

ORGANIZADORES

Jucelio Soares dos Santos
Jannayna Domingues Barros Filgueira
Rosângela Araújo de Medeiros
Geam Carlos de Araújo Filgueira
Fabio Junior Francisco da Silva
Johny Venicios Carvalho da Silva
Rômulo Rodrigues de Moraes Bezerra

anais do VI CONGRESSO SERTANEJO DE COMPUTAÇÃO

AVANÇOS EM COMPUTAÇÃO APLICADA
Robótica, IA e Educação no Sertão



ORGANIZADORES

Jucelio Soares dos Santos
Jannayna Domingues Barros Filgueira
Rosângela Araújo de Medeiros
Geam Carlos de Araújo Filgueira
Fabio Junior Francisco da Silva
Johny Venicios Carvalho da Silva
Rômulo Rodrigues de Moraes Bezerra

anais do **V** CONGRESSO SERTANEJO DE COMPUTAÇÃO

AVANÇOS EM COMPUTAÇÃO APLICADA
Robótica, IA e Educação no Sertão



DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)

A532

Anais do VI Congresso Sertanejo de Computação: Avanços em Computação Aplicada - Robótica, IA e Educação no Sertão / Organização Jucelio Soares dos Santos... [et al.]. - São Paulo: Pimenta Cultural, 2025.

Demais organizadores: Jannayna Domingues Barros Filgueira, Rosângela Araújo de Medeiros, Geam Carlos de Araújo Filgueira, Fabio Junior Francisco da Silva, Johny Venícios Carvalho da Silva, Rômulo Rodrigues de Moraes Bezerra.

Volume 1

Livro em PDF

ISBN 978-85-7221-449-0

DOI 10.31560/pimentacultural/978-85-7221-449-0

1. Computação Aplicada. 2. Educação Tecnológica. 3. Inteligência Artificial. 4. Metodologias Ativas. 5. Inovação no Sertão. I. Santos, Jucelio Soares dos (Org.). II. Filgueira, Jannayna Domingues Barros (Org.). III. Medeiros, Rosângela Araújo de (Org.). IV. Filgueira, Geam Carlos de Araújo (Org.). V. Silva, Fabio Junior Francisco da (Org.). VI. Silva, Johny Venícios Carvalho da (Org.). VII. Bezerra, Rômulo Rodrigues de Moraes (Org.). VIII. Título.

CDD 004.07

Índice para catálogo sistemático:

I. Computação Aplicada

Simone Sales • Bibliotecária • CRB: ES-000814/0

Copyright © Pimenta Cultural, alguns direitos reservados.

Copyright do texto © 2025 os autores e as autoras.

Copyright da edição © 2025 Pimenta Cultural.

Esta obra é licenciada por uma Licença Creative Commons:

Atribuição-NãoComercial-SemDerivações 4.0 Internacional - (CC BY-NC-ND 4.0).

Os termos desta licença estão disponíveis em:

<<https://creativecommons.org/licenses/>>.

Direitos para esta edição cedidos à Pimenta Cultural.

O conteúdo publicado não representa a posição oficial da Pimenta Cultural.

Direção editorial	Patricia Biegging Raul Inácio Busarello
Editora executiva	Patricia Biegging
Gerente editorial	Landressa Rita Schiefelbein
Assistente editorial	Júlia Marra Torres
Estagiária editorial	Ana Flávia Pivisan Kobata
Diretor de criação	Raul Inácio Busarello
Assistente de arte	Naiara Von Groll
Editoração eletrônica	Andressa Karina Voltolini
Estagiárias em editoração	Raquel de Paula Miranda Stela Tiemi Hashimoto Kanada
Imagens da capa	Chat GPT - AI Generator, Freepik - AI Generator
Tipografias	Acumin, Gobold
Revisão	Os organizadores
Organizadores	Jucelio Soares dos Santos Jannayna Domingues Barros Filgueira Rosângela Araújo de Medeiros Geam Carlos de Araújo Filgueira Fabio Junior Francisco da Silva Johny Venicios Carvalho da Silva Rômulo Rodrigues de Morais Bezerra

PIMENTA CULTURAL

São Paulo • SP

+55 (11) 96766 2200

livro@pimentacultural.com

www.pimentacultural.com



2 0 2 5

CONSELHO EDITORIAL CIENTÍFICO

Doutores e Doutoradas

Adilson Cristiano Habowski

Universidade La Salle, Brasil

Adriana Flávia Neu

Universidade Federal de Santa Maria, Brasil

Adriana Regina Vettorazzi Schmitt

Instituto Federal de Santa Catarina, Brasil

Aguimario Pimentel Silva

Instituto Federal de Alagoas, Brasil

Alaim Passos Bispo

Universidade Federal do Rio Grande do Norte, Brasil

Alaim Souza Neto

Universidade Federal de Santa Catarina, Brasil

Alessandra Knoll

Universidade Federal de Santa Catarina, Brasil

Alessandra Regina Müller Germani

Universidade Federal de Santa Maria, Brasil

Aline Corso

Universidade do Vale do Rio dos Sinos, Brasil

Aline Wendpap Nunes de Siqueira

Universidade Federal de Mato Grosso, Brasil

Ana Rosangela Colares Lavand

Universidade Estadual do Norte do Paraná, Brasil

André Gobbo

Universidade Federal da Paraíba, Brasil

André Tanus Cesário de Souza

Faculdade Anhanguera, Brasil

Andressa Antunes

Universidade Federal de Ouro Preto, Brasil

Andressa Wiebusch

Universidade Federal de Santa Maria, Brasil

Andreza Regina Lopes da Silva

Universidade Federal de Santa Catarina, Brasil

Angela Maria Farah

Universidade de São Paulo, Brasil

Anísio Batista Pereira

Universidade do Estado do Amapá, Brasil

Antonio Edson Alves da Silva

Universidade Estadual do Ceará, Brasil

Antonio Henrique Coutelo de Moraes

Universidade Federal de Rondonópolis, Brasil

Arthur Vianna Ferreira

Universidade do Estado do Rio de Janeiro, Brasil

Ary Albuquerque Cavalcanti Junior

Universidade Federal de Mato Grosso, Brasil

Asterlindo Bandeira de Oliveira Júnior

Universidade Federal da Bahia, Brasil

Bárbara Amaral da Silva

Universidade Federal de Minas Gerais, Brasil

Bernadette Beber

Universidade Federal de Santa Catarina, Brasil

Bruna Carolina de Lima Siqueira dos Santos

Universidade do Vale do Itajaí, Brasil

Bruno Rafael Silva Nogueira Barbosa

Universidade Federal da Paraíba, Brasil

Caio Cesar Portella Santos

Instituto Municipal de Ensino Superior de São Manuel, Brasil

Carla Wanessa do Amaral Caffagni

Universidade de São Paulo, Brasil

Carlos Adriano Martins

Universidade Cruzeiro do Sul, Brasil

Carlos Jordan Lapa Alves

Universidade Estadual do Norte Fluminense Darcy Ribeiro, Brasil

Caroline Chioquetta Lorenset

Universidade Federal de Santa Catarina, Brasil

Cassia Cordeiro Furtado

Universidade Federal do Maranhão, Brasil

Cássio Michel dos Santos Camargo

Universidade Federal do Rio Grande do Sul, Brasil

Cecilia Machado Henriques

Universidade Federal de Santa Catarina, Brasil

Christiano Martino Otero Avila

Universidade Federal de Pelotas, Brasil

Cláudia Samuel Kessler

Universidade Federal do Rio Grande do Sul, Brasil

Cristiana Barcelos da Silva

Universidade do Estado de Minas Gerais, Brasil

Cristiane Silva Fontes

Universidade Federal de Minas Gerais, Brasil

Daniela Susana Segre Guertzenstein

Universidade de São Paulo, Brasil

Daniele Cristine Rodrigues

Universidade de São Paulo, Brasil

Dayse Centurion da Silva

Universidade Anhanguera, Brasil

Dayse Sampaio Lopes Borges

Universidade Estadual do Norte Fluminense Darcy Ribeiro, Brasil

Deilson do Carmo Trindade

Instituto Federal de Educação, Ciência e Tecnologia do Amazonas, Brasil

Diego Pizarro

Instituto Federal de Brasília, Brasil

Dorama de Miranda Carvalho

Escola Superior de Propaganda e Marketing, Brasil

Edilson de Araújo dos Santos

Universidade de São Paulo, Brasil

Edson da Silva

Universidade Federal dos Vales do Jequitinhonha e Mucuri, Brasil

Elena Maria Mallmann

Universidade Federal de Santa Maria, Brasil

Eleonora das Neves Simões

Universidade Federal do Rio Grande do Sul, Brasil

Eliane Silva Souza

Universidade do Estado da Bahia, Brasil

Elvira Rodrigues de Santana

Universidade Federal da Bahia, Brasil

Estevão Schultz Campos

Centro Universitário Adventista de São Paulo, Brasil

Éverly Pegoraro

Universidade Federal do Rio de Janeiro, Brasil

Fábio Santos de Andrade

Universidade Federal de Mato Grosso, Brasil

Fabrcia Lopes Pinheiro

Universidade Federal do Estado do Rio de Janeiro, Brasil

Fauston Negreiros

Universidade de Brasília, Brasil

Felipe Henrique Monteiro Oliveira

Universidade Federal da Bahia, Brasil

Fernando Vieira da Cruz

Universidade Estadual de Campinas, Brasil

Flávia Fernanda Santos Silva

Universidade Federal do Amazonas, Brasil

Gabriela Moysés Pereira

Universidade Federal do Rio de Janeiro, Brasil

Gabriella Eldereti Machado

Universidade Federal de Santa Maria, Brasil

Germano Ehler Pollnow

Universidade Federal de Pelotas, Brasil

Geuciane Felipe Guerim Fernandes

Universidade Federal do Pará, Brasil

Geymeesson Brito da Silva

Universidade Federal de Pernambuco, Brasil

Giovanna Ofretorio de Oliveira Martin Franchi

Universidade Federal de Santa Catarina, Brasil

Handherson Leylton Costa Damasceno

Universidade Federal da Bahia, Brasil

Hebert Elias Lobo Sosa

Universidad de Los Andes, Venezuela

Helciclever Barros da Silva Sales

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Brasil

Helena Azevedo Paulo de Almeida

Universidade Federal de Ouro Preto, Brasil

Hendy Barbosa Santos

Faculdade de Artes do Paraná, Brasil

Humberto Costa

Universidade Federal do Paraná, Brasil

Igor Alexandre Barcelos Graciano Borges

Universidade de Brasília, Brasil

Inara Antunes Vieira Willerding

Universidade Federal de Santa Catarina, Brasil

Jaziel Vasconcelos Dorneles

Universidade de Coimbra, Portugal

Jean Carlos Gonçalves

Universidade Federal do Paraná, Brasil

Joao Adalberto Campato Junior

Universidade Brasil, Brasil

Jocimara Rodrigues de Sousa

Universidade de São Paulo, Brasil

Joelson Alves Onofre

Universidade Estadual de Santa Cruz, Brasil

Jónata Ferreira de Moura

Universidade São Francisco, Brasil

Jonathan Machado Domingues

Universidade Federal de São Paulo, Brasil

Jorge Eschriqui Vieira Pinto

Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Jorge Luís de Oliveira Pinto Filho

Universidade Federal do Rio Grande do Norte, Brasil

Juliana de Oliveira Vicentini

Universidade de São Paulo, Brasil

Juliano Milton Kruger

Instituto Federal do Amazonas, Brasil

Julianno Pizzano Ayoub

Universidade Estadual de Ponta Grossa, Brasil

Julierme Sebastião Morais Souza

Universidade Federal de Uberlândia, Brasil

Junior César Ferreira de Castro

Universidade de Brasília, Brasil

Katia Bruginski Mulik

Universidade de São Paulo, Brasil

Laionel Vieira da Silva

Universidade Federal da Paraíba, Brasil

Lauro Sérgio Machado Pereira

Instituto Federal do Norte de Minas Gerais, Brasil

Leonardo Freire Marino

Universidade do Estado do Rio de Janeiro, Brasil

Leonardo Pinheiro Mozdzenski

Universidade Federal de Pernambuco, Brasil

Letícia Cristina Alcântara Rodrigues

Faculdade de Artes do Paraná, Brasil

Lucila Romano Tragtenberg

Pontifícia Universidade Católica de São Paulo, Brasil

Lucimara Rett

Universidade Metodista de São Paulo, Brasil

Luiz Eduardo Neves dos Santos

Universidade Federal do Maranhão, Brasil

Maikel Pons Giralt

Universidade de Santa Cruz do Sul, Brasil

Manoel Augusto Polastrelí Barbosa

Universidade Federal do Espírito Santo, Brasil

Marcelo Nicomedes dos Reis Silva Filho

Universidade Estadual do Oeste do Paraná, Brasil

Márcia Alves da Silva

Universidade Federal de Pelotas, Brasil

Marcio Bernardino Sirino

Universidade Federal do Estado do Rio de Janeiro, Brasil

Marcos Pereira dos Santos

Universidad Internacional Iberoamericana del Mexico, México

Marcos Uzel Pereira da Silva

Universidade Federal da Bahia, Brasil

Marcus Fernando da Silva Praxedes

Universidade Federal do Recôncavo da Bahia, Brasil

Maria Aparecida da Silva Santandel

Universidade Federal de Mato Grosso do Sul, Brasil

Maria Cristina Giorgi

Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Brasil

Maria Edith Maroca de Avelar

Universidade Federal do Ouro Preto, Brasil

Marina Bezerra da Silva

Instituto Federal do Piauí, Brasil

Marines Rute de Oliveira

Universidade Estadual do Oeste do Paraná, Brasil

Maurício José de Souza Neto

Universidade Federal da Bahia, Brasil

Mauricio José de Souza Neto

Universidade Federal da Bahia, Brasil

Michele Marcelo Silva Bortolai

Universidade de São Paulo, Brasil

Mônica Tavares Orsini

Universidade Federal do Rio de Janeiro, Brasil

Nara Oliveira Salles

Universidade do Estado do Rio de Janeiro, Brasil

Neide Araujo Castilho Teno

Universidade Estadual de Mato Grosso do Sul, Brasil

Neli Maria Mengalli

Pontifícia Universidade Católica de São Paulo, Brasil

Patrícia Biegling

Universidade de São Paulo, Brasil

Patrícia Flavia Mota

Universidade do Estado do Rio de Janeiro, Brasil

Patrícia Helena dos Santos Carneiro

Universidade Federal de Rondônia, Brasil

Rainei Rodrigues Jadejiski

Universidade Federal do Espírito Santo, Brasil

Raul Inácio Busarello

Universidade Federal de Santa Catarina, Brasil

Raymundo Carlos Machado Ferreira Filho

Universidade Federal do Rio Grande do Sul, Brasil

Ricardo Luiz de Bittencourt

Universidade do Extremo Sul Catarinense, Brasil

Roberta Rodrigues Ponciano

Universidade Federal de Uberlândia, Brasil

Robson Teles Gomes

Universidade Católica de Pernambuco, Brasil

Rodiney Marcelo Braga dos Santos

Universidade Federal de Roraima, Brasil

Rodrigo Amancio de Assis

Universidade Federal de Mato Grosso, Brasil

Rodrigo Sarruge Molina

Universidade Federal do Espírito Santo, Brasil

Rogério Rauber

Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Rosane de Fatima Antunes Obregon

Universidade Federal do Maranhão, Brasil

Samuel André Pompeo

Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Sebastião Silva Soares

Universidade Federal do Tocantins, Brasil

Silmar José Spinardi Franchi

Universidade Federal de Santa Catarina, Brasil

Simone Alves de Carvalho

Universidade de São Paulo, Brasil

Simoni Urnau Bonfiglio

Universidade Federal da Paraíba, Brasil

Stela Maris Vaucher Farias

Universidade Federal do Rio Grande do Sul, Brasil

Tadeu João Ribeiro Baptista

Universidade Federal do Rio Grande do Norte, Brasil

Taiane Aparecida Ribeiro Nepomoceno

Universidade Estadual do Oeste do Paraná, Brasil

Taíza da Silva Gama

Universidade de São Paulo, Brasil

Tania Micheline Miorando

Universidade Federal de Santa Maria, Brasil

Tarcísio Vanzin

Universidade Federal de Santa Catarina, Brasil

Tascieli Feltrin

Universidade Federal de Santa Maria, Brasil

Tatiana da Costa Jansen

Serviço Nacional de Aprendizagem Comercial, Brasil

Tayson Ribeiro Teles

Universidade Federal do Acre, Brasil

Thiago Barbosa Soares

Universidade Federal do Tocantins, Brasil

Thiago Camargo Iwamoto

Universidade Estadual de Goiás, Brasil

Thiago Medeiros Barros

Universidade Federal do Rio Grande do Norte, Brasil

Tiago Mendes de Oliveira

Universidade Federal do Rio Grande do Norte, Brasil

Vanessa de Sales Marruche

Universidade Federal do Amazonas, Brasil

Vanessa Elisabete Raue Rodrigues

Universidade Estadual do Centro Oeste, Brasil

Vania Ribas Ulbricht

Universidade Federal de Santa Catarina, Brasil

Vinicius da Silva Freitas
Centro Universitário Vale do Cricaré, Brasil

Wellington Furtado Ramos
Universidade Federal de Mato Grosso do Sul, Brasil

Wellton da Silva de Fatima
Instituto Federal de Alagoas, Brasil

Wenis Vargas de Carvalho
Universidade Federal da Grande Dourados, Brasil

Yan Masetto Nicolai
Universidade Federal de São Carlos, Brasil

PARECERISTAS E REVISORES(AS) POR PARES

Avaliadores e avaliadoras Ad-Hoc

Alcidinei Dias Alves
Logos University International, Estados Unidos

Alessandra Figueiró Thornton
Universidade Luterana do Brasil, Brasil

Alexandre João Appio
Universidade do Vale do Rio dos Sinos, Brasil

Artur Pires de Camargos Júnior
Universidade do Vale do Sapucaí, Brasil

Bianka de Abreu Severo
Universidade Federal de Santa Maria, Brasil

Carlos Eduardo B. Alves
Universidade Federal do Agreste de Pernambuco, Brasil

Carlos Eduardo Damian Leite
Universidade de São Paulo, Brasil

Catarina Prestes de Carvalho
Instituto Federal Sul-Rio-Grandense, Brasil

Davi Fernandes Costa
Secretaria da Educação do Estado de São Paulo, Brasil

Denilson Marques dos Santos
Universidade do Estado do Pará, Brasil

Domingos Aparecido dos Reis
Must University, Estados Unidos

Edson Vieira da Silva de Camargos
Logos University International, Estados Unidos

Edwins de Moura Ramires
Serviço Nacional de Aprendizagem Comercial, Brasil

Elisiene Borges Leal
Universidade Federal do Piauí, Brasil

Elizabeth de Paula Pacheco
Universidade Federal de Uberlândia, Brasil

Elton Simomukay
Universidade Estadual de Ponta Grossa, Brasil

Francisco Geová Goveia Silva Júnior
Universidade Potiguar, Brasil

Indiamaris Pereira
Universidade do Vale do Itajaí, Brasil

Jacqueline de Castro Rimá
Universidade Federal da Paraíba, Brasil

Jonas Lacchini
Pontifícia Universidade Católica de Minas Gerais, Brasil

Lucimar Romeu Fernandes
Instituto Politécnico de Bragança, Brasil

Marcos de Souza Machado
Universidade Federal da Bahia, Brasil

Michele de Oliveira Sampaio
Universidade Federal do Espírito Santo, Brasil

Nívea Consuêlo Carvalho dos Santos
Serviço Nacional de Aprendizagem Comercial, Brasil

Pedro Augusto Paula do Carmo
Universidade Paulista, Brasil

Rayner do Nascimento Souza
Serviço Nacional de Aprendizagem Comercial, Brasil

Samara Castro da Silva
Universidade de Caxias do Sul, Brasil

Sidney Pereira Da Silva
Stockholm University, Suécia

Suélen Rodrigues de Freitas Costa
Universidade Federal do Espírito Santo, Brasil

Thais Karina Souza do Nascimento
Instituto de Ciências das Artes, Brasil

Viviane Gil da Silva Oliveira
Universidade Federal do Amazonas, Brasil

Walmir Fernandes Pereira
Miami University of Science and Technology, Estados Unidos

Weyber Rodrigues de Souza
Pontifícia Universidade Católica de Goiás, Brasil

William Roslindo Paranhos
Universidade Federal de Santa Catarina, Brasil

Parecer e revisão por pares

Os textos que compõem esta obra foram submetidos para avaliação do Conselho Editorial da Pimenta Cultural, bem como revisados por pares, sendo indicados para a publicação.

AGRADECIMENTO ESPECIAL

A Comissão Organizadora do VI Congresso Sertanejo de Computação (SERCOMP) agradece imensamente ao time da plataforma eSUSFeedback pelo apoio e parceria na publicação dos Anais do evento.

Através dessa colaboração, conseguimos garantir um registro técnico e científico de qualidade, acessível e organizado, valorizando o esforço de todos os autores, avaliadores e participantes do SERCOMP.

Seguimos juntos na missão de fortalecer a produção acadêmica e a inovação no Sertão!



SUMÁRIO

CAPÍTULO 1

Luanny Kelly de Almeida Leitão

Eduardo Paulino Dantas

Pedro Henrique Alves Ireneu

Jucelio Soares dos Santos

Mikaelle Oliveira Santos Gomes

Contribuições do PROGNITIVE para o Desenvolvimento de Habilidades de Programação e Robótica no Ensino Médio13

CAPÍTULO 2

Romualdo Júnior Freitas Leite

Jucelio Soares dos Santos

Análise Comparativa de Modelos de Classificação em Processamento de Linguagem Natural para Detecção de Spam em Mensagens de Texto26

CAPÍTULO 3

Roger Natan Carvalho Silva

Aldo Candeia de Albuquerque Júnior

Kaio Vinicius Cordeiro Batista

Rosangela Araújo de Medeiros

Uma Análise sobre Habilidades necessárias para Estágio Supervisionado na área da Computação40





CAPÍTULO 4

Anniely Mariah Soares de Medeiros

Roger Natan Carvalho Silva

José Samuel Pinho Clementino

Luiz Felipe Vieira da Silva

Rosângela de Araújo Medeiros

**Experimentações Didáticas
de Docentes de Computação:**

Relatos de Docência Digital

com Metodologias Ativas em Tempos Cíbridos53

CAPÍTULO 5

Jhulia Medeiros de Queiroz

Carlos Vinicius Nunes Rodrigues

Jannayna Domingues Barros Filgueira

**Desenvolvimento e Aplicação de Jogos
Educativos no Ensino de Estudantes
com Síndrome de Down**.....70

CAPÍTULO 6

Yuri dos Santos Monteiro

Geam Carlos de Araújo Filgueira

Jannayna Domingues Barros Filgueira

**Avaliação da eficácia de uma Árvore
de Decisão na detecção
de doenças cardiovasculares**..... 81

CAPÍTULO 7

Yuri dos Santos Monteiro

Geam Carlos de Araújo Filgueira

Jannayna Domingues Barros Filgueira

**Análise da performance de um modelo
de Floresta Aleatória na Identificação
de Doenças Cardiovasculares**.....94



CAPÍTULO 8

Aldo Candeia de Albuquerque Júnior

Kaio Vinicius Cordeiro Batista

Roger Natan Carvalho Silva

Rosangela de Araújo Medeiros

Um Estudo de Caso sobre a Influência do uso do ChatGPT no desempenho de Alunos e Egressos do Curso de Ciência da Computação da UEPB - Campus VII 107

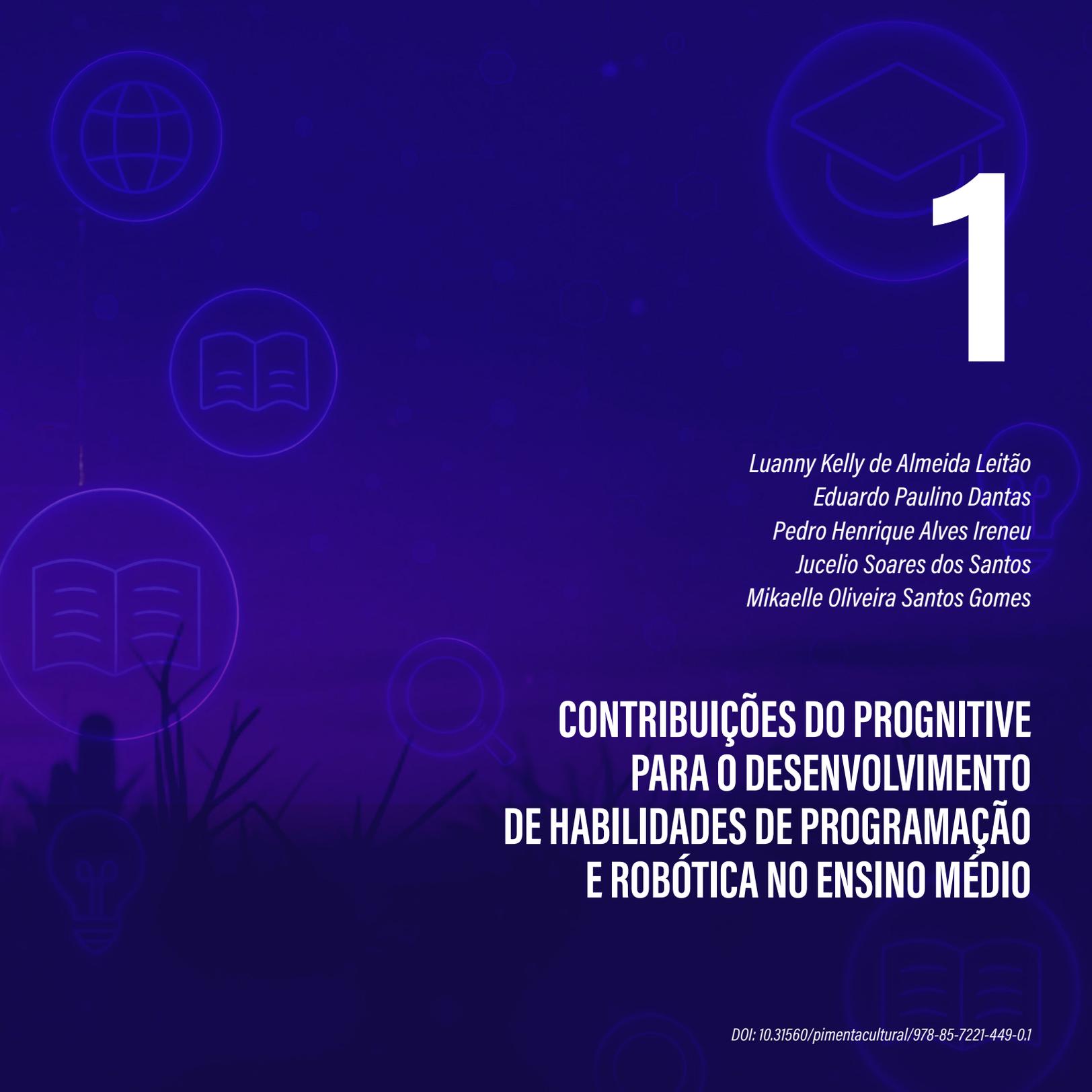
CAPÍTULO 9

Harlem Alves do Nascimento

Demetrio Gomes Mestre

Análise Comparativa de Desempenho entre Arquiteturas Monolíticas e de Microsserviços em uma Aplicação Web 122

Índice remissivo 140



1

Luanny Kelly de Almeida Leitão

Eduardo Paulino Dantas

Pedro Henrique Alves Ireneu

Jucelio Soares dos Santos

Mikaelle Oliveira Santos Gomes

CONTRIBUIÇÕES DO PROGNITIVE PARA O DESENVOLVIMENTO DE HABILIDADES DE PROGRAMAÇÃO E ROBÓTICA NO ENSINO MÉDIO

AUTORES

Luanny Kelly de Almeida Leitão

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: luanny.leitao@aluno.uepb.edu.br

Eduardo Paulino Dantas

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: eduardo.paulino.dantas@aluno.uepb.edu.br

Pedro Henrique Alves Iremeu

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: pedro.irineu@aluno.uepb.edu.br

Jucelio Soares dos Santos

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: jucelio@servidor.uepb.edu.br

Mikaelle Oliveira Santos Gomes

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: mikaelleoliveira@servidor.uepb.edu.br



ABSTRACT:

This paper analyzes the contributions of the PROGNITIVE project to developing programming and robotics skills among high school students. The project adopts a practical and personalized approach, combining theoretical lessons, hands-on activities, and individualized support while encouraging participation in competitions such as the First Tech Challenge (FTC). The results reveal significant improvements in technical skills, such as programming and logical reasoning, as well as interpersonal skills, including teamwork and leadership. Additionally, the project fosters interest in STEAM fields and strengthens the local technological ecosystem. This study highlights the positive impact of innovative educational initiatives in preparing young people for 21st-century challenges, establishing PROGNITIVE as a replicable model for technology education.



