que podem ser incluídos ou não no *Sprint Backlog* (mochila)

$c_i$  : Valorização do objeto segundo seu benefício de desenvolvimento

$c_i$ ,  $w_i$  e  $W$  são valores inteiros positivos,  $w_i \leq W$  e  $\sum_{i=1}^n w_i > W$

$$x_i = \begin{cases} 1, & \text{se o requisito está na mochila} \\ 0, & \text{caso contrário} \end{cases}$$

Fazendo dessa maneira, tem-se em (1), a função objetivo que é maximizar a escolha dos requisitos com maior índice de prioridade de desenvolvimento. Já em (2), a restrição impede que o tempo de desenvolver todos os requisitos ultrapasse o *Time box* da *Sprint* e a restrição (3) determina o domínio das variáveis. O objetivo consiste em obter os itens (solução) que fornece o máximo benefício para a próxima *Sprint*.

O modelo matemático utilizado neste trabalho foi implementado em C++ e integrado a biblioteca do *solver* Gorubi Optimizer, utilizando os pesos e prioridade de forma padrão, para otimizar as escolhas dos requisitos que irão fazer parte da próxima *Sprint* a partir do fator de priorização escolhido.

## 5 | TESTES COMPUTACIONAIS

Para validar o modelo matemático foram criadas duas instâncias de maneira aleatória, simulando situações reais, através de uma planilha de Excel, utilizando a função “=ALEATÓRIOENTRE(1;5)” para composição dos objetivos e “=ALEATÓRIOENTRE(2;16)” para estimativa de tempo de desenvolver cada requisito. Para a primeira instância apresentada na Tabela 2 foram considerados 30 requisitos e uma *Sprint* com duração de

80 horas.

	<b>B</b> ( $c_i$ )	$w_i$		<b>B</b> ( $c_i$ )	$w_i$		<b>B</b> ( $c_i$ )	$w_i$
<b>R1</b>	3	9	<b>R11</b>	1	15	<b>R21</b>	5	3
<b>R2</b>	1	5	<b>R12</b>	1	7	<b>R22</b>	4	16
<b>R3</b>	5	10	<b>R13</b>	1	11	<b>R23</b>	1	9
<b>R4</b>	2	6	<b>R14</b>	5	16	<b>R24</b>	2	14
<b>R5</b>	3	4	<b>R15</b>	5	3	<b>R25</b>	1	7
<b>R6</b>	1	7	<b>R16</b>	2	6	<b>R26</b>	4	15
<b>R7</b>	5	14	<b>R17</b>	4	14	<b>R27</b>	5	16
<b>R8</b>	5	13	<b>R18</b>	5	5	<b>R28</b>	3	6
<b>R9</b>	5	13	<b>R19</b>	4	7	<b>R29</b>	1	11
<b>R10</b>	3	8	<b>R20</b>	4	2	<b>R30</b>	2	13

Tabela 2 - Instâncias aleatória com 30 requisitos

Fonte: elaborado pelos autores

A segunda instância, apresentada na Tabela 3, foi gerada considerando 150 requisitos e uma *Sprint* com duração de 120 horas.

	<b>B</b> ( $c_i$ )	$w_i$		<b>B</b> ( $c_i$ )	$w_i$		<b>B</b> ( $c_i$ )	$w_i$		<b>B</b> ( $c_i$ )	$w_i$		<b>B</b> ( $c_i$ )	$w_i$		<b>B</b> ( $c_i$ )	$w_i$
<b>R1</b>	1	15	<b>R26</b>	1	12	<b>R51</b>	4	11	<b>R76</b>	1	11	<b>R101</b>	4	13	<b>R126</b>	1	16
<b>R2</b>	3	4	<b>R27</b>	2	8	<b>R52</b>	5	12	<b>R77</b>	3	7	<b>R102</b>	3	4	<b>R127</b>	2	8
<b>R3</b>	1	15	<b>R28</b>	1	9	<b>R53</b>	3	8	<b>R78</b>	2	3	<b>R103</b>	5	16	<b>R128</b>	2	13
<b>R4</b>	5	15	<b>R29</b>	1	14	<b>R54</b>	2	10	<b>R79</b>	3	11	<b>R104</b>	3	9	<b>R129</b>	5	10
<b>R5</b>	1	15	<b>R30</b>	1	6	<b>R55</b>	1	16	<b>R80</b>	2	3	<b>R105</b>	1	9	<b>R130</b>	3	9
<b>R6</b>	1	14	<b>R31</b>	2	15	<b>R56</b>	1	11	<b>R81</b>	4	12	<b>R106</b>	5	6	<b>R131</b>	4	12
<b>R7</b>	2	5	<b>R32</b>	1	6	<b>R57</b>	4	12	<b>R82</b>	2	11	<b>R107</b>	2	14	<b>R132</b>	4	5
<b>R8</b>	5	6	<b>R33</b>	3	8	<b>R58</b>	3	12	<b>R83</b>	4	9	<b>R108</b>	4	5	<b>R133</b>	1	8
<b>R9</b>	2	15	<b>R34</b>	4	2	<b>R59</b>	1	3	<b>R84</b>	5	7	<b>R109</b>	3	7	<b>R134</b>	4	12
<b>R10</b>	5	7	<b>R35</b>	1	13	<b>R60</b>	4	8	<b>R85</b>	5	10	<b>R110</b>	2	13	<b>R135</b>	2	14
<b>R11</b>	4	12	<b>R36</b>	4	4	<b>R61</b>	2	6	<b>R86</b>	4	13	<b>R111</b>	3	16	<b>R136</b>	1	8
<b>R12</b>	5	6	<b>R37</b>	3	3	<b>R62</b>	5	11	<b>R87</b>	2	14	<b>R112</b>	3	9	<b>R137</b>	4	6
<b>R13</b>	3	3	<b>R38</b>	3	15	<b>R63</b>	5	7	<b>R88</b>	4	8	<b>R113</b>	5	3	<b>R138</b>	3	8
<b>R14</b>	4	11	<b>R39</b>	2	14	<b>R64</b>	4	6	<b>R89</b>	1	3	<b>R114</b>	5	11	<b>R139</b>	4	16
<b>R15</b>	4	2	<b>R40</b>	2	9	<b>R65</b>	3	15	<b>R90</b>	3	4	<b>R115</b>	1	7	<b>R140</b>	3	8
<b>R16</b>	2	3	<b>R41</b>	4	2	<b>R66</b>	1	7	<b>R91</b>	5	4	<b>R116</b>	3	7	<b>R141</b>	4	15
<b>R17</b>	4	16	<b>R42</b>	4	14	<b>R67</b>	4	10	<b>R92</b>	4	15	<b>R117</b>	3	13	<b>R142</b>	1	2
<b>R18</b>	2	14	<b>R43</b>	4	5	<b>R68</b>	5	7	<b>R93</b>	4	13	<b>R118</b>	1	2	<b>R143</b>	5	5
<b>R19</b>	3	12	<b>R44</b>	3	14	<b>R69</b>	3	15	<b>R94</b>	5	13	<b>R119</b>	1	6	<b>R144</b>	3	8
<b>R20</b>	3	9	<b>R45</b>	1	9	<b>R70</b>	5	9	<b>R95</b>	5	13	<b>R120</b>	1	5	<b>R145</b>	2	12
<b>R21</b>	3	8	<b>R46</b>	2	16	<b>R71</b>	4	5	<b>R96</b>	5	16	<b>R121</b>	5	10	<b>R146</b>	5	12
<b>R22</b>	4	15	<b>R47</b>	1	7	<b>R72</b>	5	10	<b>R97</b>	4	12	<b>R122</b>	3	5	<b>R147</b>	2	9
<b>R23</b>	1	8	<b>R48</b>	5	10	<b>R73</b>	2	9	<b>R98</b>	5	11	<b>R123</b>	4	6	<b>R148</b>	5	2
<b>R24</b>	2	8	<b>R49</b>	2	14	<b>R74</b>	2	3	<b>R99</b>	3	9	<b>R124</b>	1	14	<b>R149</b>	2	14
<b>R25</b>	1	5	<b>R50</b>	1	12	<b>R75</b>	4	12	<b>R100</b>	5	9	<b>R125</b>	4	15	<b>R150</b>	5	8

Tabela 3 - Instâncias aleatória com 30 requisitos

Fonte: elaborado pelos autores

A plataforma computacional usada nos testes foi um notebook Dell, com processador i7 8ª Geração 2.00 GHz , sistema operacional Windows 10, 64 bits e Memória RAM de 8 GB.

## 6 | RESULTADOS

A Tabela 4 contém o resultado obtido pela aplicação. Para a instância com 30 requisitos e uma *Sprint* com duração de 80 horas, o software implementado retornou a solução com o tempo de 0,03 segundos e ocupando 79 (setenta e nove) horas do tempo da *Sprint*; a segunda instância com 150 requisitos e uma *Sprint* com duração de 120 horas, o software retornou a solução, com o tempo de 0,05 segundos e ocupando as 120 (cento e vinte) horas previstas para o *Sprint*.

Instância	Execução	Sprint	Solução	
30 requisitos	0,03	79/80	44	R1 R3 R4 R5 R10 R12 R15 R16 R18 R19 R20 R21 R25 R28
150 requisitos	0,05	120/120	96	R2 R7 R12 R13 R15 R16 R32 R34 R36 R37 R41 R59 R61 R64 R71 R74 R80 R89 R90 R91 R102 R106 R108 R113 R118 R120 R123 R132 R143 R148

Tabela 4 - Resultado obtido para as duas Instâncias

Fonte: elaborado pelos autores

## 7 | CONCLUSÕES

Neste trabalho foi proposto um modelo matemático para o cenário de um ciclo de desenvolvimento do *Scrum* como um Problema da Mochila 0/1 para alocar, de maneira ótima, os itens do *Sprint Backlog*, utilizando um fator de priorização para os requisitos. O modelo foi aplicado em duas instâncias geradas aleatoriamente, simulando situações reais, para alocar os requisitos na próxima *Sprint*, otimizando a escolha dos requisitos do *Sprint Backlog*. Os resultados mostraram a eficiência do modelo, pela obtenção de soluções de alocação muito rápidas e satisfatórias dos requisitos para a próxima *Sprint*, o que na prática é difícil de se obter de forma empírica.

Como continuidade deste trabalho está sendo desenvolvido uma modelagem contemplando múltiplos objetivos.



## REFERÊNCIAS

BRENNAN, K. **A Guide to the Business Analysis Body of Knowledge**. International Institute of Business Analysis. 2015.

CARVALHO, T. **Planning Poker: o que é e como funciona com o Scrum**. 2019. Disponível em: <<https://www.voitto.com.br/blog/artigo/planning-poker>>. Acesso em: 07 jun. 2020.

GAREY, M. R.; JOHNSON, D. S. (2002). *Computers and intractability*. WH Freeman.

**GUROBI OPTIMIZER REFERENCE MANUAL**. Gurobi Optimization, LLC. Disponível em: <[https://www.gurobi.com/wp-content/plugins/hd\\_documentations/documentation/9.0/refman.pdf](https://www.gurobi.com/wp-content/plugins/hd_documentations/documentation/9.0/refman.pdf)>. Acesso em: 18 jun. 2020.

KRISHNAN, A. **Techniques to Prioritize Requirements**. 2020. Disponível em: <<https://www.modernanalyst.com/Resources/Articles/tabid/115/ID/3332/Techniques-to-Prioritize-Requirements.aspx>>. Acesso em: 18 jun. 2020.

MARTELLO, S.; TOTH, P. **Knapsack Problems: Algorithms and Computer Implementations**. John Wiley & Sons, Chichester, 1990.

MEREDITH, J. R.; MANTEL, S. J. Jr. **Project Management: a managerial approach**. (4th ed.) New York: Wiley, 2000.

SCHWABER, K.; BEEDLE, M. **Agile Software Development With Scrum**. Prentice Hall. 2002.

Scrum.org. **WHAT IS SCRUM?** 2020. Disponível em: <<https://www.scrum.org/resources/what-is-scrum>>. Acesso em: 18 jun. 2020.

VITIELLO, J. **Fast Track Into Management**. 2001. *Computerworld*, 35(29), 42-43.

## MODELOS EPIDÊMICOS: PROJETO DE INICIAÇÃO CIENTÍFICA NO CONTEXTO DA COVID-19

*Data de aceite: 01/04/2021*

### Vinícius R. da Silva

Instituto Federal de Educação, Ciência e  
Tecnologia de São Paulo  
<http://lattes.cnpq.br/1594733261630907>

### Felipe Y. Hatanaka

Instituto Federal de Educação, Ciência e  
Tecnologia de São Paulo  
<http://lattes.cnpq.br/1405611803125319>

### Olavo H. Menin

Instituto Federal de Educação, Ciência e  
Tecnologia de São Paulo  
<http://lattes.cnpq.br/2068400601317329>

**RESUMO:** A pandemia do novo coronavírus (SARS-CoV-2) tem incentivado diversos grupos de pesquisa a desenvolver modelos epidêmicos capazes de auxiliar na tomada de decisões de combate à doença. Esse contexto, apesar de dramático, pode ser uma oportunidade para o ensino significativo de matemática. Apresentamos aqui os resultados preliminares de dois projetos de Iniciação Científica (IC) integrados que abordam modelagem matemática de epidemias. Um dos projetos focou em modelos compartimentais do tipo SIR (suscetível-infectado-recuperado) enquanto o outro em modelos em rede. Especificamente nesse trabalho, buscamos mostrar como o isolamento social pode ser modelado em ambas abordagens e discutir como essa estratégia de combate

influencia no desenvolvimento da epidemia. A partir dos resultados obtidos com simulações computacionais, pode-se verificar que o pico da epidemia é atrasado e que quantidade de infectados no momento do pico diminui com o aumento do isolamento social.

**PALAVRAS - CHAVE:** epidemiologia, modelos compartimentais, redes randômicas, simulações computacionais.

### EPIDEMIC MODELS: UNDERGRADUATE RESEARCH PROJECT IN THE CONTEXT OF COVID-19

**ABSTRACT:** The pandemic of the new coronavirus (SARS-CoV-2) has encouraged several research groups to develop epidemic models capable of assisting in decision-making to combat the disease. This context, while dramatic, can be an opportunity for meaningful teaching of mathematics. We present here the preliminary results of two integrated undergraduate research projects that address mathematical epidemic modeling. One project focused on SIR (susceptible-infected-recovered) compartmental models while the other on network models. Specifically in this work, we seek to show how social isolation can be modeled in both approaches and to discuss how this strategy influences the evolution of the epidemic. From the results obtained with computer simulations, it can be seen that the peak of the epidemic is delayed and that the number of infected at the time of the peak decreases with increasing social isolation.

**KEYWORDS:** epidemiology, compartmental models, random networks, computational

simulations.

## 1 | INTRODUÇÃO

A história tem sido marcada por grandes pandemias. A conquista do México pelos espanhóis, por exemplo, foi decisivamente facilitada pela dizimação de grande parte dos povos nativos causada pela varíola, trazida inadvertidamente da Europa pelos colonizadores. Já no início do século 19, as baixas causadas pela febre amarela nas tropas francesas no Haiti levaram Napoleão a vender o território da Louisiana para o recém criado Estados Unidos, impulsionando a expansão americana até o pacífico. Os meses finais da Primeira Grande Guerra, por sua vez, foram dramaticamente marcados pela pandemia de influenza, que matou mais de 50 milhões de pessoas entre 1918 e 1919. Já a partir das últimas décadas do século passado, o vírus HIV-AIDS, tornou-se uma das grandes ameaças à saúde pública, ceifando mais de 30 milhões de vidas (OLDSTONE, 2010; QUAMMEN, 2013).

Com o primeiro caso registrado no final de dezembro de 2019, na cidade chinesa de Whan (WU et al., 2020), a atual pandemia do novo coronavírus (SARS-CoV-2) atingiu, em ano, mais de 88 milhões de pessoas e deixou quase 2 milhões de mortos em todos os continentes (WHO, 2021). Considerando a não existência de vacinas, o combate à pandemia foi feito principalmente por meio de intervenções não farmacêuticas tais como isolamento de casos, fechamento de escolas e universidades, proibição de eventos e fechamento de locais que possam gerar aglomerações e até quarentena forçada e completo distanciamento social. Modelos matemáticos têm sido aplicados para avaliar tais estratégias de combate e auxiliar as autoridades de saúde nas tomadas de decisão (FERGUSON et al., 2020; FLAXMAN et al., 2020; KRAEMER et al., 2020).

Nesse sentido, o atual contexto, apesar de dramático, pode ser uma oportunidade para o ensino de matemática, conectando-a ao mundo real (DUNN; MARSHMAN, 2020). Mais especificamente durante a graduação, o processo de ensino-aprendizagem de disciplinas como cálculo e probabilidade/estatística pode tornar-se mais significativo quando problemas da vida cotidiana são abordados, especialmente se inseridos como eixos articuladores entre diferentes conteúdos (BOYCE; DIPRIMA, 2018; MORETTIN, 2010). Tal afirmação evidencia-se ainda mais em projetos de Iniciação Científica (IC). De fato, quando realizados de forma adequada, tais projetos colaboram para melhorar a auto-estima do aluno orientando, desenvolver suas habilidades práticas, refinar seu pensamento científico entre outros ganhos (PETRELLA; JUNG, 2008; MORAES; FAVA, 2000).

Apresentamos aqui os resultados preliminares de dois projetos de Iniciação Científica (IC) integrados que abordam modelagem matemática de epidemias. Um dos projetos focou em modelos compartimentais do tipo SIR (suscetível-infectado-recuperado) enquanto o outro em modelos em rede. Especificamente nesse trabalho, buscamos mostrar

como o isolamento social pode ser modelado em ambas abordagens e discutir como essa estratégia de combate influencia no desenvolvimento da epidemia. Os projetos tiveram início em março de 2020 e está sendo executado de forma remota. Foram realizadas revisões bibliográficas sobre equações diferenciais, probabilidade e modelos epidêmicos, reuniões semanais e apresentação de seminários. Também foram desenvolvidos e implementados algoritmos e realizadas simulações computacionais. Além de resultados científicos, o projeto pretende fortalecer a integração entre ensino e pesquisa, tão fundamental para uma formação acadêmica sólida do estudante.

## 2 I MODELOS EPIDÊMICOS E SIMULAÇÕES COMPUTACIONAIS

Nesta seção, descrevemos os modelos desenvolvidos e apresentados nesse trabalho. Primeiramente, discutimos o modelo compartimental SIR com isolamento social e, em seguida, um modelo baseado em rede randômica. Ambos foram desenvolvidos a partir das referências (KEELING; ROHANI, 2008; BARABÁSI; PÓSFAL, 2016; CHAPRA; CANALE, 2008).

### 2.1 Modelo compartimental SIR com isolamento social

O modelo compartimental que apresentamos aqui é baseado no clássico modelo SIR (KERMACK; MCKENDRICK, 1927), no qual a população é dividida em três compartimentos, suscetíveis (S), infecciosos (I) e recuperados (R). Sendo  $S$ ,  $I$  e  $R$  as frações da população em cada compartimento, com  $S+I+R = 1$ , a transição de indivíduos do compartimento de suscetíveis para o de infecciosos acontece de acordo com duas taxas,  $(1 - \alpha)\beta_1SI$  e  $\alpha\beta_2SI$ , com  $\beta_1 > \beta_2$ , sendo  $\alpha \in [0,1]$  a fração de suscetíveis em isolamento social. Já a taxa de transição do compartimento de infecciosos para o de recuperados é dada por  $\gamma I$ , sendo  $\gamma$  o inverso do período infeccioso. O modelo, que não considere demografia, pode ser representado como mostra a Fig. (1).

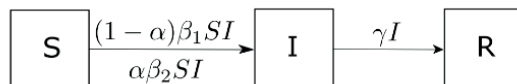


Figura 1: Diagrama do modelo compartimental SIR com isolamento social definido pelo parâmetro  $\alpha$ .

Com isso, o modelo é descrito pelo sistema de equações diferenciais

$$\frac{dS}{dt} = -(1 - \alpha)\beta_1 SI - \alpha\beta_2 SI, \quad (1)$$

$$\frac{dI}{dt} = (1 - \alpha)\beta_1 SI + \alpha\beta_2 SI - \gamma I, \quad (2)$$

$$\frac{dR}{dt} = \gamma I. \quad (3)$$

Tanto a taxa de recuperação,  $\gamma$ , quanto as taxas de transmissão,  $\beta_1$  e  $\beta_2$ , podem ser estimadas a partir de dados epidemiológicos. A primeira, como mencionado, corresponde ao inverso do período infeccioso, que pode ser obtido a partir de dados clínicos. Já para estimar as duas últimas, podemos utilizar os dados de incidência da doença nos estágios iniciais da epidemia. Para isso, considera-se que nesse período não há indivíduos em isolamento e que praticamente toda população está no compartimento de suscetíveis, ou seja,  $S \approx 1$ . Aplicando essas condições na Eq. (2), obtém-se a equação diferencial  $dI/dt = (\beta_1 - \gamma)I$ , cuja solução é , sendo  $I_0$  a fração inicial de infectados. Aplicando logaritmo em ambos os lados dessa solução, obtemos

$$\ln I(t) = \ln I_0 + (\beta_1 - \gamma)t. \quad (4)$$

Com isso, concluímos que nos estágios iniciais da epidemia o logaritmo da incidência varia linearmente como o tempo com coeficiente angular dado por  $\beta_1 - \gamma$ .

O modelo dado pelas Eqs. (1 - 3) foi implementado no software MATLAB e resolvido numericamente utilizando o método de Runge-Kutta discreto de quarta ordem com passo adaptável. As simulações foram realizadas considerando-se fixos os parâmetros mostrados na Tabela 1. A taxa de recuperação  $\gamma$  foi estimada considerando-se um período infeccioso de 5 dias e a taxa de transmissão  $\beta_1$  foi estimada a partir do ajuste de dados de incidência iniciais de três países, Brasil, EUA e Reino Unido e considerando a Eq. (4), como mostra a Fig. (2). Já a fração de isolamento assumiu os valores  $\alpha = 0,0$ ,  $\alpha = 0,2$ ,  $\alpha = 0,4$  e  $\alpha = 0,6$  e os resultados são mostrados na Fig. (3). Na Fig. (3-a) temos as frações  $S$ ,  $I$  e  $R$  em função do tempo para  $\alpha = 0$ , e na Fig. (3-b) temos a incidência  $I$  em função do tempo para vários valores de  $\alpha$ . Já a Fig. (3-c) mostra o número de casos acumulados para vários valores de  $\alpha$ .

Símbolos	Descrição	Estimativa de valores
$\beta_1$	taxa de infecção da população fora do isolamento	0,5 [dia <sup>-1</sup> ]
$\beta_2$	taxa de infecção da população em isolamento	0,1 [dia <sup>-1</sup> ]
$\gamma$	Inverso do período infeccioso	0,2 [dia <sup>-1</sup> ]

Tabela 1: Parâmetros fixo do modelo compartimental. As escolhas foram arbitrárias e a razão  $\beta_1/\beta_2$  supõe uma redução de cinco vezes a taxa de contato.

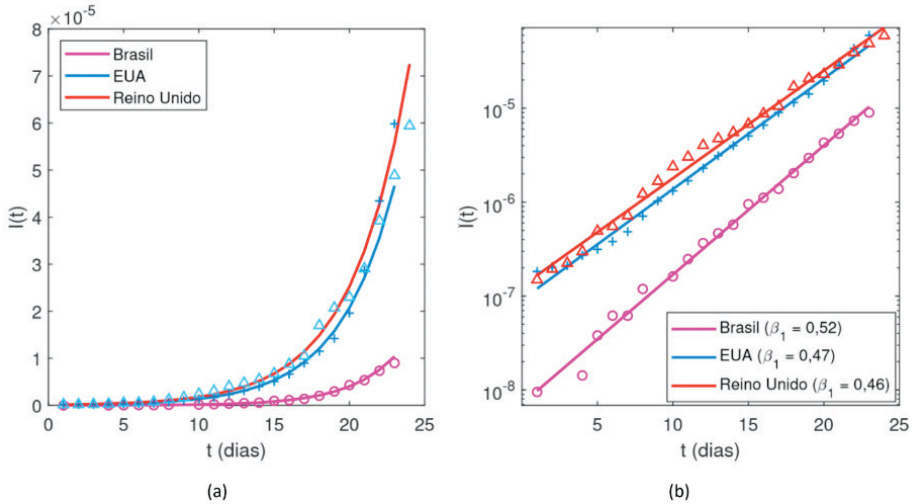


Figura 2: Incidência nos estágios iniciais da epidemia em três países, Brasil, EUA e Reino Unido em escala (a) linear e (b) logarítmica com seus respectivos valores estimados para  $\beta_1$ , considerando  $\gamma = 0,2 \text{ dia}^{-1}$ .

## 2.2 Modelo baseado em redes randômicas

Uma rede pode ser representada por um grafo  $G = (V, E)$  de  $N$  nós,  $V_1, V_2, \dots, V_{N_p}$  que podem estar conectados entre si por  $L$  links,  $E_1, E_2, \dots, E_L$ . Matematicamente, uma rede pode ser descrita por meio de uma matriz de adjacência  $\mathbf{A}$ , cujos elementos assumem os valores  $a_{ij} = 1$ , caso haja um link entre os nós  $i$  e  $j$ , e  $a_{ij} = 0$ , caso não haja link entre os nós. Em redes não direcionadas, a matriz de adjacência sempre será simétrica uma vez que o link de  $i$  para  $j$  sempre existirá de  $j$  para  $i$ , como representado na Fig. (4).

Para o modelo de redes randômicas mais usual, modelo  $G(N, p)$ , define-se um número fixo de nós,  $N$ , e uma probabilidade,  $p$ , para cada par de nós da rede possuir um link (serem conectados). Desse modo, o número médio esperado de links por nó (grau médio) é dado por

$$\langle k \rangle = p(N - 1) \quad (5)$$

O modelo epidêmico SIR em redes considera que a cada tempo discreto  $t = 0, 1, 2, 3, \dots$ , cada nó  $i$  da rede possa assumir um dos três estados,  $w_i = 0$ ,  $w_i = 1$  e  $w_i = 2$ , caso seja suscetível (S), infectado (I) ou recuperado (R), respectivamente. Inicialmente, considera-se apenas um nó infectado dentre todos os outros suscetíveis e, com o passar do tempo, a epidemia se propaga na rede, como mostra a Fig (5).

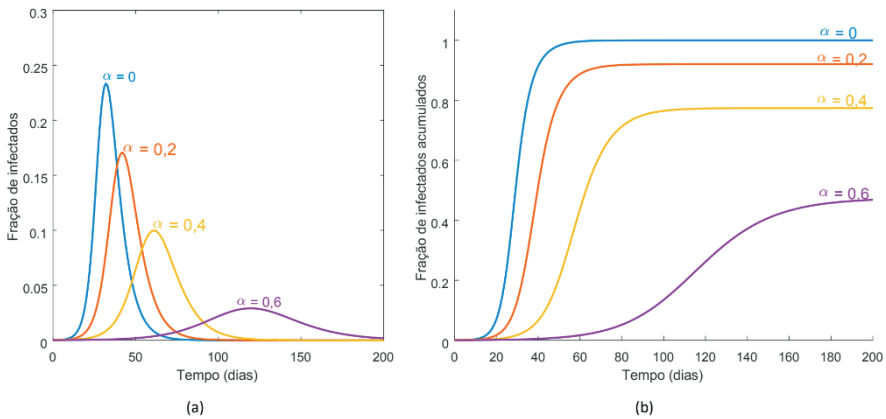


Figura 3: Resultado obtidos com o modelo compartimental para diferentes valores de  $\alpha$ , mostrando a evolução da fração de infectados (a) em cada instante de tempo e (b) acumulados

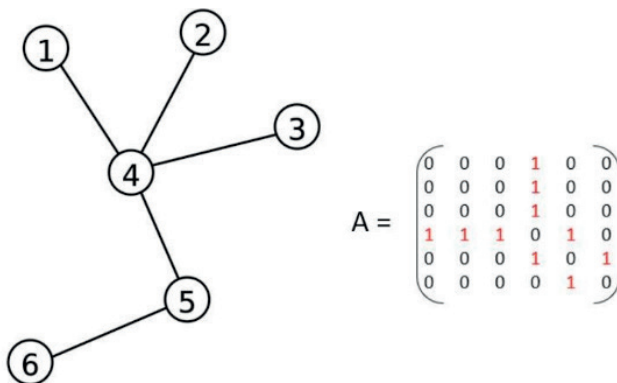


Figura 4: Rede com  $N = 6$  nós e  $L = 5$  links e sua matriz de adjacência

Nós suscetíveis só podem se infectar se estiverem conectados a um ou mais nó infectados com probabilidade

$$p^{(i)}(S \rightarrow I) = (1 - \beta)^{n_i} \tag{6}$$

onde  $\beta$  é a probabilidade de um nó suscetível ser infectado por um vizinho infeccioso e  $n_i$  é o número de vizinhos infecciosos do nó  $i$ , como representado na Fig. (6). Já pacientes infectados, possuem uma probabilidade  $\gamma$  (inverso do período infeccioso) de se recuperarem, tornando-se permanentemente imunes à doença.

Para a realização das simulações, foi desenvolvido um algoritmo utilizando o software MATLAB. Inicialmente, o algoritmo gera o grafo de uma rede randômica e, em seguida, é realizada a propagação

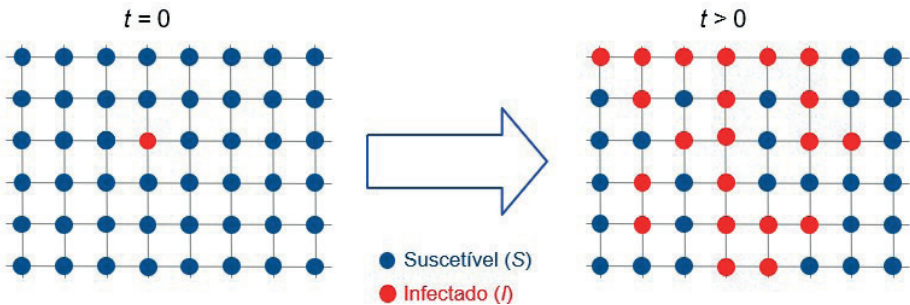


Figura 5: Status de indivíduos de uma rede regular quadrada na situação (a) inicial ( $t = 0$ ) e (b) no decorrer da epidemia ( $t > 0$ ).

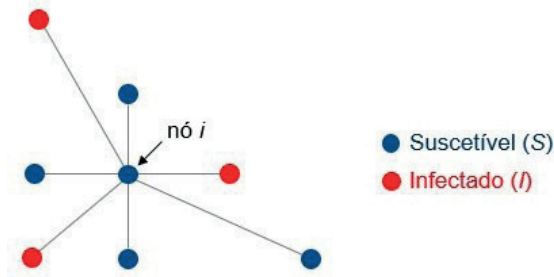


Figura 6: Vizinhos (suscetíveis e infectados) de um nó qualquer  $i$ .



epidêmica. As simulações foram realizadas adotando-se  $\beta = 0,5$ ,  $\gamma = 0,2$  e  $N = 10^4$  fixos e diferentes valores de grau médio  $\langle k \rangle$ , tornando os nós mais ou menos conectados (simulando maior ou menor isolamento social). Como exemplo, uma rede randômica com  $N = 100$  e  $\langle k \rangle = 19,8$  é apresentada na Fig. (7). Com os resultados obtidos, foi possível estudar o comportamento da epidemia com diferentes taxas de isolamento, seu pico, evolução temporal e término, como mostra a Fig. (8).

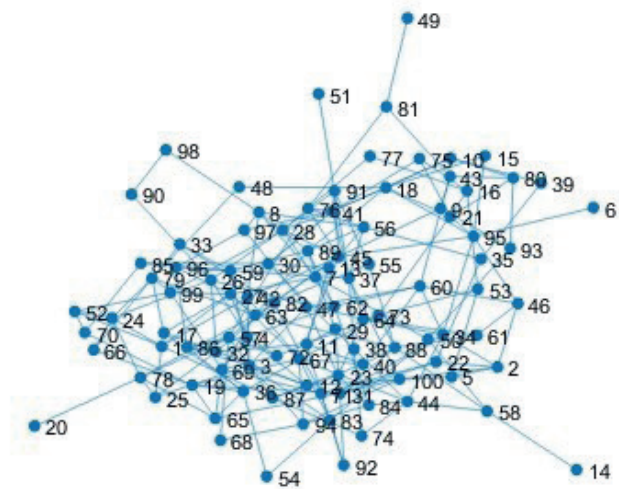


Figura 7: Grafo de uma rede randômica com  $N = 100$  e  $\langle k \rangle \approx 4,6$ .

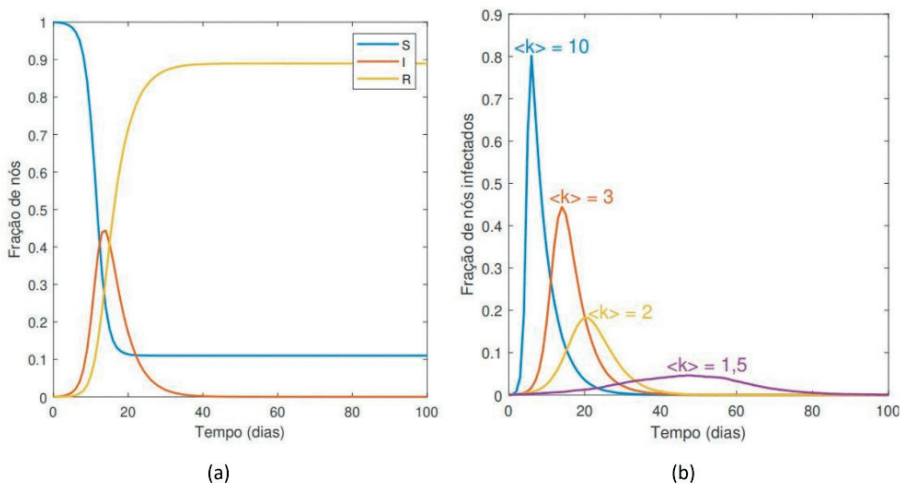


Figura 8: Resultados obtidos com o modelo em rede mostrando a evolução da fração de (a) suscetíveis, infectados e recuperados, com  $\langle k \rangle = 3$ , e (b) de infectados para diferentes valores de  $\langle k \rangle$ .

### 3 | CONCLUSÕES

Apresentamos os resultados preliminares de dois projetos de Iniciação Científica integrados que abordam modelagem matemática de epidemias. Buscamos mostrar a importância do isolamento social no combate a uma epidemia por meio de simulações computacionais baseadas em dois tipos de modelos, compartimental do tipo SIR e em rede. Os resultados mostram que o pico da epidemia é atrasado e que quantidade de infectados no momento do pico diminui com o aumento do isolamento social. Por fim, apesar do pouco tempo de execução dos projetos os alunos mostraram-se motivados e obtiveram progressos significantes e sua aprendizagem acadêmica.

### REFERÊNCIAS

BARABÁSI, A.-L.; PÓSFAL, M. **Network Science**. Cambridge: Cambridge University Press, 2016.

BOYCE, W. E.; DIPRIMA, R. C. **Equações diferenciais elementares e problemas de valores de contorno**. 10. ed. Rio de Janeiro: LTC, 2018.

CHAPRA, S. C.; CANALE, R. P. **Métodos numéricos para engenharia**. Porto Alegre: McGraw-Hill, 2008.

DUNN, P. K.; MARSHMAN, M. F. **Teaching mathematical modelling: a framework to support teachers choice of resources**. *Teaching Mathematics and its Applications*, v. 39, p. 127–144, 2020.

FERGUSON, N. M. et al. **Impact of non-pharmaceutical interventions (npis) to reduce covid-19 mortality and healthcare demand**. *Imperial College London*, 2020.

FLAXMAN, S. et al. **Estimating the number of infections and the impact of non-pharmaceutical interventions on covid-19 in 11 european countries**. *Imperial College London*, 2020.

KEELING, M. J.; ROHANI, P. **Modeling Infectious Diseases in Humans and Animals**. Princeton: Princeton University Press, 2008.

KERMACK, W. O.; MCKENDRICK, A. G. **Contributions to the mathematical theory of epidemics**. *Proceedings of the Royal Society*, v. 115A, p. 700–721, 1927.

KRAEMER, M. U. G. et al. **The effect of human mobility and control measures on the COVID-19 epidemic in China**. *Science*, v. 369, p. 493–497, 2020.

MORAES, F. F. de; FAVA, M. **A iniciação científica: muitas vantagens e poucos riscos**. *São Paulo em Perspectiva*, v. 14, p. 73–77, 2000.

MORETTIN, L. G. **Estatística básica: probabilidade e inferência**. [S.l.: s.n.], 2010.

OLDSTONE, M. B. A. **Viruses, Plagues and History**. New York: Oxford University Press, 2010.

PETRELLA, J. K.; JUNG, A. P. **Undergraduate research: importance, benefits and challenges.** *International Journal of Exercise Science*, v. 1, p. 91–95, 2008.

QUAMMEN, D. *Spillover: Animal infectious and the next human pandemic.* New York: W. W. Norton & Company Inc., 2013.

WHO. *Weekly Operational Update on COVID-19 (24 November 2020).* 2021.

WU, F. et al. **A new coronavirus associated with human respiratory disease in China.** *Nature*, v. 579, p. 265–269, 2020.

# CAPÍTULO 8

## GUIDE2BLIND: SISTEMA HÁPTICO-SONORO DE ORIENTAÇÃO PARA DEFICIENTES VISUAIS EM AMBIENTES FECHADOS - FASE 2

Data de aceite: 01/04/2021

Data de submissão: 08/03/2021

### **Lucas Rafael da Silva Martins**

Instituto Federal Sul-rio-grandense - Câmpus  
Charqueadas.

Charqueadas – RS.

<http://lattes.cnpq.br/6572656420552298>

### **Mikael Tolotti da Silva**

Instituto Federal Sul-rio-grandense - Câmpus  
Charqueadas.

Charqueadas – RS.

<http://lattes.cnpq.br/5284235895911192>

### **Bernardo Moreira**

Instituto Federal Sul-rio-grandense - Câmpus  
Charqueadas.

Charqueadas – RS.

<http://lattes.cnpq.br/5594063547802901>

### **Diego Afonso da Silva Lima**

Instituto Federal Sul-rio-grandense - Câmpus  
Charqueadas.

Charqueadas – RS.

<http://lattes.cnpq.br/6696235998719891>

### **Carlos Francisco Soares de Souza**

Instituto Federal Sul-rio-grandense - Câmpus  
Charqueadas.

Charqueadas – RS.

<http://lattes.cnpq.br/6826517859097690>

### **Luis Gustavo Fernandes dos Santos**

Instituto Federal Sul-rio-grandense - Câmpus  
Charqueadas.

Charqueadas – RS.

<http://lattes.cnpq.br/0731740921318770>

### **Carlos Arthur Carvalho Sarmanho Junior**

Instituto Federal Sul-rio-grandense - Câmpus  
Charqueadas.

Charqueadas – RS.

<http://lattes.cnpq.br/6350887705184709>

**RESUMO:** Este projeto visa construir um sistema háptico sonoro para auxiliar a navegação e locomoção de deficientes visuais em ambientes fechados. Os objetivos específicos são listados, tais como: Desenvolver um dispositivo de segurança compacto para detecção de obstáculos que execute uma resposta de desvio precisa e rápida; desenvolver um aplicativo de detecção de sinal Wi-Fi; criar um sistema de mapeamento rápido e direto para qualquer ambiente fechado; desenvolver um algoritmo de rota para localizações mapeadas; realizar simulação de mapeamento de Wi-Fi e geolocalização. Visão é o sentido que permite o fluxo mais significativo de informações interpretativas, e a perda desse sentido gera barreiras substanciais para essas pessoas em muitas atividades, como a locomoção. De acordo com o censo do IBGE de 2010, cerca de 16,6 milhões de pessoas têm algum grau de deficiência visual, das quais 506 mil são totalmente cegas. É uma grande barreira para pessoas com deficiência visual encontrar métodos para ajudá-las. Como método alternativo, o grupo busca a construção de um sistema e idealiza: utilizar sensores e micromotores elétricos de vibração para detectar e alertar a direção dos objetos durante a locomoção entre as rotas, permitindo ao usuário evitar obstáculos que podem variar

e, em paralelo, o uso de um aplicativo Android que utiliza sinais de roteadores Wi-Fi para auxiliar o usuário em sua Geolocalização por meio de trilateração, guiando-o até um destino escolhido e permitindo a navegação em áreas previamente mapeadas..

**PALAVRAS - CHAVE:** Tecnologia Assistiva; Localização em ambientes fechados; Substituição sensorial.

## GUIDE2BLIND: HAPTIC-SOUND GUIDANCE SYSTEM FOR THE VISUALLY IMPAIRED INDOORS - PHASE 2

**ABSTRACT:** This project aims to build a sound haptic system to aid the navigation and locomotion of the visually impaired indoors. Specific objectives are listed such as: Developing a compact safety device for obstacle detection that performs a precise and rapid deviation response; develop a Wi-Fi power detection application; to create a fast and straightforward mapping system for any closed environment; develop a route algorithm for mapped locations; perform simulation of Wi-Fi and Geolocation mapping. Vision is the sense that allows the most significant flow of interpretive information, and the loss of that sense generates substantial barriers for these people in many activities, such as locomotion. According to the 2010 IBGE census, around 16.6 million people have some degree of visual impairment, from which 506 thousand are entirely blind. It is an big barrier for visually impaired people to find methods to help them. As an alternative method, the group seeks the construction of a system and idealizes: using sensors and electric vibration micromotors to detect and alert the direction of objects during locomotion between routes, allowing the user to avoid obstacles that may vary and, in parallel, the use of an Android application that uses signals from Wi-Fi routers to assist the user in his Geolocation through trilateration, guiding him to a chosen destination and allowing navigation within previously mapped areas.

**KEYWORDS:** Assistive Technology, Indoor Location, Sensory Replacement.

## 1 | INTRODUÇÃO

A deficiência visual, doença caracterizada por dificultar ou não permitir uma visão nítida ao seu redor, atinge atualmente grande parte da população (SOARES, 2009). Em 2002, estimava-se atingir 162 milhões de pessoas no mundo. Em 2010, esse número torna-se alarmante no Brasil, segundo o censo do Instituto Brasileiro de Geografia e Estatística cerca de 16,6 milhões de pessoas possuem algum grau de deficiência visual, e, deste total, 506 mil sejam completamente cegos (MARQUES, 2017; RANGEL et al., 2010).

Em mundo alimentado pela contínua produção de informação, tecnologia e derivados, a perda de visão, por vezes limita o indivíduo de certas atividades como orientação, locomoção, acesso a conteúdos digitais, etc (BERSCH, 2017; GLEICK, 2013). Mesmo que o corpo humano supra, em parte, a falta de visão a partir de outros sistemas receptores de informações externas, encontramos um déficit, pois, é a visão que permite distinguirmos com certeza a fisionomia dos objetos e a profundidade de ambientes, além de proporcionar ao corpo habilidades motoras, tais como a postura firme e o equilíbrio. Sem

contar que a cegueira adquirida pode causar diversos problemas, desde psicológicos, como sentimento de incapacidade social e dificuldade de relacionar-se com demais indivíduos, até fisiológicos (RANGEL et al., 2010; SOARES, 2009).

Entretanto uma área bem popular na medicina aponta que, apesar da deficiência visual, nosso corpo busca remediar a situação a partir de outros sistemas sensoriais como a audição e o tato. A plasticidade neural é um termo que se refere ao comportamento do nosso cérebro que redistribui funções do sistema nervoso quando, por alguma circunstância, tenha lesionado o tecido nervoso. As sinapses neurais são reorganizadas já que mais de uma área no córtex é responsável pelo processamento de dados externos. No caso de substituição sensorial, cientistas e pesquisadores como Cohen et al (COHEN et al., 1997) e Sadato et al (SADATO et al., 1996) afirmam terem encontrado atividade no córtex visual primário durante a leitura de braille, confirmando a correlação do córtex occipital com o visual (RANGEL et al., 2010).

Além do sentido tátil é comprovado que, em cegos e videntes, o córtex occipital é ativado em circunstâncias que requeiram uso da imaginação. Por estes estudos, Rangel et al (RANGEL et al., 2010) afirma que nenhum dos processos neurais deve ser considerado individual, isolado ou tido como exclusivo de determinada área cerebral. Dessa maneira, pode-se supor que, devido a plasticidade cerebral, nenhum deficiente visual é menos desenvolvido que outro indivíduo, mas que as tarefas são apenas redirecionadas a outras áreas.

É válido ressaltar que a taxa de recuperação cortical está diretamente relacionada à idade em que a cegueira se manifesta no indivíduo (COHEN et al., 1997). Em um estudo de teste com um produto de substituição sensorial (SS) tátil-visual, realizado por Kastrup et al (KASTRUP et al., 2009), os resultados apresentam que pessoas com cegueira congênita ou precoce se saem melhor com relação aos demais de idade avançada e cegueira tardia.

Um termo recente nesta área de inclusão técnica é a chamada Tecnologia Assistiva (TA), tecnologia voltada para a produção de equipamentos e dispositivos que facilitem a vida dos deficientes. Como definido no decreto Nº 8.953, de 10 de janeiro de 2017, artigo 3, parágrafo único, tecnologia assistiva se caracteriza em produtos, equipamentos, dispositivos, entre outros que estejam relacionados a promoção de melhorias na qualidade de vida e autonomia de pessoas com deficiência ou mobilidade reduzida (BRASIL, 2017).

O artigo publicado por Kastrup et al (KASTRUP et al., 2009) diz que as tecnologias assistivas devem ser apenas e somente para pessoas com deficiências físicas, sensoriais ou cognitivas. Aponta também que na área educacional as TA's são, eventualmente, confundidas entre si. Um cadeirante utilizará o mesmo computador que qualquer outro colega para o mesmo objetivo e que a principal diferenciação de um sistema educacional e TA se dá pelo “rompimento de barreiras”, tanto sensoriais, como motoras ou cognitivas, que impeçam a limitação ao acesso à informação (KASTRUP et al., 2009). A tecnologia assistiva é interdisciplinar, não aplicável para uma única tarefa, e esta expande-se para

um ambiente social ajudando na locomoção, comunicação, acessibilidade, portabilidade, alunos em salas de aulas e até mesmo em sistemas de substituição sensorial (KASTRUP et al., 2009).

Incluso nas tecnologias assistivas (TA's) está o conceito de Substituição Sensorial (SS), no qual se busca alternativas para que outros sentidos possam exercer o papel do sentido com deficiência ou perda total. Para entender o sistema de substituição sensorial, deve-se ter em mente conceito de plasticidade neural - regeneração ou readaptação das/ nas sinapses neurais para que outra parte do córtex occipital possa processar os dados externos (COHEN et al., 1997; RANGEL et al., 2010; SADATO et al., 1996). A substituição sensorial se dá pelo uso de TA's para sistemas táteis e auditivos e, segundo Bach-y-Rita (KASTRUP et al., 2009), a substituição sensorial que melhor se adaptou ao cotidiano foi a escrita em braille, na qual, ao invés de utilizar-se os olhos para ler, são utilizadas as ponta dos dedos. Duas tecnologias assistivas que estão entre as mais comuns quando tratamos de deficiência visual e locomoção para SS são o “cão-guia” e/ou a “bengala”, sendo que esta última tem função de equipamento de segurança para a detecção de objetos extremamente próximos e preferencialmente estáticos. Os custos de uma bengala são relativamente baixos, ao passo que o custo de um cão guia pode ser alto, e possui a vantagem de oferecer a possibilidade de guiar o usuário para locais com distâncias médias (dentro de espaços conhecidos e previamente apresentados ao cão). No ano de 2000, em Portugal, nasceu a Associação Beira Agueira de Apoio ao Deficiente Visual (ABAADV), que presta o serviço de educação de cães-guia, e os cães-guias são entregues gratuitamente a pessoa cega, ou melhor, a preço de custo, pois alguns valores são cobrados, tais como os salários dos educadores da associação e os custos das necessidades naturais do animal, podendo levar a entrega de um cão-guia educado ao custo elevado de 17.500 euros (BADALO, 2014); em paralelo, há o fato de que um cão não é uma máquina e nem um objeto e sim um animal com necessidades e vontades próprias (ABAADV, 2018). Além disso, existem o guia humano e o GPS (Sistema de Posicionamento Global) como sistemas auxiliares de locomoção (SÁ, 2003). Um ponto importante na discussão são as limitações dessas tecnologias e suas aplicações em diversos ambientes fechados, tais como shoppings, hospitais, estabelecimentos comerciais, industriais, escritórios, prédios, residências, etc, o que impossibilita a inclusão de deficientes visuais em diversos espaços sociais (BRUMER; PAVEI; MOCELIN, 2004). Isso proporciona um pensamento de custo-benefício para esses métodos de locomoção segura e até que ponto afetam a autonomia do deficiente visual. Para os métodos citados, os ambientes fechados parecem ser os mais atingidos negativamente na questão custo-benefício, pois dificilmente podem ser navegados com cães, ou por GPS, pois além de sua imprecisão (KOTARO et al., 2015b), o GPS não permite com que o deficiente visual possa desviar de obstáculos (BADALO, 2014). Assim, dos métodos citados só restam a bengala e o guia humano para ambientes fechados, sendo que o alto custo de um guia humano impossibilita o acesso a estes

ambientes para grande parte dos indivíduos deficientes visuais. A dependência de um guia humano pode afetar na vida auto sustentada e plena do deficiente visual (BRUMER; PAVEI; MOCELIN, 2004) e a bengala é um sistema de baixo custo, porém sua simplicidade tende a não completar muitos parâmetros no auxílio da locomoção em ambientes fechados, tal qual andar em um corredor com muitos objetos de obstrução e/ou locais como um shopping com uma grande circulação de pessoas e diversos andares.

O objetivo deste projeto é mais do que criar um método auxiliar de locomoção com substituição sensorial, mas observar que sejam atendidos ótimos parâmetros de custo-benefício para a aquisição de uma plataforma de auxílio à locomoção em ambientes fechados, por parte deficientes visuais, e a melhora na qualidade de vida do indivíduo deficiente.

Hipotetiza-se o desenvolvimento de um sistema háptico-sonoro formado por um aplicativo *Android* que se comunica com o usuário pelo sentido auditivo, associado a um dispositivo de segurança que utiliza o sentido háptico, ou também chamado sentido tátil, para se comunicar e orientar o deficiente visual sobre a presença e direção de objetos, tudo isto, de forma não-invasiva, e que ambos em funcionamento durante a locomoção, possam facilitar a locomoção em ambientes fechados quando comparado com outras tecnologias de locomoção para deficientes visuais. Além disso, a utilização de um aplicativo *Android*, que é uma tecnologia popular na sociedade, e a construção do dispositivo com a possibilidade de utilização de componentes eletrônicos e de processamento de baixo custo, poderá trazer ganhos na relação custo-benefício.

Outra hipótese é que, o deficiente visual, ao utilizar o sistema, alcance uma maior autonomia quando comparado a outras tecnologias de auxílio na locomoção que no momento são baseadas no que foi visto na literatura. Para isso, acredita-se que o sistema proporcione uma boa alternativa de solução de um dos principais problemas encontrados, no caso, o propósito de se locomover em ambientes fechados. Com isso hipotetiza-se que o usuário consiga ter simultaneamente ganhos em qualidade de vida e experiência durante sua locomoção.

## **2 | OBJETIVOS E QUESTÃO PROBLEMA**

A busca por qualidade de vida é um dos fatores que caracterizam a tecnologia assistiva (BRASIL, 2017) e pode ser ampliada com a inclusão cultural e econômica de deficientes visuais. Essa inclusão está relacionada diretamente com a locomoção dos mesmos em locais sociais ou de trabalho que normalmente se encontram em ambientes fechados (BRUMER; PAVEI; MOCELIN, 2004). Algumas barreiras estão sendo quebradas ao longo das últimas décadas para a inclusão de deficientes visuais no trabalho, mas um tópico que ainda apresenta barreiras é a adaptação necessária na empresa ou local de trabalho para atender algumas necessidades do deficiente visual (BRUMER; PAVEI;



MOCELIN, 2004), e com este propósito, o método apresentado pelo grupo pretende facilitar adaptação do deficiente visual nestes locais. A questão chave então é: Como construir um sistema para auxílio à localização e orientação de deficientes visuais em ambientes fechados?

Sendo assim, o objetivo deste trabalho é construir um sistema para auxílio à localização e orientação de deficientes visuais em ambientes fechados. Utilizando substituição sensorial para a comunicação com o deficiente visual, pretende-se criar um sistema de localização utilizando de sinais de roteadores com o intuito de identificar a posição e possíveis rotas de locomoção para os usuários, com orientação sonora, e concomitantemente desenvolver um sistema de segurança através da detecção de obstáculos e com responsividade através de vibração em contato indireto com a derme do indivíduo.

### **3 | DESCRIÇÃO DE MATERIAIS E MÉTODOS**

Os objetivos do projeto levam em conta dois grandes eixos, os de locomoção que necessitam de um sistema de segurança, para evitar que o deficiente visual esbarre com algum obstáculo. E o segundo grande eixo tem em vista o entendimento de navegação e orientação em ambientes fechados, para solucionar tal problema que desenvolvemos um aplicativo para localizar e indicar rotas para o usuário.

#### **Dispositivo de segurança G2B**

A Figura 1 apresenta, no caminho iniciado na linha superior, como se dá a transmissão da informação sensorial biológica (olhos) para dados adquiridos, e depois para informação biológica (DNA) para então haver a resposta do indivíduo (movimento de reação, por exemplo) a esta informação. Na mesma imagem, na linha inferior, é mostrado como se dá a transmissão de um sensor eletrônico, que emite um sinal e recebe a resposta em um sinal de frequência, o qual é traduzido internamente como uma distância, depois convertido em vibrações em um motor vibracall junto ao corpo e só então ocorre a aquisição de dados, para encaminhar ao DNA e a resposta do usuário. Quando se trata da visão, o sinal (luz) não precisa ser emitido, mas somente recebido, com velocidade e frequência extremamente altas, traduzindo milhares de informações em segundos. Já para o sensor eletrônico precisa emitir o sinal (som, que apresenta velocidade muito menor em comparação à da luz), receber o sinal de retorno, e tem ainda um caminho longo até que as (poucas) informações sejam traduzidas pelo corpo.

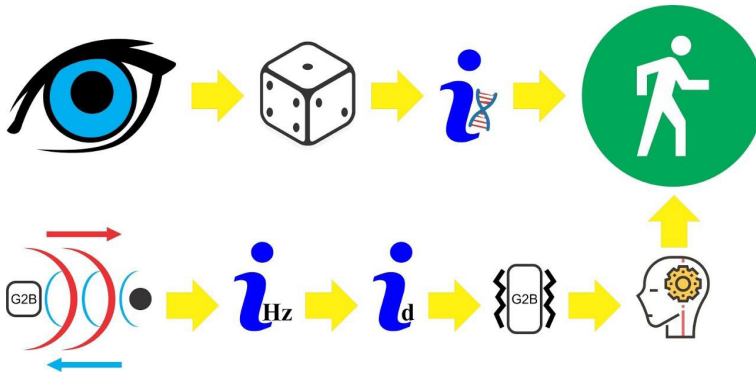


Figura 1. Comparação entre Sistema Biológico e Substituição Sensorial. Autoria própria.

Os sentidos fundamentais biológicos se unem para fazer a detecção de objetos e orientação do indivíduo de forma que este possa desviar de obstáculos, além de obter orientação espacial durante o deslocamento. Existem alguns sensores eletrônicos que permitem a detecção de objetos. Neste dispositivo foi utilizado o sensor ultrassônico HC-SR04, o qual possui alcance entre 2 e 400 cm, e funciona com a emissão de um sinal sonoro de 40 kHz, não perceptível pelo espectro sonoro humano. O sinal enviado pelo sensor atinge um objeto e a distância é calculada pela relação temporal de envio e recebimento do sinal sonoro em um receptor integrado ao sensor.

Com o intuito de atender um dos objetivos, desenvolver um dispositivo de segurança compacto para a detecção de obstáculo que efetue uma resposta de desvio rápida e nítida, foi construído um dispositivo de estrutura mecânica para alojar os componentes, circuito eletrônico e programação adequada ao circuito à detecção de objetos.

## Construção do dispositivo de segurança G2B

### *Substituição sensorial háptica*

Para o processo de substituição sensorial, foi utilizado um micromotor, o qual deverá emitir sinal de vibração para o sentido háptico (sentido do tato / recepção do sinal através da pele), gerando uma comunicação biológica. Foram distribuídos três motores de vibração no dispositivo que está em contato com o corpo, divididos de forma equivalente aos sensores, formando três conjuntos, um na parte central, um na direita e outro na esquerda do dispositivo, com cada motor respondendo respectivamente ao sensor de sua posição, possibilitando ao usuário detectar a direção do objeto através da substituição sensorial. O dispositivo inicial é apresentado na Figura 2.

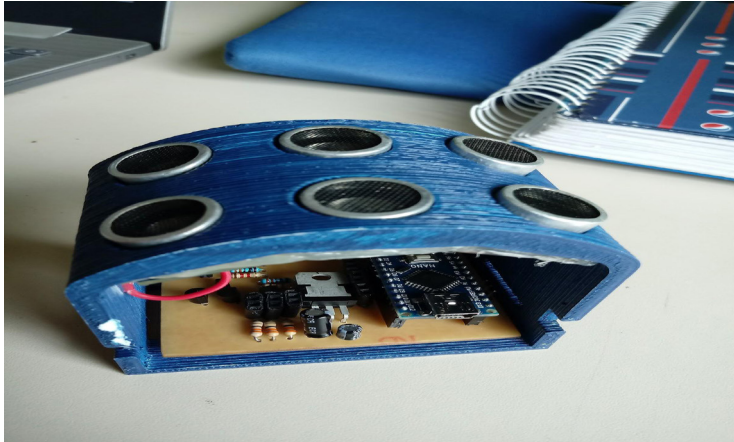


Figura 2. Dispositivo háptico-sonoro (1) para detecção de objetos. Autoria Própria.

### Circuito Eletrônico

O circuito foi calculado (apresentamos esses cálculos em nosso diário de bordo) e projetado utilizando o microcontrolador Arduino Nano, resistores com valores de 33 Ohms (3 unidades) que reduzem a tensão para aproximadamente 3.5 V (dentro da faixa de operação dos motores) e 1.2 kOhms (3 unidades) utilizados na base dos transistores BC547 e destinados a diminuir a corrente que chega aos motores (40 mA idealmente). O regulador de tensão L7805 é utilizado para rebaixar a tensão de 9 V para 5 V, e capacitores de 10  $\mu$ F (1 unidade) e 0.22  $\mu$ F (1 unidade) para auxiliar o funcionamento do regulador de tensão, reduzindo ruídos.

Logo após os cálculos e testes feitos na protoboard, foi projetado um desenho do circuito como mostra a Figura 3.

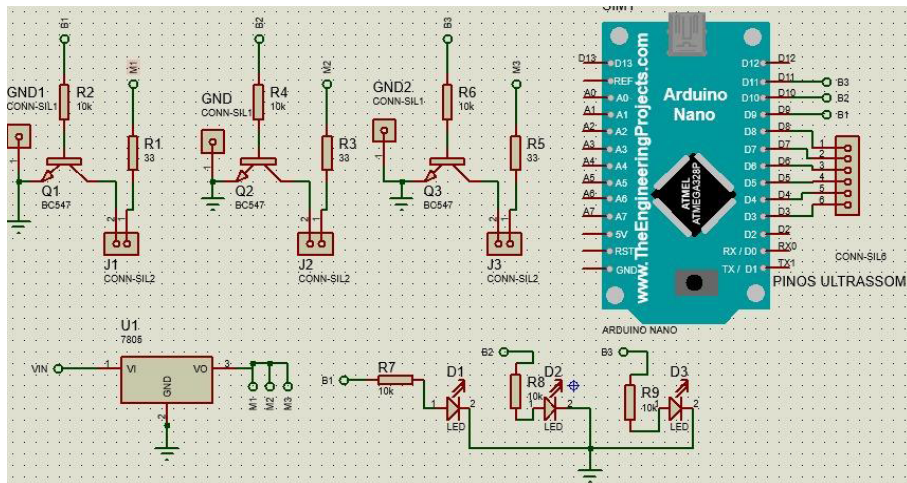


Figura 3. Visualização 2D do circuito eletrônico desenvolvido em software de simulação de circuitos.