RESUMO:

Este artigo examina as contribuições do projeto PROGNITIVE no desenvolvimento de habilidades de programação e robótica entre estudantes do ensino médio. O projeto utiliza uma abordagem prática e personalizada, integrando aulas teóricas, atividades práticas e suporte individualizado, além de promover a participação em competições como a First Tech Challenge (FTC). Os resultados demonstraram melhorias significativas nas habilidades técnicas, como programação e raciocínio lógico, e no desenvolvimento de competências interpessoais, como trabalho em equipe e liderança. Além disso, o projeto incentivou o interesse por áreas STEAM e contribuiu para o fortalecimento do ecossistema tecnológico local. Este estudo destaca o impacto positivo de iniciativas educacionais inovadoras na preparação de jovens para os desafios do século XXI, consolidando o PROGNITIVE como um modelo replicável para o ensino de tecnologia.



INTRODUÇÃO

O avanço tecnológico nas últimas décadas transformou profundamente a sociedade, tornando habilidades de programação e robótica cada vez mais essenciais em diversos setores. No contexto educacional, essas competências desempenham um papel estratégico na formação de estudantes, especialmente em áreas ligadas a ciência, tecnologia, engenharia e matemática (STEAM) (McGee, 2020). Preparar os jovens para os desafios do século XXI exige iniciativas educacionais inovadoras que integrem teoria e prática, promovendo um aprendizado significativo e alinhado às demandas contemporâneas (Manning *et al.*, 2020).

Nesse cenário, o projeto “PROGNITIVE: Promovendo Habilidades Cognitivas de Programação para Alunos do Ensino Médio, Técnico e Profissionalizante da Região de Patos-PB” foi desenvolvido com o objetivo de capacitar estudantes do ensino médio na área de programação e robótica. Por meio de uma abordagem prática e personalizada, o projeto busca não apenas ensinar conceitos técnicos, mas também despertar o interesse dos participantes por STEAM, promover o pensamento crítico e fomentar habilidades interpessoais e empreendedoras.

O projeto integra aulas teóricas, atividades práticas, suporte individualizado e participação em competições como a First Tech Challenge (FTC), proporcionando aos alunos oportunidades de aplicar o conhecimento em contextos desafiadores. Além de fortalecer a base técnica dos participantes, o PROGNITIVE visa criar um impacto positivo na comunidade, promovendo o desenvolvimento tecnológico local e preparando uma nova geração de jovens para o mercado de trabalho e para a inovação.



Este artigo analisa as contribuições do projeto, abordando sua metodologia, os resultados alcançados e as implicações para o desenvolvimento educacional e social. A análise demonstra como o projeto se destaca como uma iniciativa transformadora, alinhada às necessidades de uma sociedade cada vez mais dependente de tecnologias avançadas.

METODOLOGIA

Aplicamos o PROGNITIVE na escola Dionízio Marques de Almeida (DMA), localizado na cidade de Patos—PB. O principal objetivo do projeto foi promover o aprendizado de programação e robótica entre os estudantes do ensino médio. A metodologia adotada pelo projeto envolveu uma abordagem prática e personalizada, baseada em várias estratégias educacionais inovadoras.

Inicialmente, oferecemos aulas de programação em Java com Orientação a Objetos, integradas a conceitos de robótica (Fig. 1). Essas aulas tinham como objetivo desenvolver habilidades sólidas de programação nos estudantes, preparando-os para enfrentar desafios práticos e teóricos nesse campo em constante evolução. Além das aulas teóricas, foram realizadas atividades práticas que permitiram aos estudantes aplicar os conceitos aprendidos na prática, utilizando kits de robótica e simuladores de programação (Fig. 2).



Figura 1 – Aulas de programação em Java com Orientação a Objetos, integradas a conceitos de robótica



Fonte: Elaborado pelos autores, 2024.

Figura 2 – Atividades práticas, utilizando kits de robótica e simuladores de programação



Fonte: Elaborado pelos autores, 2024.



Propusemos exercícios para consolidar o conteúdo ensinado, promovendo a aprendizagem ativa e contribuindo para o desenvolvimento do raciocínio lógico e das habilidades algorítmicas dos estudantes. Essa abordagem permitiu uma melhor absorção dos conceitos apresentados e uma aplicação mais prática. Além disso, ferramentas digitais enriqueceram o processo de ensino-aprendizagem, fornecendo recursos interativos, tutoriais online e ambientes virtuais de aprendizado.

O projeto ofereceu suporte individualizado aos estudantes, auxiliando na resolução de dúvidas e na realização dos exercícios propostos. Esse acompanhamento personalizado garantiu uma aprendizagem mais eficiente, adaptada às necessidades de cada aluno. Além disso, foram promovidos debates em grupo, estudos de caso e desafios práticos relacionados à programação e robótica, incentivando a troca de ideias, o trabalho em equipe e o desenvolvimento de habilidades interpessoais.

Uma das ferramentas exploradas foi o Road Runner, que transforma a arena em um plano cartesiano, permitindo melhor movimentação dos robôs na etapa autônoma (Husbands, 2021). Essa abordagem possibilitou que os estudantes aprimorassem suas habilidades de programação e controle dos robôs enquanto enfrentavam desafios práticos e estimulantes. Ademais, o projeto visou preparar os estudantes para participar da FTC, uma competição internacional de robótica educacional.

A metodologia adotada pelo projeto Y baseou-se em princípios contemporâneos de ensino-aprendizagem (Ortiz-Revilla, J.; Adúriz-Bravo, A.; Greca, I. M., 2020), como a aprendizagem significativa (Bryce & Blown, 2023), a personalização do ensino, a aprendizagem ativa e o construtivismo (Vygotsky & Cole, 2018). Essa abordagem permitiu não apenas o desenvolvimento de habilidades técnicas em programação e robótica, mas também o fortalecimento de habilidades interpessoais e o estímulo ao interesse dos estudantes em STEAM.

RESULTADOS E DISCUSSÕES

O PROGNITIVE demonstrou resultados significativos no desenvolvimento de habilidades técnicas e interpessoais dos estudantes do ensino médio participantes, consolidando-se como uma iniciativa educacional inovadora voltada para o ensino de programação e robótica. Os resultados obtidos refletem o impacto positivo de estratégias pedagógicas fundamentadas na prática, no acompanhamento individualizado e no estímulo à aplicação dos conhecimentos adquiridos em contextos desafiadores.

Uma das principais conquistas do projeto foi o aumento do interesse dos estudantes em áreas STEAM. Esse crescimento foi evidenciado pela maior frequência nas atividades e pelo engajamento contínuo em explorar conteúdos além do escopo das aulas regulares. O impacto do projeto vai além da motivação imediata, incentivando os participantes a considerar carreiras em STEAM e preparando-os para um futuro acadêmico e profissional em áreas de alta demanda tecnológica.

Do ponto de vista técnico, os alunos apresentaram melhorias substanciais em habilidades de programação e raciocínio lógico. Por meio de avaliações sistemáticas, observou-se um progresso significativo na capacidade dos participantes de aplicar conceitos como orientação a objetos, estruturas de dados e lógica de algoritmos na resolução de problemas. Esse avanço foi consolidado pela abordagem prática adotada pelo projeto, que incluiu exercícios direcionados, desafios baseados em problemas reais e o uso de ferramentas como o Road Runner, utilizado para programar e controlar robôs em uma arena cartesiana. A integração entre teoria e prática mostrou-se eficiente no fortalecimento da base técnica dos alunos.



A participação em competições de robótica foi outro destaque do projeto. A equipe Legonautas, formada por estudantes do projeto, obteve o título de Vice-Campeã na Arena e recebeu o prêmio de Aliança Capitã Finalista na temporada 2023/2024 da FTC (Fig. 3 e 4). Esses resultados atestam a capacidade dos participantes de traduzir o aprendizado em desempenho competitivo, mostrando domínio técnico, criatividade e trabalho em equipe. Além disso, essas competições serviram como uma plataforma para a aplicação prática dos conhecimentos e para o desenvolvimento de habilidades interpessoais, como comunicação eficaz e colaboração.

Figura 3 – Participação dos alunos atendidos pelo projeto na Competição Nacional da FTC em Brasília



Fonte: Elaborado pelos autores, 2024.

O impacto do projeto também se estendeu à comunidade local. A capacitação de jovens em áreas como programação e robótica contribuiu para o fortalecimento do ecossistema tecnológico na região de Patos-PB, promovendo inovação e preparando uma nova



geração de profissionais. A integração entre a escola, a comunidade e o projeto foi reforçada por meio do envolvimento de famílias e outros atores locais, criando um ambiente de aprendizado colaborativo e fomentando uma educação mais conectada com as demandas reais do mercado de trabalho e da sociedade.

Figura 4 - Vice-Campeã na Arena e recebeu o prêmio de Aliança Capitã Finalista na temporada 2023/2024 da FTC



Fonte: Elaborado pelos autores, 2024.

Outro ponto relevante foi o estímulo ao desenvolvimento de habilidades interpessoais e empreendedoras. As atividades práticas, realizadas em equipe, proporcionaram oportunidades para que os participantes exercessem papéis de liderança e solucionassem problemas de forma criativa e eficiente. Esse conjunto de experiências promoveu a preparação dos estudantes para desafios futuros, tanto no ambiente acadêmico quanto no profissional, incluindo áreas como gestão de projetos e inovação tecnológica.



Por fim, o projeto destacou-se como uma iniciativa que promoveu não apenas competências técnicas, mas também o pensamento crítico e criativo dos alunos. Ao enfrentar problemas complexos e buscar soluções inovadoras, os participantes foram desafiados a desenvolver habilidades essenciais para o século XXI. Esses resultados reforçam a relevância do PROGNITIVE como um modelo educacional que combina excelência técnica com impacto social, contribuindo para a formação de cidadãos preparados para os desafios tecnológicos e socioeconômicos do futuro.

CONSIDERAÇÕES FINAIS

O PROGNITIVE demonstrou-se eficaz na promoção do ensino de programação e robótica, alcançando resultados significativos no desenvolvimento de competências técnicas e interpessoais dos estudantes do ensino médio. Destacaram-se o aumento do interesse por áreas STEAM, a melhoria nas habilidades de programação e raciocínio lógico, e o desempenho competitivo na FTC.

Além dos impactos educacionais, o projeto contribuiu para o fortalecimento do ecossistema tecnológico local, promovendo inovação e capacitando jovens para atuar em áreas estratégicas. Também incentivou habilidades interpessoais e empreendedoras, essenciais para os desafios do mercado de trabalho.

Os resultados alcançados reforçam o PROGNITIVE como uma iniciativa transformadora e replicável, com potencial para inspirar projetos educacionais similares e contribuir para o desenvolvimento tecnológico e social da comunidade atendida.

REFERÊNCIAS

BACICH, Lilian; HOLANDA, Leandro. **STEAM em sala de aula: a aprendizagem baseada em projetos integrando conhecimentos na educação básica**. Penso Editora, 2020.

BRYCE, T. G. K.; BLOWN, E. J. Ausubel's meaningful learning re-visited. **Current Psychology**, p. 1-20, 2023.

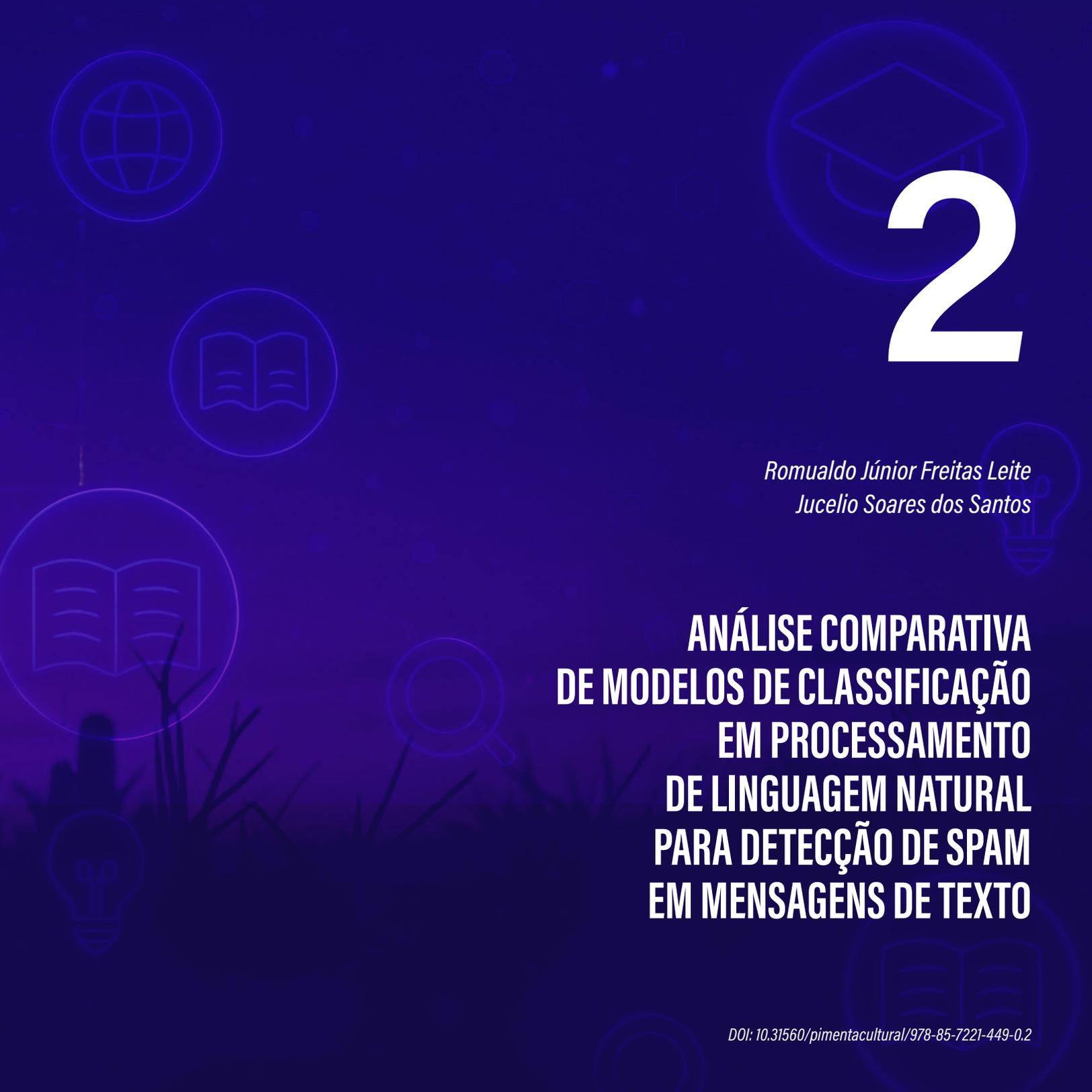
HUSBANDS, Phil. **Robots: What Everyone Needs to Know**. Oxford University Press, 2021.

MANNING, L.; SMITH, R.; CONLEY, G.; HALSEY, L. Ecopreneurial Education and Support: Developing the Innovators of Today and Tomorrow. **Sustainability**. v. 12, n. 21, p. 9228, 2020.

MCGEE, Ebony Omotola. Interrogating structural racism in STEAM higher education. **Educational Researcher**, v. 49, n. 9, p. 633-644, 2020.

ORTIZ-REVILLA, J.; ADÚRIZ-BRAVO, A.; GRECA, I. M. A framework for epistemological discussion on integrated STEAM education. **Science & Education**, v. 29, n. 4, p. 857-880, 2020.

VYGOTSKY, L.,; COLE, M. Lev Vygotsky: Learning and social constructivism. **Learning Theories for Early Years Practice**. UK: SAGE Publications Inc, p. 68-73, 2018.



2

*Romualdo Júnior Freitas Leite
Jucelio Soares dos Santos*

ANÁLISE COMPARATIVA DE MODELOS DE CLASSIFICAÇÃO EM PROCESSAMENTO DE LINGUAGEM NATURAL PARA DETECÇÃO DE SPAM EM MENSAGENS DE TEXTO

AUTORES

Romualdo Júnior Freitas Leite

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: romualdo.leite@aluno.uepb.edu.br

Jucelio Soares dos Santos

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: jucelio@servidor.uepb.edu.br



ABSTRACT:

Spam detection is vital for communication security and efficiency, given the rise of mobile and messaging services like email. Spammers increasingly target these platforms, posing risks such as phishing and scams. Excessive spam overwhelms inboxes, hindering the identification of legitimate messages. Our research focuses on employing various Natural Language Processing models to detect spam in text messages automatically. We utilized a robust dataset and found a need for balanced approaches due to message distribution disparities. We evaluated multiple algorithms, including K-Nearest Neighbors, Naive Bayes Multinomial, Support Vector Classifier, Logistic Regression, and Decision Tree Classifier. Among these, the Support Vector Classifier outperformed others, striking a balance between precision and recall, effectively identifying spam.



RESUMO:

A detecção de spam é essencial para a segurança e eficiência das comunicações, especialmente com o crescimento de serviços móveis e de mensagens, como e-mail. Spammers têm direcionado cada vez mais suas atividades para essas plataformas, gerando riscos como phishing e golpes. O excesso de spam sobrecarrega caixas de entrada, dificultando a identificação de mensagens legítimas. Nossa pesquisa concentra-se no uso de diversos modelos de Processamento de Linguagem Natural (PLN) para detectar automaticamente spam em mensagens de texto. Utilizamos um conjunto de dados robusto e identificamos a necessidade de abordagens balanceadas devido às disparidades na distribuição das mensagens. Avaliamos vários algoritmos, incluindo K-Nearest Neighbors, Naive Bayes Multinomial, Support Vector Classifier, Regressão Logística e Árvore de Decisão. Entre eles, o Support Vector Classifier apresentou o melhor desempenho, equilibrando precisão e recall, identificando spam de forma eficaz.

INTRODUÇÃO

Nos últimos anos, a revolução digital transformou profundamente a forma como nos comunicamos, remodelando a dinâmica das interações pessoais, profissionais e comerciais. O surgimento e a popularização das comunicações digitais levaram a um aumento exponencial no volume de mensagens, e-mails e interações em plataformas online. Esse crescimento, embora traga benefícios inegáveis, também acarretou desafios expressivos para a eficiência e a segurança das comunicações, tanto no âmbito individual quanto organizacional. Entre esses desafios, o spam se destaca como uma questão crítica e persistente, exigindo abordagens inovadoras e eficazes para mitigação (Kawintiranon *et al.*, 2022; Elbagir e Yang, 2019).

O termo “spam” refere-se a mensagens não solicitadas que chegam frequentemente em grandes volumes, muitas vezes com finalidades comerciais, propagandísticas ou até maliciosas. Essas mensagens invadem caixas de entrada de e-mails, mensageiros instantâneos e redes sociais, causando uma série de impactos negativos. Os efeitos do spam vão além do simples incômodo: ele pode prejudicar a produtividade dos usuários, disseminar informações enganosas ou conteúdos maliciosos, e comprometer a integridade da segurança digital. Do ponto de vista técnico, o spam é frequentemente utilizado como um vetor para ataques cibernéticos, como a distribuição de malware, tentativas de phishing e fraudes financeiras (Koggalahewa *et al.*, 2020).

Com o avanço e a sofisticação das técnicas de spam, os métodos tradicionais de triagem manual tornaram-se ineficazes e insustentáveis. Identificar manualmente mensagens indesejadas não apenas consome tempo e recursos, mas também falha em acompanhar o crescimento do volume de dados gerados. Diante disso, tornou-se essencial adotar soluções baseadas em tecnologia avançada que possam automatizar o processo de identificação e filtragem



de spam. Nesse contexto, o Processamento de Linguagem Natural (PLN) e os algoritmos de classificação surgem como ferramentas fundamentais, permitindo que sistemas computacionais analisem e compreendam o conteúdo das mensagens de maneira mais profunda e precisa (Sen *et al.*, 2020).

O PLN possibilita que as máquinas interpretem o significado, o contexto e as nuances das mensagens, o que é essencial para identificar padrões de spam que muitas vezes passam despercebidos por métodos tradicionais. Além disso, os algoritmos de classificação, treinados com grandes volumes de dados rotulados, permitem distinguir mensagens legítimas de conteúdos indesejados de forma ágil e eficaz. Esses sistemas têm demonstrado sua capacidade não apenas de reduzir o impacto do spam no cotidiano dos usuários, mas também de otimizar fluxos de comunicação e proteger sistemas contra ameaças digitais (Elbagir e Yang, 2019).

No entanto, mesmo com os avanços tecnológicos significativos, os desafios permanecem. O spam evoluiu em termos de complexidade e diversidade, utilizando estratégias sofisticadas para evitar detecção, como camuflagem de conteúdo e a utilização de linguagem ambígua. Assim, há uma demanda crescente por soluções que não apenas acompanhem essas mudanças, mas também as antecipem.

Diante desse cenário, esta pesquisa se propõe a explorar e avaliar diferentes modelos de classificação baseados em PLN, aplicados a um estudo de caso específico voltado para a detecção automática de spam em mensagens de texto. A investigação buscará comparar o desempenho de diversos algoritmos, analisando métricas como precisão, recall e eficiência computacional, a fim de identificar a abordagem mais robusta e eficaz. Ao abordar a crescente sofisticação e diversidade do spam, espera-se contribuir para o avanço do campo de detecção de spam com PLN e oferecer insights práticos para a implementação de soluções escaláveis e eficientes em ambientes reais de comunicação digital.



Além disso, o trabalho pretende discutir as implicações práticas e éticas do uso de modelos automatizados nesse contexto, como a privacidade dos dados, a transparência dos algoritmos e a aplicabilidade em diferentes cenários, desde pequenas empresas até grandes corporações. Dessa forma, esta pesquisa não apenas busca aprimorar a eficiência das comunicações digitais, mas também fortalecer a confiança e a segurança nas interações online, promovendo um ambiente mais produtivo e protegido para usuários de todo o mundo.

METODOLOGIA

A metodologia adotada abrange desde a seleção dos dados até a análise detalhada dos resultados obtidos, seguindo uma abordagem sistemática e rigorosa para garantir a validade e a confiabilidade dos resultados. Este processo foi dividido em etapas claramente definidas, descritas a seguir.

A primeira etapa consistiu na seleção cuidadosa de um conjunto de dados representativo. Optamos por utilizar um dataset publicamente disponível no Kaggle (<https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset>), criado especificamente para classificar mensagens de texto como spam ou não. A escolha deste dataset foi baseada em sua amplitude e diversidade, que refletem de forma precisa o panorama atual das comunicações digitais. A variedade de dados garante que o modelo treinado seja capaz de lidar com diferentes cenários, aumentando a robustez dos resultados.