A confecção da PCB (Printed Circuit Board) foi realizada por processo de termotransferência, método popular em confecções de circuitos eletrônicos. Foi transferida por calor a tinta do desenho da PCB para uma placa de fenolite coberta por cobre. Após isso foi a corrosão da placa com percloroeto de ferro, que remove a camada de cobre que não está coberta pela tinta, restando cobre somente nas trilhas projetadas pelo desenho. Podendo enfim soldar os componentes calculados por nós.

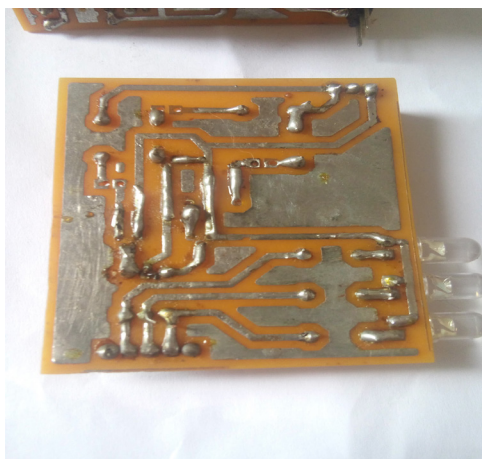


Figura 4. PCB Confeccionada através do método de termotransferência com os componentes soldados. Autoria Própria.

## Aplicativo G2B

Neste capítulo, iremos detalhar e explicar como a segunda parte do sistema Guide2Blind foi criada e como ela funciona. O sistema G2B foi criado para auxiliar deficientes visuais na localização e orientação em ambientes fechados para que eles não se percam e possam ir de um ponto a outro pelo melhor caminho possível. Quando é dito melhor caminho possível, queremos dizer direcionar a pessoa pelo menor caminho cujo qual tenha menos obstáculos e menos fluxo de pessoas, se possível.

O sistema G2B é fragmentado em duas partes. A primeira parte é voltada ao usuário, desenvolvida para dispositivos móveis, por hora *Android*, e possui uma interface simples e, que com um toque na tela e um comando de voz, o sistema consiga se comunicar com o usuário. A segunda parte é um sistema feito para *browsers*<sup>1</sup> voltada a administradores ou técnicos que atuem no sistema para fazer o mapeamento de novos lugares ou quando um lugar, já existente, sofrer alterações na planta, possa-se atualizar os dados de forma eficiente.

### Buscando distância até a roteador

Para atender os objetivos de localização dividimos o conceito de localização em dois princípios, cálculo de distância dos PA's e mapeamento:

O grupo decidiu desenvolver o sistema por perda de sinal por apresentar muita facilidade de manipulação e não possuir momentos que o sinal de estímulo pode sofrer muitas variações, gerando as imprecisões indesejadas. E para obter a distância entre o aparelho receptor do sinal Wi-Fi e roteador através da potência do sinal da antena é o utilizado Equação 1:

$$D = 10^{\frac{RSSI - TxPower}{10 * y}}$$

Equação 1. Equação de cálculo da distância do emissor de Wi-Fi através da potência.

Esta fórmula atualmente tem mostrado precisão mesmo apresentando simplicidade necessitando de poucas informações para obter a distância, mas atualmente é muito estudado qual de fato é a fórmula mais precisa para esta ação. Para entender como funciona a fórmula e se apresenta precisão para ser aplicado no aplicativo, o grupo fez um teste, calculando pelo dBm a distância e comparando com a distância real.

Entendendo estas perdas podemos utilizar um valor médio de “y” para diminuir o erro da medição. No geral o valor médio é tabelado, mas para efetuar os testes no IFSul, o grupo fez medições de 1m a 4m, e calculou o valor de perda média no câmpus. e o valor de “y” para o IFSul câmpus charqueadas foi de aproximadamente -0,99475. Para o teste

<sup>1</sup> Browser significa navegador de Internet.

da fórmula foram medidos valores de potência do sinal (RSSI) para diferentes valores de distância, então colocamos em uma tabela para identificar a precisão dos cálculos.

## A trilateração

Um dos motivos para migrar do módulo de internet conectado ao dispositivo para a utilização de dispositivos móveis foi a eficiência na coleta de dados das redes *wireless* próximas. O *Android* permite que colemos dados de potência de sinal (RSSI) e o nome das redes (SSID) em tempo contínuo sem a necessidade de se conectar à rede.

Optamos por capturar as informações a cada meio segundo em um processo paralelo e as salvamos em um vetor de objetos do tipo *Point* (criado pelo grupo). Esse vetor é enviado para a classe que cuidará da trilateração para achar a posição da pessoa. A classe que trilatera irá, primeiro, checar se realmente três roteadores foram encontrados, depois irá ordená-los da maior RSSI para a menor RSSI e então aplicará os cálculos matemáticos e irá comparar a posição X e Y da pessoa com a última posição computada.

Essa comparação serve para identificar se não houve mudanças de perda de sinal que possa interferir na precisão do cálculo. A comparação é feita calculando a distância do último ponto através da fórmula:

$$D = \sqrt{(x_2 - x_1)^2 - (y_2 - y_1)^2}$$

Equação 2. Utilização da equação de cálculo da distância do emissor de Wi-Fi através da potência. Onde: D = Distância entre 2 pontos. X2 = coordenada "X" do ponto 2. X1 = coordenada "X" do ponto 1. Y2 = coordenada "Y" do ponto 2. Y1 = coordenada "Y" do ponto 1.

Após o obter a distância, calculamos o deslocamento considerando que uma pessoa se mova a no máximo 5m/s, caso o resultado seja maior que 5m/s considera-se que tenha ocorrido um erro no cálculo de trilateração. Futuramente pretendemos utilizar o acelerômetro do celular para calcular com maior precisão a velocidade que a pessoa está se movendo para ajudar no segundo cálculo. Vale ressaltar que mesmo que não seja requisitado pelo usuário saber sua localização, o sistema continuará calculando para melhorar a precisão.

## Assistente de voz

O G2B (nome provisório do aplicativo) é, como dito anteriormente, um sistema voltado para o deficiente visual que consiste em guiá-los de um ponto a outro ajudando-os caso se percam. Esse sistema foi desenvolvido utilizando a IDE<sup>2</sup> *Android Studio* e sua interface é composta apenas por uma área clicável onde a pessoa toca na tela e dá um comando de voz.

Para a comunicação com o usuário, desenvolvemos um assistente pessoal chamado *IGOR (Indoor Guide for Orientation and Routes)*, traduzindo, guia interno para orientação e

<sup>2</sup> IDE significa Integrated Development Environment ou Ambiente de Desenvolvimento Integrado.

rotas). Seu funcionamento interno é dado após o toque na tela. A pessoa terá três segundos para falar a instrução do que ela quer fazer. Este comando de voz é interpretado através da API<sup>3</sup> da *Google Speech-to-text*<sup>4</sup>, essa API, como o próprio nome já diz em tradução literal, transcreve áudios, comandos de voz para texto (*strings*).

```
this.speechController = new STT() {
    @Override
    public void onResults(Bundle bundle) {
        super.onResults(bundle);

        final ArrayList<String> palavras = bundle.getStringArrayList(
            SpeechRecognizer.RESULTS_RECOGNITION
        );

        String discurso = "";

        if (palavras != null) {
            discurso += palavras.get(0) + " ";
        }

        String s = ResponseProcess.process(discurso, wifilist, mwf, pathDrawer, getApplicationContext());
    }
}
```

Figura 5. Trecho do código para configuração do reconhecimento de voz. Autoria Própria.

A imagem acima mostra como utilizamos a API para processar o comando de voz. Após os três segundos ouvindo o usuário, o método *onResults* é chamado as palavras ditas e transcritas do usuário ficam salvas no *ArrayList*<sup>5</sup> “palavras”. Na última linha, enviamos os comandos de voz para a classe *ResponseProcess* e lá adicionamos um mapa com cada comando de voz aceitável e suas variações/sinônimos, essas informações são salvas em um arquivo XML (imagem xml). O arquivo XML também facilitará a tradução de linguagens dos comandos de voz futuramente.

A resposta do sistema para o usuário é dada por outra API da Google chamada *Text-to-speech*<sup>6</sup> que faz o processo inverso da *Speech-to-text*, transcreve textos em palavras. Neste primeiro instante, o grupo utiliza esse sistema, entretanto, a voz é engessada de mais e, futuramente, buscaremos por alguém que grave os áudios de resposta.

3 API significa Application Programming Interface que é um conjunto de códigos pré-prontos para desenvolvimento.

4 Disponível em: <https://cloud.google.com/speech-to-text/>

5 *ArrayList* é um vetor dinâmico da linguagem de programação Java

6 Disponível em: <https://developer.android.com/reference/android/speech/tts/TextToSpeech>

```
<string name="soMuchPeopleDanger">Fique atento, há um grande fluxo de pessoas em sua trajetória</string>
<string name="soMuchObsDanger">Fique atento, várias pessoas relataram obstáculos em sua trajetória</string>
<string name="_default">Não consegui entender.</string>
```

Figura 6. Trecho do arquivo XML com as respostas e instruções aceitáveis pelo assistente.  
Autoria Própria.

## O melhor caminho

Um dos diversos comandos aceitáveis pelo nosso sistema é a pessoa perguntar o caminho até determinado local. Para isso, implementamos um sistema de rotas baseado, atualmente, no algoritmo de Dijkstra. Esse algoritmo é capaz de pegar pontos de um mapa ou grafo e retornar o menor caminho entre dois pontos (suas coordenadas e distância). Abaixo está uma representação gráfica de como o algoritmo funciona:

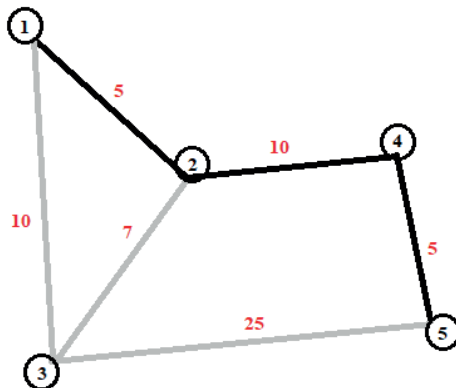


Figura 7. A figura mostra que o menor caminho do ponto 1 ao ponto 5 passa pelos pontos 2 e 4. Autoria Própria.

Atualmente, o sistema consegue apenas mostrar o caminho, mas o próximo passo é guiar a pessoa até o destino dando comandos de voz similares aos do GPS como por exemplo “siga em frente por mais 2 metros, depois vire a direita”.

## O sistema inteligente

Esta parte do sistema está na metade do processo e servirá, futuramente, para melhorar o sistema de rotas dando a melhor rota, com menos fluxo de pessoas e com menos obstáculos. Ela armazena as últimas coordenadas do usuário junto com o horário do sistema em que a coordenada foi registrada e, se houver conexão com a internet, enviará



para o nosso servidor.

Para armazenar os dados é criado um banco de dados local (MySQL Lite) e após N coordenadas armazenadas, o aplicativo sincronizará com o servidor e enviará os dados para ele. O servidor foi programado utilizando a linguagem de programação PHP com a IDE *PHPStorm*<sup>7</sup> e o banco de dados, atualmente, é feito em MySQL com o ambiente *MySQL WorkBench*<sup>8</sup>.

Esse sistema também servirá para levantamento de dados, na página do administrador possuímos uma interface que faz uma requisição aos dados do banco de dados e gera um mapa térmico com base na coordenada e horário. Para gerar o gráfico a partir dos dados vindos do banco de dados, o grupo utilizou a API em JavaScript que desenha mapas térmicos chamada *Chart & Graph*<sup>9</sup>.

A Figura 8 ilustra como o sistema mostraria os gráficos, já que ele ainda não foi lançado para o público.

Por fim, a última funcionalidade do sistema inteligente é o sistema de emergência. Se porventura o deficiente visual estiver em uma situação de risco ou sofrer algum tipo de acidente, bastará clicar na tela e dizer “socorro (nome da pessoa)”. Com isso, o sistema, tentará enviar uma mensagem para o servidor do projeto enviando sua posição e um dos representantes da equipe entrará em contato para ajudar. Depois de enviar o pedido de ajuda para o sistema, o aplicativo ligará para o número da emergência.

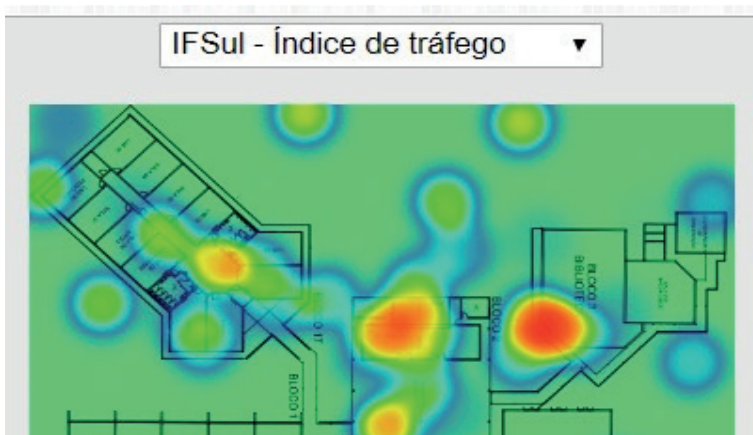


Figura 8. Os pontos mais em vermelho representam os locais com maior fluxo de pessoas. Já os pontos mais em azul, representam os pontos com menor fluxo. Autoria Própria.

7 Disponível em: <https://www.jetbrains.com/phpstorm/>

8 Disponível em: <https://www.mysql.com/products/workbench/>

9 Disponível em: <https://www.cssscript.com/category/categories/chart-graph/>

## O mapeamento

O principal foco do trabalho é localizar, orientar e guiar um deficiente visual dentro de um ambiente fechado para que ele não se perca. Entretanto para guiá-lo precisam-se de pontos de referência, rotas e um mapa interno indicando, por exemplo, as salas de aula, banheiros, laboratórios, etc. Como não encontramos nenhuma ferramenta de mapeamento que fosse maleável o suficiente para gerar os dados que precisamos para nosso sistema, decidimos por dar início a criação da nossa própria ferramenta.

Nossa ferramenta, ainda em fase de desenvolvimento, traz uma interface simples para que qualquer pessoa que queira mapear um estabelecimento consiga fazê-la sem grandes problemas. Após o mapeamento o sistema gerará automaticamente todas as informações necessárias no formato que nosso aplicativo interpreta como as de rotas, locais onde cada sala esteja e informações de roteadores.

A ferramenta foi desenvolvida para *browsers* e foi programada utilizando a *framework*<sup>10</sup> de *JavaScript* chamada *JQuery*<sup>11</sup>, *PHP* e utiliza um *framework* para os ícones chamada *Font Awesome*<sup>12</sup>. Buscamos por simplificar a interface, à esquerda, encontramos um menu de ferramentas. A primeira ferramenta adiciona uma planta de “x” largura e “y” altura. Ao inserir a planta do estabelecimento algumas ferramentas irão aparecer, a segunda, com ícone de lápis, serve para desenhar, na planta no meio da tela, onde cada estabelecimento está. Ao fechar o desenho aparecerá um polígono colorido indicando, visualmente o lugar e também, à direita da tela, aparecerá, hierarquicamente, cada estabelecimento marcado e com dois cliques poderá alterar o nome do lugar. O terceiro e quarto ícones servem para adicionar marcadores de pontos de ligação e rotas que os interligam. O motivo para separar rotas dos pontos é que na hora de gerar as rotas, o algoritmo buscará os marcadores adicionado dentro da área demarcada e considerará como uma entrada/porta.

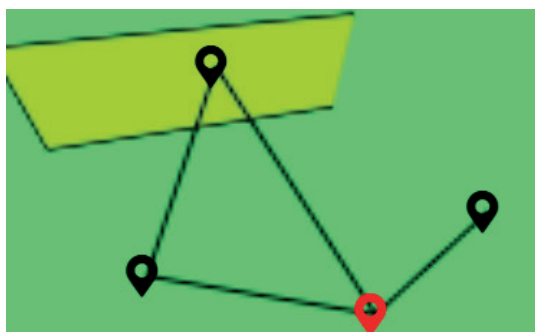


Figura 9. Exemplo de criação de rotas, marcadores, e demarcação de área. Autoria própria.

<sup>10</sup> Framework são conjuntos de códigos com funções específicas para facilitar o desenvolvimento de aplicações.

<sup>11</sup> Disponível em: <https://api.jquery.com/>

<sup>12</sup> Disponível em: <https://fontawesome.com/>

Por fim, o último botão serve para exportar os arquivos. Existem três camadas sobrepostas na grade principal para organizar a exportação dos arquivos, cada camada é um *Canvas* de *JavaScript*. Ao finalizar a marcação, o sistema lerá cada informação dos *Canvas* e gerará um arquivo de texto. Com os dados do mapeamento prontos, bastará apenas colocá-lo no aplicativo.

## 4 | RESULTADOS E DISCUSSÕES DOS RESULTADOS

O dispositivo apresentou uma “zona morta” no sistema de detecção, entre os sensores ultrassônicos, da forma que foram posicionados no mock-up. A falha possivelmente ocorreu devido ao formato dos sensores, bem como das pequenas imprecisões inerentes ao processo de construção. Essas imprecisões não podem ser corrigidas, pois para corrigi-las seria necessário colocar 3 sensores no mesmo ponto exato o que é fisicamente impossível. Não ocorreu sobreposição de sinal.

A partir do gráfico apresentado na Figura 14 é possível verificar que as medidas de sensores são imprevisíveis de 0 cm a 5 cm, e as medidas fora deste intervalo possuem uma margem de erro de 2 cm no sentido positivo. Como nosso projeto não necessita extrema precisão nas medidas de distâncias do obstáculo, esta margem de erro não é um problema nas medições.

A partir dos resultados do seguinte teste geramos o seguinte gráfico de erro de medidas do sensor ultrassônico Figura 10:

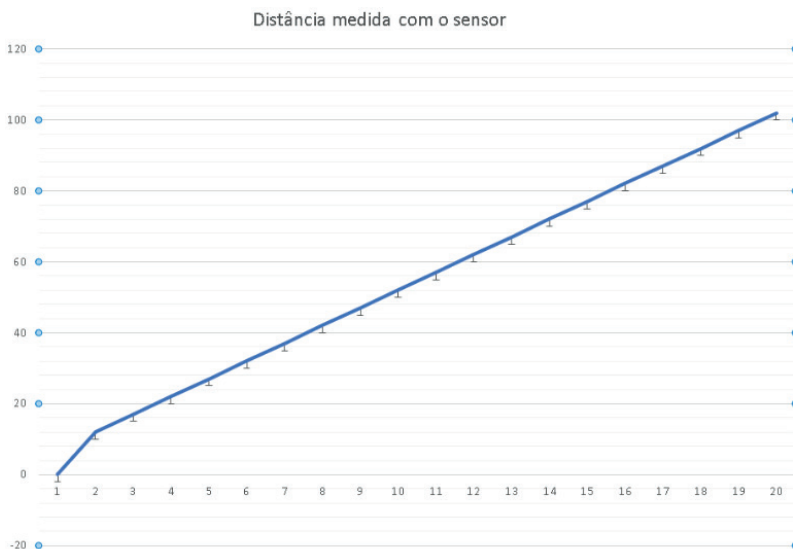


Figura 10. Gráfico de erro de dimensões do sensor ultrassônico. Autoria própria.

Os testes de comprovação da fórmula para o câmpus com a constante “y” alterada apresentou uma boa precisão em todas medidas em uma sala fechada. Como a equipe decidiu migrar para o Android, os testes de simulações que ocorreram em software apresentaram uma precisão de localização em 95% dos casos. Atualmente o grupo estuda um ambiente de teste ideal com sinais de roteadores controlados, para assim efetuar um teste do sistema na prática.

## 5 | RESULTADOS E DISCUSSÕES DOS RESULTADOS

Ao final do processo de desenvolvimento, espera-se que o circuito esteja funcional e compacto dentro do modelo. O grupo cobiça que o design do dispositivo não seja desconfortável para o usuário e que os sensores funcionem sem interferência externa para captar a distância dos obstáculos e vibrar os motores. Espera-se também que o sistema de localização indoor esteja mais preciso e otimizado, ou pelo menos funcional em primeira instância, para que o grupo possa obter algum resultado.

Com o resultado da simulação, o sistema conseguiu definir a posição da pessoa com precisão em 95% dos casos. Entretanto, o grupo está na fase de estudo do melhor ambiente de teste para identificar pontos que o aplicativo deverá ser modificado, para então aprimorá-lo. Para aumentar a precisão do sistema durante o processo de identificação de objetos, o grupo projetou uma estrutura de base para os componentes do protótipo a partir de modelagem em software de desenvolvimento 3D, seguida de impressão 3D. Estudamos o uso do ESP8266 e ele trouxe algumas desvantagens como a necessidade de utilização de outros módulos para a comunicação do usuário com o dispositivo. Com mais módulos, precisa-se de mais portas no microprocessador, o que, como consequência, encarecerá o produto final e exigirá uma remodelagem, sendo necessário inclusive a troca do microprocessador. Para evitar a inserção de novos módulos, o grupo pretende migrou as funções que seriam do ESP8266 para um aplicativo para dispositivos móveis (*Android*). e busca aprimorá-las inclusive apresentando uma integração maior com o dispositivo impresso, como incluir um canal de comunicação direto entre eles. A API “speech-to-text” apresentou precisão ao receber comandos e a resposta do aplicativo com a API “Text-to-Speech” tornou a resposta muito mais simples e rápida. Futuramente o grupo almeja coletar com precisão os dados de obstáculo para gerar dados para o sistema inteligente, além tornar a indicação da direção a ser seguida pelo dispositivo. Ao sistema de rotas deseja-se implementar um sistema de comandos para aplicativo indicar a direção deve ser seguida, com comandos do tipo “siga em 2 metros e após vire a esquerda”, “você está no seu destino”.

## REFERÊNCIAS

ABAADV. **Associação Beira Agueira de Apoio ao Deficiente Visual**. 2018.

BADALO, C. A. de O. **O Papel Do Cão-Guia Como Facilitador Da Inclusão Da Pessoa Cega Na Sociedade: Mobilidade, Segurança, Interação Social E Qualidade De Vida**. 2014.

BERSCH, R. **INTRODUÇÃO À TECNOLOGIA ASSISTIVA**. [s.l.], 2017.

BRASIL. **Decreto 8953 - Altera o Decreto nº 7.963, de 15 de março de 2013, que institui o Plano Nacional de Consumo e Cidadania e cria a Câmara Nacional das Relações de Consumo**. [s.l.]: [s.n.], 2017. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_Ato2015-2018/2017/Decreto/D8953.htm](http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2017/Decreto/D8953.htm)>. Acesso em: 14/maio/18.

BRUMER, Anita; PAVEI, Katiuci; MOCELIN, Daniel Gustavo. **Saindo da “escuridão”: perspectivas da inclusão social, econômica, cultural e política dos portadores de deficiência visual em Porto Alegre**. *Sociologias*, Porto Alegre, n. 11, p. 300-327, jun. 2004. Disponível em <[https://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1517-45222004000100013&lng=pt&tlng=pt](https://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-45222004000100013&lng=pt&tlng=pt)>. Acessos em 11 mar. 2021. <https://doi.org/10.1590/S1517-45222004000100013>.

COHEN, L. G. et al. **Functional relevance of cross-modal plasticity in blind humans**. *Nature*, [s.l.], v. 389, nº 6647, p. 180–183, 1997.

GLEICK, J. **A Informação: Uma história, uma teoria, uma enxurrada**. *Journal of Chemical Information and Modeling*, [s.l.], v. 53, nº 9, p. 1689–1699, 2013.

KASTRUP, V. et al. **O aprendizado da utilização da substituição sensorial visuo-tátil por pessoas com deficiência visual: primeiras experiências e estratégias e metodológicas**. [s.l.], 2009.

KOTARU, M. et al. **SpotFi : Decimeter Level Localization Using WiFi**. *Sigcomm 2015*, [s.l.], p. 269–282, 2015a. In: *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication - SIGCOMM '15*. New York, New York, USA: ACM Press, 2015b.

MARQUES, J. M. dos S. **Catálogo de entendimento de informações gráficas para cidadãos cegos**. [s.l.], 2017.

RANGEL, M. L. et al. **Deficiência visual e plasticidade no cérebro humano**. [s.l.], v. 12, nº 1, p. 197–207, 2010.

SÁ, E. D. **Acessibilidade: as pessoas cegas no itinerário da cidadania**. *Nossos Meios*, [s.l.], v. 24, 2003.

SADATO, N. et al. **Activation of the primary visual cortex by Braille reading in blind subjects**. *Nature*, [s.l.], v. 380, nº 6574, p. 526–528, 1996.

SOARES, A. V. **A contribuição visual para o controle postural**. *370 Rev*, [s.l.], v. 18, nº 3, p. 370–379, 2009.

## DESPACHO DE GERAÇÃO ÓTIMA ATRAVÉS DO MÉTODO DOS PONTOS INTERIORES VERSÃO PRIMAL-DUAL

Data de aceite: 01/04/2021

### Jean Ferguson Pimentel

Universidade Federal do Paraná, Pós-Graduação em Engenharia Elétrica  
Curitiba – Paraná  
<http://lattes.cnpq.br/3417083536054492>

### João Vitor Gerevini Kasper

Universidade Federal do Paraná, Pós-Graduação em Engenharia Elétrica  
Curitiba – Paraná  
<http://lattes.cnpq.br/1192062840173129>

### Juliana Almansa Malagoli

Universidade Federal do Paraná, Pós-Graduação em Engenharia Elétrica  
Curitiba – Paraná  
<http://lattes.cnpq.br/6153426592290224>

### Thelma Solange Piazza Fernandes

Universidade Federal do Paraná, Pós-Graduação em Engenharia Elétrica  
Curitiba – Paraná  
<http://lattes.cnpq.br/5939715471342143>

**RESUMO:** O conhecimento da técnica de otimização através do Método dos Pontos Interiores versão Primal-Dual é de grande importância para solucionar problemas aplicados ao sistema elétrico de potência. Neste contexto, o presente trabalho tem por objetivo apresentar uma solução ótima de despacho de geração de energia elétrica para 4 usinas térmicas, através de simulação no software MATLAB. Além disso,

o intuito é minimizar o custo de produção, minimizar as perdas e considerar em dado momento a alteração de limite de geração em um das unidades geradoras. Os resultados mostram as análises do despacho de geração, ademais, configura-se como um interessante tema de pesquisa, já que está relacionado ao sistema elétrico de potência com redução de custos e perdas em usinas térmicas.

**PALAVRAS - CHAVE:** Método Numéricos, Método dos Pontos Interiores, Otimização, Sistemas Elétricos de Potência.

### DISPATCH IN OPTIMUM GENERATION THROUGH OF INTERIOR POINTS METHOD PRIMAL-DUAL VERSION

**ABSTRACT:** The knowledge of the optimization technique through the Interior Point Method Primal-Dual version is of great importance to solve problems applied to the electric power system. In this context, the present paper aims to present an optimal solution for dispatching electricity generation for 4 plants thermal, through simulation in MATLAB software. In addition, the aim is to minimize the cost of production, minimize losses and at any given time consider changing the generation limit at one of the generating units. The results show the analysis of the generation dispatch, in addition, it is configured as an interesting research topic, since it is related to the electric power system with cost reduction and losses in thermal plants.

**KEYWORDS:** Numerical Method, Interior Point Method, Optimization, Electric Power Systems.

## 1 | INTRODUÇÃO

A energia elétrica proporciona segurança, conforto, informação e lazer para uma sociedade. Sendo então, um extraordinário indicador de desenvolvimento social e humano que nos dias atuais já faz parte da realidade da maioria das casas brasileiras. Dessa forma, a usina térmica é exclusivamente de um sistema que corresponde as somas das energias cinéticas, pelo efeito de sua temperatura (Souza, 2019), (Liu *et al.*, 2019).

O planejamento de operações do sistema elétrico de potência é um conjunto de procedimentos que usa métodos numéricos para a representação dos sistemas. Neste cenário, avalia-se os comportamentos e as alternativas para garantir o fornecimento de energia. Além disso, analisa-se os custos de operação e a solução mais econômica. Portanto, destaca-se por uma operação ótima do sistema (Rodrigues, 2007).

Dessa maneira, busca-se por aplicações de técnicas que proporcione a ótima solução para geração, transmissão e distribuição de energia elétrica. Ademais, visa sempre a melhor relação custo/benefício. Sendo assim, o uso do método dos Pontos Interiores versão Primal-Dual auxilia na busca de um despacho ótimo de geração e a redução de perdas (Probst e Oliveira, 2013). Assim, evita os desperdícios e proporciona a redução de custos.

Nesse contexto, evidencia-se a relevância cada vez maior da prestação de serviços ancilares aos sistemas de energia. A partir deles, garante-se não apenas o atendimento apropriado da demanda da carga por meio da geração e transmissão de energia, como também a estabilidade e segurança da rede (Oliveira *et al.*, 2003), (Yamagutti, 2019) e (Souza, 2019).

O presente trabalho tem por objetivo minimizar o custo de produção de energia elétrica através da otimização da distribuição da produção entre os geradores e da utilização eficiente dos recursos energéticos. Portanto, realiza-se o despacho econômico, que é o estudo da alocação ótima de uma demanda entre unidades geradoras de usinas térmicas.

Este artigo, organiza-se em 5 seções, as quais seguem os seguintes tópicos: a seção 2 apresenta o método de pontos interiores versão primal-dual, a seção 3 explana sobre a metodologia adotada, a seção 4 é designada aos resultados e discussões e por fim, a seção 5 descreve as considerações finais.

## 2 | TÉCNICA DE OTIMIZAÇÃO DE PONTOS INTERIORES VERSÃO PRIMAL-DUAL

O método de pontos interiores primal-dual é desenvolvido através do método de Newton, às condições de otimalidade desconsiderando-se as restrições de não-negatividade e incluindo uma perturbação nas condições de complementaridade. O método parte de um ponto estritamente positivo e não permite que as variáveis se tornem negativas (Oliveira *et al.*, 2003), (Probst e Oliveira, 2013).

Neste trabalho, pretende-se desenvolver um programa computacional que minimize o custo de geração e as perdas elétricas através da Técnica de Otimização de Pontos Interiores versão Primal-Dual. Tendo como objetivo atender uma carga demandada de 400 MW, através de quatro usinas térmicas geradoras que tem seus custos e capacidade de geração apresentadas na Tabela 1.

Produtor	Custo [\$/h]	Capacidade [MW]
1	$C_1(Pg_1) = 140 + 80Pg_2 + 9.10^{-3} Pg_2^2$	$0 \leq Pg_1 \leq 400$
2	$C_2(Pg_2) = 100 + 90 Pg_3 + 8.10^{-3} Pg_3^2$	$10 \leq Pg_2 \leq 400$
3	$C_3(Pg_3) = 90 + 70 Pg_3 + 8.10^{-3} Pg_3^2$	$0 \leq Pg_3 \leq 400$
4	$C_4(Pg_4) = 100 + 80 Pg_4 + 0.10^{-3}.Pg_4$	$0 \leq Pg_4 \leq 300$

Tabela 1: Custos de geração e capacidades de geração.

Onde:  $Pg$  é a potência gerada de cada usina (MW);  $C$  é o custo de geração por unidade de energia elétrica. A perda de potência ativa é expressa pela Equação (1).

$$Pl(\mathbf{Pg}) = \text{diag}(Pg) \cdot \mathbf{Bloss} \cdot \mathbf{Pg} \quad (1)$$

Sendo, que  $Pl$  é perda de potência ativa nas linhas;  $\mathbf{Bloss}$  é uma matriz ( $ng \times ng$ ) que relaciona as perdas por barra de geração:

$$\mathbf{Bloss} = \begin{bmatrix} 2 \times 10^{-4} & 1 \times 10^{-4} & 0 & 1 \times 10^{-4} \\ 1 \times 10^{-4} & 2 \times 10^{-4} & 1 \times 10^{-4} & 1 \times 10^{-4} \\ 0 & 1 \times 10^{-4} & 2 \times 10^{-4} & 1 \times 10^{-4} \\ 0 & 1 \times 10^{-4} & 2 \times 10^{-4} & 3 \times 10^{-4} \end{bmatrix}$$

O objetivo do problema é obter os valores de  $Pg_1$ ,  $Pg_2$ ,  $Pg_3$  e  $Pg_4$  que atenda a carga demandada de 400 MW para as seguintes condições:

(i) Função objetivo que minimiza custo de produção:

$$\text{Min } C(\mathbf{Pg}) = \mathbf{Pg}' \mathbf{QPg} + \mathbf{b}' \mathbf{Pg} + co \quad (2)$$

$$\sum_{i=1}^{ng} Pg_i - Pd - Pl(\mathbf{Pg}) = 0 \quad (3)$$

$$Pg \text{ min} \leq Pg \leq Pg \text{ max} \quad (4)$$



Em que:

(2) Define o valor dos custos de geração de potência ativa de cada unidade;

(3) Define a potência gerada pelas usinas deve ser capaz de atender a carga e as perdas;

(4) Define as equações de restrições operativas que cada máquina deve obedecer.

(ii) Função objetivo que minimiza as perdas elétricas:

$$C(Pg) = \mathbf{u}^t Pg$$

$$\sum_{i=1}^{ng} Pg_i - Pd - Pl(\mathbf{Pg}) = 0 \quad (6)$$

$$Pg_{\min} \leq Pg \leq Pg_{\max} \quad (7)$$

Em que:

(5) Define a nova função objetivo do problema que será a somada das potências geradas;

(6) Define que a potência gerada pelas usinas deve ser capaz de atender a carga e as perdas;

(7) Define as equações de restrições operativas que cada máquina deve obedecer.

Além disso, destaca-se a definição de outros parâmetros:  $\mathbf{Q}$  é a matriz diagonal com coeficientes quadráticos de custo;  $\mathbf{b}$  é o vetor de coeficientes lineares de custo;  $co$  é a constante de custo;  $Pd$  é o vetor de cargas nas  $nb$  barras;  $\mathbf{u}$  é o vetor unitário.

Após encontrar os valores de  $Pg_1$ ,  $Pg_2$ ,  $Pg_3$  e  $Pg_4$  para as condições, supõe-se uma alteração na capacidade máxima de geração da usina 3 para um valor de 100 MW e encontra-se novamente os valores  $Pg_1$ ,  $Pg_2$ ,  $Pg_3$  e  $Pg_4$  nessas condições, através das funções objetivo (i) e (ii).

### 3 I METODOLOGIA

Para obter a solução ótima do problema, faz-se necessário a montagem da Função Lagrangeana do problema de otimização, onde são consideradas as Equações de custo de cada usina conforme Equação (8).

$$L(\mathbf{Pg}, \lambda, \pi) = \mathbf{Pg}' \mathbf{Q} \mathbf{Pg} + \mathbf{b}' \mathbf{Pg} + co - \mu \sum_{i=1}^{ng} [\ln(\mathbf{smin}_i) + \ln(\mathbf{smax}_i)] + \lambda \cdot [Pd - \sum_{i=1}^{ng} (Pg_i)] + \pi \mathbf{max}' \cdot (\mathbf{Pg} - \mathbf{Pgmax} + \mathbf{smax}) + \pi \mathbf{min}' \cdot (-\mathbf{Pg} + \mathbf{Pgmin} + \mathbf{smin}) \quad (8)$$

As variáveis **Pg**, **smin** e **smax** são as variáveis primal do problema e as variáveis  $\lambda$ , **mmax** e **mmin** são as duais. Tendo a função Lagrangeana do problema, aplicou-se as condições de otimalidade de Karush-Kuhn-Tuck que consiste na primeira derivada:

$$\nabla_{\mathbf{Pg}} L = 2 \cdot \mathbf{Q} \mathbf{Pg} + \mathbf{b} - \lambda \cdot [\mathbf{e}] - \boldsymbol{\pi}_{min} + \boldsymbol{\pi}_{max} = 0 \quad (9)$$

$$\nabla_{\lambda} L = \mathbf{Pd} - \sum_{i=1}^{ng} (Pgi) = 0 \quad (10)$$

$$\nabla_{\boldsymbol{\pi}_{max}} L = \mathbf{Pg} - \mathbf{Pg}_{max} + \mathbf{s}_{max} = 0 \quad (11)$$

$$\nabla_{\boldsymbol{\pi}_{min}} L = -\mathbf{Pg} + \mathbf{Pg}_{min} + \mathbf{s}_{min} = 0 \quad (12)$$

$$\nabla_{\mathbf{s}_{max}} L = -\boldsymbol{\mu} \cdot \mathbf{e} + \mathbf{s}_{max} \cdot \boldsymbol{\pi}_{max} = 0 \quad (13)$$

$$\nabla_{\mathbf{s}_{min}} L = -\boldsymbol{\mu} \cdot \mathbf{e} + \mathbf{s}_{min} \cdot \boldsymbol{\pi}_{min} = 0 \quad (14)$$

Após isso, aplicou-se o método de Newton, supondo uma função  $f(z)$ , que deseja encontrar o ótimo da função partindo de um ponto inicial  $\mathbf{Z}^K$ , em direção de  $\Delta \mathbf{Z}$ . Assim,

$$\Delta \mathbf{Z}^K = - \left[ W(\mathbf{Z}^K) \right]^{-1} * f(\mathbf{Z}^K) \quad (15)$$

Sendo, que a cada iteração  $k=k+1$  atualiza  $\mathbf{Z}^k$  conforme Equação (16).

$$\mathbf{Z}^{K+1} = \mathbf{Z}^K + \Delta \mathbf{Z}^K \quad (16)$$

Onde:  $W$  é a matriz Hessiana, com a seguinte estrutura:

$$\mathbf{W} = \begin{bmatrix} 2\mathbf{Q} & -\mathbf{e} & \text{diag}(\mathbf{e}) & -\text{diag}(\mathbf{e}) & \mathbf{0} & \mathbf{0} \\ [-\mathbf{e}]^t & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \text{diag}(\mathbf{e}) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \text{diag}(\mathbf{e}) & \mathbf{0} \\ -\text{diag}(\mathbf{e}) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \text{diag}(\mathbf{e}) \\ \mathbf{0} & \mathbf{0} & \mathbf{Smax} & \mathbf{0} & \boldsymbol{\Pi} \max & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Smin} & & \boldsymbol{\Pi} \min \end{bmatrix}$$

Com isso, a cada iteração o sistema de equações a serem resolvidos é:

$$\mathbf{W} \cdot \begin{bmatrix} \Delta \mathbf{Pg} \\ \Delta \lambda \\ \Delta \pi_{\max} \\ \Delta \pi_{\min} \\ \Delta s_{\max} \\ \Delta s_{\min} \end{bmatrix} = - \begin{bmatrix} \Delta_{p_g} L \\ \Delta \lambda L \\ \Delta_{\pi_{\max}} L \\ \Delta_{\pi_{\min}} L \\ \Delta_{s_{\max}} L \\ \Delta_{s_{\min}} L \end{bmatrix}$$

A cada iteração, o sistema linear é resolvido, sendo a próxima etapa a determinação do comprimento do passo nos espaços primal e dual. De modo que, as variáveis de folga ***smin***, ***smax*** sejam todas  $\geq 0$  e os multiplicadores de Lagrange sejam: ***πmin***  $\geq 0$ , ***πmax***  $\geq 0$ . Após cada iteração, deve-se recalculer o parâmetro barreira  $\mu$ , conforme Equação (17).

$$\mu = \frac{\mathbf{s}^t \boldsymbol{\pi}}{2 \cdot \mathbf{J} \cdot \beta} \quad (17)$$

Resumidamente, o algoritmo para resolução do problema consiste em:

1. Escolha  $\mu_0$ , valores iniciais para variáveis primais e duais. Faça  $k = 0$ ;
2. Calcule o valor das condições de otimalidade (cálculo de  $\nabla z L$ );
3. Se (norma infinita de  $\nabla_z L$ )  $< tol = 10^{-6}$ , FIM, a solução é  $\mathbf{Z}^k$ .
4. Caso contrário, faça  $k = k + 1$  e vá ao Passo 4;
5. Resolução do Sistema Linear:  $\mathbf{W}^k \cdot \Delta \mathbf{Z}^k = \nabla_z L$ ;
6. Atualize todas as variáveis.
7. Atualize o parâmetro barreira  $\mu$  e retorna para o passo 2.

## 4 | RESULTADOS E DISCUSSÕES

O método proposto foi implementado no *software* computacional (MATLAB) visando testar as contingências e as alterações no problema de otimização proposto. Visando obter o resultado ótimo para o despacho de geração das 4 usinas térmicas, nos quais os custos e as perdas serão minimizados através do método dos Pontos Interiores versão Primal-Dual.

Além disso, destacam-se os estudos desenvolvidos neste trabalho:

- Caso I: a função objetivo é minimizar o custo de geração, o despacho atende uma carga de 400 MW, as perdas e os limites operativos são apresentados na

Tabela I;

- Caso II: o limite máximo da máquina 3 foi alterado de 400 MW para 100 MW, usando a mesma função objetivo anterior;
- Caso III: a função objetivo é minimizar as perdas, o despacho atende uma carga de 400 MW, as perdas e os limites operativos são apresentados na Tabela I;
- Caso IV, o limite máximo da máquina 3 foi alterado de 400 MW para 100 MW, usando a mesma função objetivo anterior.

Os valores de despacho encontrados bem como os valores de custos e das perdas para o caso I e II são apresentados nas Tabelas 2 e 3.

<b>Usina</b>	<b>Geração</b>
<i>Pg<sub>1</sub></i>	104,391 MW
<i>Pg<sub>2</sub></i>	10 MW
<i>Pg<sub>3</sub></i>	307,549 MW
<i>Pg<sub>4</sub></i>	0 MW
<b>Custo de operação</b>	<b>126.971,337 \$MW/h</b>
<b>Perdas</b>	<b>21,940 MW</b>
<b>Fator de aceleração</b>	<b>1,15</b>

Tabela 2 – Custo de operação minimizado para o Caso I.

<b>Usina</b>	<b>Geração</b>
<i>Pg<sub>1</sub></i>	203,709 MW
<i>Pg<sub>2</sub></i>	10,017 MW
<i>Pg<sub>3</sub></i>	202,891 MW
<i>Pg<sub>4</sub></i>	0,816 MW
<b>Perdas</b>	<b>17,435 MW</b>
<b>Fator de aceleração</b>	<b>1,15</b>

Tabela 3 – Perdas minimizada para o Caso II.

A alteração do limite máximo da máquina 3, provocou uma redistribuição de potência ativa entre as unidades geradoras. Com isso, o custo aumentou em relação ao caso anterior, já as perdas sofreram uma queda significativa de valor, em torno de 2 MW.

As Tabelas 4 e 5 apresentam os resultados encontrados para os casos III e IV, mas sem o custo, visto que a função objetivo usada para esses casos é a minimização das perdas.

<b>Usina</b>	<b>Geração</b>
$Pg_1$	206,535 MW
$Pg_2$	10 MW
$Pg_3$	100 MW
$Pg_4$	103,264 MW
<b>Custo de operação</b>	<b>133.024,906 \$MW/h</b>
<b>Perdas</b>	<b>19,800 MW</b>
<b>Fator de aceleração</b>	<b>1,15</b>

Tabela 4 – Custo de operação minimizado para o Caso III.

<b>Usina</b>	<b>Geração</b>
$Pg_1$	209,5493 MW
$Pg_2$	73,080 MW
$Pg_3$	99,965 MW
$Pg_4$	36,573 MW
<b>Perdas</b>	<b>19,171 MW</b>
<b>Fator de aceleração</b>	<b>1,1</b>

Tabela 5 – Perdas minimizada para o Caso IV.

A alteração da função objetivo para os casos III e IV em ambos os casos, tornou-se as perdas para as duas últimas contingências menores, quando comparadas aos casos I e II. No caso I em relação ao custo do caso III, teve uma redução de aproximadamente 22% nas perdas e na comparação com o caso IV essa redução foi de aproximadamente 4%. Nota-se, que a alteração do limite máximo de geração da máquina 3 prejudicou o valor da função objetivo.

## 5 | CONSIDERAÇÕES FINAIS

Os estudos na área de Sistemas Elétricos de Potência (SEP) estão inteiramente condicionados a restrições operativas, apenas o cálculo do estado da rede, obtido através da solução de fluxo de carga, muitas vezes não é suficiente para encontrar um ponto operativo que atenda todas as restrições do problema. Com isso, problemas de otimização mais complexos tem sido elaborados buscando modelar os fenômenos existentes no SEP de forma mais coerente.

A aplicação do método dos Pontos Interiores na versão Primal-Dual tem importância significativa no Sistema Elétrico de Potência. Neste trabalho, foi possível notar essa relevância uma vez que as alterações apresentadas para o problema, mostraram-se de

forma objetiva e direta os impactos nos custos de produção e nas perdas do sistema. Além disso, através da aplicação do método com o auxílio de um *software*, nota-se que a usina 4 tem um custo maior de produção. Porém, tem uma perda menor, já a usina 3 tem um custo de operação menor, fatos que só foram comprovados após a aplicação do método.

## REFERÊNCIAS

Liu, W. J.; Chi, M.; Liu, Z. W.; Guan, Z. H.; Chen, J.; Xiao, J. W. (2019). "***Distributed optimal active power dispatch with energy storage units and power flow limits in smart grids***", International Journal of Electrical Power & Energy Systems, Volume 105, 2019, Pages 420-428, <https://doi.org/10.1016/j.ijepes.2018.07.060>.

Oliveira, A. R. L.; Soares, S.; and Nepomuceno, L. (2003). "***Optimal active power dispatch combining network flow and interior point approaches***", in IEEE Transactions on Power Systems, vol. 18, no. 4, pp. 1235-1240, Nov. 2003, doi: 10.1109/TPWRS.2003.814851.

Probst, R. W.; Oliveira, A. R. L. (2013). "***Métodos de pontos interiores aplicados ao problema de pré-despacho de um sistema hidrotérmico***", Revista eletrônica pesquisa operacional para o desenvolvimento, V. 5, n.1, p. 14-30, Rio de Janeiro-RJ, Brasil.

Rodrigues, N. M. (2007). "***Um algoritmo cultural para problemas de despacho de energia elétrica***," Dissertação de Mestrado, Dept. Computação. Ciência da Comp., Universidade Estadual de Maringá, Maringá-PR, Brasil.

Souza, V. M. (2019). "***Pré-Despacho Ótimo da Geração Distribuída para Melhoria da Operação de Redes de Distribuição com Presença de Reguladores de Tensão no Modo Bidirecional***", Dissertação de Mestrado, Dept. Eng. Elet. Eng. Elétrica, Universidade Federal do Pará, Belém-PA, Brasil.

Yamagutti, L. C. (2019). "***Despacho ótimo de geração e controle de potência reativa no sistema elétrico de potência***", Dissertação de Mestrado, Universidade Estadual Paulista Júlio de Mesquita Filho, Ilha Solteira-SP, Brasil.

## COMBINING RAINFALL AND WATER LEVEL DATA FOR MULTISTEP HIGH TEMPORAL RESOLUTION EMPIRICAL HYDROLOGICAL FORECASTING

*Data de aceite: 01/04/2021*

*Data de submissão: 22/01/2021*

### **Cintia Pereira de Freitas**

Instituto Nacional de Pesquisas Espaciais -  
INPE  
São José dos Campos – SP  
<http://lattes.cnpq.br/9232847817073786>

### **Michael Macedo Diniz**

Instituto Federal de Educação, Ciência e  
Tecnologia de São Paulo  
São Paulo – SP  
<http://lattes.cnpq.br/3186705419723846>

### **Glauston Roberto Teixeira de Lima**

Centro Nacional de Monitoramento e Alertas de  
Desastres Naturais  
São José dos Campos – SP  
<http://lattes.cnpq.br/7867516471353997>

### **Marcos Gonçalves Quiles**

Universidade Federal de São Paulo  
São José dos Campos – SP  
<http://lattes.cnpq.br/8867164774240536>

### **Stephan Stephany**

Instituto Nacional de Pesquisas Espaciais -  
INPE  
São José dos Campos – SP  
<http://lattes.cnpq.br/1446664587151293>

### **Leonardo Bacelar Lima Santos**

Centro Nacional de Monitoramento e Alertas de  
Desastres Naturais  
São José dos Campos – SP  
<http://lattes.cnpq.br/9147853693310634>

**ABSTRACT:** This work proposes a multilayer perceptron neural network to forecast temporal series of water level at the outlet of a watershed located in a mountain region in the Brazilian state of Rio de Janeiro with up to a 2-hour antecedence, which is the time of concentration that measures the response of the watershed. This particular watershed was chosen due to previous natural disasters that affected that area, causing floods and landslides. Input data was collected by a set of hydrological monitoring stations in the considered watershed and is composed by water level and/or rainfall measures acquired with a 15-minute resolution. The neural network was implemented using the Python language, the Tensorflow library and the Keras API. Prediction results were evaluated by the Nash-Sutcliffe coefficient (NSE) and by the root mean square error, showing a good agreement between predicted and observed values of the water-level temporal series, specially when combining both water level and rainfall data. In such case, values of NSE reached 0.994 for prediction antecedence of 15 minutes, and 0.9016, for 120 minutes.

**KEYWORDS:** Watershed, Artificial Neural Network, Water Level, Hydrological Monitoring Station.

### COMBINAÇÃO DE DADOS DE PRECIPITAÇÃO E NÍVEL DE ÁGUA PARA PREVISÃO HIDROLÓGICA EMPÍRICA DE ALTA RESOLUÇÃO TEMPORAL

**RESUMO:** Este trabalho propõe uma rede neural perceptron multicamadas para previsão de séries temporais de nível de água de uma bacia

hidrográfica localizada em uma região serrana do estado do Rio de Janeiro com até 2 horas de antecedência, que é o tempo de concentração que mede a resposta da bacia hidrográfica. Essa bacia em particular foi escolhida devido a desastres naturais anteriores que afetaram aquela área, causando enchentes e deslizamentos de terra. Dados de entrada foram coletados por um conjunto de estações de monitoramento hidrológico na bacia considerada sendo composto por medidas de nível da água e/ou precipitação adquiridas com resolução de 15 minutos. A rede neural foi implementada usando linguagem Python, a biblioteca Tensorflow e a API Keras. Os resultados da previsão foram avaliados pelo coeficiente Nash-Sutcliffe (NSE) e pela raiz do erro quadrático médio, mostrando uma boa concordância entre valores previstos e observados na série temporal do nível da água, especialmente quando combinados dados de nível de água e precipitação. Nesse caso, os valores de NSE atingiram 0.994 para previsão com antecedência de 15 minutos, e 0.9016 para 120 minutos.

**PALAVRAS - CHAVE:** Bacia Hidrográfica, Rede Neural Artificial, Nível de Água, Estação de Monitoramento Hidrológico.

## 1 | INTRODUCTION

Occurrences of natural disasters are increasing every year, due to climatic changes and man-induced susceptibilities and vulnerabilities. Among these disasters are the floods, which are frequent in many countries, and have an obvious potential of causing injuries, fatalities and damages. However, such effects can be eventually mitigated by forecasts issued with a minimum antecedence. This work proposes the forecast of water level at the outlet of a watershed in the city of Nova Friburgo, located in the mountain region of the state of Rio de Janeiro in Brazil, which has been affected by floods and landslides.

A neural network is proposed to forecast temporal series of water level at the outlet of the watershed with up to a 2-hour antecedence. Input data was collected by a set of hydrological monitoring stations in the Grande river watershed. It is composed of temporal series of water level and rainfall measures acquired with a 15-minute resolution. The programming environment includes the Python language, the Tensorflow library and the Keras API. Three different sets of predictions tests were performed, using only temporal series of water level, rainfall and water level and rainfall. Results show that temporal series of water level at the watershed can be predicted with acceptable accuracy for up to two hours antecedence, using at least temporal series of water level as input for the neural network.

A multitude of forecast schemes for variables related to floods, such as water level or discharge/flow of rivers, have been proposed in the last decades. Mathematical models can be applied but depend on the knowledge of physical parameters and also accurate initial and boundary conditions. On the other hand, the so-called data-driven models are based on algorithms that can be trained using known post-mortem data in order to “learn” and thus make predictions or classifications from new data. These are the machine-learning algorithms applied to hydrological forecasting and are discussed in Section 2.



It is difficult to conclude what is the better method/approach since each prediction depends on the implementation of the proposed method, the availability of data, and the particular geographical and physical scenario. Therefore, a good method/approach is possibly the one that achieves a prediction performance that is acceptable for operational use in terms of issuing warnings of disasters for the civil defense.

The current work is a development of a former prediction approach for the water level at the outlet of the same watershed at Lima et al. (2016), using a neural network, but not applied to time series.

## 2 | LITERATURE REVIEW OF HYDROLOGICAL PREDICTION

An extensive review of flood prediction using machine learning was presented in Mosavi et al. (2018) showing multiple methods for short-term and long-term predictions. The review started with thousands of articles, but selected almost two hundred original research articles, which were then analyzed and compared. These methods are mostly based on artificial neural networks (ANN), but also included neuro-fuzzy or adaptive neuro-fuzzy inference systems (ANFIS), and support vector machine (SVM), among other methods. Different classes of neural networks are discussed. Hybrid approaches combining different methods and ensemble prediction systems were also presented. Most articles propose a given method, in order to compare it to another one that appeared before in a literature reference. According to the considered article, different variables can be predicted: there are continuous variables like river water level (stage level), river water flow, rainfall, rainfall-runoff, or categorical variables like flood, urban flood and flash flood. According to this review, there is an increasing trend in the use of neural networks, in comparison to other methods along the years, intended for an antecedence of a few hours, as in this work, but some works consider in this class an antecedence of a few days. In tropical countries, many of these events start and end in a short period of time, so hydrological forecast based on observed data is hard to be done. In the case of high temporal resolution flood prediction, multilayer perceptron (MLP) neural networks seems to be often employed, despite the difficulty of optimizing its architecture and choosing suitable activation functions. New architectures, like recurrent the neural networks (RNN, Elman (1990)) and one of their variants, the long short-term memory neural networks (LSTM Hochreiter & Schmidhuber (1997); Gers et al. (2000)), began to be more employed in recent years for the prediction of temporal series. RNNs are neural networks where the connections between neurons/nodes form a directed graph along a temporal sequence, making them suitable to deal with time series.

The review contains a general comparison of methods considering the complexity of the related algorithm, easiness of use, processing time, accuracy, and input dataset. Short-term prediction usually employs historical datasets as input, i.e. a temporal series with resolution of minutes, hours or days for each variable. In addition, the survey compare the

prediction performance of the different methods by means of the mean square error and the correlation coefficient between predicted and actual values. However, it is a 2018 survey, and new approaches have been proposed more recently, with a wide variety of approaches as exemplified Lugt & Feelders (2020); Zhu et al. (2020); Choi et al. (2020).

Besides machine learning approaches, there are deterministic approaches, based on mathematical models, but they are out of the scope of this work. As an example, besides Weather Research and Forecasting (WRF), which is a mesoscale numerical model for weather prediction, there is the WRF-Hydro modeling system ([https://ral.ucar.edu/projects/wrf\\_hydro/overview](https://ral.ucar.edu/projects/wrf_hydro/overview)), an ensemble of models that include the WRF itself, land and hydrological models, developed by the National Center for Atmospheric Research (NCAR). WRF-Hydro was written in Fortran-90, and it was parallelized to be executed in clusters and supercomputers. Users can select a given area of the USA territory and ask for hydrological predictions, like streamflow for the watersheds contained in that area (<https://water.noaa.gov/tools/nwm-image-viewer>). However, it is stated as an experimental service, and it is difficult to know if its accuracy can be assessed.

Early warning centers have been interested in empirical hydrological models as well. Lima et al. (2016) proposed an ANN for estimating the level on the station Conselheiro Paulino in the Mountain Region of Rio de Janeiro, Brazil and reached a NSE index of approximately 0.98 (for prediction antecedence of 15 min) and 0.87 (for 120 min). On the other hand, Silva et al. (2016) developed an ANN to forecast the discharge in the river “rio Claro” in Caraguatuba, São Paulo, Brazil and reached an efficiency in their alerts about 0.91 considering a critical flow threshold.

The work Zhu et al. (2020) is about forecasting streamflow (volume of water per time) in the huge watershed of the Chinese Yangtze river using only a few hydrological monitoring stations on this river or its tributaries. Authors propose to use a hybrid approach based on a LSTM neural network embedding Gaussian process regression in order to perform a probabilistic daily streamflow forecasting, which yields a varying prediction interval for the time series. They state that the proposed approach associates the inductive biases of the LSTM network while retaining the non-parametric, probabilistic property of the Gaussian process. They performed predictions with other approaches such as a “pure” LSTM or an ensemble using standard neural networks, stating that this hybrid LSTM is more suitable for water resources management and planning.

Another recurrent neural network is proposed in Lugt & Feelders (2020) for predicting water level temporal series at some hydrological monitoring stations using a RNN that has a encoder/decoder with exogenous variables (precipitation and flow rates). These stations are placed in rivers around a Dutch polder, a piece of land below sea level that require water level management by pumping. This particular RNN is said to predict more accurately captures high temporal resolution fluctuation of the water level.

In Choi et al. (2020), authors compare some implementations of machine learning

algorithms to predict water level temporal series in a Korean wetland, with emphasis in water level peaks. They tested algorithms like neural network, decision tree random forest and support vector machine concluding that the random forest had the better prediction performance. Input data was composed of upstream water level temporal series and some meteorological variables like rainfall, temperature and wind.

### 3 I DATA AND METHOD

The area of study of this work is the Grande River watershed, located in the mountainous region of the state of Rio de Janeiro, in Nova Friburgo city (Figure 1). This is an area susceptible to natural disasters like floods or landslides, with a potential of causing serious harm to the local population (Lima et al., 2016). The Grande river watershed was delineated using the software TerraHidro (Abreu et al., 2012) and a 30-meter resolution Digital Elevation Model (DEM), obtained by an on-board radar of the Shuttle Radar Topography Mission (SRTM). TerraHidro is a free software for distributed hydrological modeling written in C++ developed by National Institute for Space Research (INPE) and written in the C++ language. TerraHidro employs INPE's TerraView GIS for manipulating maps, which includes the Terralib library and a spatial database.

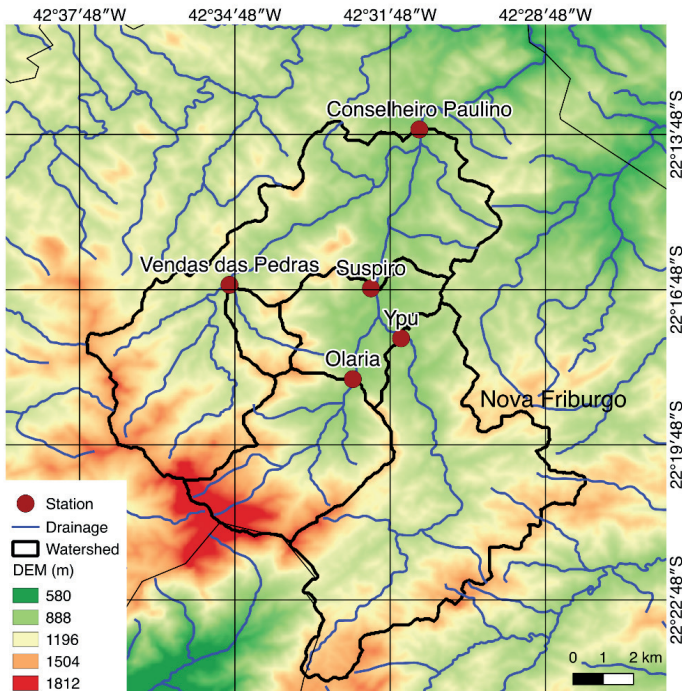


Figure 1 - Area of study - Grande river watershed.