Após a seleção do conjunto de dados, procedemos com o processamento e a limpeza dos dados. Essa etapa foi fundamental para preparar o conjunto de dados para as análises subsequentes, envolvendo técnicas como a remoção de *stopwords* (palavras comuns que não agregam significado relevante), tokenização



(divisão de textos em unidades menores, como palavras ou frases) e *stemming* (redução das palavras às suas raízes). Essas técnicas foram aplicadas com o objetivo de reduzir a dimensionalidade do conjunto de dados e simplificar sua estrutura, facilitando a identificação de padrões relevantes durante a análise.

Com os dados devidamente processados, passamos à fase de execução da pesquisa. Nesta etapa, aplicamos algoritmos de aprendizado de máquina e técnicas de Processamento de Linguagem Natural (PLN) para classificar as mensagens como spam ou não. Os modelos utilizados foram treinados e avaliados com base em métricas específicas, como precisão, *recall* e F1-score, que fornecem uma medida abrangente de sua eficácia na classificação das mensagens.

Após a aplicação dos algoritmos e a execução dos experimentos, analisamos os resultados obtidos em detalhes. Essa análise incluiu a interpretação das métricas de desempenho dos modelos de classificação e a identificação de padrões e tendências nos dados. Esse passo foi essencial para compreender a eficácia dos modelos em detectar spam e para identificar áreas potenciais de melhoria. Além disso, a análise detalhada permitiu explorar as nuances dos resultados, como o impacto de características específicas das mensagens na classificação.

Finalmente, identificamos e discutimos as principais ameaças à validade dos resultados. Entre os desafios abordados, destacam-se questões como a generalização dos resultados, possíveis vieses nos dados e limitações inerentes aos modelos utilizados. Por exemplo, o dataset utilizado pode não refletir completamente as características de mensagens em outros contextos ou idiomas, o que limita a aplicabilidade universal dos modelos treinados. Reconhecer e discutir essas ameaças é essencial para garantir a confiabilidade e a robustez dos resultados da pesquisa.

Adicionalmente, a metodologia incluiu uma abordagem iterativa, em que ajustes foram feitos ao longo do processo, tanto no



pré-processamento dos dados quanto na seleção e parametrização dos algoritmos. Essa flexibilidade foi fundamental para otimizar o desempenho dos modelos e assegurar que os resultados obtidos fossem relevantes e confiáveis.

Por fim, além de gerar insights sobre a classificação de spam, os resultados desta pesquisa oferecem subsídios para aprimorar modelos futuros, sugerindo possíveis melhorias no pré-processamento de dados e na escolha de algoritmos, assim como recomendações para minimizar os impactos de vieses e limitações. Assim, esta metodologia não apenas atende aos objetivos do estudo, mas também estabelece uma base sólida para pesquisas futuras na área de detecção automática de spam.

RESULTADOS E DISCUSSÕES

A análise exploratória desempenhou um papel fundamental na compreensão das características do conjunto de dados utilizado para a detecção de spam. A análise inicial destacou uma significativa disparidade entre as classes “ham” (não spam) e “spam”, com 4516 mensagens classificadas como “ham” e 653 como “spam”. Essa discrepância reflete um problema comum em aplicações reais, onde a classe majoritária domina o conjunto de dados, dificultando o aprendizado do modelo em relação à classe minoritária.

Essa assimetria apresenta desafios específicos para a modelagem, incluindo o risco de o algoritmo priorizar a classe dominante, resultando em vieses que comprometem a eficácia da detecção de spam. Para lidar com esse problema, foi aplicado um processo de balanceamento de dados, utilizando técnicas como *oversampling* da classe minoritária, *undersampling* da classe majoritária ou uma combinação dessas abordagens. Métodos avançados, como o uso



do Synthetic Minority Oversampling Technique (SMOTE), poderiam ser integrados em futuros experimentos para criar instâncias sintéticas e melhorar ainda mais a representatividade da classe minoritária.

Além disso, foram realizadas análises descritivas detalhadas para explorar as características textuais das mensagens. As mensagens apresentaram, em média, 79 caracteres, 18 palavras e 2 frases, refletindo uma estrutura relativamente curta. No entanto, a presença de valores extremos no comprimento das mensagens indicou uma alta variabilidade no conjunto de dados. Essa variabilidade pode introduzir ruído nos algoritmos de aprendizado de máquina, dificultando a extração de padrões relevantes. Para minimizar esses efeitos, o pré-processamento incluiu normalização do comprimento das mensagens, por meio de técnicas de truncamento ou *padding*, garantindo que todas as entradas apresentassem dimensões uniformes.

O pré-processamento também incluiu etapas como remoção de *stopwords*, tokenização e *stemming*. A remoção de *stopwords* visou eliminar palavras irrelevantes que não contribuem para a detecção de padrões de spam, enquanto a tokenização dividiu o texto em unidades menores (tokens) para facilitar a análise. O *stemming* reduziu palavras às suas raízes, reduzindo a dimensionalidade do conjunto de dados e aprimorando a capacidade dos modelos de identificar padrões relevantes.

Cinco algoritmos de aprendizado de máquina foram implementados para a tarefa de classificação: K-Nearest Neighbors (KNN), Naive Bayes Multinomial, Support Vector Classifier (SVC), Regressão Logística e Árvore de Decisão. Esses algoritmos foram selecionados por sua robustez e por sua aplicação consolidada em problemas de classificação. A avaliação foi conduzida utilizando métricas como *accuracy*, *precision*, *recall* e F1-Score, garantindo uma análise abrangente do desempenho dos modelos. Os resultados obtidos estão apresentados na Tabela 1.



Tabela 1 – Resultados de desempenho dos algoritmos de classificação de *spam*

Algorithms	Accuracy	Precision	Recall	F1-Score
K-Nearest Neighbors	0.894446	1.000000	0.115741	0.207469
Naive Bayes Multinomial	0.953855	1.000000	0.613426	0.760402
Support Vector Classifier	0.969881	0.954930	0.784722	0.861499
Regressão Logística	0.935618	0.890196	0.525463	0.660844
Árvores de Decisão	0.936723	0.755668	0.694444	0.723764

Fonte: Elaborado pelos autores, 2024.

Os resultados apresentados na Tabela 1 indicam o desempenho dos modelos nos critérios estabelecidos. A métrica *accuracy* indica a proporção de previsões corretas em relação ao total de previsões realizadas. O algoritmo SVC apresentou o maior valor de *accuracy* (0.969881), seguido pelo Naive Bayes Multinomial (0.953855). Esses resultados destacam a capacidade desses algoritmos de realizar uma classificação geral precisa no conjunto de dados desbalanceado.

A *precision* mede a proporção de mensagens corretamente classificadas como *spam* em relação ao total de mensagens classificadas como *spam*. O KNN e o Naive Bayes Multinomial alcançaram valores perfeitos (1.000000), indicando que esses algoritmos minimizaram falsos positivos, ou seja, classificaram poucas mensagens “ham” como *spam*.

O *recall*, que indica a proporção de mensagens de *spam* corretamente identificadas pelo modelo em relação ao total de mensagens de *spam*, apresentou variação significativa entre os algoritmos. O SVC obteve o maior valor (0.784722), evidenciando sua eficácia na recuperação de mensagens de *spam*.

O F1-Score combina *precision* e *recall* em uma única métrica para avaliar o desempenho geral. O SVC novamente se destacou, com o maior valor (0.861499), seguido pelo Naive Bayes Multinomial



(0.760402). Esse resultado demonstra que o SVC conseguiu equilibrar bem a precisão e a recuperação, superando os outros modelos na tarefa de classificação.

Os resultados desta pesquisa demonstram que o algoritmo SVC apresentou o melhor desempenho geral, destacando-se como a abordagem mais eficaz para a tarefa de detecção de spam no contexto analisado. Contudo, algumas limitações precisam ser consideradas para a interpretação dos resultados e para a aplicação prática dos modelos desenvolvidos. Uma das principais limitações está relacionada à generalização dos resultados. O conjunto de dados utilizado, embora representativo para o contexto avaliado, pode não refletir a diversidade encontrada em outros cenários, como mensagens provenientes de diferentes idiomas, plataformas ou com variações culturais e regionais. Essa falta de diversidade nos dados pode restringir a aplicabilidade dos modelos treinados, sugerindo a necessidade de validações adicionais com conjuntos de dados mais amplos e heterogêneos.

Outra questão importante é o impacto das técnicas de balanceamento de dados. Embora abordagens como *oversampling* e *undersampling* tenham sido eficazes neste estudo, a escolha de métodos específicos pode influenciar o desempenho dos algoritmos. Métodos mais avançados, como o Synthetic Minority Oversampling Technique (SMOTE) ou estratégias baseadas em aprendizado generativo, poderiam ser explorados para avaliar se há melhorias adicionais na capacidade de generalização dos modelos. Além disso, algumas limitações inerentes aos algoritmos utilizados devem ser destacadas. O K-Nearest Neighbors (KNN), por exemplo, apresentou baixo desempenho em termos de *recall*, indicando que é menos eficaz em cenários onde a recuperação de mensagens de spam é prioritária.

Por fim, outro aspecto a ser considerado é a aplicabilidade prática dos modelos. Apesar do excelente desempenho do SVC em termos de métricas como *accuracy* e F1-Score, o custo computacional



associado a esse algoritmo pode ser elevado em cenários com grandes volumes de dados ou quando a atualização do modelo em tempo real é necessária. Podendo exigir ajustes nos parâmetros ou o uso de algoritmos alternativos em aplicações que demandam maior eficiência computacional.

CONSIDERAÇÕES FINAIS

Este estudo avaliou a aplicação de técnicas de aprendizado de máquina e PLN para a detecção de spam em mensagens de texto. O SVC apresentou o melhor desempenho entre os cinco algoritmos avaliados, destacando-se nas métricas *accuracy*, *recall* e *F1-Score*. O pré-processamento e o balanceamento de dados foram etapas fundamentais para lidar com a assimetria no conjunto de dados e otimizar a análise. Apesar dos resultados positivos, limitações como a generalização para outros contextos e o custo computacional de alguns modelos apontam para a necessidade de pesquisas futuras, incluindo a exploração de redes neurais profundas e conjuntos de dados mais diversificados. Este trabalho fornece uma base sólida para o desenvolvimento de sistemas escaláveis, contribuindo para o avanço da detecção de *spam* e propondo novas direções para pesquisas futuras.

REFERÊNCIAS

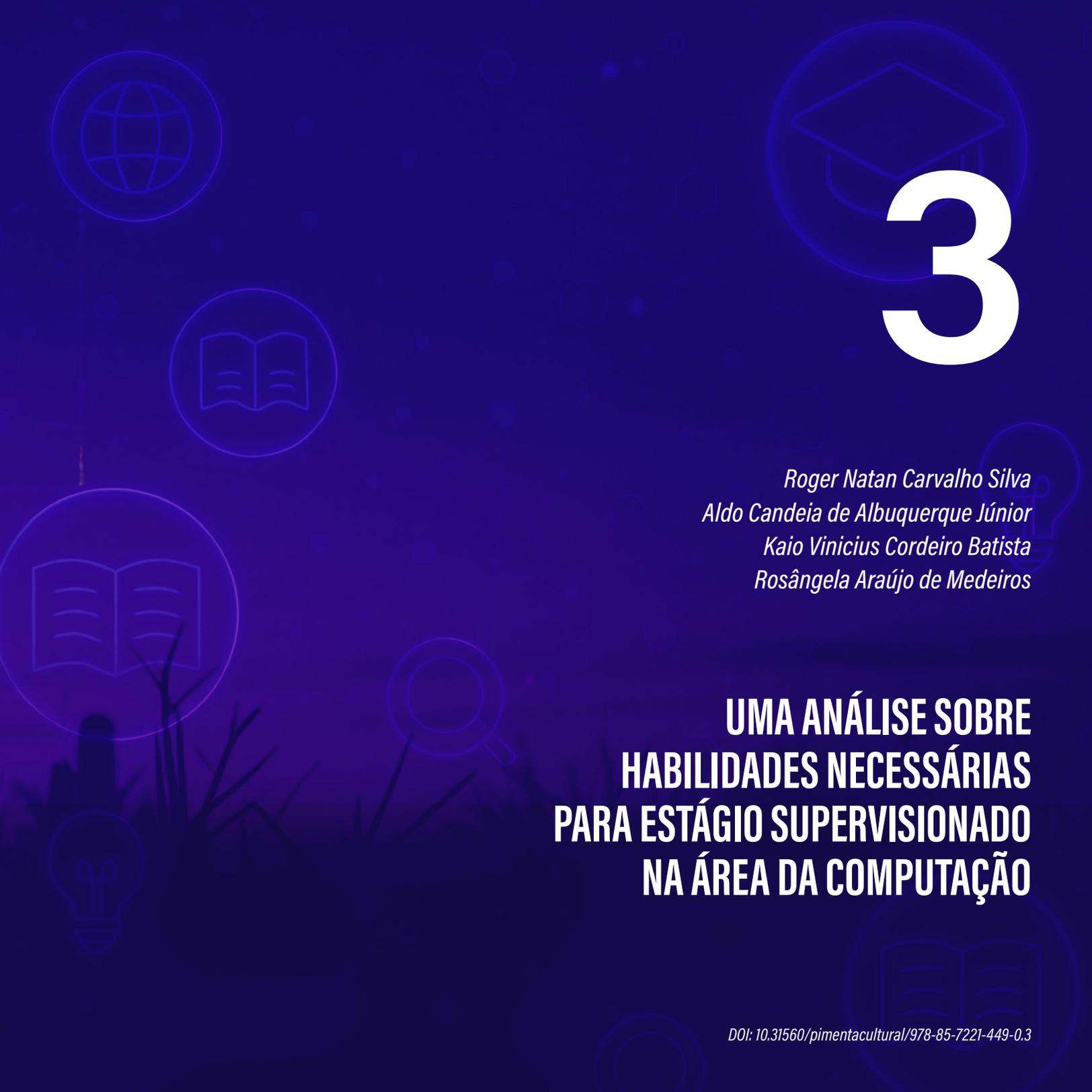
ELBAGIR, S.; YANG, J. Twitter sentiment analysis using natural language toolkit and vader sentiment. *In: Proceedings of the international multiconference of engineers and computer scientists*. [S.l.: s.n.], 2019. v. 122, p. 16.

JUNNARKAR, A. *et al.* E-mail spam classification via machine learning and natural language processing. *In: IEEE. 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*. [S.l.], 2021. p. 693-699.

KAWINTIRANON, K.; SINGH, L.; BUDAK, C. **Traditional and context-specific spam detection in low resource settings**. Machine Learning, Springer, v. 111, n. 7, p. 2515–2536, 2022.

KOGGALAHEWA, D. N.; XU, Y.; FOO, E. Spam detection in social networks based on peer acceptance. *In: Proceedings of the Australasian Computer Science Week Multiconference*. [S.l.: s.n.], 2020. p. 1–7.

SEN, P. C.; HAJRA, M.; GHOSH, M. Supervised classification algorithms in machine learning: A survey and review. *In: Emerging Technology in Modelling and Graphics: Proceedings of IEM Graph 2018*. [S.l.]: SPRINGER, 2020. p. 99–111.



3

*Roger Natan Carvalho Silva
Aldo Candeia de Albuquerque Júnior
Kaio Vinicius Cordeiro Batista
Rosângela Araújo de Medeiros*

UMA ANÁLISE SOBRE HABILIDADES NECESSÁRIAS PARA ESTÁGIO SUPERVISIONADO NA ÁREA DA COMPUTAÇÃO

AUTORES

Roger Natan Carvalho Silva

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: roger.silva@aluno.uepb.edu.br

Aldo Candeia de Albuquerque Júnior

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: roger.silva@aluno.uepb.edu.br

Kaio Vinicius Cordeiro Batista

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: kaio.batista@aluno.uepb.edu.br

Rosângela de Araújo Medeiros

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: rosangelamedeiros@servidor.uepb.edu.br



ABSTRACT:

The main objective of this paper was to give an analysis of opinions from people who work in the field of Computing about the main soft skills and hard skills required by the professional field, relating to supervised internship. For carrying out this research, which is similar to a case study, five (5) representatives from different sectors, such as managers, business people and professors who were previously internship coordinators were interviewed. It was found out that soft skills, related to communication, are more relevant than technical skills.



RESUMO:

O objetivo principal deste artigo foi fazer uma análise das opiniões de pessoas que trabalham na área de Computação sobre as principais soft skills e hard skills exigidas pelo campo profissional, relacionadas ao estágio supervisionado. Para a realização desta pesquisa, que se assemelha a um estudo de caso, foram entrevistados cinco (5) representantes de diferentes setores, como gestores, empresários e professores que já foram coordenadores de estágio. Foi constatado que as soft skills, relacionadas à comunicação, são mais relevantes do que as técnicas.

INTRODUÇÃO

A aplicação da Ciência da Computação (CC) e suas subáreas cresceram exponencialmente em todos setores humanos, alterando as relações sociais e, principalmente, a esfera empresarial. A Computação em Nuvem e Inteligências Artificiais, por exemplo, são alguns campos da Computação que possibilitaram a evolução de diversos processos, seja a partir da automação de tarefas, das inovações no gerenciamento de *softwares*, ou na diminuição dos custos operacionais em empresas, comércios e indústrias (Rinaldi, 2017).

As tecnologias digitais tornaram-se componentes vitais às empresas, participando dos processos de tomadas de decisão e aplicadas em níveis estratégicos (Santi, 2017), além de outras demandas. Nesse sentido, o Brasil ocupa a 12ª posição no *ranking* mundial de investimentos na área, sendo responsável por 36,5% dos investimentos em tecnologia digital da América Latina e ocupando um lugar de liderança dos avanços em desenvolvimento de *softwares* e, conseqüentemente, dos sistemas de informação (ABES, 2023).

Dado o crescimento de mercado dessa área no Brasil e no mundo, torna-se necessária a formação de profissionais qualificados e capazes de lidar com as problemáticas presentes nas conjunturas empresariais. Stevenson e Starkweather (2010) apontaram uma relevância significativa das *soft skills* (aptidões emocionais e comunicativas) - em relação às *hard skills* (qualificações técnicas), mencionadas pelos líderes executivos da área, na análise de currículos dos profissionais para contratação.

Uma das maneiras de habilitar graduandos de Computação para atuar neste contexto durante o percurso acadêmico – predecessor do mercado de trabalho – pode ser a participação em atividades de estágio supervisionado, definido como “um ato educativo dentro do ambiente laboral que visa o aprimoramento das



competências necessárias ao cenário profissional” (Brasil, 2008; Castro e Silva, 2012).

Considerando este panorama, o presente trabalho tem como objetivo apresentar uma análise de opiniões de pessoas que atuam na área de Computação sobre principais *soft skills* e *hard skills* exigidas no estágio supervisionado. Nesta perspectiva, este estudo tem a devida importância, visto que pode nortear estudantes de Computação acerca de requisitos necessários para a realização de estágio supervisionado e a egressos que também devem se adaptar às tendências de mercado.

Este artigo pode ser caracterizado como de natureza primária (Wazlawick, 2021), que se aproxima de um estudo de caso, desenvolvido a partir da realização de entrevistas com docente, donos e representantes de empresas do entorno do campus VII - Patos - da Universidade Estadual da Paraíba (UEPB).

Ao longo do artigo, adotou-se a seguinte estrutura: nesta primeira seção, uma introdução sobre a temática. A segunda parte se destina a apresentar os conceitos e autores relacionados. Na aba seguinte, sobre a metodologia, define-se a tipologia de pesquisa e a coleta de dados. Posteriormente, nos resultados e discussão, expõe-se a análise das informações obtidas. E por fim, têm-se as considerações finais, remetendo todo o percurso do estudo e as contribuições futuras

REFERENCIAL TEÓRICO

O termo *Hard Skills* se refere ao conhecimento diretamente aplicável, ou seja, de caráter técnico. Na área da computação essas habilidades podem ser mencionadas como domínio em outro idioma, conhecimentos em lógica de programação, domínio



em desenvolvimento de *software*, análise de dados, entre outros. Em resumo, o domínio de ação das HS é pragmático (Geithner; Menzel, 2016).

Além dos conhecimentos importantes para a execução de suas funções como profissionais de TI, existem também as chamadas *Soft Skills* (SS). As SS são o conjunto de competências caracterizadas pelas capacidades de comunicação, inteligência emocional, trabalho em equipe, participação social, entre outras (Hendarman; Cantner 2018). Dois podem ser os fatores principais para o desenvolvimento das SS, ao longo do tempo – sob a perspectiva do amadurecimento– e da prática social de cada indivíduo.

Apesar desses conceitos apresentarem em suas morfologias, vocábulos que se referem a suavidade (*soft*) ou dureza (*hard*) dos atributos, não se pode afirmar, por sua vez, que as HS, são mais difíceis de se adquirir, nem considerar as SS como dispensáveis. Uma das principais habilidades necessárias a um gerente de *software* é a capacidade de se comunicar entre as diferentes hierarquias e audiências que compõem o processo de desenvolvimento de aplicações (Zielinski, 2005).

Pedrosa e Silva (2019) examinam a relevância das SS na descrição de vagas para gerentes de projetos de TI no Brasil. Entre suas observações, constatou-se que liderança, comunicação e resolução de problemas foram competências altamente valorizadas e destacadas. Ao longo do trabalho, os autores ainda mencionam que as HS são fundamentais para o desempenho técnico, mas que as SS contribuem para o sucesso dos projetos e das organizações.

Ao ingressar na Universidade os estudantes têm a oportunidade de aprender conhecimentos necessários para atuar e profissionalizar-se no mercado de trabalho. O estágio supervisionado é uma das possibilidades oferecidas aos graduandos durante sua formação.



Dessa maneira, todo conjunto de compreensão teórica adquirida em sala de aula, estudos dirigidos e participação dos alunos nas atividades da Universidade (em projetos de extensão e pesquisa) podem ser trabalhados de maneira prática ao decorrer da prática acompanhada (Castro; Salva, 2012).

Parte da literatura que discute sobre esta prática profissional acompanhada refere-se ao estágio em licenciatura em computação, ao qual é previsto como obrigatório em seus respectivos cursos. Ainda que voltada para a docência, o desenvolvimento do estágio supervisionado pode habilitar competências em comum a outras áreas da computação. Silva e Barbosa *et al.* (2023) destacam entre as competências trabalhadas o aprimoramento das competências com tecnologias, gestão em grupo, capacidade de observação e reflexão e a comunicação.

METODOLOGIA

O presente trabalho caracteriza-se como de natureza primária (Wazlawick, 2021), porque busca dados novos no campo empírico de atuação profissional. A partir da classificação de Wazlawick (2021), tipificada como uma pesquisa de objetivo exploratório-descritivo e, quanto aos procedimentos técnicos, aproxima-se de um estudo de caso, cujo objeto de análise foram opiniões de gestores, empresários e um docente da área, que já foi coordenador de estágio em curso de Computação.

A coleta de dados ocorreu através de entrevistas, com roteiro composto por seis (6) questões semi-estruturadas. Foram realizadas por meio de reuniões síncronas, tanto de modo presencial quanto online (por videoconferência) e outras de forma assíncrona, via



aplicativo de comunicação instantânea. A escolha do grupo investigado foi o contato prévio com estudantes ou docentes do curso de Computação da UEPB - Campus VII, bem como foi considerada a disponibilidade e agenda dos entrevistados.

Para análise das respostas, foram transcritas as opiniões dos entrevistados e extraídos dados relevantes ao escopo da pesquisa. Utilizou-se uma abordagem quanti-qualitativa (Wainer, 2007), na qual elementos semelhantes foram agrupados para discussão dos resultados, considerando *soft* e *hard skills* necessárias ao estagiário.

Importante mencionar que a codificação para identificar de forma anônima os diferentes entrevistados foi feita a partir da sigla E (para entrevistado), seguida de um número, conforme pode ser visualizado no decorrer da seção que contempla resultados e discussão.

RESULTADOS E DISCUSSÕES

A primeira questão do roteiro da entrevista era voltado para caracterizar o ramo da Computação no qual atuavam os entrevistados, e a partir das respostas, verificou-se que um (1) que atuava na universidade, como docente no curso de Computação (E1), três (3) representantes de empresas atuavam no campo de desenvolvimento de *software* - em setores como análise de dados (E2), gestão de patrimônio público (E3) e gestão de saúde pública (E4) - e um (1) gerenciava uma entidade privada de serviço social, sem fins lucrativos, focado em gestão (E5).

Outra questão do roteiro de entrevista era focada em identificar *soft skills* necessárias para ingressar em vagas de estágio supervisionado e as respostas foram organizadas no Quadro 1, disponível a seguir.



Quadro 1 – Soft Skills mencionadas pelos entrevistados

Soft Skills mencionadas	Entrevistados
Capacidade de trabalhar em equipe	E1, E2, E3, E4, E5
Comunicação	E1, E2, E3, E4, E5
Autonomia	E1, E2
Confiabilidade	E2
Interpretação	E3

Fonte: Elaborado pelos autores, 2024.

Stevenson e Starkweather (2010) apresentam um panorama de diversas habilidades na área de TI reforçando a importância das *soft skills* em relação às *hard skills*. Através do Quadro 1, observou-se que, em concordância ao trabalho das referidas autoras, as principais aptidões requeridas foram a capacidade de trabalhar em equipe e a comunicação.

Outro fator apontado por E2 e reforçado pelo E1 foi a construção de *network* com contatos na área de Computação, como uma preocupação que deve ser pensada desde o início da graduação. “É a partir do *network* que é possível sobreviver às mudanças de mercado, se você [o estudante] quer se desenvolver na área [...] precisa construir uma *network*” (EP2). Ademais, é importante também que os estudantes estabeleçam contatos através de redes sociais profissionais como *LinkedIn*, conectando-se a empresários e empresas em que buscam atuar.