The data for this study was collected from five hydrological monitoring stations of the State Environmental Institute (INEA), as listed in Table 1. These stations are located in Nova Friburgo, being the Conselheiro Paulino station located in the pour point (outlet) of the Grande river watershed. All the stations provide rainfall and water level data, with a 15-minute temporal resolution. Data was collected between December 1, 2011 and March 24, 2013, amounting to 46,080 samples for each station.

<b>Name</b>	<b>Longitude</b>	<b>Latitude</b>
Conselheiro Paulino	42°31'12.49"W	22°13'42.47"S
Olaria	42°32'31.96"W	22°18'31.83"S
Suspiro	42°32'05.36"W	22°16'46.43"S
Venda das Pedras	42°34'53.51"W	22°16'42.47"S
Ypu	42°31'35.41"W	22°17'45.09"S

Table 1 - Hydrological monitoring stations used in this study.

The related data-driven model was implemented in the Python language using Tensorflow, an open library developed by Google using the Keras API (Chollet et al., 2015), which is also open, modular and user-friendly. This model is based on a Multilayer Perceptron (MLP) neural network. Such network architecture is composed of multiple layers that allow to solve complex problems related to non-linearly separable datasets in several areas of study as Meteorology, Medicine, Cryptography, etc. (Rumelhart et al., 1986). The Keras API is employed in much complex neural networks than the MLP, such as convolutional and recurrent neural networks, typical of Deep Learning implementations, but it was adopted in this work anticipating the use of such neural networks in the ensuing work. On the implemented MLP network, the input layer has one neuron for each input parameter, one hidden layer with 5 neurons and the output layer has one neuron, as shown in Figure 2. It employs the Rectified Linear Unit (ReLU) activation function for the hidden layer, and the linear function for the output layer. The ReLU function is a non-linear activation function, commonly used in Deep Learning. The advantage of this activation function over the other ones is not to activate all the neurons at the same time since, when an input is negative, its value is converted to 0, not activating the corresponding neuron. This MLP neural network with a single hidden layer using the ReLU activation function achieved a good prediction performance, as shown ahead, and thus a more complex multiple-hidden layer MLP neural network was not required.

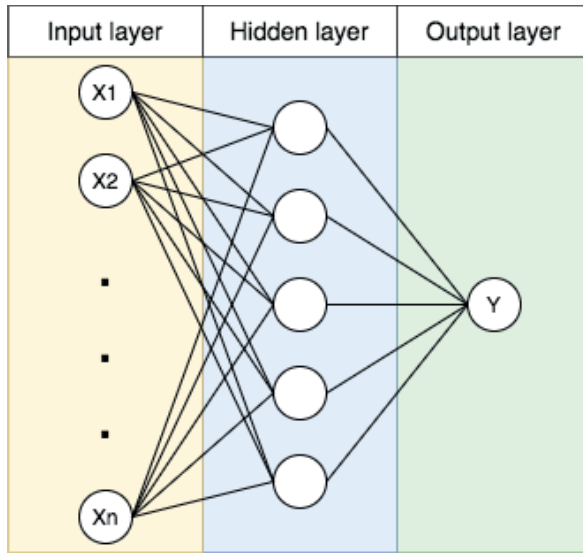


Figure 2- The architecture of the proposed MLP network.

The adopted loss function for the MLP network was the mean absolute error (MAE) and the optimization algorithm was Adam (Adaptive Moment Estimation), used instead of the classical stochastic gradient descent procedure to update the network weights. The training process was limited by a total of 100 epochs. The source code and the dataset used in this work are available for download on the git repository ([https://github.com/cpfreitas/redes\\_neurais](https://github.com/cpfreitas/redes_neurais)).

This work employs two metrics to evaluate the prediction performance of the proposed neural network. The first one is the Nash–Sutcliffe model efficiency coefficient (NSE), commonly used for hydrological models, and defined as:

$$NSE = 1 - \frac{\sum_{t=1}^T (Q_p^t - Q_o^t)^2}{\sum_{t=1}^T (Q_o^t - \bar{Q}_o)^2},$$

where  $\bar{Q}_o$  is the mean of the actual discharges in the time interval  $[1, T]$  that contains  $T$  discrete values,  $Q_p^t$  is the predicted discharge at time  $t$ , and  $Q_o^t$  is the actual discharge at time  $t$ . The NSE is equivalent to the correlation coefficient between predicted and actual data. In this work, instead of the water discharge, the NSE was calculated in terms of the water level. The better the prediction, the higher is the NSE above the threshold value of 0.5, but it is difficult to reach the top value of 1.0 that corresponds to a perfect prediction. The second metric is the well known root mean square error (RMSE), which is the standard deviation of the prediction errors. It is also calculated here for the discrete times in the interval  $[1, T]$ , and its value should be close to zero for a good prediction. RMSE is defined

by the following equation:

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (Q_p^t - Q_o^t)^2}{T}},$$

where  $Q_p^t$  and  $Q_o^t$  are respectively the predicted and observed discharge values at time  $t$ , as in the former equation for NSE.

## 4 | RESULTS AND DISCUSSION

In this work, a neural network is proposed to predict a water level temporal series at the outlet of the watershed using as input temporal series of water level and/or rainfall at different hydrological monitoring stations. The prediction is done for antecedences ranging from 15 to 120 minutes.

Input data was split as follows: 70% for the training, 17.5% for the validation, and 12.5% for the test of the neural network. Data in the period from December 2011 to January 2013 with randomly sorted samples was used for training and validation, while data from the months of February and March 2013 was used for test.

Three set of tests were performed, based on different combinations of the available input datasets. The first test employs only water level data, the second one uses only rainfall data, while the third one inputs both water level and rainfall data. All tests were devised for the prediction of the water level at the Conselheiro Paulino station for up to 2 hours antecedence using as input the datasets from the five hydrological monitoring stations described above.

Predicted and actual/observed temporal series of the water level at the Conselheiro Paulino station with an antecedence of 15 and 120 minutes is shown in Figure 3 for a period of nearly 5,000 time intervals, corresponding to about 52 days of the months of February and March 2013. More detailed curves of the same predicted and actual/observed temporal series are shown in Figure 4 for time intervals 60 to 600 and in Figure 5 for time intervals 2500 to 2800, but in this case, for all five antecedences from 15 to 120 minutes. Input data included water level and rainfall temporal series. These figures show a good agreement between predicted and actual values.

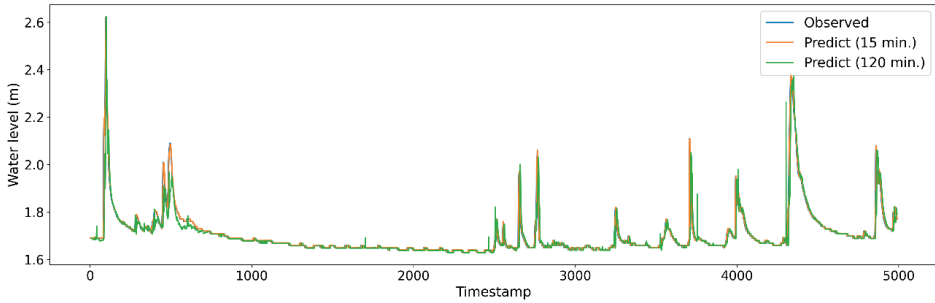


Figure 3 - Observed and predicted values of water level at the outlet for 15 and 120 minutes of antecedence considering all the 5,000 time intervals.

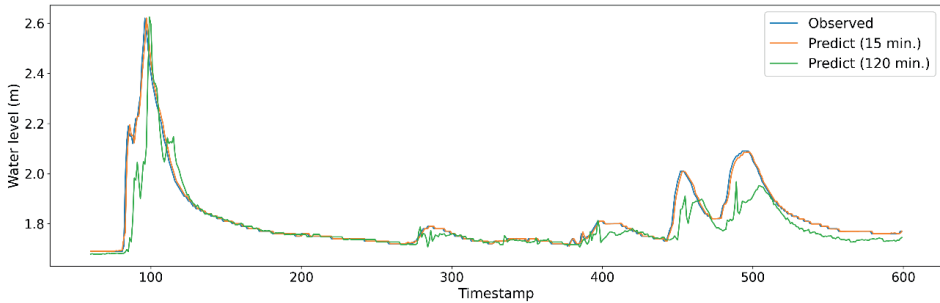


Figure 4 - Observed and predicted values of water level at the outlet for 15 and 120 minutes of antecedence considering time intervals 60 to 600.

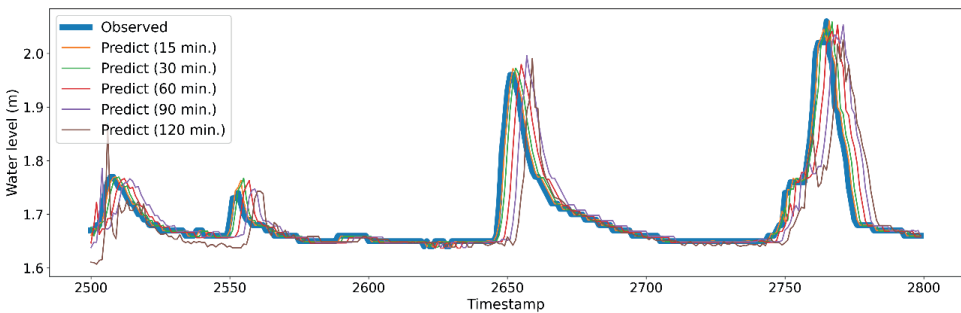


Figure 5 - Observed and predicted values of water level at the outlet for 15, 30, 60, 90 and 120 minutes of antecedence considering time intervals 2500 to 2800.

The agreement between observed and predicted values is substantially better for 15 than for 120 minutes antecedence as show in Figure 3. The mismatch is more accentuated when the variation of the water level is high, especially when the water level is increasing.

As seen in Figure 5, the prediction curve is delayed in relation to the observed one, with the delay increasing as the antecedence of prediction increases. However, the delay is rather small, being all the observed water level peaks at the outlet reproduced by the predictions.

The performance of the three sets of predictions can be evaluated in Table 2 for the NSE and Table 3 for the RMSE. Both tables show the values of these two metrics for antecedences from 15 to 120 minutes and for the three different sets of input time series: only water level, only rainfall and water level and rainfall.

According to Table 2 and 3, the test with only rainfall inputs did not succeed, showing very low values of NSE and higher values of RMSE in comparison to other two sets of predictions. When only water level inputs were considered, the performance was good, the RMSE did not exceed 0.0066 and the NSE was at least 0.8858. Adding rainfall inputs besides the water level ones, the performance slightly improves for all regarded antecedences, raising these values to 0.0063 (RMSE) and 0.9016 (NSE). As expected, on these two sets of tests, the quality of the prediction decreases as the antecedence increases.

<b>Antecedence</b>	<b>Water level</b>	<b>Rainfall</b>	<b>Water level + Rainfall</b>
<b>15 min.</b>	0.9941	0.0692	0.9944
<b>30 min.</b>	0.9827	0.082	0.9842
<b>45 min.</b>	0.9682	0.1007	0.9715
<b>60 min.</b>	0.9537	0.1135	0.9589
<b>75 min.</b>	0.9381	0.1202	0.947
<b>90 min.</b>	0.9182	0.1343	0.9301
<b>105 min.</b>	0.902	0.1377	0.9159
<b>120 min.</b>	0.8858	0.1406	0.9016

Table 2 - Values of NSE for the different predictions.

<b>Antecedence</b>	<b>Water level</b>	<b>Rainfall</b>	<b>Water level + Rainfall</b>
<b>15 min.</b>	0.0013	0.0353	0.0013
<b>30 min.</b>	0.0022	0.0352	0.0022
<b>45 min.</b>	0.0031	0.0351	0.0031
<b>60 min.</b>	0.0039	0.0351	0.0038
<b>75 min.</b>	0.0047	0.0349	0.0046
<b>90 min.</b>	0.0054	0.0348	0.0052
<b>105 min.</b>	0.006	0.0347	0.0057
<b>120 min.</b>	0.0066	0.0346	0.0063

Table 3 - Values of RMSE (in meters) for the different predictions.



The scatter plot between the predicted and observed values of the water level (in meters) at the outlet of the watershed, for 15 and 120 minutes of antecedence, is shown in Figure 6. In general, the models for 15 and 120 minutes of antecedence have no bias to overestimate or underestimate the prediction. The mean errors (mean difference between observed and predicted values) of both models were close to zero, 0.0002 and 0.0044 respectively, which confirms this fact.

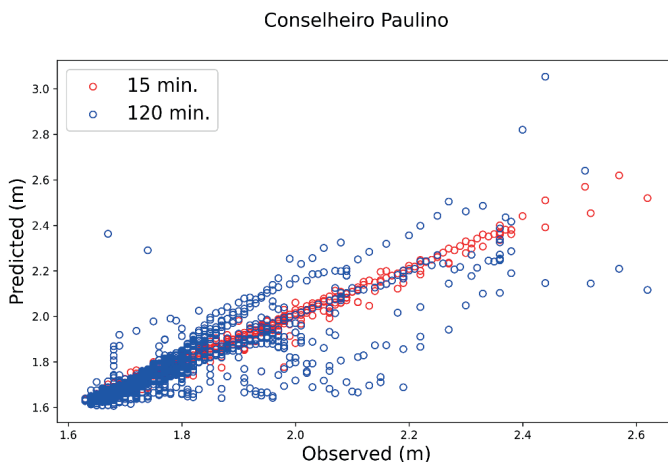


Figure 6 - Scatter plot between predicted and observed values of water level at the outlet of the watershed for 15 and 120 minutes of prediction antecedence.

## 5 | CONCLUSIONS

This work proposed the use of a neural network to predict temporal series of water level at the outlet of a watershed with antecedences ranging from 15 to 120 minutes using as input data temporal series of (i) water level or (ii) rainfall or (iii) both, collected by five hydrological monitoring stations in the watershed.

The neural network was trained and validated using temporal series of water level and rainfall data comprising 14 months of data with a temporal resolution of 15 minutes and tested using 2 months of data. A good agreement was found between the predicted and actual time series for an antecedence of 15 minutes using the neural network trained with water level and rainfall data. The quality of the predictions was assessed with the NSE and RMSE metrics. Naturally, higher antecedences resulted in prediction with lower accuracy, but NSE was above 0.94 for up to a 75-minute antecedence using water level and rainfall data, or up to 60-minute antecedence using only water level data. In both cases, RMSE was very low. On the other hand, predictions using only rainfall data did not succeed.

The proposed approach was tested for a case of study, showing the possibility of its

use in an operational scenario. As future work, we propose to compare these results to the results obtained by other hydrological prediction models, and to perform further prediction tests for other Brazilian watersheds. In addition, it is intended to perform an operational test of this forecasting tool using data from the network of hydrological monitoring stations of the Brazilian Centre for Monitoring and Early Warnings of Natural Disasters (CEMADEN).

## ACKNOWLEDGEMENTS

Authors thank the following funding: São Paulo Research Foundation (FAPESP) grant#2015/50122-0, DFG-IRTG grant #1740/2 and CNPq grant #420338/2018-7.

## REFERENCES

- Abreu, E.; Rosim, S.; Rennó, C.; Oliveira, J. R.; Jardim, A.; Ortiz, J.; Dutra, L., **Terrahidro - A Distributed Hydrological System to delimit large basins**, International Geoscience and Remote Sensing Symposium (IGARSS), 2012.
- Choi, C.; Kim, J.; Han, H.; Han, D.; Kim, H. S., **Development of Water Level Prediction Models Using Machine Learning in Wetlands: A Case Study of Upo Wetland in South Korea**, Water, vol 12, no. 93, 2020.
- Chollet, F. and others, **Keras**. Available: <https://keras.io>, 2015.
- de Lima, G. R. T.; Santos, L. B. L.; de Carvalho, T. J.; Carvalho, A. R.; Cortivo F. D.; Scofield, G. B.; Negri, R. G., **An operational dynamical neuro-forecasting model for hydrological disasters**, Modeling Earth Systems and Environment, vol 2, 94, 2016.
- Silva, M. R.; Santos, L. B. L.; Scofield, G. B.; Cortivo, F. D., **Utilização de Redes Neurais Artificiais em alertas hidrológicos: estudo de caso na bacia do rio Claro em Caraguatatuba-SP**. Anuário do Instituto de Geociências (UFRJ. IMPRESSO), v. 39, p. 23-31., 2016.
- Elman, J. L., **Finding Structure in Time**. Cognitive Science, vol 14, 179-211, 1990.
- Gers, F. A.; Schmidhuber, J.; Cummins, F., **Learning to Forget: Continual Prediction with LSTM**. Neural Computation, vol 12, no. 10, 2451-2471, 2000.
- Hochreiter, S.; Schmidhuber, J., **Long short-term memory**. Neural Computing, vol 9, no. 8, 1735-80, 1997.
- Lugt, B.; Feelders, A., **Conditional Forecasting of Water Level Time Series with RNNs**. AALTD@PKDD/ECML, 2020
- Mosavi, A.; Ozturk, P.; Chau, K. W., **Flood prediction using machine learning models: literature review**, Water, vol 10, no. 11: 1536, 2018.

Rumelhart, D. E.; Hinton, G. E.; Williams, R. J., **Learning representations by back-propagating errors**, *Nature*, vol 323, 533–536, 1986.

Zhu, S.; Luo, X.; Yuan, X.; Xu, Z., **An improved long short-term memory network for streamflow forecasting in the upper Yangtze River**, *Stoch Environ Res Risk Assess*, vol 34,1313–1329, 2020.

## CLASSIFICAÇÃO DE PASSAGEIROS DOMÉSTICOS DE LINHAS AÉREAS UTILIZANDO REDES NEURAIS ARTIFICIAIS DO TIPO MLP

*Data de aceite: 01/04/2021*

*Data de submissão: 04/01/2021*

### **Sidnei Gouveia Junior**

Universidade Estadual de Montes Claros,  
Programa de Pós-Graduação em Modelagem  
Computacional e Sistemas  
Montes Claros – Minas Gerais  
<http://lattes.cnpq.br/0716670135458157>

### **Narciso Ferreira dos Santos Neto**

Universidade Estadual de Montes Claros,  
Departamento de Administração  
Montes Claros – Minas Gerais  
<http://lattes.cnpq.br/5703335372799074>

### **Nilton Alves Maia**

Universidade Estadual de Montes Claros,  
Departamento de Ciências da Computação  
Montes Claros – Minas Gerais  
<http://lattes.cnpq.br/3101079034762740>

**RESUMO:** Este experimento investigou o potencial da aplicação de uma rede neural artificial (Multilayer Perceptron) para a classificação de perfis de passageiros no setor aéreo brasileiro, mais precisamente na conexão do par de aeroportos mais relevante em várias métricas do país. A partir de uma revisão bibliográfica sobre o assunto, foram elencados dois perfis distintos de passageiros, cada um com características específicas, em relação ao seu comportamento de consumo. Após alguns ajustes e testes na ligação escolhida, foi estabelecida uma arquitetura e parametrização

satisfatórias para a classificação dos passageiros com dados variáveis que podem ser adquiridos pelas empresas aéreas no momento da compra da passagem. O resultado dos experimentos revelou-se muito promissor, o que possibilitaria a futura aplicação desse método pela indústria aérea.

**PALAVRAS - CHAVE:** Planejamento Aéreo, Segmentação de Demanda Aérea, Gerenciamento de Receita, Rede Neural Artificial.

### AIRLINES DOMESTIC PASSENGERS CLASSIFICATION USING ARTIFICIAL NEURAL NETWORKS OF THE TYPE MLP

**ABSTRACT:** This experiment investigated the potential of applying an artificial neural network (Multilayer Perceptron) for the classification of passenger profiles within the Brazilian airline industry, more precisely on the connection of the most relevant airport pair in several metrics in the country. From a bibliographical review on the subject, it was listed two different passenger profiles each with specific characteristics, in relation to their consumption behavior. After some adjustments and tests on the chosen network, a satisfactory topology and parameterization was established for the classification of the passengers with variable data that can be acquired by the airlines at the time of ticket purchase. The result of the experiments turned out to be very promising, which would enable the future application of this method by airline industry.

**KEYWORDS:** Airline Planning, Airline Demand Segmentation, Revenue Management, Artificial Neural Network.

## 1 | INTRODUÇÃO

Independentemente do mercado de negócios de uma empresa, torna-se importante identificar os motivos pelos quais um consumidor adquire um produto ou serviço, não apenas do ponto de vista das estratégias de marketing, fidelização do mesmo ou experiência do consumo. Esses fatores são relevantes, mas entender os motivos que o levaram a tomar a decisão de consumo está intrinsecamente ligado ao planejamento da produção e das operações da empresa. Já que a determinação dos custos de produção que garantem a sustentabilidade das operações é relativa às receitas esperadas; portanto, entender o comportamento da demanda é um ponto chave na projeção de receita. Os diferentes padrões e razões pelas quais eles foram levados ao consumo, estabelecem a relação de valor para o cliente e o preço pelo qual está disposto a pagar.

No contexto de uma companhia aérea, a área de planejamento responsável por esse processo é chamada de Revenue Management. E o entendimento clássico do gerenciamento de receita pode ser simplificado em três linhas amplas para melhor entendimento, a saber: demanda (passageiros), oferta (assentos) e preço, que é a relação entre demanda e oferta. Segundo Reyes (2006), estabelecer a maior receita possível para um determinado voo é o resultado do equilíbrio desse relacionamento.

Segundo Talluri e van Ryzin (2004), o gerenciamento de receita difere-se do simples gerenciamento do inventário de assentos de uma aeronave em um voo. Esses autores veem em uma constituição conceitual que precisa entender alguns aspectos, como a natureza multidimensional da demanda em dimensões como locais, canais de aquisição de serviços e adiantamentos; a ligação entre efeitos não diretos das decisões relacionadas ao gerenciamento da demanda; condições favoráveis à implementação da gestão e compreensão da heterogeneidade da demanda; a necessidade de tornar o serviço mais flexível, o preço cobrado como sinal de qualidade e o monitoramento constante de dados e informações do setor. Portanto, o gerenciamento de receita é visto para eles como o gerenciamento de valor do produto.

Conforme Kayser (2008), o Revenue Management, na prática, para que seja eficaz, precisa de uma previsão da demanda por esses dois conjuntos diferentes de passageiros: negócios e lazer. Os viajantes a negócios são caracterizados por compras próximas à partida, com notável sensibilidade ao horário e à frequência dos voos, tendendo à aversão à maioria das restrições tarifárias e à propensão de pagar por tarifas mais altas. Por outro lado, os passageiros a lazer geralmente planejam com antecedência e podem comprar bilhetes no início do período de vendas de reservas, não são sensíveis à programação ou à frequência do voo e não são tão sensíveis às restrições tarifárias e, portanto, estão propensos a comprar bilhetes com preços mais baixos. Assim, todo o processo de segmentação e precificação de tarifas será em torno de subconjuntos derivados desses dois conjuntos majoritários, que determinam a captura de receita diretamente por meio de

seu comportamento. Embora o gerenciamento de receita às vezes possa ser confuso entre viajantes a lazer e a negócios, eles desempenharam um papel importante na alocação e sustentabilidade financeira das companhias aéreas em todo o mundo.

Sendo que a importância em identificar o motivo da viagem está diretamente associada ao preço que um passageiro está disposto a pagar pelo serviço de transporte aéreo. Logo, o motivo da viagem implica padrões de consumo muito complexos para a elaboração de algum modelo matemático convencional, que os interpreta e resulta na identificação objetiva do motivo da viagem.

A rede neural artificial é uma técnica computacional que apresenta um modelo inspirado na estrutura neural natural que simula a capacidade do cérebro humano de adaptar ou aprender, generalizar, agrupar e organizar dados. Que estão incluídos em unidades simples e adaptáveis chamadas neurônios. O modelo de rede artificial usado neste trabalho é do tipo MLP (Multi-layer Perceptron). A MLP é uma generalização do perceptron, sendo constituída de uma camada de entrada, uma ou mais camadas ocultas de nós computacionais e uma camada de saída. Nesse tipo de rede, cada uma das camadas tem funções específicas. Assim, cada neurônio computa uma soma ponderada de suas entradas e passa essa soma na forma de uma função não-linear limitada. Em nível de mesoestrutura, tem-se duas ou mais camadas com conexão feedforward (Vieira & Bauchspiess, 1999). De acordo com Duarte (2009), adicionando-se uma ou mais camadas intermediárias (ocultas), aumenta-se o poder computacional de processamento não-linear e armazenagem da rede.

Assim, este trabalho tem como objetivo avaliar a eficiência do uso de Redes Neurais Artificiais do tipo MLP para a classificação do tipo de viagem aérea realizada pelos passageiros. Para isso, utilizou-se dados reais de uma Pesquisa de Origem e Destino do Transporte Aéreo de Passageiros de 2014, onde também se incluíam dados relativos ao motivo da viagem dos indivíduos.

## 2 | MATERIAIS E MÉTODOS

A Pesquisa de Origem e Destino do Transporte Aéreo de Passageiros, foi realizada em 2014, e teve como principal objetivo aprimorar, complementar e detalhar o sistema de informações da Aviação Civil no Brasil. Sendo realizada em quatro etapas distintas, coletou informações em mais da metade dos aeroportos que operavam voos regulares no Brasil em 2012. Registra-se que, em termos de fluxo de passageiros, esse grupo de aeroportos correspondia quase que pela totalidade do número total de passageiros embarcados naquele ano. Segundo a Empresa Brasileira de Logística (2014) a pesquisa subsidiou informações coletadas, que foram utilizadas como base para o estudo do comportamento atual e futuro do transporte aéreo.

## 2.1 Dados Utilizados

Os dados utilizados para este experimento foram coletados em 2014 pela “Pesquisa de Origem e Destino do Transporte Aéreo de Passageiros. Para viabilizar o experimento, foram utilizados os registros de passageiros que voavam entre o par dos aeroportos: Aeroporto de Congonhas, localizado na cidade de São Paulo, e Aeroporto Santos Dumont, localizado na cidade do Rio de Janeiro. Segundo a Agência Nacional de Aviação Civil (2014), essa foi a principal rota doméstica aérea do país, tanto em número de operações quanto em passageiros transportados. Dos 1.937 entrevistados para esta rota, 1.036 foram selecionados por fornecerem informações suficientes para a condução do experimento. Das sete perguntas selecionadas, uma é o motivo da viagem, as outras perguntas são relativas ao voo selecionado, se tem ou não uma conexão, data e horário do voo, antecedência da compra e preço pago pelo bilhete. Para o motivo da viagem entre as seis respostas possíveis, foram organizados dois grupos que correspondem a viagens a negócios ou a lazer, como é mostrado na Figura 1.



Figura 1 – Motivos específicos da viagem e perfil do passageiro.

## 2.2 Análise dos Dados

Antes do experimento, os dados foram analisados, para verificar se o seu comportamento se assemelhava com o descrito na revisão literária, presente na introdução deste artigo. O primeiro passo foi analisar se os dados de antecedência da compra e o preço do bilhete tinham uma relação mínima com o motivo da viagem. O gráfico apresentado na Figura 2 permite analisar essa relação:

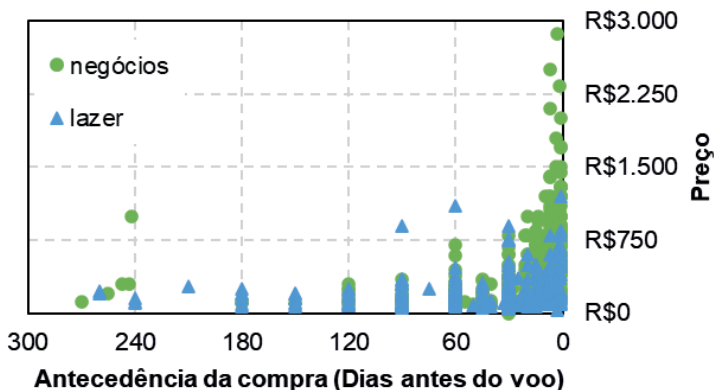


Figura 2 – Motivos específicos da viagem e perfil do passageiro.

O motivo da viagem corresponde ao valor de saída da rede, onde foi utilizada a seguinte codificação binária;  $y = 1$  para negócios e  $y = 0$  para lazer. O comportamento de consumo desses grupos de passageiros descrito na literatura pode ser observado no gráfico elaborado com os dados selecionados na pesquisa. Outro ponto importante no comportamento da demanda descrito na literatura, que caracteriza passageiros de negócios é a preferência sensibilidade notável ao horário e à frequência dos voos. A Figura 3 mostra o gráfico de preferência dos passageiros viajando a negócio.

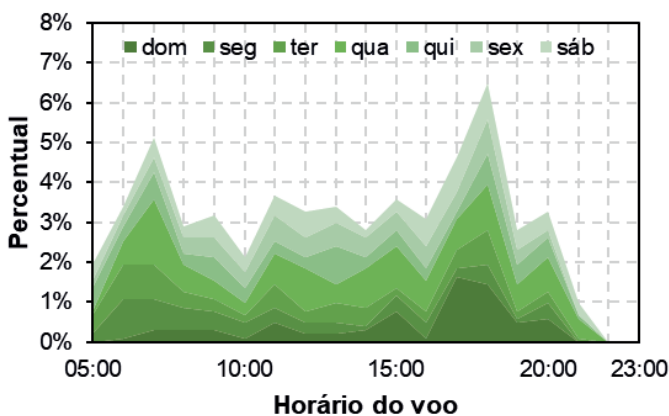


Figura 3 – Preferência dos passageiros viajando a Negócio.