Hendarman e Cantner (2018) destacam em seu trabalho que as *HS* e *SS* estão amplamente conectadas ao ambiente de trabalho e não são dois pólos opostos, alinhando esse pensamento ao comentário do E2 “o que as empresas que contratam estagiários estão interessadas são pessoas que sejam confiáveis e interessadas”,



podemos inferir que formar conexões é imprescindível para participar de seleções de estágio supervisionado, e que demonstrar interesse partir da participação em atividades acadêmicas, tais como congressos, palestras e projetos de extensão ou pesquisa — atividades que refletem o domínio das competências mencionadas anteriormente — pode ser fundamental para adquirir uma vaga no mercado de trabalho.

Em relação às habilidades técnicas (*hard skills*) na escolha de candidatos às vagas de estágio supervisionado, importante para formação profissional (Brasil, 2008), a análise das respostas indicou que alguns dos conhecimentos exigidos foram relativos à lógica de programação, bancos de dados e programação web. Quanto a esta questão, o empresário E2 comentou sobre o uso de recursos gratuitos acessíveis aos estudantes, como as plataformas *HackerRank*, *LeetCode* e *Geek for Geek*.

Sobre as *hard skills*, é importante ressaltar que estão atreladas à adaptabilidade e pró-atividade, posto que as linguagens de programação são variadas e o mercado explora diferentes demandas. Foi nesse sentido, que o entrevistado E3 apresentou uma visão fundamental acerca dos domínios técnicos/conceituais em Computação, ao afirmar que: “programador não é aquele que tem domínio sobre uma linguagem de programação, programador é aquele que conhece bem a lógica de programação e pode, assim, programar em qualquer linguagem”.

Além dessas questões referentes às capacidades conceituais específicas da área de Computação, e essenciais para o estágio supervisionado, outra habilidade citada pelo E1 foi o domínio da língua inglesa, inclusive para atuar em empresas nacionais, que podem ter conexões estrangeiras. Portanto, foi possível observar a relevância das *HS* mencionadas por Pedrosa e Silva (2019).



CONSIDERAÇÕES FINAIS

O objetivo principal foi atingido parcialmente, tendo em vista o reduzido número de entrevistas realizadas, por conta do tempo disponível e da agenda complexa de outros possíveis seis (6) entrevistados. No entanto, avalia-se que este trabalho tem contribuição para estudantes e egressos de Computação, pois enfatiza a importância das aptidões na comunicação e no trabalho em equipe, bem como o domínio da língua inglesa, a adaptabilidade e a construção de *network* em atividades acadêmicas.

Considerando as lacunas deste trabalho, pode-se realizar trabalhos futuros envolvendo um *corpus* maior de entrevistados, de outros ramos da CC, para realização de um estudo comparativo. Além disso, pode-se ampliar a discussão sobre a importância de um mentor durante o estágio, bem como identificar as modalidades em que a atividade acontece, híbrida e remota, e a discussão sobre suas divergências.

Outro ponto relevante a se mencionar é que, durante as entrevistas, as perguntas foram conduzidas sem preparar previamente os candidatos à pesquisa para participação, o que pode ter resultado na baixa distinção entre as SS, sendo mencionadas apenas aquelas mais acessíveis e comuns na memória dos participantes. Uma sugestão pode ser acompanhar o questionário com uma ampla lista de aptidões das HS e SS.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE EMPRESAS DE SOFTWARE (ABES). **Mercado brasileiro de software**: panorama e tendências. São Paulo: Gráfica Corset, 2023.



BRETA, A. **Core skills**: nem soft, nem hard, 10 habilidades essenciais para um mundo em transformação. São Paulo: Teya, 2010.

BRASIL. **Lei nº 11.788, de 25 de setembro de 2008**. Brasília, 2008.

CASTRO, A. T. K. A.; SALVA, S. Estágio como Espaço de Aprendizagem Profissional da Docência no Curso de Pedagogia. *In*: Seminário de Pesquisa em Educação da Região Sul, 9, 2012. Caxias do Sul. **Anais [...]**. Caxias do Sul: UCS, 2012.

Geithner, S.; Menzel, D. *Effectiveness of Learning Through Experience and Reflection in a Project Management Simulation. Simulation and Gaming. ISAGA*, [s. l.], v. 47, n. 2, p. 228, 2016.

HENDARMAN, A. F.; CANTNER, U. *Soft skills, hard, skills, and individual innovativeness. Eurasian bus rev*, [s. l.], v. 8, n.7, 2018.

RINALDI, M. A. **Competências profissionais essenciais de um gestor de tecnologia da informação**. 2017. Dissertação (Mestrado em Administração) - Faculdade de Administração, Universidade Federal da Paraíba, João Pessoa, 2008.

SANTI, B. X. **Competências básicas do profissional de computação em nuvem no Brasil**: Um estudo baseado em anúncios de emprego. 2017. Monografia (Especialização em Gerência Empresarial) — Universidade de Taubaté, São Paulo, 2017.

STEVENSON, D. H.; STARKWEATHER, J. *PM critical competency index: IT execs prefer soft skills. International Journal of Project Management*, [s. l.], v. 28, n. 7, 2010.

WAINER, J. **Métodos de pesquisa quantitativa e qualitativa para a ciência da computação**. Campinas-SP: Instituto de Computação da UNICAMP, 2007.

WAZLAWICK, R. S. **Metodologia de Pesquisa para Ciência da Computação**. 3 ed. Rio de Janeiro: LTC/GEN, 2021.

ZIELINSKI, D. *Soft skills, hard truths. Training*. [s.l.], v. 42, 7, 2005.



4

Anniely Mariah Soares de Medeiros

Roger Natan Carvalho Silva

José Samuel Pinho Clementino

Luiz Felipe Vieira da Silva

Rosângela de Araújo Medeiros

EXPERIMENTAÇÕES DIDÁTICAS DE DOCENTES DE COMPUTAÇÃO:

**RELATOS DE DOCÊNCIA DIGITAL
COM METODOLOGIAS ATIVAS
EM TEMPOS CÍBRIDOS**

AUTORES

Anniely Mariah Soares de Medeiros

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: anniely.medeiros@aluno.uepb.edu.br

Roger Natan Carvalho Silva

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: roger.silva@aluno.uepb.edu.br

José Samuel Pinho Clementino

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: jose.pinho@aluno.uepb.edu.br

Luiz Felipe Vieira da Silva

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: luiz.vieira@aluno.uepb.edu.br

Rosângela de Araújo Medeiros

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: rosangelamedeiros@servidor.uepb.edu.br



ABSTRACT:

This article has the objective to present an analysis of reports of teaching practices guided by active methodologies using digital technologies of a Computing Science course at the Universidade Estadual da Paraíba, campus VII, Patos, that were collected in a project titled "Didactic Experimentation in Higher Education in Cybrid Times." In this sense, it is characterized as a case study, realized with five (5) university professors of the aforementioned course, carrying out a qualitative-quantitative analysis where the responses were recorded and transcribed. The use of eight (8) active methodologies was verified, complemented with the implementation of twelve (12) digital technologies.



RESUMO:

Este artigo tem como objetivo apresentar uma análise de relatos de práticas de ensino orientadas por metodologias ativas utilizando tecnologias digitais de um curso de Ciência da Computação da Universidade Estadual da Paraíba, campus VII, Patos, que foram coletados em um projeto intitulado "Experimentação Didática no Ensino Superior em Tempos Cíbridos". Nesse sentido, caracteriza-se como um estudo de caso, realizado com cinco (5) professores universitários do referido curso, realizando uma análise quali-quantitativa onde as respostas foram gravadas e transcritas. Foi verificado o uso de oito (8) metodologias ativas, complementadas com a implementação de doze (12) tecnologias digitais.

INTRODUÇÃO

A evolução das tecnologias digitais no século XXI engloba um processo cultural de indistinção entre a vida online e *offline*, visto que estamos conectados às ferramentas digitais a todo momento, configurando uma vida *onlife*, conforme define Floridi (2015). Tal contexto também foi analisado por Anders (2001) na virada do século, que definiu a intersecção entre o estado de vida material, o tecnológico e o ciberespaço como ciberidismo.

Nesse sentido, este cenário afeta a vida das pessoas, em todos aspectos, pois modifica nossas formas de ser, sentir, estar e aprender (Medeiros, 2020) e as práticas humanas nas diferentes áreas estão sendo cada vez mais questionadas e rompidas. Impõe-se então a necessidade de mudanças, o que inclui o ambiente educacional, visto que o modelo de ensino transmissivo de séculos atrás não tem apresentado recursos que atenda a diferentes estilos e formas de aprender dos estudantes que vivenciam esse contexto híbrido e *onlife*.

As formas de aprender no século XXI devem considerar as práticas e vivências digitais dos estudantes, propondo formas de ensinar que utilizem recursos tecnológicos em processos de ensino e de aprendizagem norteados por metodologias mais propícias a engajar esses estudantes (Bacich; Moran, 2018), que vivem essa lógica da fluidez, da ubiquidade (Lemos, 2005), da hiperconexão, conforme aponta Medeiros (2020).

Essas metodologias são definidas como ativas, compondo um rol de técnicas e métodos de ensino nos quais o foco é o desenvolvimento da autonomia, sendo o estudante o protagonista da aprendizagem, realizando atividades práticas, significativas e que envolvem colaboração e participação efetiva. Diversos estudos (Bacich; Moran, 2018; Alves, 2023; Silva, 2023) descrevem uma variedade de metodologias ativas, mas neste trabalho optou-se por enfatizar algumas,



a saber: Sala de aula invertida, aprendizagem baseada em projetos, gamificação e algumas outras.

Defende-se, então, que tais práticas de ensino devem estar presentes na educação da era digital, incluindo o ensino superior e, neste caso, deve fazer parte da atuação de docentes em cursos de graduação da área de Computação, para incentivar o desenvolvimento de habilidades fundamentais nesses tempos híbridos, possibilitando uma formação que prepare os futuros profissionais para o mercado de trabalho.

Assim, conhecer experimentações didáticas pautadas em metodologias ativas é essencial para pensarmos o ensino de Computação sob uma perspectiva que estimula a participação e o engajamento estudantil, sendo um fator preponderante para garantir a permanência do estudante em uma área reconhecida com altos índices de evasão, tal como identifica Albuquerque *et al.* (2023).

Considerando este cenário, o objetivo do artigo é apresentar uma análise de relatos de práticas docentes norteadas por metodologias ativas com uso de tecnologias digitais, coletadas em entrevistas com cinco professores do curso de Computação da Universidade Estadual da Paraíba, do campus VII - Patos. Foi realizada uma pesquisa que se aproxima de um estudo de caso (Wazlawick, 2021), com entrevistas norteadas por um roteiro de seis (6) questões semiestruturadas.

Este artigo estrutura-se em quatro seções: esta primeira aborda a problemática, o objetivo, bem como a necessidade de pensar sobre a docência no curso de Computação, a partir de metodologias ativas atreladas ao uso das tecnologias digitais. Na seção seguinte se desenvolveu a descrição da metodologia quanto a sua tipologia e a descrição do estudo. Na terceira seção são apresentados os resultados e discussão. Ademais, nas considerações finais, ocorreu a retomada dos resultados atingidos pelo trabalho, assim como as lacunas encontradas e as implicações do artigo sobre trabalhos futuros.



REFERENCIAL TEÓRICO

O processo de evolução das tecnologias vem configurando a cultura digital, definida por Costa (2008) e Kenski (2018) como um processo de revolução cultural que se constitui cada vez mais pela interseção entre a vida *online* e *offline*, na qual não existe dicotomia entre a vida presencial e a digital (Floridi, 2015), configurando a vida *onlife*.

Esse conceito dialoga com o termo *cibridismo*, definido por Anders (2001), que explicita ser um espaço híbrido, constituído por uma mistura entre os espaços físicos e os simbólicos, como se o ser humano, a partir do momento que a tecnologia digital foi sendo disseminada e estabelecida em nossa sociedade, construiu uma interconexão entre ele - o humano - e a máquina, entre a vida física e a virtual (Hamdan, 2009).

Nesse sentido, Hamdan (2009) afirma que a conexão digital por interfaces tende a se tornar cada vez mais móvel, cotidianidade e invisível, de modo que essa vida e realidade cívica seja naturalizada pelas pessoas, nas diferentes esferas e experiências sociais, inclusive na educação.

Neste panorama, é notável que os métodos de ensino transmissivos não se adequam mais à estrutura socialmente delimitada por todas essas alterações apresentadas nos tempos digitais, e assim está posta a necessidade de integrar o ambiente de ensino à essa lógica do digital.

Dessa forma, abordar a formação docente conectada à cultura digital resulta na necessidade de debater e propor um processo formativo tal como afirma Medeiros (2020), de maneira que o docente seja preparado para explorar metodologias que aprimorem o desenvolvimento do discente imerso na sociedade *onlife* e cívica.



Essa proposição metodológica, enquanto um conjunto de diretrizes orientadoras do processos de ensino e de aprendizagem, delimita formas de abordagem e estratégias para a execução pedagógica (Bacich; Moran, 2018), caminhando para um modo de ensinar que prioriza a ação e o engajamento estudantil, com metodologias ativas.

De acordo com Bacich e Moran (2018), tais metodologias caracterizam-se pelo enfoque na utilização de métodos ativos, divergentes do convencional pré-estabelecido e criativos que, por meio da aprendizagem ativa, centra-se no protagonismo do aluno, colaborando para a construção efetiva e significativa do conhecimento.

METODOLOGIA

O presente trabalho caracteriza-se, quanto à natureza, como uma pesquisa primária, segundo Wazlawick (2021), em função de propor uma análise oriunda de dados novos, coletados na realidade. No que tange aos objetivos, converge em uma produção exploratória e quanto aos procedimentos técnicos, aproxima-se de um estudo de caso (Wazlawick, 2021), utilizando como instrumento de coleta de dados a realização de entrevistas presenciais.

TIPOLOGIA DA PESQUISA

A pesquisa caracteriza-se quanto à natureza, como uma pesquisa primária (Wazlawick, 2021), em função de apresentar conhecimento novo a partir de um estudo de campo, visto que analisa experimentações didáticas relativas às metodologias ativas com uso de tecnologias digitais no ensino de Computação.



No que tange aos objetivos, é uma produção exploratória-descritiva – a partir da classificação de Wazlawick (2021) – desenvolvida a partir do levantamento de dados por entrevistas com os docentes da Universidade Estadual da Paraíba (UEPB) - Campus VII. A utilização da abordagem exploratória possibilitou uma investigação abrangente do tema em estudo e os dados foram analisados de forma qualitativa.

COLETA E ANÁLISE DE DADOS

Ao decorrer desse estudo, foram realizadas entrevistas com um grupo de 5 docentes do curso de Computação da UEPB - Campus VII, durante os meses de março e abril do ano de 2024 que utilizavam metodologias ativas associadas às tecnologias digitais, para ministrar suas disciplinas.

As entrevistas foram realizadas presencialmente, norteadas por um roteiro de seis (6) questões semi estruturadas, construído coletivamente. Foram gravadas em vídeo, com autorização dos docentes entrevistados e transcritas em texto, para posterior análise qualitativa.

Ocorreram no contexto do projeto de extensão, intitulado “Experimentações didáticas no ensino superior em tempos híbridos” também conhecido como “Formação Docente Digital”, fundamentado em metodologias ativas (Alves, 2023; Camargo; Daros, 2018; Bacich; Moran, 2018) e que tinha como objetivo organizar e propor momentos formativos para docentes universitários no contexto dos tempos híbridos.

Para análise das respostas, foram criadas planilhas nas quais se identificou repetições de ideias nas falas transcritas, transformadas em palavras-chave, inspirando-se em uma análise qualitativa.



Durante essa fase do estudo, buscou-se tratar dos seguintes eixos temáticos, relativos às perguntas do roteiro da entrevista:

- Quais as metodologias ativas foram adotadas pelos docentes?
- Quais as tecnologias digitais eram utilizadas pelos docentes?
- Quais os pontos positivos das metodologias ativas associadas às tecnologias digitais no ensino?

Importante mencionar que foi utilizada a classificação das ferramentas digitais definida por Bacich e Moran (2019) para analisar os recursos citados pelos docentes entrevistados. Dessa maneira, foi possível identificar em quais grupos de uso essas tecnologias se enquadraram.

RESULTADOS E DISCUSSÕES

As falas dos docentes investigados sobre as experimentações didáticas foram organizadas e analisadas considerando o objetivo deste estudo, que enfocou o uso de metodologias ativas com tecnologias digitais. Assim, o cerne da análise girou em torno de quais metodologias ativas e tecnologias digitais foram mais citadas.

Considerando o roteiro de questões para caracterizar o grupo investigado, a primeira pergunta estava relacionada às disciplinas ministradas pelos docentes entrevistados. As disciplinas ministradas pelos docentes entrevistados eram as seguintes: Métodos Avançados de Programação, Análise e Projeto de Sistema, Algoritmos, Programação Web, Engenharia de Software I, Engenharia de Software II, Teoria dos Grafos.



O grupo investigado foi composto por cinco docentes e, para garantir o sigilo e a integridade dos participantes da pesquisa, além de facilitar a análise das respostas, adotou-se um código para pontuar as especificidades das falas, sendo cada docente atribuído um número ao lado da letra P (pesquisado/professor): P1, P2, P3, P4 e P5.

AS METODOLOGIAS ATIVAS NO GRUPO INVESTIGADO

O Quadro 1, disponível a seguir, apresenta as metodologias mencionadas pelos cinco (5) docentes entrevistados, utilizadas durante suas aulas. Verificou-se que a metodologia ativa mais mencionada foi a *Aprendizagem por projetos*, seguida da *Gamificação*.

Quadro 1 – Metodologias Ativas utilizadas em um Ambiente Educacional

Metodologias ativas	Docentes
Aprendizagem baseada em projetos	P1, P2, P3, P5
Gamificação	P3, P4, P5
Aprendizagem baseada em problemas	P2, P5
Atividades de forma cooperativa	P2, P5
Do it yourself	P3
Sala de Aula invertida	P4
Fóruns de discussão	P5
Aprenda no seu ritmo	P5

Fonte: Elaborado pelos autores, 2024.

Os dados expostos no Quadro 1 podem demonstrar que essas metodologias tendem a ser adaptáveis ao contexto do ensino



de Computação, pelo fato dos discentes na área apresentarem maior engajamento com o universo digital e cívico (Anders, 2001), pela natureza do curso, estimulando a participação ativa, em concordância com as proposições de Witt e Kemczinski (2020).

O docente P1 relatou que, de acordo com metodologias ativas utilizadas, o feedback dos estudantes: *"Em geral, era bastante positivo, eles gostam de desafios, eles gostam de projetos, eles gostam de estar ativamente construindo coisas. Eles gostam de ver o resultado das coisas que eles constroem. Isso é satisfatório."* demonstrando o interesse dos alunos na utilização da metodologia baseada em projetos.

Além disso, é necessário comentar que o docente P5 também realizava pesquisa de sua prática, no sentido de experimentos didáticos, visto que explorava e comparava diversas metodologias ativas associadas às tecnologias, conduzindo verdadeira experimentações no contexto de suas aulas, tanto que suas respostas aparecem em diferentes menções apresentadas no Quadro 1.

O conceito de metodologias ativas envolve o protagonismo discente (Bacich; Moran, 2018), foi mencionado em diferentes falas, posto que alguns docentes indicaram que os métodos ativos contribuíam para estimular o interesse e imersão dos alunos nas aulas e compreensão do conteúdo.

Nesta perspectiva, o docente P1 evidenciou que esse tipo de ensino *"permite que o aluno consiga explorar possibilidades, consiga testar e avaliar criticamente como ele vai conseguir resolver certos desafios"*, relativos ao desenvolvimento do pensamento computacional e ao entendimento da Programação, vivenciando habilidades para abstração de códigos de Programação.

Outro docente - P2 - enfatizou que as metodologias ativas propiciaram o desenvolvimento de *softs skills* necessárias à



formação profissional, considerando os requisitos e demandas do mercado de trabalho. Assim, explorou a importância da prática docente em Computação ser baseada em projetos e problemas da vida social e/ou cotidiana. Esse também foi um ponto tratado pelo docente P1, quando pontuou que: *"Isso vai permitir que ele [o estudante] esteja preparado, quando sair daqui, para os diferentes desafios que ele vai encontrar lá fora"*.

Foi mencionado também pelo docente P5 que a implementação desses métodos de ensino ativo pode acarretar a diminuição dos índices de evasão no curso de Ciência da Computação, ideia que está correlacionada ao trabalho apresentado por Albuquerque *et al.* (2023).

AS TECNOLOGIAS DIGITAIS UTILIZADAS PELO GRUPO INVESTIGADO

Apresenta-se no Quadro 2 as tecnologias digitais utilizadas pelos cinco (5) docentes entrevistados, no qual verifica-se uma variedade de ferramentas que representam a diversidade das disciplinas ministradas pelos docentes do curso de Ciência da Computação da UEPB - Campus VII. Analisou-se também que foram mais mencionadas ferramentas digitais voltadas para prática de conceitos da área e como repositórios, sendo poucas as tecnologias digitais utilizadas para interação.

Bacich e Moran (2018) apresentam um contexto das tecnologias no ensino de literatura que podem ser adotadas de três (3) maneiras: as Tecnologias como repositórios de conhecimento, Ferramentas de interação e Ferramentas para a produção de novos conhecimentos.



Quadro 2 – Tecnologias digitais utilizadas pelo grupo investigado

Tecnologias digitais	Docentes
Ferramentas para a produção de novos conhecimentos	
Ferramentas de prototipação: Figma, Draw.io e Sketch	P1, P2, P5
Ambiente de Desenvolvimento Integrado (IDEs)	P2, P5
Editores de programação web	P2
Robótica	P5
Ferramentas como repositório de conhecimento	
Materiais instrucionais (documentos, arquivos produzidos pelos docentes e pelos estudantes, como <i>Google Docs</i> , apresentações de slides, <i>Jambord</i>)	P3, P4, P5
<i>Google Classroom</i>	P1, P3
Code.org	P5
Ferramentas de repositório de conhecimento e de interação	
Ferramentas de avaliação em Programação: <i>Screen Programming</i>	P5
Ferramentas de planejamento e organização de trabalho/projetos: <i>Notion</i> e <i>Trello</i>	P3

Fonte: Elaborado pelos autores, 2024.

Nesse sentido, em conformidade ao pensamento de Bacich e Moran (2018), podem ser observadas que as tecnologias adotadas pelos entrevistados possibilitaram a produção de novos conhecimentos, sendo mencionadas por P2 e P5, como IDEs, Editores de programação Web e as ferramentas de prototipação.

Percebeu-se, a partir da análise, que tais ferramentas podem contribuir para o desenvolvimento de habilidades necessárias ao profissional do mercado de trabalho, bem como os recursos voltados para produção de novos conhecimentos. Sobre elas, o docente P2 indicou que “*apresentam funcionalidades que facilitam a criação de projetos novos e a elaboração de um sistema desde o início, já*



organizando os pacotes e as classes nos devidos lugares”. Nesse sentido, a utilização dessas ferramentas e tecnologias em sala de aula pode preparar os estudantes para os prováveis desafios do mercado de trabalho.

Quanto ao uso de ferramentas para organização de materiais/ conteúdos de aulas, enquanto repositórios, foram pelo docente P3: *Classroom, Trello e Notion*, e que poderiam também ser exploradas como ferramentas de interação, embora não tenham sido citados esse intuito. Ademais, somente um professor - P5 - mencionou o uso de uma ferramenta digital voltada especificamente para avaliação da aprendizagem de programação, o *Screen Programming*. A avaliação mediada por recursos digitais em um viés das metodologias ativas é um processo teórico e prático a ser construído.

CONSIDERAÇÕES FINAIS

O objetivo deste estudo foi atingido a partir da coleta e análise de dados, posto que averiguou-se relatos de práticas docentes norteadas por metodologias ativas com uso de tecnologias digitais, de um curso de Computação da Universidade Estadual da Paraíba, do campus VII - Patos, e assim pode contribuir para ilustrar que o curso de Computação, mesmo sendo da área da Ciências Exatas, com elementos teóricos de abstração consideráveis, pode ser encaminhado a partir de experimentações didáticas conectadas e integradas ao contexto *onlife* (Floridi, 2015).

Identificou-se, a partir da análise das falas, que as metodologias ativas podem colaborar para um ambiente de ensino nas áreas de Exatas propício aos processos de ensino e de aprendizagem mais dinâmicos, relevantes e significativos, voltados para desenvolver a autonomia dos estudantes.



Como lacuna, não foram abordados os desafios aos docentes para implementação das metodologias ativas em sala de aula; relacionadas à ausência de preparo no contexto das metodologias ativas em sala de aula durante a formação desses profissionais de ensino. Além disso, as possíveis questões de acessibilidade e deserto digital também não foram investigadas, e é uma demanda urgente, visto que parte dos alunos brasileiros não têm acesso contínuo às tecnologias digitais.

Trabalhos futuros devem aprofundar essas lacunas, com estudos de caso sobre as práticas docentes, levando em consideração uma análise das dificuldades ao adotar essas metodologias ativas. Nesse sentido, apresentar mecanismos para que os docentes possam se profissionalizar com as tecnologias digitais e metodologias ativas em sala de aula.

REFERÊNCIAS

- ANDERS, Peter. Towards an architecture of mind. **Simpósio CAiia-STAR: Parâmetros extremos**. Novas dimensões da interactividade, 11-12 Jul. 2001.
- ALVES, Havana Diogo. Análise da aplicação de metodologias ativas em disciplina de Lógica de Programação. *In: ENCONTRO NACIONAL DE COMPUTAÇÃO DOS INSTITUTOS FEDERAIS (ENCOMPIF)*, 10., 2023, João Pessoa. **Anais [...]**. Porto Alegre: Sociedade Brasileira de Computação, 2023. p. 85-92.
- ALBUQUERQUE JÚNIOR, Aldo Candeia *et al.* Um estudo de caso sobre evasão no curso presencial de computação da universidade estadual da Paraíba - Campus VII - Patos - PB. *In: CONGRESSO NACIONAL DE EDUCAÇÃO (CONEDU)*, 9, 2023. João Pessoa. **Anais [...]**. João Pessoa: Centro de Convenções de João Pessoa, 2023.
- BACICH, Lilian; MORAN, José Manuel (Orgs.). **Metodologias ativas para uma educação inovadora**: abordagem teórico-prática. Porto Alegre: Penso, 2018.
- CAMARGO, Fausto; DAROS, Thuine. **A sala de aula inovadora**: estratégias pedagógicas para fomentar o aprendizado ativo. Porto Alegre: Penso, 2018.



COSTA, Rogério da. **A cultura digital**. 3. ed. São Paulo: Publifolha, 2008.

FLORIDI, Luciano. **The Onlife Manifesto: Being Human in a Hyperconnected Era**. London: Springer, 2015.

HAMDAN, Camila. **Realidade híbrida**. 2009. 100 f. Dissertação (Mestrado em Artes). Universidade de Brasília, Brasília, 2009.

KENSKI, Vani Moreira. **Cultura digital** (verbetes). *In*: MILL, Daniel (org.). **Dicionário Crítico de Educação e Tecnologia e de Educação a Distância**. Campinas: Papyrus, 2018. p. 139-143.

MEDEIROS, Rosângela de Araújo. **Ser docente universitário em tempos digitais: (trans)formar é preciso**. 2020. 280 f. Tese (Doutorado em Educação). Faculdade de Educação, Universidade de São Paulo, 2020.

MISHRA, Punya. **Considering Contextual Knowledge: The TPACK Diagram Gets an Upgrade**. *Journal of Digital Learning in Teacher Education*, 35, 76-78, 2019.

MISHRA, Punya; KOEHLER, Matthew J. Introducing technological pedagogical content knowledge. *In*: ANNUAL MEETING OF THE AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, **Anais** [...]. 2008. p. 16.

MISHRA, Punya; KOEHLER, Mathew. Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. **Teachers College Record**, v. 108, n. 6, June 2006, p. 1017-1054.

SILVA, Luiz Felipe Vieira da *et al.* Um estudo sobre as demandas de formação relativas a TPACK com docentes de cursos superiores de computação em tempos híbridos. *In*: CONGRESSO NACIONAL DE EDUCAÇÃO (CONEDU), 9, 2023. João Pessoa. **Anais** [...]. João Pessoa: Centro de Convenções de João Pessoa, 2023.

WAZLAWICK, Raul Sidnei. **Metodologia de Pesquisa para Ciência da Computação**. 3 ed. Rio de Janeiro – RJ: Gen, 2021.

WITT, Diego Teixeira; KEMCZINSKI, Avaniilde. **Metodologias de aprendizagem ativa aplicadas à computação: uma revisão da literatura**. *Informática na educação: teoria e prática*, v. 23, n. 1, jan./abr., 2020.



5

*Jhulia Medeiros de Queiroz
Carlos Vinicius Nunes Rodrigues
Jannayna Domingues Barros Filgueira*

DESENVOLVIMENTO E APLICAÇÃO DE JOGOS EDUCACIONAIS NO ENSINO DE ESTUDANTES COM SÍNDROME DE DOWN

AUTORES

Jhulia Medeiros de Queiroz

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: jhulia.queiroz@aluno.uepb.edu.br

Carlos Vinicius Nunes Rodrigues

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: carlos.vinicius.soares@aluno.uepb.edu.br

Jannayna Domingues Barros Filgueira

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: jannayna@servidor.uepb.edu.br



ABSTRACT:

This research aims to apply multidisciplinary educational games, through the teaching of information technology, in order to assist, instruct, and promote the inclusion of students with Down Syndrome in educational and work areas, highlighting the importance of using games and basic computer skills in the daily lives of volunteers.



RESUMO:

Esta pesquisa tem como objetivo aplicar jogos educativos multidisciplinares, por meio do ensino de informática, com o objetivo de auxiliar, instruir e promover a inclusão de alunos com Síndrome de Down nos ambientes educacional e de trabalho, destacando a importância da utilização de jogos e conhecimentos básicos de informática no cotidiano dos voluntários.

INTRODUÇÃO

“Inclusão é a nossa capacidade de entender e reconhecer o outro e, assim, ter o privilégio de conviver e compartilhar com pessoas diferentes de nós.” Essa afirmação de Maria Teresa (Mantoan, 2016), pesquisadora e professora da Unicamp, encapsula a essência da jornada empreendida pelo projeto. Neste estudo, exploramos uma parceria transformadora entre um projeto de extensão da Universidade Estadual da Paraíba (UEPB) e a Associação de Pais e Amigos dos Excepcionais (APAE) de Patos, focada no desenvolvimento integral de estudantes com Síndrome de Down no primeiro ano do ensino fundamental.

Ao transcender os métodos tradicionais de ensino, este projeto não apenas busca estimular o progresso acadêmico, mas também promover habilidades cognitivas, linguísticas, sociais e emocionais por meio de metodologias lúdicas nas disciplinas de Português e Matemática. Além de ser uma ferramenta educacional, configura-se como um instrumento de inclusão, autonomia e capacitação, fortalecendo valores de diversidade e equidade na comunidade.

A colaboração entre o projeto de extensão da UEPB e a APAE de Patos exemplifica o potencial transformador da união de esforços na educação e na sociedade em geral. Com a participação ativa de aproximadamente 95 alunos da APAE, este projeto visa proporcionar um ambiente de aprendizado estimulante e acolhedor para todos os envolvidos.

Além disso, o Instituto Federal da Paraíba - Campus Patos (IFPB) desempenha um papel crucial ao contribuir com sua experiência e recursos. A estreita colaboração entre o projeto “Brincar de Aprender” e o IFPB demonstra o compromisso das instituições em promover a inclusão educacional e o desenvolvimento positivo dos estudantes com Síndrome de Down.



Este artigo examina os jogos educativos desenvolvidos e as aulas de revisão aplicadas aos alunos da APAE, destacando a importância de adaptar as atividades às necessidades individuais dos participantes. Por meio da análise de métricas como as reações dos estudantes, interação com a tecnologia e processo de ensino, busca-se fornecer subsídios para aprimorar as práticas pedagógicas e promover uma abordagem inovadora na Educação Especial.

Dessa forma, o projeto emerge como um exemplo inspirador de como a cooperação entre diferentes instituições pode criar oportunidades de aprendizado significativas e impactantes para todos os envolvidos, promovendo uma sociedade mais inclusiva e justa.

METODOLOGIA

O projeto “Brincar de Aprender” adota uma metodologia que se fundamenta na integração entre teoria e prática, visando uma abordagem holística e eficaz para o desenvolvimento integral dos estudantes com necessidades específicas. A interação entre teoria e prática neste projeto é essencial. Os conhecimentos pedagógicos e tecnológicos são aplicados de forma integrada na criação de atividades lúdicas e interativas. A teoria fornece as bases conceituais e metodológicas para o planejamento e a execução das atividades, enquanto a prática permite a aplicação efetiva desses conhecimentos em um ambiente real de ensino-aprendizagem.

As estratégias de ação adotadas envolvem aplicação de jogos educativos, cuidadosamente desenvolvidos para atender às necessidades individuais dos discentes. Estes jogos são projetados com o propósito de promover a interação, o raciocínio lógico, a criatividade e o desenvolvimento das habilidades cognitivas, linguísticas, sociais e emocionais. Tal abordagem se fundamenta na premissa de



que a informática, além de ser uma ferramenta de nível social, cultural e de entretenimento, também pode trazer diversos benefícios para a educação. Por meio dela, é possível promover uma aprendizagem de forma mais dinâmica e prática, ao mesmo tempo em que se estimula o aprimoramento dessas competências nos alunos (Santos e Pequeno, 2011).

Adicionalmente, as atividades são planejadas e conduzidas de forma colaborativa, envolvendo não apenas os alunos, mas também os professores, monitores e orientadores da APAE. Essa abordagem colaborativa permite uma rica troca de conhecimentos e experiências entre os diferentes atores envolvidos no projeto, enriquecendo assim o processo de ensino-aprendizagem e contribuindo para a construção de um ambiente educacional mais inclusivo e participativo.

A metodologia adotada no projeto “Brincar de Aprender” combina teoria e prática de forma integrada, utilizando estratégias de ação colaborativas e materiais específicos para promover o desenvolvimento integral dos estudantes com necessidades específicas. Essa abordagem visa proporcionar uma experiência educacional significativa e inclusiva, contribuindo para o fortalecimento das habilidades e a autonomia dos participantes.

RESULTADOS E DISCUSSÕES

Ao decorrer desta pesquisa, um total de sete jogos foram criados utilizando a ferramenta Scratch(2023), todos aplicados em laboratório com os alunos da APAE durante as aulas do projeto. Esta ferramenta gratuita foi originalmente concebida para o público jovem, entretanto, sua utilização vem se estendendo para uma variedade de faixas etárias, permitindo uma ampla gama de projetos educacionais,



proporcionando assim, o uso da programação de maneira acessível a todos os públicos. Considerando os desafios cognitivos, sociais, sensoriais e motores enfrentados pelos discentes da instituição, a maioria demonstrou interesse e apreciação pelas atividades propostas.

Entre os jogos criados, podemos destacar o jogo intitulado “Sequência Numérica” ilustrado na na Figura 1, que tem como objetivo principal aprimorar noções de matemática básica e pôr em prática os conhecimentos a respeito dos números de um a dez. Neste jogo, o aluno deve auxiliar um personagem a realizar a contagem de seu material escolar. Para progredir e avançar todas as fases, o estudante deve relacionar corretamente o número à quantidade de objetos presentes na tela.

Figura 1 - Fases do jogo sobre sequência numérica



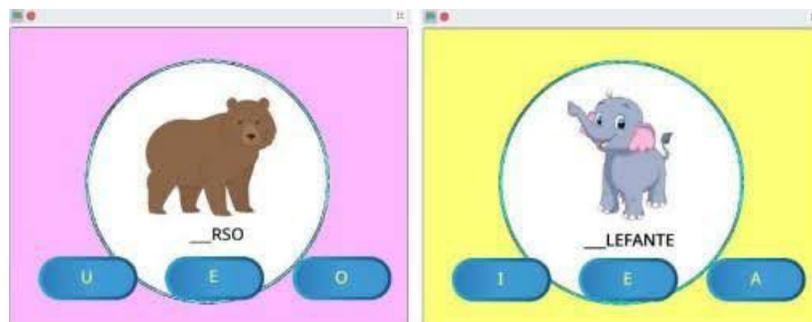
Fonte: Elaborado pelo autor, 2024.

O jogo Formação de Palavras com Vogais (Figura 2) foi desenvolvido com o objetivo de estimular a interação dos alunos com o computador, ao mesmo tempo em que promove o aprendizado de conceitos relacionados à disciplina de Língua Portuguesa. Para avançar para a próxima etapa, o aluno deve selecionar a vogal correta que completa o nome da imagem. Ao acertar, uma nova fase é iniciada, apresentando outra imagem e repetindo o processo com uma vogal diferente, incentivando o aprendizado progressivo.



Além dele, também foi aplicado aos alunos o jogo Quebra-Cabeça, representado na Figura 3, que tem como propósito fortalecer suas habilidades cognitivas, como raciocínio lógico e resolução de problemas. Para concluir o jogo, é necessário que o estudante, ao clicar para iniciar, posicione atentamente as peças em seus respectivos lugares a fim de completar a imagem para assim cumprir o objetivo do jogo.

Figura 2 - Fases do jogo Formação de palavras com vogais



Fonte: Elaborado pelo autor, 2024.

Figura 3 - Jogo Quebra-Cabeça



Fonte: Elaborado pelo autor, 2024.

É importante destacar que todos os jogos foram utilizados durante as aulas com o apoio dos participantes do projeto em questão, considerando as dificuldades específicas que muitos alunos



tinham em ler frases completas. Portanto, a aplicação foi conduzida com auxílio de recursos visuais, como slides interativos, uma breve explicação sobre o conteúdo e a temática dos jogos. Após esta etapa de preparação, os discentes foram orientados a resolverem os desafios propostos por conta própria.

Ao final, ficou evidente que grande parte da turma demonstrou interesse pelos jogos e foi notável o progresso no ensino obtido por eles. Sendo assim, foi determinado que os jogos tiveram seus objetivos alcançados e, conseqüentemente, contribuíram para o aprimoramento da aprendizagem dos discentes. Com base nos resultados obtidos e na relevância dos recursos tecnológicos na educação de pessoas com deficiência, é essencial promover a divulgação de tecnologias assistivas, a fim de garantir sua adoção e utilização tanto por alunos quanto por professores (Poker *et al.*, 2012).

CONSIDERAÇÕES FINAIS

Considerando os impactos pedagógicos deste projeto, visa-se a sua contribuição para a introdução de metodologias tecnológicas e inclusivas no cotidiano dos estudantes com necessidades específicas, auxiliando os alunos no ambiente escolar e promovendo a fixação dos conhecimentos compreendidos em sala de aula.

Por fim, pretende-se continuar o projeto e o desenvolvimento de novos jogos direcionados às necessidades específicas dos estudantes. Com a futura implementação de novos jogos na sala de aula, espera-se fazer um questionário para a coleta de dados referentes à opinião dos alunos acerca dos jogos, bem como o quantitativo do conteúdo absorvido por eles. Além disso, tem-se o objetivo futuro de estimular esses alunos a desenvolver seus próprios jogos digitais, utilizando a ferramenta scratch, com o intuito de aproximar esses



indivíduos de tecnologias atuais e que promovem a gamificação inclusiva no ensino especial, bem como promover a inclusão e abrir oportunidades desses alunos no ramo trabalhista.

REFERÊNCIAS

BRASILIA. Secretaria de Educação Especial. MEC. **Política nacional de educação especial na perspectiva da educação inclusiva**. 2008. Disponível em: <<http://portal.mec.gov.br/arquivos/pdf/politicaeducespecial.pdf>>. Acesso em: 28 mar. 2024.

LISBOA, M.F.L.S. **A deficiência e o preconceito: uma visão histórica e atual sobre a pessoa com deficiência**. 2020. Disponível em: <<https://revistas.fucamp.edu.br/index.php/cadernos/article/view/2245/1393>>. Acesso em: 28 de set. 2023.

MANTOAN, Maria Teresa. **Inclusão: A escola está preparada para ela? - Brasil Escola**. Disponível em: <<https://meuartigo.brasilecola.uol.com.br/educacao/inclusao-escola-esta-preparada-par-a-ela.htm>>. Acesso em: 01 mai. 2024.

MORELLATO, C. FELIPPIM, M.C.T. PASSERINO, L.M. GELLER, M. **Softwares educacionais e a educação especial**: refletindo sobre aspectos pedagógicos. Disponível em: <<http://www.seer.ufrgs.br/renote/article/download/13887/7803>>. Acesso em: 16 mar. 2024.

POKER, R.B. NAVEGA, M.T. PETITTO, S. **Acessibilidade na escola inclusiva**: tecnologias, recursos e o atendimento educacional especializado. Disponível em: <https://www.marilia.unesp.br/Home/Publicacoes/af-v4_colecao_poker_navega_petitto_2012-pcg.pdf>. Acesso em: 24 mar. 2024.

SANTOS, L.P. PEQUENO, R. **Novas tecnologias e pessoas com deficiências: a informática na construção da sociedade inclusiva?** Disponível em: <<http://books.scielo.org/id/6pdyn/pdf/sousa-9788578791247-04.pdf>>. Acesso em: 02 mar. 2024.

SCRATCH. **Imagine, Program, Share**. Disponível em: <<https://scratch.mit.edu/>>. Acesso em: 20 set. 2023.

SILVA, J.L.D. NUNES, I.D. GOMES, L.L. **Experiência com ensino de programação e aplicação de jogos digitais educativos**. Congresso Regional sobre Tecnologias na Educação. UFRN, Campina Grande-PB, 2016.

6

*Yuri dos Santos Monteiro
Geam Carlos de Araújo Filgueira
Jannayna Domingues Barros Filgueira*

AVALIAÇÃO DA EFICÁCIA DE UMA ÁRVORE DE DECISÃO NA DETECÇÃO DE DOENÇAS CARDIOVASCULARES

AUTORES

Yuri dos Santos Monteiro

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: yuri.monteiro@aluno.uepb.edu.br

Geam Carlos de Araújo Filgueira

Instituto Federal de Educação Ciência e Tecnologia.

E-mail: geam.filgueira@ifpb.edu.br

Jannayna Domingues Barros Filgueira

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: jannayna@servidor.uepb.edu.br



ABSTRACT:

This study aims to assess the performance of a machine learning model – Decision Tree, in the task of diagnosing cardiovascular diseases with a focus on strokes. A dataset containing clinical information of patients, including risk factors, medical history and test results, the Heart Disease - UCI, was used. The model was assessed considering metrics such as accuracy (78%), recall (78%), precision (70%) and F1-score (78%). This study highlights the effectiveness of the proposed model due to the value of such metrics. In addition, a visual analysis of the model's performance was performed using the ROC Curve, whose area under the curve was calculated at 0.79, thus corroborating its effectiveness.



RESUMO:

O presente estudo visa averiguar o desempenho de um modelo de aprendizado de máquina - Árvore de decisão, na tarefa de diagnosticar doenças cardiovasculares com foco em AVCs. Foi utilizado um conjunto de dados contendo informações clínicas de pacientes, incluindo fatores de risco, histórico médico e resultados de exames, o Heart Disease - UCI. O modelo foi aferido considerando métricas como acurácia (78%), recall (78%), precisão (70%) e F1-score (78%). Este estudo evidencia a eficácia do modelo proposto devido ao valor de tais métricas. Além disso, foi realizada uma análise visual da performance do modelo por meio da Curva ROC, cuja área sob a curva foi calculada em 0.79, corroborando assim a sua efetividade.



INTRODUÇÃO

Hoje, o número de incidência de enfermos com doenças cardiovasculares é uma preocupação de caráter mundial. Este fenômeno transcende fronteiras geográficas e abrange diversas camadas sociais, desencadeando uma série de implicações tanto para os sistemas de saúde quanto para os domínios socioeconômicos. A crescente prevalência destas enfermidades emerge como uma consequência complexa do envelhecimento demográfico, mudanças nos padrões comportamentais e uma interação intrincada entre fatores de risco, tais como obesidade, sedentarismo e tabagismo (Roth *et al.* 2020).

As doenças cardiovasculares são a principal causa de morte no mundo, sendo responsáveis por aproximadamente 17,9 milhões de mortes por ano, segundo dados da Organização Mundial da Saúde (OMS, 2023). Este aumento na incidência de doenças cardíacas requer estratégias de diagnóstico e gestão clínica inovadoras e eficazes, enfatizando a importância da adoção de tecnologias para melhorar os resultados clínicos e reduzir a mortalidade associada a estas doenças.

As Inteligências Artificiais surgem como uma alternativa para auxiliar clínicas, instituições e profissionais da área da saúde, bem como auxiliar no diagnóstico precoce destas patologias. Ao integrar algoritmos avançados de aprendizado de máquina e análise de dados, as IA têm o potencial de processar grandes conjuntos de informações clínicas de forma rápida e precisa, identificando padrões complexos e fornecendo insights valiosos para a tomada de decisões (Harrison, 2019).

Neste contexto, o algoritmo de árvore de decisão é um modelo bastante eficaz para tal demanda, como mostrado em (Nery *et al.* 2023) ostentando a eficácia deste modelo para não só diagnosticar o paciente, mas também identificar o tipo de enfermidade



cardiovascular. Este estudo tem como finalidade apresentar um modelo de árvore de decisão criado utilizando a linguagem Python, capaz de identificar doenças cardiovasculares.

Baseando-se nas métricas de desempenho, como a área sob a curva ROC, acurácia, precisão, recall e F1-score conforme explicado por (Géron, 2019) o modelo será avaliado e sua eficácia pode ser comprovada. Esse tipo de iniciativa representa um avanço significativo no campo da saúde digital, demonstrando como a integração de técnicas de inteligência artificial e análise de dados pode melhorar substancialmente a capacidade diagnóstica dos profissionais de saúde, possibilitando uma intervenção precoce e eficaz na gestão das doenças cardiovasculares.

REFERENCIAL TEÓRICO

Nos últimos anos, tem sido observado um aumento preocupante na incidência de doenças cardíacas em escala global, representando um desafio considerável para os sistemas de saúde ao redor do mundo (Roth *et al.* 2020). Esse fenômeno tem despertado uma série de preocupações tanto na esfera acadêmica quanto na clínica, destacando a necessidade premente de uma resposta abrangente e imediata.

A distribuição das mortes por doenças cardiovasculares é profundamente desproporcional, afetando de maneira mais severa países de baixa e média renda. Mais de três quartos das mortes relacionadas a problemas cardiovasculares ocorrem nessas regiões, onde o acesso limitado a cuidados de saúde de qualidade, bem como a prevenção e tratamentos eficazes, agrava ainda mais a situação (Roth *et al.*, 2020).

Além disso, a maioria das mortes prematuras causadas por doenças crônicas não transmissíveis ocorre em países de baixa e



média renda, sendo uma parte significativa dessas mortes atribuídas a doenças cardiovasculares (Roth *et al.*, 2020). Esses dados evidenciam a vulnerabilidade dessas populações e a necessidade urgente de intervenções direcionadas para mitigar os fatores de risco associados.

Nesse contexto, os avanços tecnológicos, especialmente no campo da aprendizagem de máquina, oferecem perspectivas promissoras para aprimorar tanto o diagnóstico quanto o tratamento das enfermidades cardiovasculares. Machine learning, um ramo da inteligência artificial, dedica-se ao desenvolvimento de algoritmos e modelos que capacitam sistemas computacionais a aprender e realizar previsões ou tomar decisões baseadas em grandes volumes de dados (Géron, 2019).

Diferentemente dos métodos convencionais de programação, onde instruções explícitas são codificadas para cada tarefa, o machine learning permite que os computadores reconheçam padrões e deduzam regras a partir de extensos conjuntos de dados. Esse processo é viabilizado por meio de diversas técnicas, como redes neurais, árvores de decisão e regressão, que são treinadas e ajustadas para maximizar a precisão e a eficiência em tarefas específicas.

Estudos recentes têm apontado o potencial significativo de algoritmos de aprendizagem de máquina, como os baseados em árvores de decisão, na previsão e identificação precoce de doenças cardíacas como (Nery *et al.* 2023). No entanto, para avaliar efetivamente a eficácia desses modelos, é essencial compreender métricas de desempenho como precisão, recall e F1-score. Essas métricas se tornam ainda mais cruciais em face de desafios de desbalanceamento de classes, onde a distribuição das classes no conjunto de dados pode ser desigual (Harrison, 2019).

Além disso, técnicas de avaliação como a curva ROC e a área sob a curva (AUC) fornecem insights valiosos sobre a capacidade discriminatória e os padrões de erro dos modelos de aprendizagem



de máquina. Essas ferramentas são essenciais para uma compreensão mais profunda do desempenho e da confiabilidade dos modelos em contextos clínicos (Géron, 2019).

Por fim, a linguagem de programação Python, aliada à sua biblioteca Scikit-learn, destaca-se como uma ferramenta fundamental para o desenvolvimento e implementação desses modelos inteligentes devido à sua versatilidade, facilidade de uso e ampla gama de funcionalidades disponíveis (Harrison, 2019). Assim, a integração de avanços tecnológicos, métricas de avaliação robustas e ferramentas de implementação eficientes desempenha um papel crucial na busca por soluções eficazes no enfrentamento das doenças cardiovasculares.

MATERIAIS E MÉTODOS

As informações obtidas neste artigo provêm de pesquisas empíricas quantitativas. A pesquisa é quantitativa devido à sua abordagem sistemática e objetiva na coleta e análise de dados numéricos, que permitem a avaliação quantificável dos algoritmos, o que pode levar em consideração a precisão dos algoritmos no diagnóstico de doenças cardíacas.

Os dados foram obtidos do repositório UCI Machine Learning, originalmente coletados de bancos de dados de doenças cardíacas em Cleveland, Hungria, Suíça e VA Long Beach. Foi utilizado um dos subconjuntos, o banco de dados Cleveland Heart Disease¹, que contém 13 atributos amplamente utilizados em estudos de aprendizado de máquina para prever a presença de Acidente Cascular Cerebral em pacientes, como idade, sexo, pressão arterial e achados eletrocardiográficos. Os dados clínicos foram obtidos através de exames médicos realizados em 303 pacientes com suspeita de cardiopatia.



A pesquisa é quantitativa por sua abordagem sistemática e objetiva na coleta e análise de dados numéricos, permitindo uma avaliação precisa e mensurável dos atributos do modelo de árvore de decisão, que poderá comprovar a qualidade do algoritmo no diagnóstico de doenças cardiovasculares. Conforme as orientações de (Géron, 2019) a árvore de decisão foi criada utilizando a linguagem de programação *Python*, juntamente com as bibliotecas *scikit-learn*, *pandas* e *matplotlib*. Tal escolha metodológica é congruente com as melhores práticas vigentes na comunidade científica, assegurando a reprodutibilidade e a qualidade dos resultados obtidos no contexto da pesquisa em análise.

Esse modelo foi aplicado à tarefa de previsão de doenças cardiovasculares, específico para o Acidente Vascular Cerebral. A análise visa identificar se o modelo é apropriado para auxiliar clínicas, hospitais e profissionais da saúde, considerando a importância da rapidez e confiabilidade no diagnóstico precoce.

RESULTADOS E DISCUSSÕES

A avaliação do desempenho do modelo baseou-se no estudo de (Nery *et al.* 2023). A área sob a curva ROC e a exibição e interpretação das matrizes de confusão foram examinadas. Além disso, os modelos foram ponderados com base em métricas de desempenho, acurácia, precisão, recall e F1-score que surgem a partir da matriz de confusão.