Este perfil de passageiros viajando a negócios é caracterizado por uma maior preferência de voos com partidas programadas às 7:00 e 18:00, representados pelos picos observados na Figura 3. Por outro lado, os passageiros a lazer não são tão sensíveis à



programação ou à frequência do voo, como pode ser visto no gráfico apresentado na Figura 4.

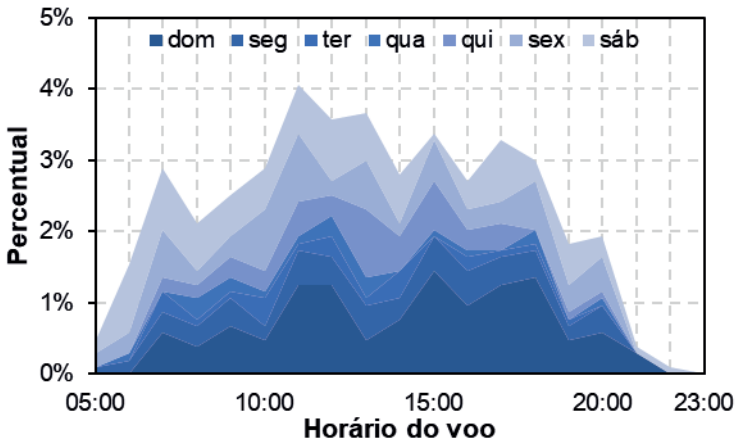


Figura 4 – Preferência dos passageiros viajando a Lazer.

Na Figura 4 pode ser observada a dispersão na distribuição dos passageiros a lazer por voos em horários intermediários àqueles preferidos por viajantes a negócios.

### 2.3 Topologia da Rede Neural Artificial

A rede MLP utilizada neste trabalho foi implementada no software Python 3.7 com o pacote Sklearn. Todas as execuções foram feitas em um notebook DELL I15-5566-A50P com um processador Intel Core i7-7500U 2.90GHz com 8GB de RAM no sistema operacional Windows 10 64bits. A arquitetura da MLP é formada por seis entradas, três camadas escondidas com três neurônios em cada uma delas e uma saída. A Figura 5 mostra a arquitetura da rede MLP.

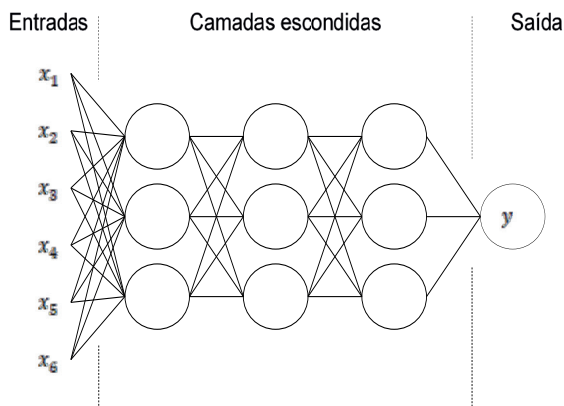


Figura 5 – Arquitetura da Rede Neural Artificial Utilizada.

A função de ativação utilizada nos neurônios da MLP foi a tangente hiperbólica. A tangente hiperbólica é definida pela equação (1):

$$\tanh(x) = 2 \times \sigma(2x) - 1 \quad (1)$$

onde  $\sigma(x)$ , é a função sigmóidal definida pela equação (2):

$$\sigma(x) = \frac{e^x}{1 + e^x} \quad (2)$$

A rede utilizou uma taxa de aprendizagem inicial de 0,01 adaptável que a cada duas épocas consecutivas sem variação tolerável no erro dividia a taxa de aprendizagem por 5. Tolerância da variação do erro de  $10^{-8}$ , por 1.000 épocas consecutivas. O número máximo de iterações suportado pelo algoritmo é 10.000 épocas. Outro ponto importante foi a necessidade de normalização dos valores assumidos por  $x_n$  a uma escala entre 0 e 1. Já que os intervalos de valores assumidos pelas variáveis de entrada, onde  $x_n \in \mathbb{N}$ , eram discrepantes entre si, conforme mostrado na Tabela 2.

$x_n$	Intervalo
$x_1$	[0, 1, 2, 3, ..., 360]
$x_2$	[1, 2, 3, ..., 9999]
$x_3$	[0,1]
$x_4$	[1, 2, 3, 4, 5, 6, 7]
$x_5$	[5, 6, 7, ..., 23]
$x_6$	[1, 2, 3, ..., 12]

Tabela 2: Intervalo dos valores assumidos por

Para a normalização dos dados, utilizou-se do método min-max onde o valor normalizado de cada membro do conjunto é dado pela seguinte fórmula da equação (3):

$$z = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (3)$$

onde min e max são respectivamente o menor e o maior valor do intervalo de x.

Para avaliar a capacidade de treinamento da rede, os dados foram selecionados aleatoriamente com 70% de probabilidade de serem utilizados para treinamento e 30% de probabilidade de serem utilizados para teste. O pseudocódigo da rede MLP é apresentado na Figura 6:

Algoritmo 1: Pseudocódigo da rede MLP	
1.	<b>Rede_MLP()</b>
2.	<b>Carregar</b> dados
3.	dados_treinamento = dados × 0,7
4.	dados_teste = dados × 0,3
5.	taxa_aprendizagem = 0,01
6.	épocas = 0
7.	<b>Enquanto</b> alteração ≤ 1000 <b>ou</b> tolerância_erro ≤ 1 <sup>e-8</sup> <b>faça</b> :
8.	<b>forward</b> (dados_treinamento)
9.	<b>backward</b> (taxa_aprendizagem)
10.	épocas += 1
11.	<b>fim enquanto</b>
12.	testar_rede(dados_teste)
13.	<b>Fim</b>

Figura 6 – Pseudocódigo da Rede MLP.

## 2.4 Métricas de Avaliação da Classificação

Os resultados da classificação realizada pela rede MLP foram avaliados levando em conta os valores da Acurácia, Precisão, *Recall* e *F1-score*. A Acurácia tem como objetivo identificar o percentual de amostras classificadas corretamente, e é calculada a partir da soma do número de verdadeiros positivos e verdadeiros negativos, dividido pelo total de dados da amostra. A Precisão é uma medida de quão exata foi a classificação. Ela tem como objetivo identificar a taxa de amostras que foram classificadas como positivas que são realmente positivas. O *Recall* é a medida de quão correta foi a predição para a classe. Ele tem como objetivo identificar a taxa de amostras positivas que foram classificadas como positivas. O *F1-score* é a *média harmônica entre a precisão e o recall*. A adoção dessas métricas permite uma melhor compreensão dos resultados do experimento além do entendimento de quão acurada a rede foi em classificar os dados.

### 3 | RESULTADOS

Nesta seção são apresentados os resultados da classificação do tipo de viagem realizada pelos passageiros utilizando a rede MLP. A classificação foi avaliada levando em conta a Acurácia, Precisão, Recall e F1-score. Para a validação do experimento foram realizadas dez execuções do classificador. A Figura 6 apresenta um box-plot com a distribuição dos resultados obtidos com para a Acurácia.

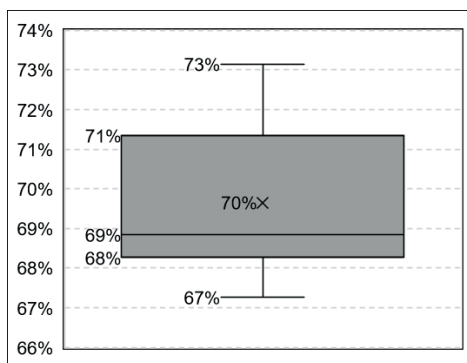


Figura 6 –Box-plot dos resultados de acurácia.

Observando-se a Figura 6, nota-se que foi obtida uma Acurácia média de 70% para as execuções. A Figura 7 apresenta um box-plot com a distribuição dos resultados da precisão, recall e score para as classes Negócios e Lazer.

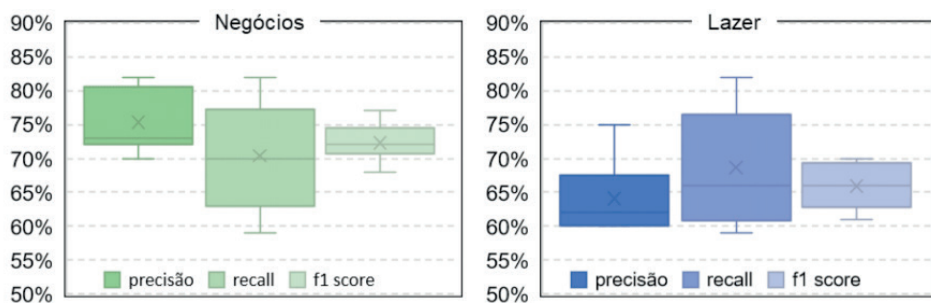


Figura 7 –Box-plot'sdos resultados das métricas de qualidade, por tipo de passageiro.

Pode ser observado na Figura 7 que a rede MLP obteve um resultado melhor na classificação de passageiros viajando a negócios do que a lazer. Para viajantes a negócios, a precisão média foi de 75%, contra uma precisão média de 74% para viajantes a lazer. Em ambas as figuras foi possível constatar a maior aptidão da rede MLP em classificar

passageiros a negócios, podendo se considerar certa simetria na distribuição dos valores de recall para as duas classes, e conseqüentemente um melhor desempenho do score para a classe de Negócios.

## 4 | CONCLUSÕES

Este trabalho teve como objetivo avaliar a eficiência do uso de Redes Neurais Artificiais do tipo MLP para a classificação do tipo de viagem aérea realizada pelos passageiros. Durante as execuções da rede MLP foi obtida uma Acurácia média de 70%. A rede MLP obteve um resultado melhor na classificação de passageiros viajando a negócios do que a lazer. Para viajantes a negócios, a precisão média foi de 75%, contra uma precisão média de 64% para viajantes a lazer. Os resultados iniciais obtidos com esse trabalho para a classificação do tipo de passageiro são promissores. Em um serviço essencialmente ofertado para massas, como é o caso da indústria de transporte aéreo, a identificação do motivo pelo qual um determinado cliente está adquirindo aquele serviço fornecido simplesmente com base no seu comportamento de aquisição, pode vir a ser uma vantagem estratégica considerável para o prestador de serviço que detém o uso dessa tecnologia. Em termos práticos, essa informação quando usada para a precificação do serviço pode fazer com que o preço cobrado seja mais próximo do preço máximo ao qual o passageiro está disposto a pagar por aquela viagem. Outro ponto é que a percepção da qualidade do serviço ofertado pode ser melhorada, com base nas expectativas de cada um desses grupos para aquele serviço.

Sugere-se para trabalhos futuros o emprego de redes neurais artificiais em um universo maior de voos e conseqüentemente de passageiros, além da adoção de mais variáveis de entrada, que possam contribuir na melhoria da acurácia dos resultados já obtidos nesse trabalho, além de comparações com outros métodos de classificação de dados.

## REFERÊNCIAS

ANAC, Agência Nacional de Aviação Civil (2014). **Anuário do Transporte Aéreo 2014**. Brasília, DF: ANAC, 2014.

Bittencourt, G. (1998). **Inteligência artificial - Ferramentas e Teorias**. Editora: Universidade Federal de Santa Catarina, 1998.

Braga, A. P., Carvalho, A. P. L. F., Ludermir, T. B. (2000) **Redes Neurais Artificiais: Teoria e Aplicações**. Rio de Janeiro, RJ: LTC, 2000.

Duarte, V. A. (2009). **MP-DRAUGHTS - Um Sistema Multiagente de Aprendizagem Automática para Damas Baseado em Redes Neurais de Kohonen e Perceptron Multicamadas**. Dissertação, Universidade Federal de Uberlândia, Uberlândia, MG.

EPL, Empresa de Planejamento e Logística (2014). **Pesquisa de origem/destino do transporte aéreo de passageiros, Relatório Técnico**. Brasília, DF. EPL, 2014.

Haykin, S. (1994). **Neural Networks: A Comprehensive Foundation**, Macmillan College Publishing, 1994.

Kayser, M. R. (2008). **RM Methods for Multiple Fare Structure Environments**. Thesis (Master of Science Department of Civil and Environmental Engineering, Massachusetts Institute of Technologies, Cambridge, MA. 2008.

Leong, Lai-Ying, Teck-Soon H., Voon-Hsien L., and Keng-Boon O. “**An SEM–artificial-neural-network analysis of the relationships between SERVPERF, customer satisfaction and loyalty among low-cost and full-service airline.**” *Expert Systems with Applications* 42.19 (2015): 6620-6634.

Otwell, K., S. S. Hornby, and Hutchison, W. (1994) “**A Large-Scale Neural Network Application for Airline Seat Allocation,**” an invited talk at the World Congress on Neural Networks, 1994.

Reyes H. M. (2006). **Hybrid Forecasting for Airline Revenue Management in Semi-Restricted Fare Structures**. Thesis (Master of Science). Department of Civil and Environmental Engineering, Massachusetts Institute of Technologies, Cambridge, MA. 2006.

Slack, N., Chambers, S., Johnston, R. and Betts, A. (2008), **Operations and Process Management**, FT Prentice-Hall, London, UK. 2008.

Srisaeng, P., Baxter, G., Wild, G. (2015). **Using an artificial neural network approach to forecast Australia’s domestic passenger air travel demand**. *World Review of Intermodal Transportation Research*. 2015 5. 281. 10.1504.

Vieira, Z. P., & Bauchspiess, A. (1999). **Implementação do Servocontrole Auto-Sintonizado em Tempo-Real Utilizando Rede Perceptron Multicamadas**. IV Congresso Brasileiro de Redes Neurais, 308-313.

# CAPÍTULO 12

## APRENDIZADO CONJUNTO APLICADO NA PREDIÇÃO DO MERCADO DE AÇÕES BRASILEIRO

Data de aceite: 01/04/2021

Data de submissão: 08/01/2021

### Alvaro Pedroso Queiroz

Universidade Tecnológica Federal do Paraná  
Cornélio Procópio – Paraná  
<http://lattes.cnpq.br/7863634128031457>

### Giovani Volnei Meinerz

Universidade Tecnológica Federal do Paraná  
Cornélio Procópio – Paraná  
<http://lattes.cnpq.br/0746723713646857>

### Érica Ferreira de Souza

Universidade Tecnológica Federal do Paraná  
<http://lattes.cnpq.br/8904855809524041>

**RESUMO:** A previsão do mercado de ações é considerada uma tarefa complexa, envolvendo áreas interdisciplinares como estatística, economia e computação. Com o avanço das técnicas de *machine learning*, tal prática se tornou um recurso bastante utilizado para a maximização de lucros do mercado. É crescente o número de investidores e pesquisadores desse âmbito em cenário nacional, no entanto, após pesquisa bibliográfica realizada, não foram encontrados trabalhos que utilizam modelos que fazem combinação de técnicas de *machine learning*. Este trabalho apresenta o *Stock Market Ensemble Predictor* (SMARTER), um modelo de aprendizado de máquina que combina diferentes técnicas de regressão, desenvolvido com o objetivo de realizar análises preditivas

sobre dados históricos do mercado de ações brasileiro, visando aumentar a acurácia da precisão por meio da combinação dos resultados de múltiplas abordagens. Obteve-se com as técnicas *OLS* e *Bayesian Ridge*, aliado a técnica de *Voting*, o maior coeficiente de determinação  $R^2$  médio dentre as combinações testadas. O valor alcançado foi de 0,914864, formando um modelo mais confiável, aumentando a acurácia da precisão do resultado.

**PALAVRAS - CHAVE:** Bolsa de valores. Previsão. Aprendizado do computador.

### ENSEMBLE LEARNING APPLIED IN THE PREDICTION OF THE BRAZILIAN STOCK MARKET

**ABSTRACT:** Forecasting the stock market is considered a complex task, involving interdisciplinary areas such as statistics, economics and computing. With the advancement of machine learning techniques, this practice has become a widely used resource for maximizing market profits. The number of investors and researchers in this field is growing in the national scenario, however, after a bibliographic research, no studies were found using models that combine machine learning techniques. This paper presents the *Stock Market Ensemble Predictor* (SMARTER), a machine learning model that combines different regression techniques, developed with the objective of performing predictive analyzes on historical data from the Brazilian stock market, aiming to increase the accuracy of precision through combining the results of multiple approaches. With the *OLS* and *Bayesian Ridge* techniques, combined with

the Voting technique, the highest average coefficient of determination  $R^2$  among the tested combinations was obtained. the value reached was 0.914864, forming a more reliable model, increasing the accuracy of the precision of the result.

**KEYWORDS:** Stock exchanges. Forecasting. Machine learning.

## 1 | INTRODUÇÃO

O mercado de ações é o ambiente no qual empresas de capital aberto e que estejam listadas na bolsa, negociam frações de seu patrimônio através de operações de compra e venda. Essas frações denominam-se ações, que correspondem à participação que o investidor tem em uma empresa, representando um direito sobre os ativos e lucros dessa companhia (NETO, 2018, p. 294-295). O setor acionário brasileiro está concentrado em uma bolsa de valores denominada Brasil, Bolsa, Balcão (B3).

A previsão da oscilação do mercado de ações é considerada uma tarefa complexa para muitos analistas da área de finanças quantitativas, tendo características interdisciplinares, envolvendo áreas como estatística, economia e computação (PASUPULETY et al., 2019).

Com o avanço dos algoritmos de *machine learning* nos domínios financeiros, essa prática se tornou um dos principais recursos na análise técnica atualmente, podendo estipular modelos que possam maximizar os lucros do mercado (LIU, 2017, p. 228-229).

Em termos estatísticos, segundo Pahwa e Agarwal (2019), existem métodos de regressão linear que utilizam valores e atributos e estabelecem um relacionamento entre eles, sendo muito utilizado devido a sua simplicidade e eficácia na previsão. Enquadrando esses conceitos no aprendizado de máquina, é possível adaptar a mesma técnica utilizando recursos para treinamento de um classificador que prediz o valor do rótulo com determinada precisão.

Dessa forma, existe uma gama de técnicas de *machine learning* presentes em diferentes métodos, capazes de realizar previsões através de dados. Logo, surge a indagação de que tipo de técnica deve ser escolhida dentre as muitas existentes. Uma das alternativas é o emprego do *ensemble learning*, em que é possível estabelecer um modo de aprendizado capaz de realizar a união de várias técnicas de *machine learning* (ZHANG & MA, 2012, p. 1-2).

De acordo com Polikar (2009), um sistema baseado em *ensemble* é obtido por meio da combinação de vários modelos que, dentre suas aplicações, pode ser utilizado principalmente para melhorar o desempenho de um modelo ou reduzir a probabilidade de uma má seleção em um modelo ruim.

Assim, ponderado o impacto e a crescente quantidade de estudos publicados sobre o mercado acionário aliado a métodos de *machine learning*, este trabalho tem por objetivo desenvolver um modelo de análise preditiva baseado em métodos de aprendizado conjunto (*ensemble learning*), sendo denominado de *Stock Market Ensemble Predictor* (SMARTER),



pra aplicação sobre dados históricos do mercado de ações brasileiro, visando aumentar a acurácia da previsão por meio da combinação dos resultados de múltiplas abordagens.

## 2 | MATERIAIS E MÉTODOS

Para a realização deste trabalho de pesquisa, foram utilizadas algumas tecnologias, tais como: MongoDB, sistema gerenciador de banco de dados (SGBD) responsável pelo armazenamento dos dados coletados; linguagem de programação Python; Jupyter Notebook, como ambiente de desenvolvimento integrado; biblioteca Scikit-learn, para análise preditiva dos dados; e as bibliotecas do projeto *open-source* Anaconda, para utilização de ferramentas importantes para a realização da ciência de dados, como a manipulação dos dados com o Pandas, a utilização de estruturas de dados e manipulação dessas com Numpy e Scipy, a criação de gráficos e visualizações de dados em geral com o Matplotlib e o Seaborn, bem com a utilização de funções matemáticas com o Math.

Dessa forma, para a criação dos modelos, definiu-se um *workflow*, em que se estipulou um conjunto de etapas que sistematicamente transforma e processa os dados com a finalidade de criar soluções preditivas, sendo organizado pelos seguintes processos: análise do problema, preparação dos dados, adequação ao problema, pré-processamento, treinamento, teste e avaliação do modelo.

Na etapa de **análise do problema**, focou-se no conhecimento do problema em questão, na realização de estudos e pesquisas em artigos que permitiram situar conclusões e observações já formalizadas a respeito desse âmbito e conseqüentemente proporcionar uma definição de resultado esperado.

Em seguida, iniciou-se a etapa de **preparação dos dados**, que foram coletados a partir das cotações históricas da B3 disponíveis em seu site. Os dados foram armazenados no SGBD e em seguida realizou-se uma análise exploratória, com a finalidade de obter uma visão geral, como a identificação de atributos que possuíam valores constantes, os que possuíam valores ausentes, além de obter informações estatísticas destas, para ajudar em tomadas de decisões.

Sendo assim, a quantidade total coletada foi de **2.245.476 registros** em um intervalo temporal compreendido entre 01/01/2015 à 31/05/2019. No entanto, como o presente trabalho visa o mercado acionário, só interessam os registros relacionados aos tipos de mercado **à vista e fracionário**. Do total coletado, cerca de **36,4%** correspondem a esses tipos de mercado, obtendo um conjunto de dados a ser trabalhado com **817.789 registros**.

O mercado de ações possui diferentes formas de influências no preço, o que pode gerar oscilações para determinadas ações. Dessa forma, na etapa de **adequação ao problema**, foi realizada uma divisão em diferentes cenários e intervalos temporais, a fim de verificar como os modelos se comportam em tais configurações.

Para a realização dos testes dos modelos, foram definidas ações de **5 empresas**,

baseado em sua importância para o mercado acionário (as mais negociadas, com os maiores volumes). São elas: ABEV3, BBDC4, ITUB4, PETR4 e VALE3.

Ademais, foram estabelecidos **cenários** distintos, compostos por diferentes **períodos**, com a ideia de testar o comportamento dos modelos em **intervalos temporais** com diferentes quantidades de registros.

Na sequência, tais intervalos foram divididos em duas situações distintas quanto a volatilidade, obtendo cenários com **alta volatilidade** (períodos em que possui altas variações no preço da ação) e **volatilidade normalizada** (períodos com pouca variação no preço da ação).

Para que os modelos de *machine learning* tenham um bom desempenho, é necessário que esses recebam dados com qualidade e úteis para a predição, evitando possíveis problemas com dados ruidosos, redundantes, perdidos, dentre outros (KOTSIANTIS et al., 2007). Para tal, os dados selecionados precisaram passar por um **pré-processamento** antes de serem utilizados para a realização do treinamento.

Sendo assim, foram selecionados os seguintes atributos: **Código de negociação**, para a identificação das ações; **Data do pregão** (data em que a ação foi destinada a fechamento de negócios de compra e venda), para a formulação dos intervalos temporais; **Preço de abertura**, **Preço máximo**, **Preço mínimo** e **Volume total**, para serem os atributos preditores; **Preço de fechamento**, como o atributo a ser predito.

Como os atributos preditores apresentavam unidade de medida diferentes (os preços estão em reais e o volume total em quantidade), essa diferença poderia ocasionar problemas em alguns algoritmos, fazendo-os enviesar para as variáveis com maior ordem de grandeza (SKIENA, 2017, p. 103-104). Logo, a **padronização** desses dados se tornou necessária, utilizando a fórmula *z-score* representada pela equação (1):

$$Z_i = \frac{x_i - \mu}{\sigma} \quad (1)$$

Onde a variável  $x_i$  representa o valor do atributo,  $\mu$  é equivalente à média da distribuição do atributo e  $\sigma$  o desvio padrão associado. Essa fórmula foi aplicada para cada valor presente em cada atributo, deixando todos em uma mesma escala.

Após toda a preparação realizada, os dados estavam prontos para serem utilizados para a criação dos modelos. Nessa etapa, o foco principal foi o **treinamento** destes, utilizando diversas técnicas e métodos, a fim de realizar combinações e aperfeiçoamentos, conquistando melhores resultados.

Na realização do treinamento, é muito importante que o **conjunto de treino** seja diferente do **conjunto de teste**, evitando a ocorrência de *overfitting* (LIU, 2017, p. 17). Assim, uma prática interessante é a **validação cruzada**, a qual consiste do treinamento e teste em diferentes subconjuntos de dados, a fim de entender como o modelo é generalizado (BROWNE, 2000).

Como os dados possuem características temporais, a divisão dos dados foi estabelecida conforme a técnica **Time Series Split** que, a cada iteração utiliza a quantidade de dados que possui datas de pregões correspondentes aos períodos definidos para a realização do treino do modelo e o próximo período utilizado para teste, respeitando os padrões de tempo, ou seja, evitando que dados futuros sejam utilizados para previsão em tempos passados, como explicitado na Figura 1.

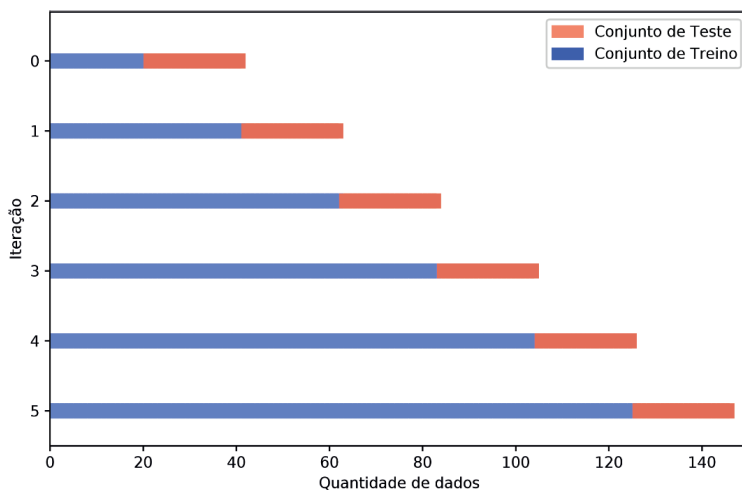


Figura 1 - Divisão de dados para treinamento e teste

Fonte: Próprio autor (2020).

O **conjunto de treino** é variado, sendo composto pela quantidade de **21 datas de pregão**, correspondentes a um intervalo temporal de **1 mês**. A cada iteração, os dados do próximo mês são agrupados ao conjunto de treinamento até que se obtenha um conjunto de treino equivalente a **6 meses**, ou seja, **126 datas de pregão**.

Já o **conjunto de teste** é composto pelos dados que possuem as **próximas 21 datas de pregão** do conjunto de treinamento, ou seja, sempre é testado com os dados do **mês posterior**.

Por fim, na etapa de **avaliação**, estipulou-se uma métrica de avaliação que emite valores úteis para mensurar a eficácia dos modelos testados, a fim de realizar comparações e utilizar como parâmetro para melhorias. Logo, utilizou-se o **coeficiente de determinação  $R^2$**  para a verificação de como cada modelo se ajustou aos dados, ou seja, o quanto o modelo é capaz de generalizar e ser efetivo com dados desconhecidos e, assim, realizar comparações de comportamento entre as diferentes técnicas utilizadas.

Tal métrica possui como valor de coeficiente um limite superior igual a 1, ou seja, quanto mais próximo o **coeficiente** estiver do **valor 1**, significa, a princípio, que os dados estão próximos à reta de regressão (MARTINS, 2018).

### 3 | RESULTADOS E DISCUSSÃO

Esta seção discutirá os resultados alcançados pelos modelos desenvolvidos após a aplicação dos procedimentos descritos.

Na Tabela 1, encontra-se a média dos coeficientes de determinação  $R^2$ , obtidos pelos modelos com base em diferentes técnicas, em cenários com alta volatilidade e volatilidade normal.

Técnica	Alta volatilidade	Volatilidade Normal
Voting (Ordinary Least Squares + Bayesian Ridge)	0,902867	<b>0,926860</b>
Ordinary Least Squares	0,902983	<b>0,926633</b>
Bayesian Ridge	0,901908	<b>0,927001</b>
Bagging (Bayesian Ridge)	0,900691	<b>0,927055</b>
Bagging (Ordinary Least Squares)	0,903929	<b>0,922914</b>
Boosting (Ordinary Least Squares)	0,898316	<b>0,923812</b>
Stacking (Ordinary Least Squares + Bayesian Ridge)	0,890688	<b>0,928051</b>
Boosting (Bayesian Ridge)	0,892249	<b>0,922674</b>

Tabela 1 - Coeficiente de determinação  $R^2$  médio dos modelos em diferentes cenários de volatilidade

Fonte: Próprio autor (2020).

Percebe-se que dentre os modelos citados, todos possuem maiores valores de coeficiente de determinação  $R^2$  nos cenários com volatilidade considerada normal.

Para tanto, os resultados alcançados para intervalos com alta volatilidade, ainda assim são bem expressivos e com bastante relevância, indicando que mesmo com a alta variação de preço da ação, os modelos conseguem ter eficácia na previsão.

Portanto, é importante que um modelo de *machine learning* tenha seu desempenho satisfatório em ambos os cenários considerados (WENG et al.,2018). Sendo assim, a Tabela 2 apresenta a média dos melhores coeficientes de determinação  $R^2$  para os modelos desenvolvidos pelas diferentes técnicas, considerando os dois cenários de volatilidade.

Técnica	Coefficiente de determinação R <sup>2</sup>
<b>Voting (Ordinary Least Squares + Bayesian Ridge)</b>	<b>0,914864</b>
Ordinary Least Squares	0,914808
Bayesian Ridge	0,914455
Bagging (Bayesian Ridge)	0,913873
Bagging (Ordinary Least Squares)	0,913422
Boosting (Ordinary Least Squares)	0,911064
Stacking (Ordinary Least Squares + Bayesian Ridge)	0,909370
Boosting (Bayesian Ridge)	0,907462

Tabela 2 - Coeficiente de determinação R<sup>2</sup> médio dos modelos em ambos os cenários de volatilidade

Fonte: Próprio autor (2020).

Nota-se que a combinação das técnicas *Ordinary Least Squares* (OLS) e *Bayesian Ridge*, por meio de votação, alcança um resultado médio com valores superiores aos obtidos de forma individual.