Estas ferramentas de avaliação foram escolhidas pela sua importância na determinação da capacidade dos modelos em distinguir entre classes, em minimizar erros de classificação e em maximizar a fiabilidade dos resultados. Através destas abordagens, é fornecida uma avaliação abrangente e rigorosa do desempenho dos



modelos de aprendizagem de máquina, destacando as suas capacidades e limitações.

Primeiramente, na análise da eficiência do modelo de machine learning de Árvore de decisão com base em suas métricas de desempenho, é essencial examinar a matriz de confusão. Essa matriz fornece uma discriminação detalhada de dos acertos e erros em cada classe de saída, o que é crucial para o cálculo das métricas de desempenho selecionadas. A Tabela 1 apresenta a matriz de confusão do modelo.

Tabela 1 - Matriz de confusão

		Valor Previsto	
		Não	Sim
Valor Real	Não	23	6
	Sim	9	23

Fonte: Elaborado pelo autor, 2024.

Ao analisar a matriz de confusão, observa-se que o desempenho da árvore de decisão foi satisfatório, apresentando um equilíbrio notável na minimização de falsos positivos e falsos negativos. Esse comportamento sugere que o modelo se mostra eficaz na tarefa proposta, indicando a capacidade de classificação correta, tanto das ocorrências positivas quanto das negativas.

Os modelos foram ponderados com base nas métricas de acurácia, recall, precisão e F1-score. Essas métricas fornecem uma visão da eficiência dos modelos. A acurácia mede a proporção de previsões corretas em relação ao total de casos, enquanto o recall avalia a capacidade do modelo de identificar todas as instâncias positivas. A precisão, por sua vez, indica a proporção de previsões positivas corretas em relação ao total de casos positivos, e o F1-score oferece um equilíbrio entre recall e precisão. A tabela 2 mostra as métricas do modelo.



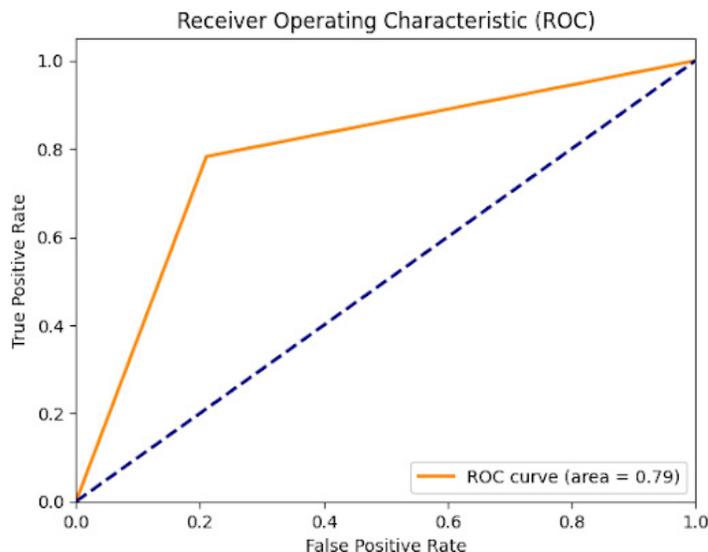
Tabela 2 – Métricas de desempenho

Métrica	Valor
Acurácia	0,78
Recall	0,78
Precisão	0,70
F1-score	0,78

Fonte: Elaborado pelos autores, 2024.

Com base na tabela de métricas da árvore de decisão, o modelo apresentou um desempenho consistente, com resultados variando entre 70% e 80% em todas as métricas avaliadas. Esses resultados indicam que o modelo demonstrou uma aptidão satisfatória para a tarefa de classificação, corroborando a eficácia preconizada por Géron (2019). Além disso, foi calculada a Área Sob a Curva (AUC) ROC, conforme ilustrado na Figura 1, a fim reforçar ainda mais a avaliação da capacidade discriminativa do modelo.

Figura 1 - Área sob a curva de ROC



Fonte: Elaborado pelo autor, 2024.



A linha tracejada representa a curva ROC de um classificador puramente aleatório, enquanto um classificador ideal é caracterizado por estar o mais distante possível dessa linha, preferencialmente na direção do canto superior esquerdo. A excelência de um classificador é refletida por uma Área Sob a Curva ROC (AUC) próxima de 1, indicando uma capacidade de discriminação perfeita, enquanto um classificador puramente aleatório tem uma AUC de 0,5 (Géron, 2019). Portanto, a eficácia da árvore de decisão proposta é evidenciada pela sua AUC de 0,79, sugerindo uma capacidade de previsão significativa e consistente.

CONSIDERAÇÕES FINAIS

A aplicação de algoritmos de aprendizado de máquina na previsão de doenças cardiovasculares, com especial foco em acidentes vasculares cerebrais (AVCs), destaca a relevância dessas tecnologias como ferramentas auxiliares nas decisões clínicas. O modelo analisado, uma Árvore de Decisão, evidenciou a capacidade de diagnóstico precoce, demonstrando-se uma abordagem promissora para a identificação antecipada de riscos e a melhoria dos desfechos clínicos.

Diante da crescente importância da aprendizagem de máquina no contexto das doenças cardiovasculares, é essencial explorar modelos adicionais, como Floresta Aleatória, Support Vectors Machine (SVM) e K-Nearest Neighbors (KNN), além das árvores de decisão mencionadas. Também é fundamental avaliar a eficácia desses modelos na classificação não só de doenças cardíacas em geral, mas também de condições específicas. Essas análises comparativas podem enriquecer nosso entendimento e contribuir para o desenvolvimento de ferramentas diagnósticas mais precisas.



Portanto, insta-se à comunidade acadêmica e clínica a dedicar esforços significativos na condução de estudos comparativos entre diferentes modelos de aprendizagem de máquina, bem como na investigação de sua eficácia na classificação de doenças cardiovasculares específicas. Essa abordagem mais abrangente e detalhada não apenas enriquecerá o entendimento sobre a aplicabilidade desses métodos na prática clínica, mas também poderá fornecer insights valiosos para o desenvolvimento de ferramentas diagnósticas e terapêuticas mais precisas e personalizadas no contexto das doenças cardiovasculares.

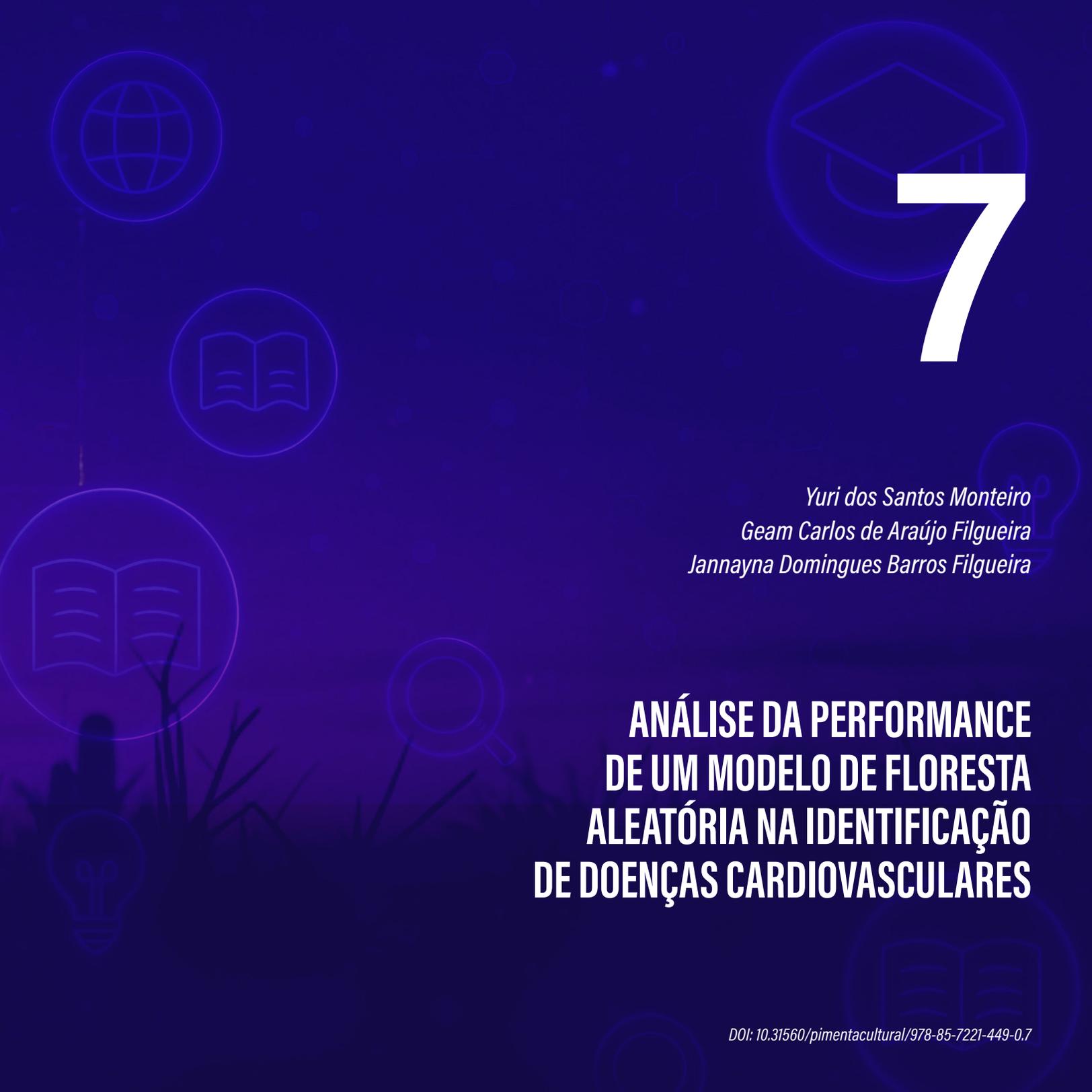
REFERÊNCIAS

GÉRON, Aurélien. **Mãos à obra: aprendizado de máquina com Scikit-Learn & TensorFlow**. Alta Books, 2019.

HARRISON, Matt. **Machine learning pocket reference: working with structured data in python**. O'Reilly Media, 2019.

NERY, Barbara. *et al.* **Comparação De Modelos De Machine Learning Para Diagnóstico De Doenças Cardiovasculares: Um Estudo sobre Acidente Vascular Cerebral, Arritmia Cardíaca E Doença Arterial Coronariana**. Revistaft, 2023.

ROTH, Gregory A. *et al.* Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. **Journal of the American college of cardiology**, v. 76, n. 25, p. 2982-3021, 2020.



7

*Yuri dos Santos Monteiro
Geam Carlos de Araújo Filgueira
Jannayna Domingues Barros Filgueira*

ANÁLISE DA PERFORMANCE DE UM MODELO DE FLORESTA ALEATÓRIA NA IDENTIFICAÇÃO DE DOENÇAS CARDIOVASCULARES

AUTORES

Yuri dos Santos Monteiro

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: yuri.monteiro@aluno.uepb.edu.br

Geam Carlos de Araújo Filgueira

Instituto Federal de Educação Ciência e Tecnologia.

E-mail: geam.filgueira@ifpb.edu.br

Jannayna Domingues Barros Filgueira

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: jannayna@servidor.uepb.edu.br



ABSTRACT:

The present study investigated the performance of a machine learning model - Random Forest - in the task of diagnosing cardiovascular diseases. To enable this analysis, a dataset containing clinical information from numerous individuals, obtained from the Kaggle platform, was used. The model was evaluated based on the following metrics: accuracy (90%), recall (87%), precision (90%), and F1-score (88%). This paper demonstrates the effectiveness of the proposed model, as evidenced by the high values of its metrics. Additionally, a visual assessment of the model's performance was conducted using the ROC Curve, where the area under the curve was determined to be 0.79, further reinforcing its effectiveness.



RESUMO:

O presente estudo investigou o desempenho de um modelo de aprendizado de máquina - Floresta Aleatória, na tarefa de diagnosticar doenças cardiovasculares. Para viabilizar tal análise, foi usado um conjunto de dados contendo informações clínicas de inúmeros indivíduos, obtido na plataforma kaggle. O modelo foi ponderado com base nas métricas: acurácia (90%), recall (87%), precisão (90%) e F1-score (88%). Este artigo comprova a eficácia do modelo proposto, mostrado pelo elevado valor de suas métricas. Adicionalmente, uma avaliação visual do desempenho do modelo foi conduzida através da Curva ROC, onde a área sob a curva foi determinada como 0.79, fortalecendo a validade de sua eficácia.



INTRODUÇÃO

Atualmente, a incidência global de enfermidades cardiovasculares é uma preocupação de magnitude internacional, transcendendo fronteiras geográficas e impactando varias esferas socioeconômicas. Conforme defendido por (Roth *et al.*, 2020) esse fenômeno é multifacetado, decorrente do envelhecimento demográfico, transformações nos comportamentos individuais e uma intrincada interação de fatores de risco, como obesidade, sedentarismo e tabagismo.

As doenças cardiovasculares representam a principal causa de mortalidade global, resultando em mais de 17 milhões de mortes anualmente, de acordo com a Organização Mundial da Saúde (OMS, 2023). Esse aumento na prevalência de doenças cardíacas demanda a implementação de estratégias eficientes para diagnóstico e gestão clínica, ressaltando a importância de incorporar tecnologias que possam melhorar os resultados clínicos e diminuir a mortalidade associada a essas condições.

Neste cenário desafiador, as Inteligências Artificiais (IA) surgem como uma promissora ferramenta para colaborar com clínicas, instituições e profissionais da saúde, visando o diagnóstico precoce e a gestão eficaz das enfermidades cardiovasculares. A incorporação de algoritmos avançados de aprendizado de máquina e análise de dados permite às IA processar extensos conjuntos de dados clínicos de forma ágil e precisa, identificando padrões complexos e oferecendo informações relevantes para embasar decisões clínicas (Harrison, 2019).

Neste viés, o algoritmo de Floresta Aleatória emerge como uma poderosa ferramenta, conforme demonstrado por Santos. (2022), destacando sua eficácia tanto no diagnóstico do paciente quanto na identificação precisa do tipo de enfermidade cardiovascular. Com o propósito de contribuir para esse avanço, o presente



estudo propõe um modelo de Floresta Aleatória implementado em Python, destinado à identificação de doenças cardiovasculares.

A avaliação da eficácia do modelo fundamenta-se em métricas de desempenho robustas, como a Área Sob a Curva ROC, acurácia, precisão, recall e F1-score, conforme discutido por Géron (2019). Essa análise oferece uma base sólida para validar o modelo proposto e evidenciar sua utilidade prática. Essas iniciativas representam avanços significativos na área da saúde digital, evidenciando o potencial da integração de técnicas de inteligência artificial e análise de dados para aprimorar substancialmente a capacidade diagnóstica dos profissionais de saúde e permitir intervenções precoce e eficazes na gestão das doenças cardiovasculares.

REFERENCIAL TEÓRICO

Nos últimos anos, tem sido observado um aumento preocupante na incidência de doenças cardíacas em escala global, representando um desafio considerável para os sistemas de saúde ao redor do mundo (Roth *et al.*, 2020). Esse fenômeno tem despertado uma série de preocupações tanto na esfera acadêmica quanto na clínica, destacando a necessidade premente de uma resposta abrangente e imediata.

A distribuição das fatalidades por doenças cardiovasculares é extremamente desigual, impactando de forma mais intensa as nações de baixa e média renda. Mais de 75% das mortes vinculadas a problemas cardíacos acontecem nessas áreas, onde a falta de acesso a serviços de saúde de qualidade, além da escassez de medidas preventivas e tratamentos eficazes, agrava ainda mais o cenário (Roth *et al.*, 2020).

Além disso, a maior parte das mortes prematuras provocadas por doenças crônicas não transmissíveis ocorre em nações de baixa



e média renda, com uma fração substancial dessas mortes atribuída a problemas cardiovasculares (Roth *et al.*, 2020). Esses dados evidenciam a fragilidade dessas populações e a necessidade urgente de intervenções específicas para atenuar os fatores de risco associados.

Nesse contexto, os avanços tecnológicos, particularmente no campo do aprendizado de máquina, oferecem oportunidades promissoras para aprimorar tanto o diagnóstico quanto o tratamento das doenças cardiovasculares. O aprendizado de máquina, uma área dentro da inteligência artificial, se concentra na criação de algoritmos e modelos que capacitam sistemas computacionais a aprender e a fazer previsões ou tomar decisões com base em grandes volumes de dados (Géron, 2019).

Estudos recentes têm apontado o potencial significativo de algoritmos de aprendizagem de máquina, como os baseados em Floresta Aleatória, na previsão e identificação precoce de doenças cardíacas (Santos, 2022). No entanto, para avaliar efetivamente a eficácia desses modelos, é essencial compreender métricas de desempenho como precisão, recall e F1-score. Essas métricas se tornam ainda mais cruciais em face de desafios de desbalanceamento de classes, onde a distribuição das classes no conjunto de dados pode ser desigual (Harrison, 2019).

Ademais, métodos de avaliação como a curva ROC e a Área Sob a Curva (AUC) oferecem informações importantes sobre a capacidade dos modelos de aprendizado de máquina em diferenciar entre classes e seus padrões de erro. Essas ferramentas são fundamentais para uma análise mais detalhada do desempenho e da confiabilidade dos modelos em contextos clínicos (Géron, 2019).

Além disso, a crescente integração de tecnologias de processamento de dados e algoritmos avançados com registros de saúde eletrônicos está proporcionando uma nova dimensão à análise e gestão das doenças cardiovasculares. A capacidade de correlacionar



dados clínicos com modelos preditivos pode não apenas melhorar a precisão dos diagnósticos, mas também permitir intervenções mais personalizadas e oportunas, ampliando o impacto positivo dessas tecnologias na saúde pública global (Harrison, 2019).

Por fim, a linguagem de programação Python, aliada à sua biblioteca Scikit-learn, destaca-se como uma ferramenta fundamental para o desenvolvimento e implementação desses modelos inteligentes devido à sua versatilidade, facilidade de uso e ampla gama de funcionalidades disponíveis (Harrison, 2019). Assim, a integração de avanços tecnológicos, métricas de avaliação robustas e ferramentas de implementação eficientes desempenha um papel crucial na busca por soluções eficazes no enfrentamento das doenças cardiovasculares.

MATERIAIS E MÉTODOS

As informações apresentadas neste artigo são derivadas de estudos empíricos quantitativos. A pesquisa é considerada quantitativa devido à sua abordagem objetiva e sistemática na coleta e análise de dados numéricos, o que permite uma avaliação mensurável dos algoritmos, levando em conta a precisão dos mesmos no diagnóstico de doenças cardíacas. Explorando os principais atributos para avaliação de desempenho de um algoritmo de floresta aleatória desenvolvido.

Os dados foram fornecidos pela plataforma kaggle, advindos de pesquisas telefônicas anuais que coletam dados sobre o estado de saúde dos residentes dos EUA realizada pelo Sistema de Vigilância de Fatores de Risco Comportamentais (BRFSS), tal base de dados apresenta 39 variáveis preditoras e uma variável de classificação, tal como: Índice de Massa Corporal (IMC), consumo de álcool, idade, raça e tempo de sono, além de uma variável de classificação, tendo ao todo 319795 instancias.



Esta pesquisa adota uma abordagem quantitativa devido à sua metodologia sistemática e objetiva na coleta e análise de dados numéricos. Isso permite uma avaliação precisa e mensurável dos atributos do modelo de floresta aleatória, contribuindo para a validação da eficácia do algoritmo no diagnóstico de doenças cardiovasculares. Seguindo as diretrizes de (Géron, 2019), a implementação da floresta aleatória foi realizada utilizando a linguagem de programação Python, juntamente com as bibliotecas scikit-learn, pandas e matplotlib. Essa escolha metodológica está alinhada com as melhores práticas da comunidade científica, garantindo a reprodutibilidade e a qualidade dos resultados obtidos na pesquisa em questão.

RESULTADOS E DISCUSSÕES

A avaliação do desempenho desse modelo de floresta aleatória foi fundamentada no estudo de Santos (2022) O modelo foi aplicado com o propósito de prever doenças cardiovasculares em um contexto amplo. Foram analisadas a área sob a curva ROC, bem como a exibição e interpretação das matrizes de confusão. Além disso, os modelos foram ponderados com base em métricas de desempenho, incluindo acurácia, precisão, recall e F1-score, que são derivadas da matriz de confusão.

Esses meios de avaliação foram selecionadas devido à sua relevância na determinação da capacidade dos modelos de aprendizagem de máquina em diferenciar classes, reduzir erros de classificação e maximizar a confiabilidade dos resultados. Através dessas abordagens, é possível realizar uma análise detalhada e rigorosa do desempenho dos modelos de aprendizado de máquina, destacando suas vantagens e limitações.



Esta análise busca determinar se o modelo é suficientemente robusto para fornecer suporte às clínicas, hospitais e profissionais de saúde, especialmente considerando a crítica importância da precisão e rapidez no diagnóstico precoce dessas condições.

Para começar, ao analisar a eficácia do modelo de machine learning de Floresta aleatória com base em suas métricas de desempenho, é fundamental examinar a matriz de confusão. Essa matriz oferece uma visão detalhada dos acertos e erros em cada classe de saída, sendo crucial para o cálculo das métricas de desempenho escolhidas. A Tabela 1 exibe a matriz de confusão do modelo.

Tabela 1 - Matriz de confusão

		Valor Previsto	
		Não	Sim
Valor Real	Não	57189	1178
	Sim	4943	649

Fonte: Elaborado pelo autor, 2024.

Ao examinar a matriz de confusão, nota-se que o desempenho do modelo foi satisfatório, exibindo um equilíbrio notável na redução de falsos positivos e falsos negativos. Esse comportamento indica que o modelo é eficaz na tarefa proposta, demonstrando uma capacidade adequada de classificar corretamente tanto as ocorrências positivas quanto as negativas.

Os modelos foram avaliados com base nas métricas de acurácia, recall, precisão e F1-score, que proporcionam uma compreensão da eficiência dos modelos. A acurácia reflete a proporção de previsões corretas em relação ao total de casos, enquanto o recall mede a capacidade do modelo de identificar todas as instâncias positivas. A precisão mostra a proporção de previsões positivas corretas em relação ao total de casos positivos, e o F1-score oferece um equilíbrio entre recall e precisão. A Tabela 2 apresenta as métricas do modelo.



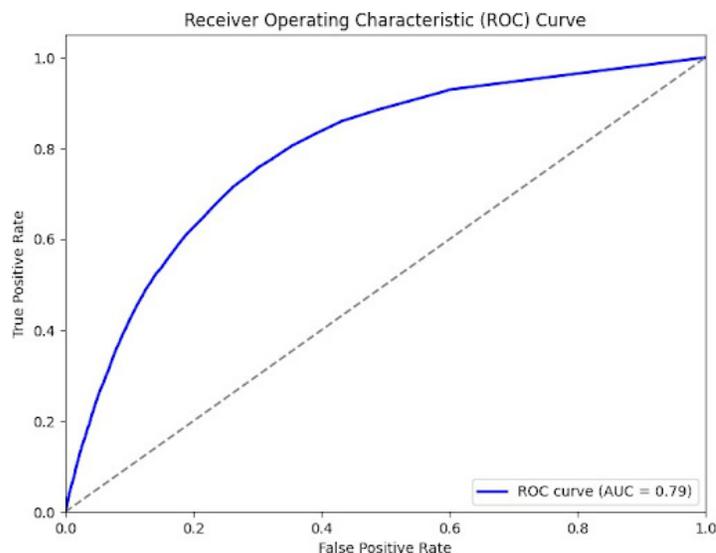
Tabela 2 - Métricas de desempenho

Métrica	Valor
Acurácia	0.90
Recall	0.87
Precisão	0.90
F1-score	0.88

Fonte: Elaborado pelos autores, 2024.

De acordo com a tabela de métricas, o modelo mostrou um desempenho estável, com valores variando entre 87% e 90%. Dessa forma, a eficácia do algoritmo foi evidenciada pelos altos valores das suas métricas, conforme discutido por Santos (2022). Além disso, foi calculada a Área Sob a Curva (AUC) ROC, conforme exibido na Figura 1, com o intuito de reforçar a avaliação da capacidade discriminativa do modelo.

Figura 1 - Área sob a curva de ROC



Fonte: Elaborado pelo autor, 2024.



A linha tracejada representa a curva ROC de um classificador puramente aleatório, o qual faz previsões de forma arbitrária. Por outro lado, um classificador ideal é caracterizado por estar o mais afastado possível dessa linha, idealmente posicionado no canto superior esquerdo do gráfico (Géron, 2019). A qualidade de um classificador é quantificada pela Área Sob a Curva ROC (AUC), onde um valor próximo de 1 indica uma capacidade de discriminação perfeita, enquanto um classificador puramente aleatório alcança uma AUC de 0,5 (Harrison, 2019). Dessa forma, a eficácia do algoritmo proposto é evidenciada pelo seu valor de AUC de 0,79 e pelo afastamento significativo da linha tracejada, indicando uma capacidade de previsão significativa e consistente.

CONSIDERAÇÕES FINAIS

A aplicação de algoritmos de aprendizado de máquina supervisionado na previsão de doenças cardiovasculares, destaca a relevância dessas tecnologias como ferramentas auxiliares nas decisões clínicas. O modelo analisado evidenciou a capacidade de diagnóstico precoce, demonstrando-se uma abordagem promissora para a identificação antecipada de riscos e a melhoria dos desfechos clínicos.

Diante da crescente relevância da aprendizagem de máquina no contexto das doenças cardiovasculares, é imperativo aprofundar a investigação em modelos suplementares, tais como Convolutional Neural Networks (CNN), Support Vector Machines (SVM) e K-Nearest Neighbors (KNN), além das Florestas Aleatórias. É também essencial proceder à avaliação meticulosa da eficácia desses modelos não apenas na classificação de doenças cardíacas em geral, mas também na identificação de condições específicas. Tais análises comparativas tendem a enriquecer nosso discernimento,



contribuindo de forma significativa para o avanço no desenvolvimento de ferramentas diagnósticas de maior precisão.

Neste sentido, conclama-se à comunidade acadêmica e clínica a empreender esforços substanciais na condução de estudos comparativos entre distintos modelos de aprendizagem de máquina, bem como na investigação de sua eficácia na classificação de doenças cardiovasculares específicas. Esta abordagem ampla e detalhada não apenas ampliará nosso conhecimento sobre a aplicabilidade prática desses métodos, mas também poderá proporcionar insights inestimáveis para o aprimoramento de ferramentas diagnósticas e terapêuticas mais precisas e individualizadas no contexto das doenças cardiovasculares.

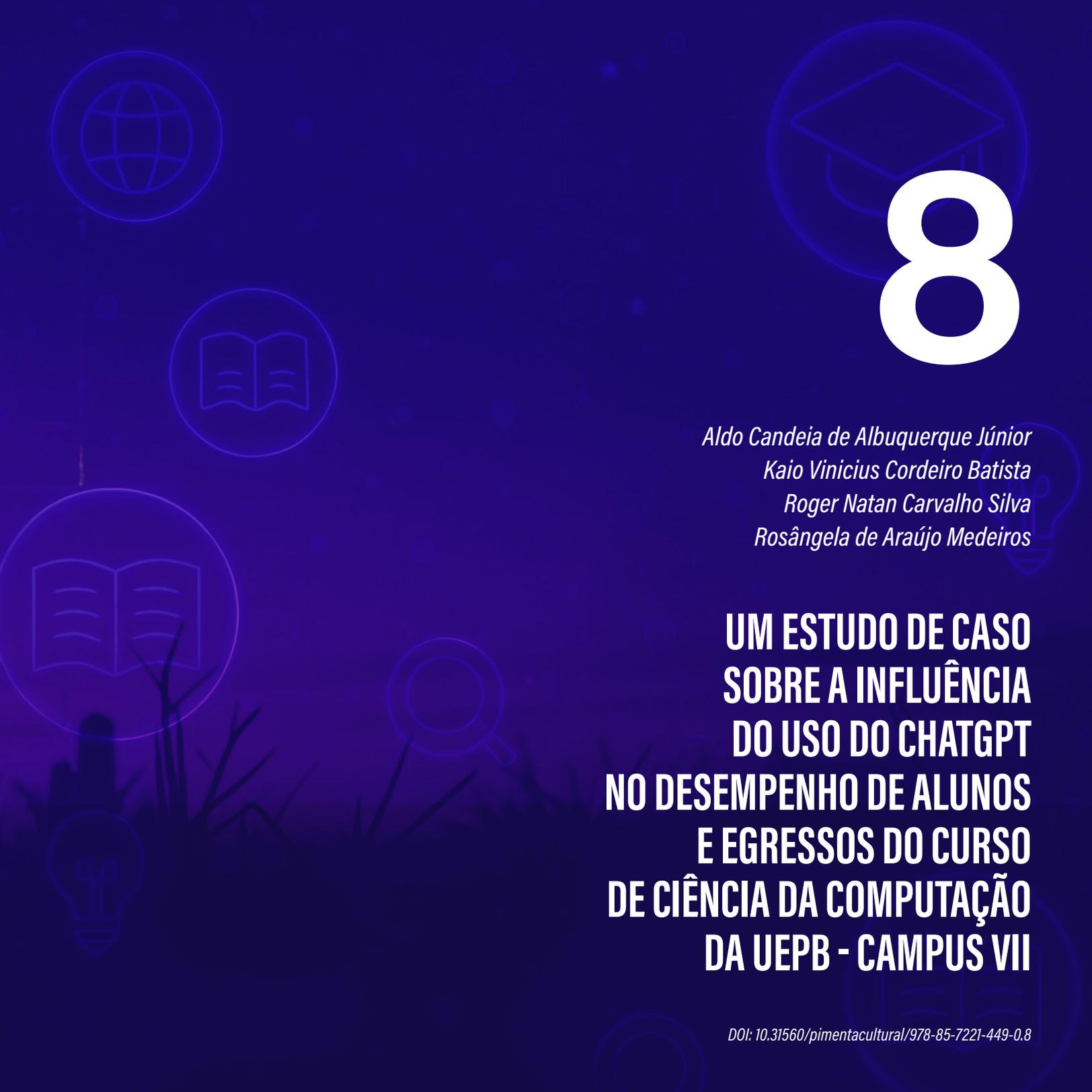
REFERÊNCIAS

GÉRON, Aurélien. **Mãos à obra: aprendizado de máquina com Scikit-Learn & TensorFlow**. Alta Books, 2019.

HARRISON, Matt. **Machine learning pocket reference: working with structured data in python**. O'Reilly Media, 2019.

ROTH, Gregory A. *et al.* Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. **Journal of the American college of cardiology**, v. 76, n. 25, p. 2982-3021, 2020.

SANTOS, Bruno Bruno Silva dos. **Uma análise exploratória de dados e o uso de aprendizado de máquina para classificação de doenças cardiovasculares**. 2022. Trabalho de Conclusão de Curso. Universidade Federal do Rio Grande do Norte.



8

Aldo Candeia de Albuquerque Júnior

Kaio Vinicius Cordeiro Batista

Roger Natan Carvalho Silva

Rosângela de Araújo Medeiros

UM ESTUDO DE CASO SOBRE A INFLUÊNCIA DO USO DO CHATGPT NO DESEMPENHO DE ALUNOS E EGRESSOS DO CURSO DE CIÊNCIA DA COMPUTAÇÃO DA UEPB - CAMPUS VII

AUTORES

Aldo Candeia de Albuquerque Júnior

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: roger.silva@aluno.uepb.edu.br

Kaio Vinicius Cordeiro Batista

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: kaio.batista@aluno.uepb.edu.br

Roger Natan Carvalho Silva

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: roger.silva@aluno.uepb.edu.br

Rosângela de Araújo Medeiros

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: rosangelamedeiros@servidor.uepb.edu.br



ABSTRACT:

This paper aimed to compare performances of a group of students who used ChatGPT as a complementary study tool in the subject Programming Language I of a Computer Science course at UEPB - Campus VII related to those who did not have the same access to this technological resource throughout this discipline, including graduated students. For that, a descriptive case study was carried out with data collection through questionnaires aimed to its audience. Results showed superior performance for the group of students who did not use Artificial Intelligence in such an academic environment.



RESUMO:

Este artigo teve como objetivo comparar o desempenho de um grupo de estudantes que utilizaram o ChatGPT como ferramenta complementar de estudo na disciplina de Linguagem de Programação I de um curso de Ciência da Computação na UEPB - Campus VII, em relação àqueles que não tiveram o mesmo acesso a esse recurso tecnológico ao longo da disciplina, incluindo estudantes já graduados. Para isso, foi realizado um estudo de caso descritivo, com coleta de dados por meio de questionários direcionados ao público-alvo. Os resultados mostraram um desempenho superior para o grupo de estudantes que não utilizou Inteligência Artificial nesse ambiente acadêmico.



INTRODUÇÃO

Desde seu surgimento, as Inteligências Artificiais (IA) vêm oportunizando praticidade, automação e eficiência tecnológica na realização de tarefas antes exclusivas ao ser humano. No entanto, nada se compara ao *boom* dos últimos anos gerado pelo aperfeiçoamento dessas atividades. Ferramentas dedicadas à geração de vídeos, áudios, imagens e textos têm se tornado cada vez mais populares e acessíveis a diversos públicos com variados interesses, inclusive no campo de estudos e produção acadêmica.

No campo da educação, tem se proliferado o uso de ferramentas de IA geradoras de texto, como o *ChatGPT* (Lo, 2023). Desenvolvido e lançado pela empresa norte americana *OpenAI* em 30 de novembro de 2022, o *ChatGPT* é um modelo de IA definida como *Generative Pre-Trained Transformer*, do inglês, "Transformador Generativo Pré-treinado", baseada em *Deep Learning* (Barbosa; Portes, 2023).

Ao realizar uma consulta acerca de um determinado tema, o *ChatGPT* expõe resultados às perguntas do usuário, diferentemente de ferramentas de busca tradicionais que respondem em formato de *links*. Além disso, como aponta Barbosa e Portes (2023), o *ChatGPT*, em sua versão mais completa, também pode sugerir textos, músicas, receitas e até códigos de programação.

Nessa perspectiva, o *ChatGPT* pode tornar-se um aliado no momento do estudo de conteúdos da área de Computação, para complementar atividades acadêmicas dos estudantes nas aulas e até interferir no seu desempenho em Programação, fazendo-se necessário identificar diferenças no aproveitamento acadêmico de alunos(as) observados antes e depois do advento da ferramenta.

Em um estudo realizado por Husain (2023), com foco especial em instrutores de Programação e seus estudantes em uso do *ChatGPT*, pode-se observar a integração do *ChatGPT* no aprendizado



de programação, sendo utilizado tanto pelos alunos – auxiliando na resolução de problemas de programação – quanto pelos professores – sugerindo ideias para elaboração de conteúdos relacionados ao tema.

Yilmaz e Yilmaz (2023) também produziram um estudo sobre este tema com estudantes do ensino superior, que apontarem benefícios e desvantagens do uso do *ChatGPT* em atividades de Programação. A maioria indicou que facilita o aprendizado, porém um grupo de alunos afirmou que o uso da IA poderia afetar negativamente suas habilidades de pensamento, visto que as respostas eram obtidas de maneira rápida e fácil, sem necessidade de esforço e raciocínio.

Tendo em vista este cenário e esta possibilidade, o objetivo deste trabalho foi comparar o desempenho de um grupo de alunos que tiveram acesso ao *ChatGPT* como ferramenta e estudos em relação ao desempenho de outro grupo de egressos que não o utilizou como material de consulta na disciplina de Linguagem de Programação I (LPI) do curso de Ciência da Computação da UEPB - Campus VII.

Este estudo apoiou-se nas metodologias de pesquisa para Ciência da Computação descritas por Wazlawick (2021), caracterizando-se como uma pesquisa de natureza primária, aproximando-se de um estudo de caso comparativo e para coleta de dados, utilizou-se a aplicação de questionários para um grupo de estudantes matriculados e alunos egressos de um curso de Computação.

Deste modo, o artigo está estruturado em quatro seções, a saber: esta primeira, a introdução, que contempla problemática, objetivo e resumo da metodologia. Na segunda seção, é apresentado o percurso metodológico, que engloba a tipologia de pesquisa segundo Wazlawick (2021), e o processo de coleta e análise de dados.

Na seção seguinte, é explorado o referencial teórico que embasou a produção da pesquisa, tratando da história e conceitos da Inteligência Artificial e do seu uso no contexto acadêmico.



Na quarta seção, são expressos os resultados e discussão, com a análise dos dados coletados através do questionário. Por fim, tem-se as considerações finais e as referências.

METODOLOGIA

Este trabalho dispôs-se a comparar o desempenho de alunos e egressos na disciplina de LPI do curso de Ciência da Computação da UEPB - Campus VII, considerando suas opiniões sobre o uso do *ChatGPT* na aprendizagem de Programação, bem como as médias daqueles que puderam utilizar tal ferramenta de estudos complementares relacionados às médias dos que não tiveram o mesmo acesso à Inteligência Artificial.

Para tanto, propôs-se um estudo de natureza primária, conforme as proposições de Wazlawick (2021), no qual realiza-se entrevistas, experimentos, questionários, observações, entre outras formas de capturar os dados no contexto aplicado do cotidiano, para que se defina um novo conhecimento a partir de uma teoria construída para explicá-lo. Quanto aos procedimentos técnicos, de acordo com Wazlawick (2021), aproximou-se de um estudo de caso exploratório, visto que buscou explorar o desempenho de estudantes e egressos, considerando o uso do *ChatGPT* na disciplina de Linguagem de Programação I.

Dois (2) questionários com perguntas semi-estruturadas foram desenvolvidos na plataforma *Google Forms*: um (1) para o grupo de alunos matriculados (com 16 questões) e outro para o grupo de egressos do curso (com 12 questões). Em suma, os questionários traziam questões abertas e de múltipla escolha – estas com uso da escala *Likert* (Likert, 1932) para mensuração das respostas – relacionadas às metodologias de ensino dos professores, métodos de estudos



complementares e uso de IA geradora de texto, e para a obtenção de respostas, foram divulgados nos grupos de *Whatsapp* das turmas e do curso geral, com ajuda de professores e da coordenação.

Como um dos principais parâmetros de comparação, questionou-se aos alunos sobre suas médias finais na disciplina de LPI. Uma média final em qualquer disciplina cursada no referido curso de Computação é obtida a partir de duas (2) notas construídas ao longo de um período letivo: as notas das unidades 1 e 2, respectivamente. Aplicando-se um cálculo comum de média do tipo $(\text{unidade 1} + \text{unidade 2}) / 2$, pode-se, então, obter a média final de cada estudante.

REFERENCIAL TEÓRICO

O primeiro vislumbre do que viria a se consolidar como conceito de Inteligência Artificial é datado de 1950, a partir de um artigo intitulado *Máquinas de Computação e Inteligência (Computing Machinery and Intelligence)* escrito por Alan Turing, no qual o autor inicialmente se propunha a questionar se máquinas poderiam pensar (Turing, 1950).

Um dos conceitos foi definido por McCarthy *et al.* (2004), quando afirma que a IA pode ser entendida como uma ciência ou engenharia usada para criar dispositivos conscientes, especialmente programas de computador inteligentes, e que está relacionada à tarefa similar de usar máquinas para compreender a mente humana.

Desde o teste de Turing, que avaliava uma conversação entre um ser humano e uma máquina e tomava como conclusão final a percepção do ser humano em questão sobre o seu interlocutor – sendo considerado bem sucedido quando o ser humano não conseguia distinguir se seu interlocutor era humano ou máquina – as IAs evoluíram em muitos aspectos. E, hoje, muitas IAs conseguem passar nesse teste (Ludermir, 2021).



A evolução nesta área tem sido avassaladora e há décadas, diferentes tipos têm sido explorados e aperfeiçoados. Na segunda década do século XXI, as IAs podem se diferir em três (3) tipos: IA Focada, IA Generalizada e IA Superinteligente (Ludermir, 2021):

1. A IA Focada ou IA Fraca corresponde a algoritmos especializados em resolução de problemas em áreas específicas na qual os sistemas podem armazenar grandes quantidades de dados e os algoritmos podem realizar as tarefas complexas para as quais foram desenvolvidos, como nos sistemas de recomendação de plataformas de *streaming*, e o próprio *ChatGPT* (Ludermir, 2021).
2. Os algoritmos da IA Generalizada ou IA Forte podem alcançar o mesmo nível de capacidade dos seres humanos para realização de diversas tarefas. Em geral, esse tipo de IA usa aprendizado de máquina como ferramenta. Um exemplo de tarefa em que essa IA é utilizada é a Visão Computacional (Ludermir, 2021).
3. A IA Superinteligente possui algoritmos para resolução de tarefas mais capacitadas do que os seres humanos em quase todos os aspectos. No entanto, ainda se desconhece a existência de IAs superinteligentes e não se sabe se ainda virão a existir (Ludermir, 2021).

Considerando tantas aplicações e possibilidades, os usos da IA tem se ampliado para diferentes áreas da vida humana, inclusive no campo da educação, tanto nos sistemas de ensino online, para a personalização do ensino – no qual se cria um ambiente de estudo adaptado a cada estudante individualmente, permitindo que progrida em seu próprio ritmo – até a identificação de problemas de aprendizagem – quando se monitora o progresso dos alunos e fatores externos que possam atrapalhá-los, ou levar a evasão, possibilitando intervenções mais eficazes por parte dos professores e das instituições de ensino (Costa Júnior *et al.*, 2023).



Esse uso também pode acontecer por parte de docentes e estudantes, na produção de aulas e na resolução de tarefas. Segundo Lo (2023), algumas IA podem ser úteis para docentes, auxiliando na geração de materiais de curso e tarefas de avaliação, como também na tradução de conteúdos; já para os discentes, é capaz de responder às suas dúvidas, resumir informações, verificar conceitos, prestar assistência e fornecer *feedbacks* ao redigir redações, entre outros.

O exemplo mais comum deste tipo de IA, utilizada em ambientes de aprendizado é o *ChatGPT*, que é um sistema de processamento de linguagem natural (*Natural Language Processing - NLP*) baseado em um modelo de aprendizagem profunda projetado para gerar conversações semelhantes às humanas, sendo capaz de compreender um contexto e retornar respostas apropriadas, que podem ser em diferentes idiomas e diferentes estilos, como formal, informal ou até mesmo humorístico (Deng; Lin, 2022).

Ainda assim, estudos apontam que para textos acadêmicos e pesquisas científicas, o *ChatGPT* pode apresentar conteúdo superficial, impreciso ou incorreto, além de infrações éticas como plágio e falta de transparência de dados. É possível também a ocorrência de imprecisão de citações e referências – quando o *ChatGPT* menciona fontes inexistentes, por exemplo – (Sallam, 2023).

No contexto específico de ensino de Programação, em um estudo de caso realizado por Silva Junior *et al.* (2023), concluiu-se que a utilização do *ChatGPT* como ferramenta de estudos complementares pode trazer vantagens no aprendizado, visto que esta IA fornece suporte, *feedback* imediato e exemplos práticos de codificação. Contudo, deve-se considerar os desafios relacionados a esse uso, pois seu emprego excessivo possibilita a limitação do desenvolvimento de habilidades críticas e independentes dos alunos, além de diminuir a interação humana direta, cujo papel é essencial para ensino e promoção de tais competências.



Outras pesquisas vêm sendo publicadas nesta perspectiva, como o trabalho de Husain (2023), no qual se averiguou o uso do *ChatGPT* com instrutores de Programação e estudantes e pode-se verificar que mesmo existindo restrições e inquietações dos professores, é possível explorar o uso do *ChatGPT* no ensino de Programação, que pode auxiliar docentes na produção de aulas e também favorecer estudantes na resolução de problemas de Programação.

Outro trabalho que aborda essa temática, desenvolvido por Yilmaz e Yilmaz (2023), investigou a opinião de estudantes universitários acerca das vantagens e desvantagens do uso do *ChatGPT* nesta área. Identificaram que a maioria dos estudantes investigados indicaram que o emprego da ferramenta nos estudos reduzia o tempo gasto em tarefas que poderia ser utilizado para outros processos de pensamento algorítmico para resolução eficaz de problemas. Também indicou que pode interferir nas habilidades cognitivas e de raciocínio.

RESULTADOS E DISCUSSÕES

Com a aplicação dos questionários, obteve-se no total 102 respostas entre os estudantes matriculados e 12 advindas de estudantes egressos, que ingressaram a partir do ano de 2016 e concluíram ou deixaram o curso até o ano de 2023. Assim, conseguiu-se traçar uma análise correlacionando a disciplina de LPI com a utilização do *ChatGPT*, na perspectiva dos matriculados que tiveram acesso a esta ferramenta em relação aos graduandos e ex-alunos que não usaram o *ChatGPT* em LPI.

Quanto aos alunos com vínculo ainda vigente à instituição, quando questionados sobre como cursaram a disciplina de LPI, 62% dos 102 respondentes descreveram que foi muito tranquila ou



tranquila. Sobre a linguagem de programação utilizada, a maioria, ou seja, 93% cursaram LPI empregando C ou C++.

Continuando a análise sob a ótica dos matriculados, ponderaram que a forma de ensino baseada em *slides*, lousa e demonstração de códigos foi a mais aplicada pelos docentes. Além disso, 84,3% disseram que realizavam estudos complementares fora do horário de aula (nos quais recorreu-se mais ao método de assistir videoaulas em plataformas como o *Youtube*), com 48% afirmando que essas aprendizagens tiveram muitíssima importância para seu desempenho em LPI.

Com respeito ao contato às IAs geradoras de texto enquanto discentes de LPI, 60,8% dos estudantes até então inscritos no curso de Ciência da Computação acessaram o ChatGPT, ao passo que 37,3% não usufruíram destas ferramentas e apenas 2% empregaram outras que não fossem o recurso da *OpenAI*, como *Gemini* e *Bing*. Ademais, quando perguntados acerca da finalidade de uso destas tecnologias, a maioria escolheu a opção de resposta “para garantir um reforço na compreensão dos conteúdos”, conforme averiguado por Husain (2023).

Por sua vez, na perspectiva dos ex-alunos, oito (8) dos 12 respondentes mostraram que é possível considerar o desenrolar da disciplina de LPI como muito ou somente tranquilo. Metade das pessoas abordadas cursaram a matéria com a linguagem de programação C.

Assim como os graduandos, os egressos igualmente afirmaram que os professores utilizaram mais a metodologia baseada em *slides*, lousa e demonstração de códigos. Neste sentido, dez (10) dos 12 interrogados responderam que realizavam estudos complementares fora das aulas em sala (recorrendo muitas vezes a materiais audiovisuais no *YouTube*), sete (7) dos 12 concordaram como muitíssimos importantes estes recursos para os seus desempenhos na matéria



de LPI e todos os 12 assinalaram como “Não” quanto a ter acesso às IAs, visto que se formaram antes do lançamento destas ferramentas.

Verificando todas as médias finais dos alunos em LPI, conforme mostrado na Tabela 1, calculou-se e encontrou-se a média das médias finais para cada grupo investigado, percebe-se que este valor entre os discentes que não utilizaram o *ChatGPT* e que estão com a matrícula em dia se difere de forma superior dos que empregaram este recurso tecnológico numa margem de aproximadamente 1 ponto, e se assemelha mais com a dos egressos, apesar de também levar vantagem.

Tabela 1 – Relação entre a quantidade de alunos e a média das médias finais na disciplina de LPI

Quantidade de alunos/médias	Alunos matriculados	Alunos egressos
Quantidade de alunos - uso do <i>ChatGPT</i>	62	-
Média das médias finais - uso do <i>ChatGPT</i>	7,68	-
Quantidade de alunos - sem uso do <i>ChatGPT</i>	38	12
Média das médias finais - sem uso do <i>ChatGPT</i>	8,62	8,08

Fonte: Elaborado pelos autores, 2024.

Ao analisar os dados obtidos com os questionários e as medidas estatísticas, percebe-se que o uso do *ChatGPT* como ferramenta complementar de estudos na cadeira de LPI não interfere positivamente no desempenho medido pelas notas alcançadas pelos discentes, já que os matriculados sem o contato com a IA se sobressaíram em relação aos demais grupos.

Esta análise pode indicar que a ferramenta deve ser utilizada de maneira consciente e adequada, assim como Lo (2023) especifica, porque a correlação identificada evidencia que usar o *ChatGPT* supostamente impacta negativamente o desempenho acadêmico, pois a aprendizagem possivelmente fica prejudicada, tal como indicado no estudo de Yilmaz e Yilmaz (2023).

CONSIDERAÇÕES FINAIS

Comparados os dados, constatou-se que o desempenho dos alunos entre matriculados e egressos não usuários do *ChatGPT* alcançou maior performance em relação ao grupo de alunos auxiliados pelo uso da IA como ferramenta de estudos complementares, o que indica que não interfere positivamente nos estudos de programação no contexto deste estudo de caso específico. Logo, o objetivo proposto foi atingido.

Houve alguns desafios no trajeto deste estudo, como a baixa participação dos alunos egressos na resolução do questionário dificultou a construção de uma amostra significativa em termos de quantidade. No entanto, uma boa parte dos alunos matriculados que responderam ao questionário faziam parte de períodos avançados e, conseqüentemente, também não puderam aproveitar as funcionalidades da IA, o que equilibrou a base de dados.

É inegável a possibilidade do uso das IAs poder contribuir para o aprendizado e conseqüentemente melhorar o desempenho dos estudantes, logo, é importante que estudos como este ganhem novas páginas e ramificações. Entretanto, é essencial evidenciar também que não está sendo feita uma defesa e apologia ao uso indiscriminado da IA nos estudos de Programação, tampouco demonizar a ferramenta. A ideia é que este trabalho possa favorecer uma reflexão sobre o uso da IA pelos estudantes de uma forma responsável em um contexto de aprendizagem.

Faz-se interessante a ideia da realização de uma nova etapa para este estudo, de caráter experimental, no qual uma turma possa ser dividida em dois grupos, na realização de atividades com e sem o uso do *ChatGPT* e assim serem acompanhados, com registro do passo a passo por um docente ao longo de todo um semestre letivo.



Também é possível pensar na realização de um estudo bibliográfico, de natureza terciária, sobre o tema, tendo em vista que será crescente a investigação nesta linha. Afinal, ampliar a pesquisa para novas áreas do ensino de Computação faz-se necessário, em um momento tão inovador, crescente e irreversível da presença e expansão desta tecnologia na vida dos estudantes e no cenário da educação.