Dessa forma, aplicou-se um estudo sobre o seu comportamento de uma forma mais aprofundada, buscando identificar os principais motivos que possibilitaram um desempenho robusto ao realizar combinações de técnicas de regressão distintas, sendo destacado na Figura 2.

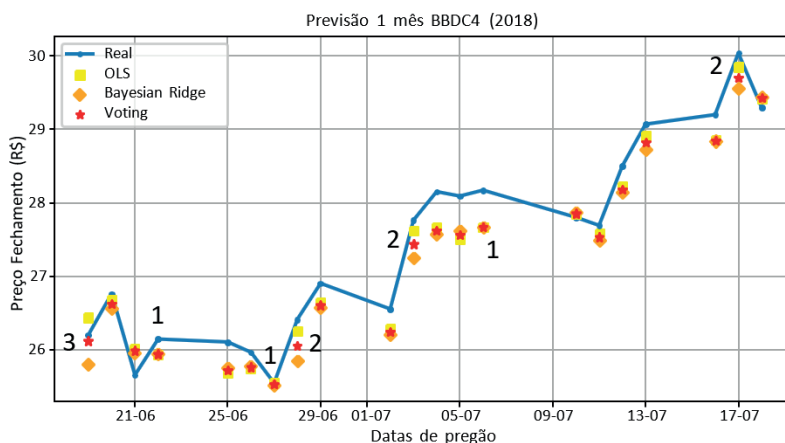


Figura 2 - Previsão de preço de fechamento pelos modelos para a ação BBDC4 utilizando 21 datas de pregão para treinamento e 21 datas de pregão para teste.

Fonte: Próprio autor (2020)

Conforme ilustrado, é possível estabelecer três situações distintas que ocorrem nas predições dos modelos gerados pelas técnicas citadas. A **primeira situação** é quando as previsões são bem **semelhantes**, obtendo os mesmos resultados ao serem combinados.

Já a **segunda situação**, ocorre quando um determinado modelo realiza uma previsão que mais se **aproxima do valor real** em comparação a previsão do outro modelo. Assim, a combinação ao realizar a intermediação, diminui a taxa de erro que o modelo com menos precisão teve, proporcionando o **equilíbrio** de suas fraquezas individuais, sendo útil para os diferentes tipos de comportamentos desenvolvidos em ações distintas.

Por fim, pode-se observar uma **terceira situação**, em que se possui uma **incongruência** de previsão entre os dois modelos utilizados (um modelo realiza previsão **acima do valor real** e o outro realiza previsão **abaixo do valor real**) e a combinação atua como um **intermediador**, obtendo valores mais próximos do real.

Vale destacar que a forma de combinação que a biblioteca utilizada implementa a técnica de *Voting*, é através do cálculo da **média** das previsões individuais dos vários regressores de base utilizados, formando a previsão final. Posto isso, foi de extrema importância a combinação de técnicas que geraram modelos com altos desempenhos, influenciando diretamente na melhora provocada com a utilização dessa combinação.

Assim sendo, devido aos resultados promissores alcançados na utilização da combinação das técnicas **OLS** e **Bayesian Ridge** empregando a técnica de *Voting*, tal combinação foi escolhida para ser a base do modelo **SMARTER**.

## 4 | CONCLUSÃO

Este trabalho desenvolveu o SMARTER, um modelo de análise preditiva baseado em métodos de aprendizado conjunto, aplicados sobre dados históricos do mercado de ações brasileiro, com o objetivo de aumentar a acurácia da previsão da oscilação dos preços das ações.

Após a realização da análise e avaliação dos resultados obtidos, constatou-se que este trabalho de pesquisa conseguiu alcançar o objetivo proposto, ao realizar a combinação das técnicas de regressão mais adequadas para a formação de um novo modelo de aprendizado de máquina, sendo os resultados obtidos por esse, apresentar melhores resultados médios estabelecidos através do coeficiente de determinação  $R^2$  comparados aos demais modelos desenvolvidos.

Portanto, o modelo desenvolvido proporciona benefícios de aumentar a qualidade de generalização, obtendo uma melhor definição do comportamento das ações e consequentemente atingindo previsões mais confiáveis.

Tal modelo pode ser incorporado a uma plataforma que utiliza dados extraídos em tempo real, tendo potencial de ser uma ferramenta de grande utilidade tanto para investidores e também pesquisadores da área.

## REFERÊNCIAS

BROWNE, M. W. **Cross-Validation Methods**. Journal of Mathematical Psychology, v. 44, n. 1, p. 108–132, 2000. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0022249699912798>. Acesso em: 01 jun. 2020.

KOTSIANTIS, S.; KANELLOPOULOS, D.; PINTELAS, P. **Data Preprocessing for Supervised Learning**. World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, v. 1, p. 4104–4109, 2007. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.104.8413&rep=rep1&type=pdf>. Acesso em: 10 jun. 2020.

LIU, Y. **Python Machine Learning By Example**. Packt Publishing, 2017.

MARTINS, E. **Coeficiente de determinação**. Rev. Ciência Elem, v. 6, n. 1, p. 24. Disponível em: <https://rce.casadasciencias.org/rceapp/art/2018/024/>. Acesso em: 10 jul. 2020.

NETO, A. A. **Mercado Financeiro**. 14. ed. São Paulo: Atlas, 2018.

PAHWA, K.; AGARWAL, N. **Stock Market Analysis using Supervised Machine Learning**. IEEE, 2019. Disponível em: <https://ieeexplore.ieee.org/document/8862225>. Acesso em: 20 mar. 2020.

PASUPULETY, U. et al. **Predicting Stock Prices using Ensemble Learning and Sentiment Analysis**. IEEE, 2019. Disponível em: <https://ieeexplore.ieee.org/document/8791689>. Acesso em: 20 mar. 2020.

POLIKAR, R. **Ensemble learning**. Scholarpedia, v. 4, n. 1, p. 2776, 2009. Disponível em: [http://www.scholarpedia.org/article/Ensemble\\_learning](http://www.scholarpedia.org/article/Ensemble_learning). Acesso em: 27 abr. 2020.

SKIENA, S. S. **The Data Science Design Manual**. 1st. Springer Publishing Company, Incorporated, 2017.

WENG, B. et al. **Predicting short-term stock prices using ensemble methods and online data sources**. Expert Systems with Applications, v. 112, p. 258–273, 2018. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0957417418303622>. Acesso em: 17 mar. 2020

ZHANG, C.; MA, Y. **Ensemble Machine Learning: Methods and Applications**. Springer Publishing Company, Incorporated, 2012.

## INFORMATIZAÇÃO DE PROCESSOS GERENCIAIS EM UM SETOR DE ASSISTÊNCIA ESTUDANTIL: ESTUDO DE CASO NO IFMG – CAMPUS BAMBUÍ

*Data de aceite: 01/04/2021*

### **Eduardo Cardoso Melo**

Instituto Federal de Minas Gerais (IFMG)  
BambuÍ/MG

ORCID: <https://orcid.org/0000-0002-1323-5859>

### **Gabriel da Silva**

Instituto Federal de Minas Gerais (IFMG)  
BambuÍ/MG

ORCID: <https://orcid.org/0000-0001-9784-5517>

**RESUMO:** O emprego de recursos tecnológicos computacionais pode ser considerado como um dos principais meios propulsores para diversas transformações ocorridas na sociedade. Em se tratando de gestão organizacional, é quase impensável que empresas busquem atuação diversificada sem o apoio de ferramentas computacionais. Este artigo objetiva apresentar como foi o processo de construção de um software cuja premissa inicial seria contribuir para a informatização de processos internos e com a gestão de um setor de Assistência Estudantil, no caso do IFMG – Campus Bambuí. O gerenciamento das informações deste setor pesquisado era feito apenas com o uso de documentos de texto e planilhas eletrônicas, dificultando sobremaneira a integração de dados. As etapas para a produção do software tiveram como base a adaptação de propostas da metodologia ágil Scrum, com vistas a gerar e disponibilizar um produto final com funcionalidades aderentes às necessidades

do público em questão. Após a implantação do software foi possível notar que os processos de tomada de decisão dos gestores passaram a contar com o importante auxílio de informações geradas pelo sistema. O compartilhamento de informações entre os servidores públicos também foi facilitado.

**PALAVRAS - CHAVE:** Sistemas de Informação, Informatização de processos, Desenvolvimento Ágil, Institutos Federais, Auxílio na tomada de decisão.

### COMPUTERIZATION OF MANAGEMENT PROCESSES IN A STUDENT ASSISTANCE SECTOR: CASE STUDY AT IFMG – CAMPUS BAMBUÍ

**ABSTRACT:** The use of computer technology resources can be considered as one of the main drivers for several transformations that have taken place in society. In organizational management context, it is almost unthinkable for companies to seek diversified performance without the support of computational tools. This article aims to present how the software construction process was, whose initial premise would be to contribute to the computerization of internal processes and to the management of a Student Assistance sector, in the case of IFMG - Campus Bambuí. The management of information in this sector was done only with the use of text documents and spreadsheets, making data integration extremely difficult. The steps for the production of the software were based on the adaptation of proposals from the agile methodology called Scrum, with a view to generating and making available a final product



with functionalities according to the needs of the target public. After the implantation of the software, it was possible to aid the decision-making processes of managers now have the important support information generated by the system. Information sharing between public employees was also facilitated.

**KEYWORDS:** Information Systems, Computerization Processes, Agile Development, Federal Institutes, Aid in Decision Making.

## 1 | INTRODUÇÃO

A Tecnologia da Informação (TI) tem participado diretamente das principais transformações ocorridas no mundo, em especial a partir do final do século XX, sendo facilitadora de diversas mudanças que impactaram os indivíduos tanto em aspectos profissionais quanto pessoais. Quando tais transformações envolvem o ambiente organizacional é possível destacar ainda mais o papel da TI, haja visto que por meio de seus recursos as empresas passam a ter condições de otimizar seus processos internos, incrementar a produtividade de seus colaboradores e obter resultados mais satisfatórios (VILELA et al., 2016; SILVA et al., 2016). As empresas modernas de variadas áreas de atuação dependem dos recursos da TI, como os Sistemas de Informação (SI), para a realização de seus negócios. O nível de dependência das empresas para com sistemas informatizados é alto e em contínuo crescimento, especialmente se for considerada a evolução frequente dos recursos tecnológicos (MARCHIORI e MAINARDES, 2015. ROZA, 2017).

Quando é analisada a utilização de recursos de TI no gerenciamento de unidades educacionais no Brasil, em especial aquelas públicas, torna-se importante ressaltar que mesmo que avanços tenham sido notados recentemente, o contexto geral ainda indica um considerável conjunto de possíveis melhorias (FREITAS, CRESMASCO JÚNIOR e SOUZA, 2015). A ampliação do uso de recursos tecnológicos na administração pública precisa ser incentivada para que as pessoas, sejam elas os próprios servidores públicos ou qualquer indivíduo que necessite de um serviço, possam obter benefícios advindos do suporte prestado pela tecnologia (ARAÚJO, PINHEIRO JÚNIOR e ABREU, 2015; ASSIS, 2016; SZIMANSKI et al., 2015; ZORZAL, 2015).

Atualmente diversas ferramentas tecnológicas são utilizadas no ambiente organizacional do Campus Bambuí do Instituto Federal de Minas Gerais (IFMG), em especial *softwares* para auxiliar na gestão de seus setores e no acesso à informação por parte da comunidade acadêmica (CGTI, 2020). Em sua estrutura organizacional, o Campus Bambuí dispõe de um setor específico para atendimento de demandas dos alunos, denominado Coordenadoria de Assistência Estudantil (CAE). Sua principal atuação consiste na oferta de diversificados serviços para a comunidade acadêmica, tais como alimentação no Restaurante Universitário, atendimentos especializados nas áreas de Odontologia, Medicina, Nutrição, Psicologia e Assistência Social, atividades de esporte e

lazer e residência na Moradia Estudantil (CAE, 2020).

Apesar de oferecer estes e outros serviços para uma comunidade de cerca de 2.500 alunos e estender o atendimento do Restaurante Universitário a aproximadamente 410 servidores e trabalhadores terceirizados, a CAE não contava, até o ano de 2019, com um *software* integrado capaz de auxiliar os servidores que lá atuam na condução e registro de suas atividades, bem como na automatização e otimização de rotinas internas. Toda a gestão documental do setor era feita em formulários de papel, documentos de texto e planilhas eletrônicas, praticamente impossibilitando qualquer demanda de integração de dados e, em especial, a produção de informações relevantes capazes de auxiliar os gestores nos processos de tomadas de decisão. Outra situação frequente era a falta de controle de acesso aos dados do setor, pois os documentos elaborados pelos servidores ficavam disponíveis em uma pasta compartilhada na rede interna sem qualquer análise de perfil sobre o usuário que a acessava. Agrega-se a este contexto um incêndio ocorrido no *Data Center* do campus que ocasionou a perda do equipamento onde os documentos da CAE eram salvos remotamente; quase todo o histórico de dados do setor foi perdido nesta ocasião. Por fim, a Moradia Estudantil possui uma série de rotinas internas que, até então, eram realizadas manualmente, como a chamada diária dos alunos residentes e o controle de saídas dos mesmos.

A proposta deste trabalho foi compreender as reais necessidades gerenciais do setor de Assistência Estudantil do IFMG – Campus Bambuí para, com base neste entendimento, planejar, desenvolver e implantar um *software* integrado e personalizado que disponibilizasse funcionalidades capazes de auxiliar os servidores nas suas tarefas cotidianas, melhorando o atendimento ao público e possibilitando uma gestão íntegra e confiável das informações.

## 2 | FUNDAMENTAÇÃO TEÓRICA

Esta seção visa apresentar os conceitos fundamentais diretamente relacionados com o tema abordado no presente trabalho, o qual envolve estudos sobre Sistemas de Informação e sua contribuição para a melhoria de processos administrativos na gestão pública.

### Sistemas de Informação

Mattos (2017) descreve o conceito de Sistemas de Informação como sendo sistemas com especialização em processar e comunicar dados por meio de módulos interligados entre si em uma rede comum. De Sordi e Meireles (2018) também enfatizam o caráter abrangente dos Sistemas de Informação a partir do momento que eles envolvem e integram em sua utilização vários atores e componentes, porém nada disso seria possível sem o devido tratamento aos dados brutos que são inseridos nos sistemas.

Uma das definições mais aceitas sobre os Sistemas de Informação é apresentada

por Laudon e Laudon (2015), ao afirmarem que eles podem ser visualizados como um conjunto de módulos que atuam relacionados (normalmente de forma direta) com o intuito principal de auxiliar o gerenciamento organizacional e os processos de tomada de decisão nas empresas.

Gomes Filho (2018) ressaltam o atual momento da tecnologia e sua ampla utilização nas mais variadas áreas da sociedade, indicando que o emprego de Sistemas de Informação é viável na maioria das situações que requerem melhorias organizacionais. Entretanto, é fundamental que sejam conduzidos estudos contínuos para averiguar se o Sistema de Informação se mantém alinhado com os objetivos da organização e de seus usuários ao longo do tempo.

## **Tecnologia da Informação (TI) na Gestão Pública**

Uma das áreas de estudo que abordam a aplicação de inovadoras tecnologias da informação no contexto da Gestão Pública é conhecida como “Governo Eletrônico (*e-government*)”. Seu foco é compreender como as novas ferramentas tecnológicas podem auxiliar os membros da administração pública na melhoria de vários aspectos da gestão, notadamente aquele que trata sobre o relacionamento do governo com os cidadãos, empresas e unidades de sua própria estrutura (DIAS, SANO e MEIRELES, 2019).

No entendimento de Rabelo *et al.* (2019), o governo eletrônico precisa ser analisado com uma visão ampla que permita compreendê-lo como um novo modo do governo cumprir o papel do Estado na sociedade de maneira efetiva, ao invés de somente ser considerado como a aplicação de tecnologias para automatizar processos. Brognoli e Ferenhof (2020) apresentam um amplo estudo sobre como vem sendo conduzida a transformação digital no governo brasileiro desde o ano 2000.

## **3 | METODOLOGIA**

Trata-se de um trabalho com abordagem predominantemente qualitativa e de natureza aplicada. De acordo com Freitas e Prodanov (2013, p. 70), “a pesquisa qualitativa não requer uso de métodos e técnicas estatísticas, sendo o ambiente natural a fonte direta para coleta de dados e o pesquisador o instrumento-chave”. Silveira e Córdova (2009) caracterizam a pesquisa aplicada como sendo aquela cujo foco é a solução de problemas específicos, objetivando construir novos conhecimentos práticos a respeito do que é estudado. No que se refere aos procedimentos, esta pesquisa pode ser categorizada como experimental e estudo de caso.

Em relação ao gerenciamento do projeto de desenvolvimento do *software*, algumas propostas da metodologia ágil *Scrum* foram adaptadas e utilizadas, tais como a definição dos *backlogs* (produto, versões e *sprints*) e cerimônias para acompanhamento e discussões sobre o projeto. As atividades desta pesquisa foram realizadas, em sua

maioria, no Laboratório de Sistemas Computacionais (LSC) do Grupo de Pesquisa em Sistemas Computacionais (GPSisCom) do IFMG – Campus Bambuí. Neste ambiente, os participantes tiveram acesso a todos os recursos necessários para o desenvolvimento das atividades planejadas, desde computadores, internet e servidor próprio para aplicações *web*, até salas para reuniões diversas e testes no *software* desenvolvido.

### **Primeira etapa: levantamento e documentação dos requisitos**

A primeira etapa da pesquisa foi iniciada no segundo semestre de 2018 e baseou-se no estudo de como era o funcionamento da CAE naquele momento. Diversas reuniões presenciais com entrevistas não estruturadas foram realizadas com os gestores e servidores deste setor, sendo feita a devida documentação de todos os assuntos tratados para eventuais consultas futuras. Outra atividade conduzida foi a análise de documentos e formulários internos do setor utilizados nas tarefas cotidianas, promovendo assim uma ampla visão de quais informações eram importantes para os gestores e para o sistema como um todo. Naquele momento a equipe da pesquisa identificou a necessidade que o *software* funcionasse como uma aplicação *web*, permitindo o acesso remoto dos usuários e sendo escalável em termos de infraestrutura de armazenamento e processamento.

A partir do conhecimento obtido na primeira etapa dos trabalhos, as atividades para levantamento dos requisitos para o *software* foram iniciadas. Para o refinamento e entendimento complementar dos requisitos foram conduzidas novas reuniões e entrevistas não estruturadas junto aos servidores da CAE. A documentação final dos requisitos foi feita como Estórias de Usuário.

Com o intuito de proporcionar uma visão gráfica das principais funcionalidades disponíveis no sistema para seus usuários, um Diagrama de Caso de Uso geral foi elaborado (Figura 1) e apresentado aos envolvidos com o projeto para validação do escopo que viria a ser desenvolvido.

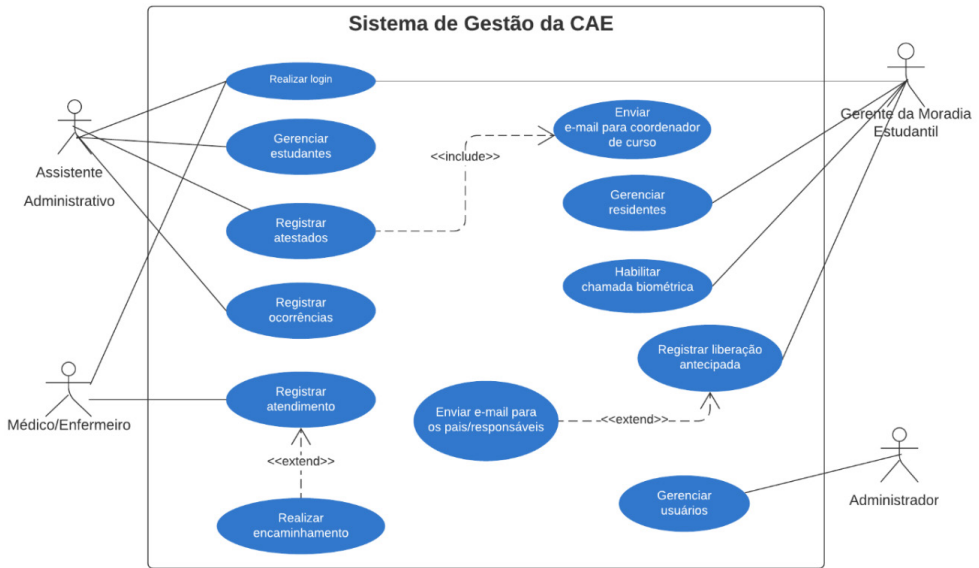


Figura 1: Diagrama de Caso de Uso geral do sistema

Fonte: Os autores (2021)

A partir da definição inicial das funcionalidades do sistema e do entendimento dos requisitos para o seu desenvolvimento, foi elaborado o modelo dos dados na ferramenta *MySQL Workbench* versão 6.3, sendo que sua implementação foi efetivada no Sistema Gerenciador de Banco de Dados (SGBD) *MySQL Community Server Edition* versão 5.7.

A definição das tecnologias para desenvolvimento do sistema foi baseada nos conhecimentos prévios da equipe e no fato de serem de uso livre. A produção do *front-end* da aplicação *web* foi executada com o apoio das tecnologias *HyperText Markup Language* (HTML 5), *Cascading Style Sheets* (CSS 3), *JavaScript*, *Bootstrap 4*, *jQuery 3* e *Asynchronous Javascript and XML* (AJAX), enquanto a produção do *back-end* foi baseada em *PHP 7*, *framework CodeIgniter 3* e servidor *Apache 2.5*.

## Segunda etapa: desenvolvimento da primeira versão do *software*

Os recursos relacionados com o acesso ao sistema fizeram parte da versão inicial desenvolvida (*Sprint 1*), sendo o *login* a primeira funcionalidade implementada pela equipe. Os usuários podem realizar o acesso ao sistema informando seu CPF e uma senha previamente cadastrada pelo Administrador do ambiente. Para aumentar a segurança em relação ao acesso ao sistema, todas as senhas são armazenadas criptografadas no banco de dados e é feita uma limitação das tentativas de acesso não autorizadas para cada usuário. Dentro do escopo de acesso ao sistema, foi desenvolvido também o recurso de gestão dos usuários, cujo objetivo principal é permitir ao Administrador gerenciar as

informações de todas aquelas pessoas que utilizam o sistema.

Considerando a priorização das funcionalidades definidas no momento da criação do *backlog* do projeto, a segunda funcionalidade implementada foi o recurso de gestão dos estudantes que utilizam os serviços da CAE. Nesta rotina o usuário precisa cadastrar desde dados simples, como nome e números de documentos, até dados mais específicos que serão utilizados pelos gestores da CAE como, por exemplo, questões de saúde do estudante.

Em seguida foi implementada a funcionalidade para gestão dos atestados médicos entregues pelos estudantes e validados pela CAE. Os usuários podem inserir dados específicos do atestado como a sua duração e o médico emissor e, caso seja necessário, anexá-lo para futuras consultas. Uma vez que se trata de dados sensíveis e particulares, apenas o Administrador pode ter acesso aos documentos que porventura tenham sido anexados. Outra funcionalidade desenvolvida nesta ocasião foi a de registro de ocorrências (disciplinares ou não) de estudantes, situação rotineira no cotidiano da CAE. As ocorrências podem ser classificadas em públicas ou particulares, sendo essas últimas acessíveis apenas pelo usuário que as cadastrou e pelo Administrador do sistema. A última funcionalidade entregue na versão inicial do sistema foi o gerenciamento de atendimentos feitos pela Enfermaria do *campus*, local onde estudantes e também servidores são recebidos para consultas diversas.

O próprio ambiente de desenvolvimento foi utilizado para a realização dos testes iniciais na primeira versão elaborada, sendo tais tarefas cumpridas pela equipe da pesquisa. Ao mesmo tempo, optou-se por disponibilizar uma cópia local do sistema para alguns servidores da CAE testarem as funcionalidades desenvolvidas. Foi realizado um treinamento prévio para todos os usuários que fizeram parte deste momento de testes iniciais. Pequenos ajustes foram solicitados pelos servidores da CAE, os quais foram prontamente realizados pela equipe da pesquisa a fim de que a primeira versão fosse considerada como pronta para uso em produção. Neste momento, o setor de Tecnologia da Informação do *campus* disponibilizou um ambiente próprio para hospedagem do sistema, cuja configuração foi concluída e permitiu que o sistema passasse a ser utilizado oficialmente pela CAE no início de 2019.

### **Terceira etapa: desenvolvimento da segunda versão do *software***

O desenvolvimento da segunda versão do sistema (*Sprint 2*) foi iniciado tão logo a primeira entrou em produção, focando a implementação de funcionalidades específicas para auxiliar o gerenciamento dos processos internos da Moradia Estudantil. Ela possuía alguns processos internos relacionados com o controle de presença dos residentes, até então dependentes da ação de seus servidores, cujos dados eram preenchidos em formulários de papel. O primeiro deles a ser projetado no sistema foi a liberação antecipada, rotina onde o servidor preenche os dados de determinado estudante, que ficará dispensado de

se apresentar na Moradia durante um período específico e mediante devidas justificativas, como viagem em feriados prolongados para a cidade dos pais ou participação em eventos externos ao campus.

O segundo processo interno da Moradia Estudantil implementado no sistema foi o controle de frequência dos estudantes residentes. Inicialmente foi criado um serviço no SGBD *MySQL* para que, assim que um novo dia é iniciado no servidor, um registro de frequência seja gerado para todos os residentes com a situação “Falta”. A proposta é que os estudantes fiquem com falta até que comprovem a presença na Moradia, o que é feito durante um período de tempo parametrizado pelo gestor do local, ocorrendo normalmente durante o início da noite. Foi implementada também uma rotina para que os usuários Administradores possam consultar a geração de listas de chamadas, possibilitando analisar diariamente se os registros de frequência foram gerados corretamente pelo SGBD.

Ainda dentro do escopo do controle de frequência, a equipe da pesquisa implementou a funcionalidade de chamada dos residentes. Como a identificação dos estudantes deveria ser feita com base na leitura da impressão digital, inicialmente o cadastro de dados dos moradores foi alterado para permitir o armazenamento de duas impressões digitais para cada estudante. O gestor da CAE optou pela padronização de utilizar o dedo indicador de cada mão. Esta modificação permitiu que a rotina de chamada fosse construída tendo como foco a simplicidade. O estudante deve apenas informar seu CPF e, em seguida, posicionar no leitor biométrico a digital de um de seus dedos previamente cadastrados para que o sistema realize as devidas validações (se o estudante é residente da Moradia, se está suspenso por algum motivo, se já havia indicado a presença anteriormente) e altere a situação da chamada para “Presente”. O leitor biométrico adquirido pelo campus e utilizado no sistema foi o *Nitgen Fingkey Hamster III HFDU06S*.

Encerrado o desenvolvimento das funcionalidades da segunda versão, a mesma foi validada em um ambiente de testes similar ao de produção. Enquanto os cadastros gerais foram testados por alguns membros da equipe da CAE, a rotina de controle de frequência foi avaliada durante dois meses envolvendo quatro quartos da Moradia Estudantil. Em seguida, a versão em produção foi atualizada e disponibilizada com as novas funcionalidades.

## 4 | RESULTADOS E DISCUSSÃO

Desde a implantação da primeira versão do sistema em produção, ocorrida no início de 2019, os servidores da CAE interromperam o uso de formulários em papel ou eletrônicos (por meio de documentos de texto ou planilhas eletrônicas) daquelas funcionalidades já implementadas e testadas. Optou-se pela utilização plena do sistema para que a rotina antiga de trabalho não impactasse na familiarização dos usuários com a nova ferramenta. Para que isto fosse possível, treinamentos foram realizados com os servidores e houve um acompanhamento frequente dos gestores da CAE desde o início do projeto.

Em relação aos aspectos técnicos do sistema produzido, a equipe de pesquisa preferiu desenvolver um projeto visual padronizado e simplificado, de maneira que os usuários não tivessem dificuldades para encontrar as funcionalidades disponíveis nem enfrentassem problemas na utilização das rotinas. Neste sentido, todas as interfaces gráficas do sistema apresentam o mesmo padrão de formato gráfico, considerando-se desde aspectos de cores e tipografia até os elementos inseridos nos formulários. Na Figura 2 é possível visualizar a página para manutenção dos dados dos estudantes cadastrados no sistema e constatar a simplicidade dos elementos gráficos selecionados.

ID	Nome	CPF	Curso atual	Última atualização	Ações
4	Abel Carvalho Fernandes	160.118.345-05	Bacharelado em Engenharia da Computação	29/10/2019 23:33:07	[Ícone de edição] [Ícone de telefone] [Ícone de coração] [Ícone de compartilhamento] [Ícone de impressão]
2	Clara Guimarães Melo	222.222.222-22	Bacharelado em Administração	12/09/2018 19:54:29	[Ícone de edição] [Ícone de telefone] [Ícone de coração] [Ícone de compartilhamento] [Ícone de impressão]
5	Fernanda Cristina Carvalho Guimarães	444.555.666-77	Bacharelado em Medicina Veterinária	29/10/2019 23:38:25	[Ícone de edição] [Ícone de telefone] [Ícone de coração] [Ícone de compartilhamento] [Ícone de impressão]
7	Josefrido Cardoso	122.122.122-12	Bacharelado em Engenharia de Alimentos	29/10/2018 10:48:57	[Ícone de edição] [Ícone de telefone] [Ícone de coração] [Ícone de compartilhamento] [Ícone de impressão]
1	Lucca Guimarães Melo	123.456.789-10	Técnico em Manutenção Automotiva Subsequente	05/11/2019 00:39:00	[Ícone de edição] [Ícone de telefone] [Ícone de coração] [Ícone de compartilhamento] [Ícone de impressão]

Figura 2: Página para manutenção dos dados dos estudantes

Fonte: Os autores (2021)

Para melhorar a experiência dos usuários, foi implementada no sistema a proposta de utilização de janelas modais para manipulação dos dados. Quando uma página é carregada, a listagem de registros relacionados é exibida; caso o usuário acione opções como alteração ou inclusão de registros, a janela com o formulário de dados é exibida de maneira sobreposta à listagem, ficando esta bloqueada até que a janela modal seja fechada. A Figura 3 ilustra esse comportamento do sistema ao apresentar a janela modal para alteração dos dados de um estudante fictício.

Figura 3: Alteração dos dados de um estudante por meio de uma janela modal

Fonte: Os autores (2021)



Conforme exposto anteriormente, o processo de registro de frequência foi idealizado para ser de fácil manuseio, em especial por parte dos estudantes. A Figura 4 apresenta a página para realização das chamadas diárias; nela é possível visualizar o campo para preenchimento pelo estudante de seu CPF e um botão para iniciar o processo de leitura da impressão digital.



Figura 4: Página para registro de frequência com auxílio de leitor de impressão digital

Fonte: Os autores (2021)

Uma das grandes possibilidades do sistema está relacionada à geração e obtenção de informações sobre os registros inseridos no banco de dados, o que até então não era possível ou, no melhor dos casos era muito complexo, em função da utilização de documentos de texto e planilhas eletrônicas dispersas entre os usuários do setor. Durante o primeiro ano de utilização do sistema (2019), foram registrados eletronicamente 1.125 atestados médicos, 797 ocorrências, 551 atendimentos no setor de enfermagem e 985 liberações antecipadas de estudantes que residem na Moradia Estudantil. Tais números não podem ser comparados com dados de anos anteriores em função da baixa confiabilidade dos indicadores disponíveis na CAE, até então. Em relação à Moradia, o processo de controle de frequência (e suas rotinas auxiliares) já passou a fazer parte do cotidiano de estudantes e servidores vinculados à CAE. Neste sentido, o sistema desenvolvido trouxe importantes ganhos de produtividade, haja vista que o processo de chamada foi automatizado e a emissão de relatórios gerenciais é feita a qualquer momento, não necessitando de nenhum tipo de preenchimento manual por parte dos servidores.

Finalmente, um dos grandes propósitos da pesquisa foi alcançado, pois a partir da

implantação e utilização frequente e correta do sistema pelos servidores da CAE, tornou-se possível a emissão de um relatório consolidando todas as informações sobre determinado estudante.

## 5 | CONCLUSÕES

Os resultados obtidos a partir da execução desta pesquisa evidenciam que a gestão pode ser beneficiada pela utilização de recursos tecnológicos, tais como *softwares*, desde que os mesmos apresentem a devida adequação e aderência ao ambiente em questão e que, tanto gestores quanto outros tipos de usuários estejam dispostos a apoiar uma forma de trabalho mais dinâmica.

Percebeu-se, entre outras referências positivas, que os problemas de gerenciamento das informações da Assistência Estudantil do IFMG – Campus Bambuí são agora melhor tratados com o apoio do sistema construído, pois todas as funcionalidades desenvolvidas estavam estritamente de acordo com as demandas dos servidores que lá atuam. Ademais, foi possível notar que os processos de trabalho foram facilitados, mesmo ao considerar-se naturais dificuldades advindas do uso de uma nova ferramenta e de novos métodos de trabalho. O compartilhamento de informações passou a ser possível, independentemente do servidor que as detém, pois mesmo com os perfis definidos pelo sistema os dados básicos podem ser acessados por aquelas pessoas que realmente necessitam deles para a tomada de alguma decisão.

Outros benefícios obtidos pela utilização do sistema foram a redução na quantidade de papel utilizado no cotidiano do setor e o acesso mais rápido às informações, seja de um estudante específico ou daquelas consolidadas de determinados indicadores. A integração dos dados também foi facilitada pelo sistema, pois ao utilizar uma base de dados centralizada e padronizada, a geração de informações fica mais simples de ser processada, neste caso em formato de relatórios gerenciais cujo modelo foi definido previamente pelos servidores da CAE. A rotina de *backup* desta base de dados em nuvem permite também que, caso ocorra algum novo incidente com a infraestrutura de TI do *campus*, os dados não se percam.

Uma dificuldade enfrentada durante a execução das atividades da pesquisa que merece menção refere-se às repetidas mudanças do coordenador da Assistência Estudantil do IFMG – Campus Bambuí, evento ocorrido por três vezes até o presente momento. Tal fato criou situações que atrasaram o desenvolvimento de algumas funcionalidades do sistema, pois até que os novos gestores tomassem conhecimento do mesmo e indicassem sugestões de ajustes, as tarefas de implementação ficaram pendentes para evitar possíveis retrabalhos.

## REFERÊNCIAS

ARAÚJO, B. D. L.; PINHEIRO JÚNIOR, F. M. L.; ABREU, F. P. **Gerenciamento de aquisições de soluções de tecnologia da informação em uma universidade pública do Ceará**. Revista Gestão em Análise, v. 4, n. 1, p. 43-53, 2015.

ASSIS, W. C. **Governança de TI na administração Pública Federal: aprimorando a governança de TI no setor público brasileiro**. 2016. Disponível em: <http://repositorio.uniceub.br/jspui/handle/235/8064>. Acesso em: 17 fev. 2021.

BROGNOLI, T. S.; FERENHOF, H. A. **Transformação Digital no governo brasileiro: desafios, ações e perspectivas**. Navus - Revista de Gestão e Tecnologia, [S.l.], v. 10, p. 01-11, jan/dez. 2020. ISSN 2237-4558. Disponível em: <http://navus.sc.senac.br/index.php/navus/article/view/989/pdf>. Acesso em: 24 fev. 2021.

CAE. **Coordenadoria de Assistência Estudantil (CAE) do IFMG – Campus Bambuí**. Disponível em: <http://bambui.ifmg.edu.br/portal/cae>. Acesso em: 04 mar. 2021.

CGTI. **Sistemas de Informação da Coordenadoria de Gestão de Tecnologia da Informação do IFMG – Campus Bambuí**. Disponível em: <http://www.bambui.ifmg.edu.br/portal/cgti/sistemas-de-informacao>. Acesso em: 04 mar. 2021.

DIAS, T. F.; SANO, H.; MEDEIROS, M. F. M. **Inovação e tecnologias da comunicação e informação na administração pública**. Brasília: ENAP, 2019.

FREITAS, A. L. S.; CREMASCO JÚNIOR, D.; SOUZA, M. P. **Transição da gestão de tecnologia da informação para a governança eletrônica na Universidade Federal de Rondônia**. 2015. Disponível em: <https://repositorio.ufsc.br/handle/123456789/136120>. Acesso em: 15 fev. 2021.

FREITAS, E. C. de; PRODANOV, C. C. **Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico**. 2. ed. Novo Hamburgo: Feevale, 2013.

GOMES FILHO, A. C. **Gestão de Tecnologia da Informação: Abordagem para o Setor Público**. 2018. Disponível em: [http://repositorio.unicentro.br:8080/jspui/bitstream/123456789/1314/1/GOMES\\_Gestao\\_de%20TI\\_abordagem\\_para\\_setor\\_publico.pdf](http://repositorio.unicentro.br:8080/jspui/bitstream/123456789/1314/1/GOMES_Gestao_de%20TI_abordagem_para_setor_publico.pdf). Acesso em: 23 fev. 2021.

LAUDON, K.; LAUDON, J. **Sistemas de Informações Gerenciais**. São Paulo: Pearson, 2015.

MARCHIORI, D. M.; MAINARDES, E. W. **A Relação Entre os Fatores Criadores do Technostress e a Qualidade Percebida de Serviços de Tecnologia da Informação**. ENCONTRO SEMINÁRIOS EM ADMINISTRAÇÃO (SEMEAD), v. 18, 2015.

MATTOS, A. C. M. **Sistemas de Informação: uma visão executiva**. 2. ed. São Paulo: Saraiva, 2017.

RABELO, C. L. A. et al. **A participação da sociedade no governo eletrônico sob a perspectiva da democracia digital**. Revista Argumentum-Argumentum Journal of Law, v. 13, p. 225-255, 2019. Disponível em: <http://ojs.unimar.br/index.php/revistaargumentum/article/view/1093/687>. Acesso em: 15 fev. 2021.

ROZA, R. H. **Revolução informacional e os avanços tecnológicos da informática e das telecomunicações**. Ciência da Informação em Revista, v. 4, n. 3, p. 3-11, 2017.

SILVA, R. T. et al. **Gestão por Processos em Tecnologia da Informação**: um estudo em uma empresa de médio porte. Revista de Administração do UNIFATEA, v. 12, n. 12, 2016.

SILVEIRA, D. T.; CÓRDOVA, F. P. **A pesquisa científica**. Porto Alegre: UFRGS, 2009.

SORDI, J. O.; MEIRELES, M. **Administração de Sistemas de Informação**. 2. ed. São Paulo: Saraiva, 2018.

SZIMANSKI, F. et al. **Reestruturando a Área de Tecnologia da Informação**: um relato de experiência na administração pública. In: Anais do XI Simpósio Brasileiro de Sistemas de Informação. SBC, 2015. p. 475-482.

VILELA, B. D. et al. **Coordenação em cadeias de suprimentos**: o papel da tecnologia da informação e da gestão orientada por processos. Exacta, v. 14, n. 4, p. 645-660, 2016.

ZORZAL, L. **Transparência das informações das universidades federais**: estudo dos relatórios de gestão à luz dos princípios de boa governança na Administração Pública Federal. 2015. Disponível em: <http://repositorio.unb.br/handle/10482/19202>. Acesso em: 15 fev. 2021.

Data de aceite: 01/04/2021

**Pedro Henrique Senkiio Cardoso**

Universidade Anhembi Morumbi  
São Paulo, Brasil

**RESUMO:** Os mascotes virtuais encontraram seu nicho dentro do gênero “simulação” nas principais plataformas de distribuição de jogos, ao lado de títulos como *The Sims*, que também utiliza algoritmos especializados em simular emoções em seus personagens. Porém há uma grande diferença entre a simulação de emoções de *The Sims* e de mascotes virtuais como *Talking Tom* ou *Pou*. Este artigo discute quais são essas diferenças, levantando os pontos positivos e negativos da complexidade na inteligência artificial dos mascotes virtuais mais populares e seu impacto na experiência dos jogadores.

### THE SIMULATION OF EMOTIONS ON VIDEO GAMES

**ABSTRACT:** Virtual pets have found their place on game distribution platforms under the “simulation” genre, beside games like *The Sims*, that also use emotion simulation algorithms to rule their character’s behaviour. However, there is a huge difference between the emotion simulation of *The Sims* and virtual pets like *Talking Tom* and *Pou*. This paper discusses what are these differences, considering the positive and negative side of complexity within the most popular virtual pets AI and its impacts on player experience.

**KEYWORDS:** Artificial Intelligence, Emotions, Games.

### 1 | INTRODUÇÃO

Dentro das principais plataformas distribuidoras de jogos como *Google Play*, *Apple Store* e *Steam*, o gênero “simulação” tem seu lugar ao lado de outras categorias mais populares como “ação” e “aventura”. O rótulo “simulação” implica em conteúdo similar a atividades do mundo real como cozinhar, pilotar veículos, jardinagem entre outros tópicos mundanos. Apesar deste vínculo com o mundo real, os jogos do gênero “simulação” podem tomar liberdades para se aproximar do entretenimento em detrimento da fidelidade da simulação. [1]

Entre os títulos que constituem o gênero “simulação” destacam-se os jogos que simulam interações sociais como *The Sims*, *Talking Tom*, *Pou* e *Boo*. Nestes exemplos o jogador interage com agentes inteligentes especializados em simular necessidades fisiológicas como saciedade, energia e saúde. Com base nestas variáveis o agente inteligente demonstra diferentes emoções ao jogador por meio de expressões faciais, tom de voz e linguagem corporal. A complexidade da simulação de emoções destes agentes inteligentes pode ser medida pela quantidade de variáveis que impactam o algoritmo.

Enquanto as simulações de emoções mais simples usam como parâmetro poucas variáveis como as necessidades do agente inteligente, as mais complexas usam mais parâmetros como traços de personalidade, experiências recentes e até mesmo o contexto, criando interações complexas.

O benefício de simulações complexas com muitas variáveis é evitar que as respostas do agente inteligente se tornem previsíveis, expondo a máquina debaixo dos panos e comprometendo a imersão do jogador. [2] Apesar disso, um sistema complexo demais pode ter o mesmo efeito ao ser interpretado como “aleatório” ou “acaso” pelo usuário. [3]

A demonstração de emoções é um dos principais elementos que caracterizam um mascote virtual. [4] Com o intuito de evitar que este sistema importante se torne muito simples e previsível, ou muito complexo e imprevisível, é possível recorrer à utilização de modelos teóricos que sintetizam o tópico complexo das emoções, o deixando mais tangível e compreensível. Com esta ferramenta, os desenvolvedores de jogos utilizam uma referência mais confiável do que recorrer apenas às suas experiências pessoais.

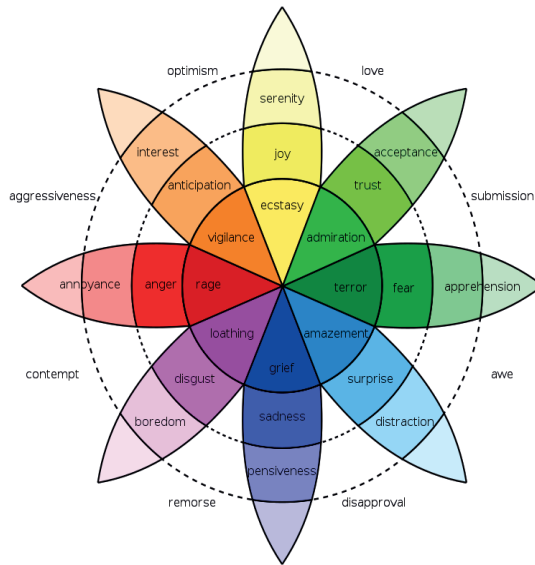
O uso de experiências anedóticas no desenvolvimento destes sistemas os carrega de ideias que só fazem sentido para os jogadores que compartilham a mesma cultura dos desenvolvedores. Isto pode ser desejável caso o público alvo seja constituído apenas destes jogadores. Um exemplo de sistema impactado pela cultura dos desenvolvedores é *The Sims*, em que o jogador tenta suprir todas as necessidades de seu personagem para que este fique sempre contente: Em outras culturas, o equilíbrio entre emoções é mais valorizado que o foco nas emoções positivas. [5]

## 2 | SINTETIZANDO EMOÇÕES

Existem diversos modelos que elucidam e definem o conceito de “emoção” de formas diferentes. [6] Entre eles, destaca-se o modelo do psicólogo Robert Plutchik, definindo as emoções em dez postulados:

1. O conceito de emoção é aplicável em todo grau evolucionário, se aplicando a todo animal.
2. Emoções têm origem evolutiva que se diversificou em várias expressões em espécies diferentes.
3. Emoções tiveram um papel evolutivo, ajudando organismos a lidar com problemas impostos pelo ambiente.
4. Apesar de se expressar de formas diferentes em espécies diferentes, há elementos em comum que podem ser identificados.
5. Há um número pequenos de emoções primevas, primárias ou básicas.
6. Todas outras emoções são combinações ou estados derivados de emoções primárias.

7. Emoções primárias são constructos hipotéticos, estados idealizados que só podem ser caracterizados por meio de diferentes tipos de evidência.
8. Emoções primárias podem ser definidas em pares opostos e exclusivos.
9. Toda emoção varia em grau de similaridade uma da outra.
10. Cada emoção pode se manifestar em níveis diferentes de intensidade.

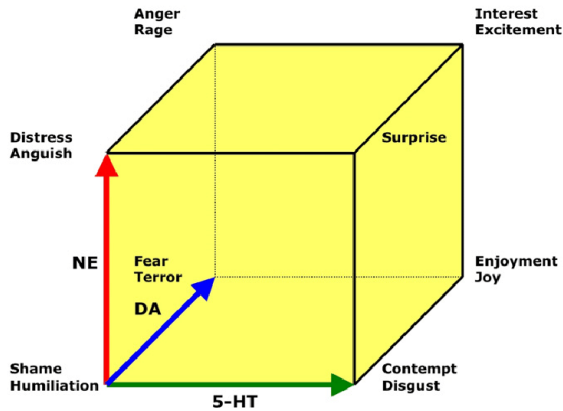


### Modelo visual das emoções: Roda de Plutchik

Este modelo apresenta as emoções básicas em posição diametralmente oposta às suas contrapartes, se desdobrando em suas versões mais brandas.

O modelo de Plutchik foi organizado com base na evolução das espécies em sua obra *“A Psychoevolutionary Theory of Emotions”*, considerando a utilidade das emoções para a sobrevivência, como por exemplo o medo contribuindo como proteção. [7]

Antes do modelo de Plutchik, o psicólogo Silvan Tomkins organizou as emoções de uma forma diferente em sua obra *“Affect Theory”*. [8] que foi utilizado por Hugo Lövhelm e associada à atividade dos neurotransmissores do cérebro humano, tornando as emoções observáveis. Essa associação deu origem ao modelo visual em forma de cubo em que cada vértice representa uma emoção e os vetores representam a atividade de neurotransmissores, mapeando as emoções de forma que é possível navegar entre todas elas alterando apenas três variáveis: Noradrenalina (vetor vermelho), dopamina (vetor azul) e serotonina (vetor verde). [9]



Cubo de Lövheim: O vetor “NE” corresponde à Noradrenalina, “DA” corresponde à dopamina e S-HT corresponde à Serotonina.

Estes modelos podem ser aplicados na elaboração das inteligências artificiais especializadas em simular emoções ao compor uma máquina de estados em que cada estado corresponde a uma emoção. Apesar dos modelos não serem totalmente compatíveis, há muitas similaridades que podem contribuir com a elaboração de um sistema especializado em simular emoções.

### 3 | APLICANDO O MODELO

Como discutido na introdução, um jogo do gênero “simulação” não precisa ser fiel à realidade como um simulador. O objetivo da aplicação destes modelos no algoritmo do mascote virtual é de criar a ilusão de complexidade emocional em um agente inteligente. Portanto os modelos serão utilizados como a base para criar algo novo, mais sintético e palatável.

Desenvolvido em conjunto com este artigo para ilustrá-lo, o jogo “Virtual Tamer” utiliza como base do algoritmo de simulação de emoções a roda de Plutchik e o cubo de Lövheim apresentados acima.

#### 3.1 Máquina de Estados

A máquina de estados utiliza três parâmetros, assim como o Cubo de Lövheim, para navegar pelas emoções. Estes parâmetros foram adaptados substituindo os neurotransmissores pelas necessidades básicas do mascote virtual: Serotonina representada pelo atributo “saciedade”, Dopamina representada pelo atributo “energia” e Noradrenalina representada pelo atributo “coragem”.

É possível inferir que a roda de Plutchik trata as emoções como quatro variáveis independentes, mas que quando somadas resultam em uma emoção complexa compatível com uma posição no Cubo de Lövheim. Portanto, foram atribuídas sub-emoções para



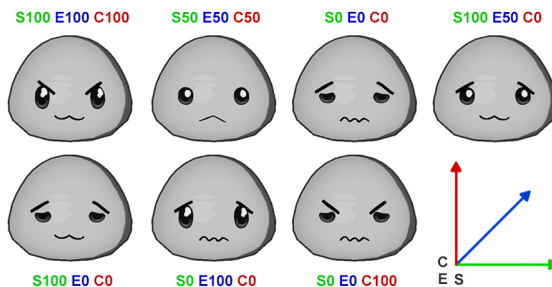
cada vetor do cubo baseado nas emoções de Plutchik: Cada vetor retorna o valor “baixo”, “neutro” e “alto” representando a magnitude do vetor.

Este sistema permite navegar entre emoções através da modificação das necessidades do mascote, portanto criando uma resposta adequada conforme o mascote interage com o jogador e com o ambiente de forma dinâmica. Porém esta resposta ainda não é visível ao jogador, para transmitir o *feedback* da máquina de estados é necessário um sistema de representação visual.

### 3.2 Representação Visual

Para expressar adequadamente a emoção resultante da máquina de estados ao jogador foi necessário demonstrar o impacto de cada uma das três variáveis principais de acordo com as suas respectivas intensidades.

A fim de otimizar este processo, o rosto do personagem foi dividido em três partes que se complementam, cada parte é impactada apenas por uma das variáveis. A saciedade é responsável pela boca do mascote virtual. A energia é responsável pelos olhos. A coragem é responsável pelas sobrancelhas.



Expressões faciais do mascote virtual.

Os valores representam a magnitude dos vetores. Verde representa a saciedade, azul representa a energia e vermelho representa a coragem.

Isoladamente cada variável transmite uma ideia através da área do rosto que comanda. As três variáveis juntas transmitem a emoção equivalente à sua posição na máquina de estados. Dessa forma, apenas olhando para o mascote o jogador já pode reconhecer se há algo de errado com as suas necessidades. Depois de alguma prática, o jogador pode até mesmo saber exatamente qual é a necessidade deficiente apenas observando a expressão facial do mascote.

## 4 | PERSONALIDADE

Apesar de ser regido por apenas três variáveis, a máquina de estados usa outros parâmetros para decidir quais estímulos afetam o agente inteligente tanto no sentido do vetor quanto na intensidade. Assim surge a oportunidade de adicionar uma personalidade ao agente inteligente. Uma vez que cada estímulo passa por esse filtro com o poder de amplificar, neutralizar ou inverter o input recebido.

O conceito de personalidade é sintetizado pelo Myer Briggs Type Indicator, um teste de personalidade que atribui características ao realizador. **[10]** São quatro características que juntas classificam o realizador em um de dezesseis grupos de personalidade. Existem controvérsias quanto a validade deste teste devido à inconsistência dos resultados cada vez que o teste é aplicado no mesmo indivíduo quando comparado à testes semelhantes. **[11]** Apesar deste problema na identificação das características que compõe a personalidade, este teste de personalidade é o mais utilizado.

Este problema não afeta o uso do modelo de Myer Briggs na composição do algoritmo para um agente inteligente, porque não há características a serem identificadas. Logo o sistema pode utilizar apenas a parte do teste que associa as características para atribuir uma personalidade ao mascote virtual. Com base neste modelo, foram definidas quatro características com pares opostos. A personalidade do agente inteligente é definida por estas quatro escolhas: Introverso ou Extroverso; Realista ou Idealista; Racional ou Emocional; Espontâneo ou Organizado.

Cada uma dessas escolhas afeta o modo com que o agente inteligente reage aos estímulos do jogador e do ambiente.

## 5 | CONCLUSÃO

Inteligências artificiais especializadas na simulação de emoções para fins de entretenimento podem se beneficiar de biomimetismo e modelos teóricos que sintetizam o conceito de emoção. Porém é importante pontuar que o papel da IA emotiva é apenas de criar uma ilusão para convencer o usuário, sem compromisso com a fidelidade do modelo teórico apesar do benefício da verossimilhança. Portanto deve haver um equilíbrio entre fidelidade e clareza do sistema para manter a imersão do jogador, tendo o entretenimento como fim principal.

## REFERÊNCIAS

[1] NARAYANASAMY, Viknashvaran & WONG, Kok Wai & FUNG, Chun Che & RAI, Shri “Distinguishing games and simulation games from simulators.” <link>

[2] CHENG, Kevin & CAIRNS, Paul “Behaviour, realism and immersion in games” <link>

[3] MILLINGTON, Ian & FUNGE, John “Artificial Intelligence for Games” page 19

[4] KAPLAN, Frederic “Artificial attachment: Will a robot ever pass ainsworth’s strange situation tests” index 2.4 <link>

[5] PARKER, Clifton B. “European Americans embrace positive feelings, while Chinese prefer a balance of feelings” <link>

[6] ORTONY, A. & TURNER, T. J. “What’s basic about basic emotions?” Psychological Review 97(3), 315-331 <link>

[7] PLUTCHIK, Richard “A psychoevolutionary theory of emotions” <link>

[8] TOMKINS, Silvan & DEMOS, C “Exploring affect: The selected writings of Silvan S. Tomkins”

[9] LÖVHEIM, Hugo “A new three-dimensional model for emotions and monoamine neurotransmitters” <link>

[10] MYERS, I. B. “The Myers-Briggs Type Indicator: Manual” <link>

[11] FURNHAN, Adrian “The big five versus the big four: the relationship between the Myers-Briggs Type Indicator (MBTI) and NEO-PI five factor model of personality” <link> **Referência de Formato:** SBGAMES <link>

## **SOBRE O ORGANIZADOR**

**ERNANE ROSA MARTINS** - Doutor em Ciência da Informação com ênfase em Sistemas, Tecnologias e Gestão da Informação, na Universidade Fernando Pessoa, em Porto/Portugal. Mestre em Engenharia de Produção e Sistemas, possui Pós-Graduação em Tecnologia em Gestão da Informação, Graduação em Ciência da Computação e Graduação em Sistemas de Informação. Professor de Informática no Instituto Federal de Educação, Ciência e Tecnologia de Goiás - IFG (Câmpus Luziânia) ministrando disciplinas nas áreas de Engenharia de Software, Desenvolvimento de Sistemas, Linguagens de Programação, Banco de Dados e Gestão em Tecnologia da Informação. Pesquisador do Núcleo de Inovação, Tecnologia e Educação (NITE), certificado pelo IFG no CNPq. ORCID: <https://orcid.org/0000-0002-1543-1108>. Personal homepage: <https://ernane.martins.wordpress.com/>

## ÍNDICE REMISSIVO

### A

Alfabetização 6, 34, 35, 36, 37, 38, 44, 45

Algoritmo Genético 5, 18, 46, 47, 50, 51, 52, 54, 55, 56, 57

Algoritmos 17, 19, 70, 130, 132, 151

Android 4, 79, 82, 87, 88, 89, 94

Aprendizado do computador 129

Aprendizagem 34, 35, 36, 38, 45, 69, 76, 124, 127

### C

Classificação 5, 8, 18, 24, 118, 120, 125, 126, 127

Computador 24, 80, 129

Correlação 5, 24, 25, 30, 31, 32, 80

### D

Dados 2, 13, 14, 15, 16, 17, 18, 19, 20, 24, 25, 26, 27, 29, 30, 34, 46, 47, 51, 53, 56, 63, 71, 80, 81, 83, 87, 88, 91, 92, 93, 94, 105, 106, 118, 119, 120, 121, 122, 125, 127, 129, 130, 131, 132, 133, 136, 138, 140, 141, 143, 144, 145, 146, 147, 148, 158

### E

Educação 1, 35, 36, 37, 45, 68, 81, 105, 158

Evolução Diferencial 5, 46, 47, 50, 52, 54, 55, 56, 57

### F

Framework 1, 2, 5, 59, 60, 61, 76, 92, 143

### G

Gamificação 36, 38

### H

Hardware 19, 20

### I

Inteligência Artificial 5, 24, 151

### J

Jogo 34, 35, 38, 39, 40, 41, 42, 44, 45, 154

### L

Linguagem de programação 35, 36, 89, 91, 131

## **M**

Machine Learning 21, 25, 107, 108, 116, 129, 130, 132, 134, 137

Método dos Pontos Interiores 7, 96

Método Numéricos 96

Modelagem 5, 13, 15, 16, 18, 20, 21, 26, 27, 29, 30, 57, 59, 66, 68, 69, 76, 94, 118

Modelo 5, 16, 17, 18, 19, 20, 24, 25, 26, 27, 30, 31, 32, 33, 46, 47, 48, 57, 59, 60, 63, 64, 66, 70, 71, 72, 73, 75, 94, 120, 129, 130, 131, 132, 133, 134, 136, 143, 148, 152, 153, 154, 156

Modelos Compartimentais 68, 69

## **N**

Network 23, 33, 68, 76, 104, 105, 106, 107, 108, 109, 110, 111, 112, 115, 116, 117, 118, 128

## **O**

Otimização 6, 7, 18, 19, 46, 47, 50, 51, 52, 54, 55, 56, 57, 59, 64, 96, 97, 98, 99, 101, 103, 140

## **P**

Perceptron 5, 105, 107, 110, 118, 120, 127, 128

Previsão 5, 6, 13, 14, 15, 16, 17, 18, 19, 20, 21, 24, 33, 105, 106, 119, 129, 130, 131, 133, 134, 135, 136

Primal-Dual 7, 96, 97, 98, 101, 103

Programação 5, 6, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 45, 63, 84, 89, 91, 119, 123, 131, 158

## **R**

Rede Neural Artificial 5, 106, 118, 120, 123, 124

Redes Randômicas 68, 72

Regressão Linear 16, 20, 130

## **S**

Scratch 6, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45

Scrum 5, 59, 60, 61, 66, 67, 138, 141

Simulações Computacionais 5, 68, 70, 76

Sistemas Elétricos de Potência 96, 103

Softwares 38, 60, 63, 139, 148

Sprint 7, 59, 60, 61, 62, 63, 64, 65, 66, 143, 144

Substituição Sensorial 79, 80, 81, 82, 83, 84, 95

## T

Tecnologia 35, 59, 62, 68, 79, 80, 82, 95, 105, 127, 139, 141, 144, 149, 150, 158

Tecnologias Assistivas 80, 81

Twitter 21, 24, 25, 26, 27, 33

# TECNOLOGIAS, MÉTODOS E TEORIAS NA ENGENHARIA DE COMPUTAÇÃO 2

[www.atenaeditora.com.br](http://www.atenaeditora.com.br) 

[contato@atenaeditora.com.br](mailto:contato@atenaeditora.com.br) 

[@atenaeditora](https://www.instagram.com/atenaeditora) 

[www.facebook.com/atenaeditora.com.br](https://www.facebook.com/atenaeditora.com.br) 

 **Atena**  
Editora

Ano 2020



# TECNOLOGIAS, MÉTODOS E TEORIAS NA ENGENHARIA DE COMPUTAÇÃO 2

[www.atenaeditora.com.br](http://www.atenaeditora.com.br) 

[contato@atenaeditora.com.br](mailto:contato@atenaeditora.com.br) 

[@atenaeditora](https://www.instagram.com/atenaeditora) 

[www.facebook.com/atenaeditora.com.br](https://www.facebook.com/atenaeditora.com.br) 

 **Atena**  
Editora

Ano 2020