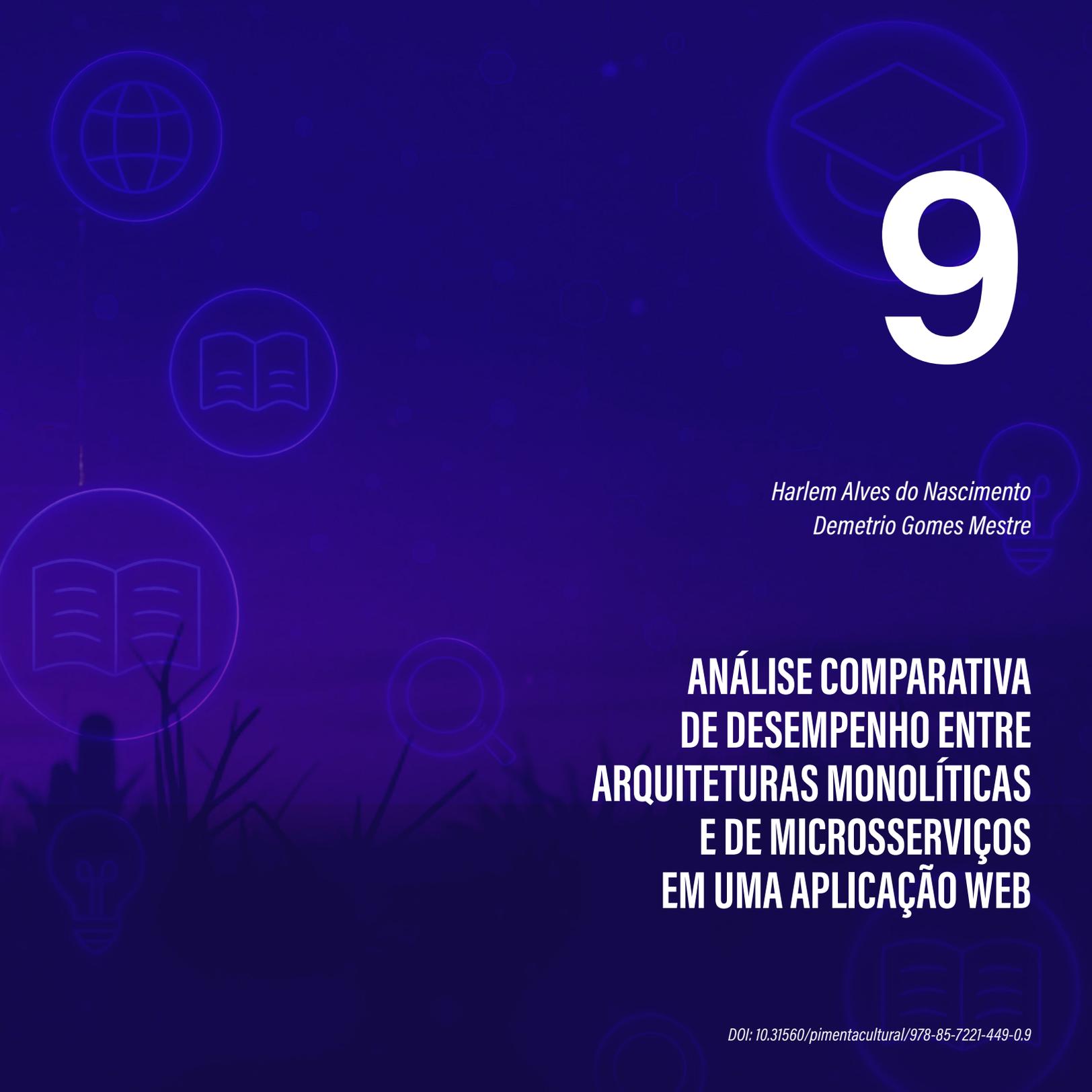
REFERÊNCIAS

BARBOSA, Lucia Martins; PORTES, Luiza Alves Ferreira. A Inteligência Artificial. **Revista Tecnologia Educacional**, Rio de Janeiro, n. 236, p. 16-27, 2023.

COSTA JÚNIOR, João Fernando *et al.* A inteligência artificial como ferramenta de apoio no ensino superior. **Rebena - Revista Brasileira de Ensino e Aprendizagem**, [S. l.], v. 6, p. 246-269, 2023. Disponível em: <https://rebena.emnuvens.com.br/revista/article/view/111>. Acesso em: 21 jun. 2024.

DENG, Jianyang; LIN, Yijia. (2022). *The benefits and challenges of ChatGPT: An overview*. **Frontiers in Computing and Intelligent Systems**, v. 2, n. 2, p. 81-83. Disponível em: <https://drpress.org/ojs/index.php/fcis/article/view/4465>. Acesso em: 26 jun. 2024.

HUSAIN, Anas Jebreen Atyeh. *Potentials of ChatGPT in computer programming: Insights from programming instructors*. **Journal of Information Technology Education**.



9

*Harlem Alves do Nascimento
Demetrio Gomes Mestre*

ANÁLISE COMPARATIVA DE DESEMPENHO ENTRE ARQUITETURAS MONOLÍTICAS E DE MICROSERVIÇOS EM UMA APLICAÇÃO WEB

AUTORES

Harlem Alves do Nascimento

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

Email: harlem.nascimento@servidor.uepb.edu.br

Demetrio Gomes Mestre

Universidade Estadual da Paraíba - UEPB, Campus Patos – PB.

E-mail: demetrio.mestre@servidor.uepb.edu.br



ABSTRACT:

This study compares monolithic and microservices architectures in a chat application to analyze scalability, resource efficiency, and response time differences. The methodology involved developing a monolithic application and converting it to microservices, with load tests simulating realistic scenarios. The results indicated that the microservices architecture allows selective scaling, while the monolithic architecture requires global application scaling. Efficiency in resource use depended on workload. As for latency and response time, monolithic technology stood out in some areas, while microservices were superior in others, with an impact due to the number of internal API calls.



RESUMO:

Este estudo compara as arquiteturas monolítica e de microsserviços em uma aplicação de chat para analisar diferenças em escalabilidade, eficiência no uso de recursos e tempo de resposta. A metodologia envolveu o desenvolvimento de uma aplicação monolítica e sua conversão para microsserviços, com testes de carga simulando cenários realistas. Os resultados indicaram que a arquitetura de microsserviços permite escalabilidade seletiva, enquanto a arquitetura monolítica exige escalabilidade global da aplicação. A eficiência no uso de recursos depende da carga de trabalho. Quanto à latência e ao tempo de resposta, a tecnologia monolítica se destacou em algumas áreas, enquanto os microsserviços foram superiores em outras, com impacto devido ao número de chamadas internas de API.



INTRODUÇÃO

O crescimento explosivo da internet tem levado a um aumento massivo na geração de dados, exigindo que aplicações *web* sejam projetadas para lidar com volumes cada vez maiores de tráfego. De acordo com Statista (2021), a quantidade de dados gerados anualmente passou de dois zettabytes em 2010 para 64,2 zettabytes em 2020, com previsões para atingir 120 *zettabytes* até 2023 e 181 *zettabytes* até 2025. Dessa forma, é crucial que aplicações e serviços *web* garantam a eficiência e capacidade de processamento suficientes para suportar o aumento desse volume de dados.

Para lidar com esse problema, a escolha da arquitetura de software se torna um fator crítico no desenvolvimento de aplicações *web*. Como mencionado por Bosch (1999), a arquitetura de *software* afeta diretamente o desempenho, a robustez e a capacidade de manutenção de um sistema. No contexto da comparação entre arquiteturas monolíticas e de microsserviços, essa decisão estratégica influencia não apenas a eficiência do desenvolvimento, mas também a capacidade de lidar com cenários de alta carga.

O propósito deste estudo é comparar o desempenho de duas abordagens arquiteturais distintas para o desenvolvimento de *software*: a arquitetura monolítica e a de microsserviços. O objetivo é fornecer percepções para decisões informadas no cenário do desenvolvimento de sistemas avaliando o comportamento de um aplicativo *web* construído com essas duas abordagens.

REFERENCIAL TEÓRICO

Essa seção aborda os principais autores e estudos, contextualizando o tema e destacando as bases conceituais que



sustentam a pesquisa, incluindo discussões sobre o modelo cliente-servidor, as arquiteturas monolítica e de microsserviços, além dos trabalhos relacionados.

O modelo cliente-servidor, concebido na década de 1970 pela divisão de pesquisa Xerox PARC, tornou-se comum em aplicações baseadas em Protocolo de Controle de Transmissão (TCP) ou Protocolo da Internet (IP), conforme afirmado por Rosen e Shklar (2009). Nesse modelo, os servidores, também chamados de serviços, aguardam solicitações de programas clientes para, em seguida, processá-las. Embora os navegadores web sejam exemplos típicos de clientes, os servidores também podem atuar como clientes ao fazer solicitações a outros servidores.

Uma aplicação monolítica é caracterizada por um sistema em que todos os componentes e módulos são executados em um único processo, sem possibilidade de operação independente (Dragoni *et al.*, 2017). Embora essa arquitetura ofereça simplicidade na gestão, encapsulando todos os componentes em um único local, ela apresenta desafios significativos em aplicações complexas a longo prazo.

De acordo com Fowler e Lewis (2014), a arquitetura de microsserviços divide uma aplicação em pequenos serviços independentes, cada um operando de forma autônoma em seu próprio processo e comunicando-se por meio de Interfaces de Programação de Aplicativos (API) leves. Essa abordagem permite que cada serviço utilize diferentes linguagens de programação e bancos de dados, sendo substituídos e atualizados de forma independente. Além disso, cada componente pode gerenciar sua própria base de dados, adotando o conceito de Persistência Poliglota, onde serviços podem utilizar diferentes tecnologias de banco de dados.

Foram realizadas pesquisas sobre publicações relevantes com o objetivo de avaliar as contribuições existentes na área de



interesse deste trabalho. Foram selecionados três artigos que abordavam temas correlatos, utilizando a *IEEE Xplore* para fontes em inglês e o Google Acadêmico para trabalhos em português do Brasil.

Al-Debagy e Martinek (2018) compararam arquiteturas monolíticas e de microsserviços, mostrando que a monolítica teve melhor desempenho com menos usuários, mas as duas se igualaram com mais usuários. A configuração com *onsul* superou a com *Eureka*, ferramentas usadas para a descoberta de serviços e gerenciamento de configuração, em *throughput*¹.

Duarte (2017) analisou as vantagens e desvantagens das arquiteturas monolítica e de microsserviços implementando um sistema de gerenciamento de cinemas. O estudo destacou que a arquitetura monolítica oferece melhor tempo de resposta, enquanto microsserviços são mais eficazes em projetos maiores, apesar do maior tempo de comunicação.

Amaral e Carvalho (2017) concluíram que monólitos são mais eficientes para aplicações simples, enquanto microsserviços são melhores para escalabilidade e isolamento funcional, embora com maiores custos de infraestrutura.

METODOLOGIA

Segundo Wazlawick (2009), este estudo pode ser categorizado como experimental quantitativo, pois envolve a coleta e análise de dados numéricos para avaliar e comparar o desempenho

1 Throughput refere-se à quantidade de trabalho ou dados que um sistema consegue processar em um determinado intervalo de tempo.



de diferentes arquiteturas de software. Para iniciar a comparação, foi desenvolvida uma aplicação monolítica e, logo após, foi criada uma nova versão a partir de sua decomposição em microsserviços, baseado no modelo proposto por Fowler e Lewis (2014).

Foi desenvolvido um serviço de bate-papo *web* nos dois modelos de arquitetura. O *chat* permite aos usuários o registro e edição de perfil, *login* e *logout*, conversas privadas e públicas, além de notificações de novas mensagens. Através do *Docker*² e *Kubernetes*³, foi configurado o ambiente de execução para garantir recursos controlados e condições de teste idênticas para ambas as aplicações.

Para o desenvolvimento, foi utilizado o *Java* com o *framework Spring Boot* para o *backend*, *Javascript* com *ReactJS* para o *frontend* e *MySQL* para a persistência de dados. A máquina *host* é equipada com um processador *Xeon E5 2620V3* de 6 núcleos físicos e 12 *threads*, e possui 16 *gigabytes* de Memória de Acesso Aleatório (RAM) tendo como sistema operacional o *Ubuntu 22.04.3 LTS*. Destes recursos, foram destinados 6000 *milicores*⁴ e 10 *gigabytes* de memória RAM para o *Kubernetes*, restando 6000 *milicores* e 8 *gigabytes* para o Sistema Operacional. Nos recursos restantes, será executado o *Kubernetes*, *Docker* e o *Apache JMeter*.

- 2 O *Docker* permite empacotar e executar aplicativos em ambientes isolados chamados contêineres, que são leves, auto suficientes e seguros. Isso facilita a execução de vários contêineres em uma mesma máquina e o compartilhamento consistente desses ambientes de trabalho (*Docker*, 2023).
- 3 O *Kubernetes* (K8s) é um sistema de código aberto que automatiza o gerenciamento de aplicações em contêineres, simplificando sua administração ao agrupar os contêineres de forma lógica (*Kubernetes*, 2023).
- 4 Os *milicores* são representados pela unidade "m" e são uma medida que quantifica a capacidade de processamento da CPU. Essa métrica expressa a fração de um núcleo de CPU que um contêiner pode utilizar. Por exemplo, se for alocado "500m" (ou 0,5 núcleos), isso significa que o contêiner pode usar 50% da capacidade de processamento de um único núcleo da CPU.



Tabela 1 – Arquitetura da Aplicação Monolítica

Serviço	Solicitado		Limite	
	CPU (m)	Memória (m)	CPU (m)	Memória (m)
Backend	1000	1024	5000	8192
Database	500	500	3000	1024
Proxy	50	100	200	250
FrontEnd	50	250	150	500
Total	1600	1874	8350	9966

Fonte: Elaborado pelos autores, 2024.

A Tabela 1 mostra a distribuição de recursos da arquitetura monolítica no *Kubernetes*, onde os seis núcleos do processador são representados por 6000 *milicores*. Como a arquitetura monolítica centraliza todos os serviços em uma única aplicação (*BACKEND*), esta recebeu a maior parte dos recursos de hardware.

Conforme apresentado na Tabela 2, a arquitetura da aplicação é composta por serviços especializados que, na arquitetura monolítica, estavam centralizados em uma única aplicação. Esses serviços incluem autenticação (*Auth*), gerenciamento de perfis (*User*), envio de e-mails (*email-worker*), direcionamento de solicitações e segurança (*API Gateway*), processamento de imagens (*image-worker*) e comunicação em tempo real (*Message*) utilizando *RabbitMQ*. Para otimizar a distribuição dos 6000 *milicores*, que eram utilizados na aplicação monolítica, foi necessário realizar diversos testes durante a construção do novo esquema a fim de otimizar a distribuição de recursos.

Tabela 2 – Arquitetura da aplicação baseada em microsserviços

Serviço	Solicitado			Limite			
	CPU (m)	Mem. (m)	Max. de Réplicas	CPU mínimo (m)	CPU máximo (m)	Memória mínima (m)	Memória máxima (m)
frontend	50	250	1	150	150	700	700
proxy	50	100	1	200	250	700	700
auth	200	250	3	1000	1000	500	1500
user	200	250	3	1000	3000	500	1500
email-worker	200	250	3	1000	3000	500	1500
api-gateway	200	250	3	1000	3000	500	1500
image-worker	200	250	3	1000	3000	1000	3000
message	200	250	3	1000	3000	500	1500
rabbitmq	100	250	1	250	250	600	600
auth-db	100	250	1	250	250	600	600
user-db	100	250	1	250	250	600	600
email-db	100	250	1	250	250	600	600
image-db	100	250	1	250	250	600	600
message-db	100	250	1	250	250	600	600
Total	2000	3350	24	8100	18150	8900	14900

Fonte: Elaborado pelos autores, 2024.

A principal métrica considerada é a quantidade de recursos solicitados, pois são alocados garantidamente para os serviços. Os limites são fornecidos apenas se houver recursos disponíveis, dessa forma, mesmo que a máquina *host* possua 6000 *milicores*, todos não serão alocados para um único serviço, pois parte destes podem estar em uso por outros processos da arquitetura alvo dos testes.

Para avaliar o desempenho de cada arquitetura, realizamos uma série de testes de carga usando o *Apache JMeter*. Os testes envolveram operações como *login*, consulta de mensagens, conexão



via *WebSocket*⁵ e persistência de imagens de perfil, simulando situações de uso real para ambas as arquiteturas.

O ambiente de teste foi criado para ser o mais imparcial possível, com *containers Docker* gerenciados pelo *Kubernetes*, garantindo alocações específicas de recursos de *hardware*. Os testes foram projetados para avaliar aspectos como escalabilidade, eficiência de recursos e desempenho em diferentes condições de carga.

Os testes foram conduzidos com tempos específicos e quantidades de usuários determinados a partir de testes prévios para avaliar o máximo estresse na máquina *host*. O teste de *login*, com 2000 usuários em 100 segundos, verificou a eficiência da autenticação.

Para a consulta de mensagens, com 6000 usuários autenticados em 100 segundos, foi avaliada a recuperação de mensagens públicas. A conexão *WebSocket*, responsável pela interação em tempo real com o servidor, foi executada com 3000 usuários durante 100 segundos. Por fim, o teste de persistência de imagens de perfil, envolvendo 1000 usuários em 100 segundos, analisou a capacidade de processamento de alterações de imagem.

Esses testes foram selecionados devido à sua importância crítica para o funcionamento da aplicação e permitiram uma avaliação do desempenho em condições de carga intensa.

RESULTADOS E DISCUSSÕES

A **Figura 1** apresenta um resumo quantitativo dos resultados obtidos nos testes realizados na aplicação monolítica. No teste

5 O Protocolo **WebSocket**, conforme descrito por Fette e Melnikov (2011), é um modelo de comunicação bidirecional de baixa latência entre cliente e servidor

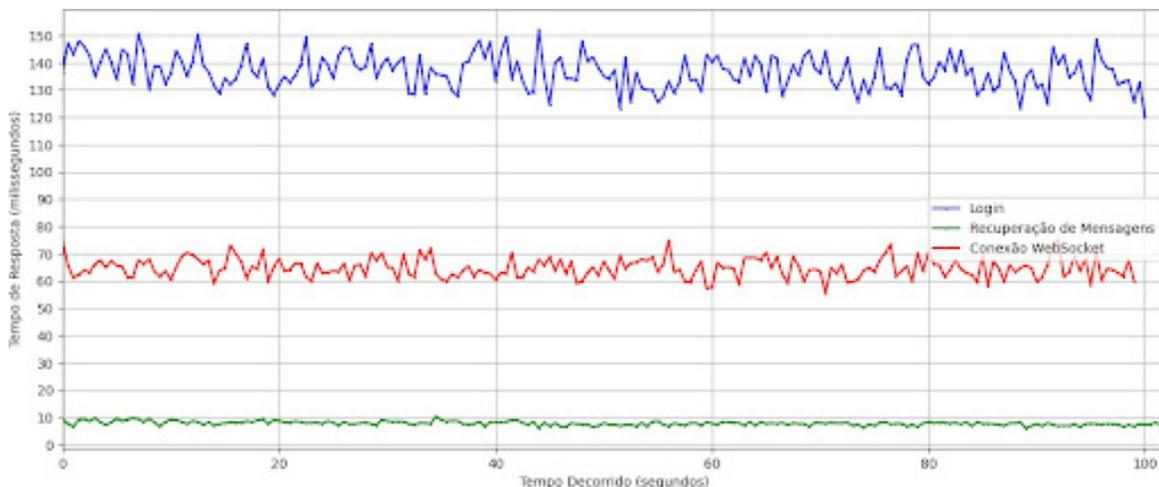


de *login*, a arquitetura monolítica registrou um tempo de resposta médio de 136 milissegundos, com um pico de 152,18 milissegundos. O consumo de recursos foi moderado, totalizando 2536 *milicores* de uso da Unidade Central de Processamento (CPU) e 897 *megabytes* de memória.

Na avaliação do tempo de recuperação de mensagens, a arquitetura monolítica apresentou um tempo de resposta médio de 7 milissegundos e um pico de 10,27 milissegundos. O consumo de recursos durante essa etapa foi de 632 *milicores* de CPU e 1002 *megabytes* de memória.

O teste de conexão *WebSocket* na arquitetura monolítica revelou um tempo de resposta médio de 64 milissegundos, com um pico de 75,13 milissegundos. A utilização de recursos para essa funcionalidade totalizou 1073 *milicores* de CPU e 984 *megabytes* de memória.

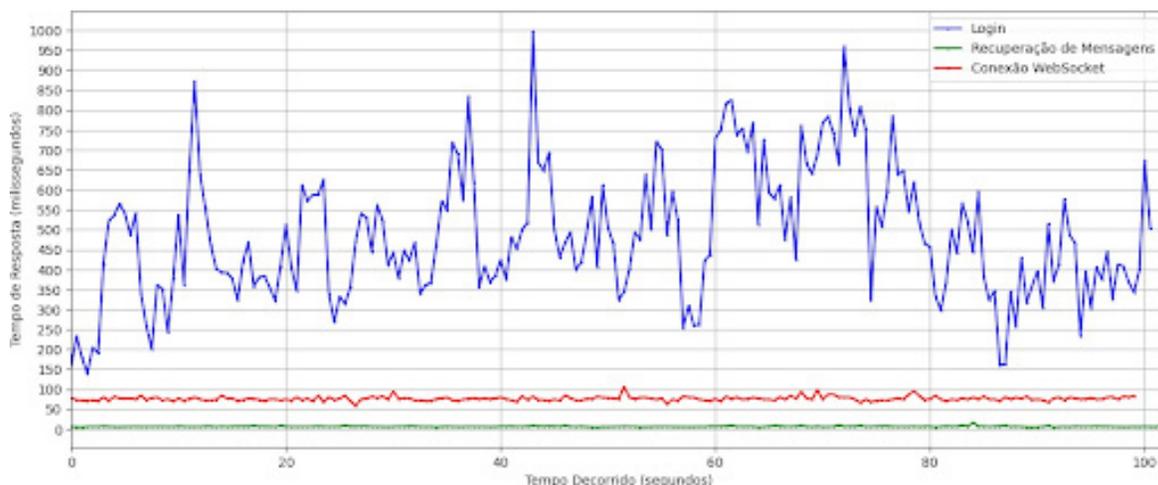
Figura 1 - Testes na aplicação monolítica



Fonte: Elaborado pelo autor, 2024.



Figura 2 - Testes na aplicação baseada em microsserviços



Fonte: Elaborado pelo autor, 2024.

A Figura 2 apresenta um resumo quantitativo dos resultados obtidos nos testes realizados na aplicação baseada em microsserviços. No teste de *login*, a arquitetura de microsserviços registrou uma média de tempo de resposta de 491 milissegundos, com um pico de 998,5 milissegundos. O consumo de recursos para esse teste foi de 2796 *milicores* de CPU e 4173 *megabytes* de memória.

No teste de recuperação de mensagens, a arquitetura de microsserviços apresentou um tempo de resposta médio de 6 milissegundos, com um pico de 16,23 milissegundos. Durante essa etapa, a utilização de recursos foi de 714 *milicores* de CPU e 4062 *megabytes* de memória.

O teste de conexão *WebSocket* na arquitetura de microsserviços indicou um tempo de resposta médio de 76 milissegundos, com um pico de 106,8 milissegundos. O consumo de recursos para essa funcionalidade totalizou 1721 *milicores* de CPU e 4222 *megabytes* de memória.



Os resultados dos testes revelam diferenças significativas entre as arquiteturas monolítica e de microsserviços. No teste de login, a arquitetura monolítica demonstrou tempos de resposta mais rápidos, indicando uma eficiência superior na autenticação dos usuários, além de exigir menos recursos em comparação com a arquitetura de microsserviços.

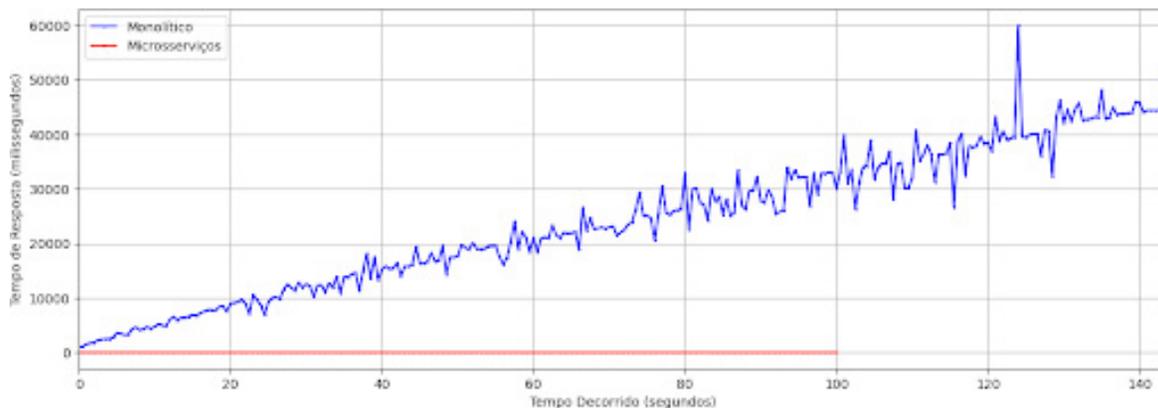
A latência adicional na arquitetura de microsserviços ocorre principalmente devido às chamadas internas entre o servidor de dados do usuário (*user*) e o servidor de autenticação (*auth*) para que só depois seja retornado ao *API Gateway*.

Por outro lado, na recuperação de mensagens, ambas as arquiteturas apresentaram desempenho semelhante em termos de tempo de resposta, mas a arquitetura monolítica revelou um consumo de recursos mais moderado. Isso se deve ao fato do micro serviço de mensagens depender apenas do serviço de usuários para fornecer uma resposta, dessa forma, os dados são recuperados da base de dados diretamente.

Na conexão *WebSocket*, a arquitetura monolítica também exibiu tempos de resposta mais ágeis e utilizou menos recursos do sistema em comparação com a arquitetura de microsserviços. Essas discrepâncias podem ser atribuídas à latência de comunicação interna entre os serviços da arquitetura de microsserviços, onde cada chamada interna requer um tempo adicional para ser processada, impactando diretamente o tempo de resposta e o consumo de recursos.



Figura 3 - Testes na rota de persistência de imagens



Fonte: Elaborado pelo autor, 2024.

A Figura 3 mostra os resultados do teste de persistência de imagens. A arquitetura monolítica teve um tempo de resposta médio de 24,577 segundos, chegando a um pico de 61,791 segundos, com um consumo de 5029 *milicores* de CPU e 2945 *megabytes* de memória. Por outro lado, a arquitetura de microserviços teve um tempo de resposta médio de 19 milissegundos e um pico de 152 milissegundos, com um consumo total de 3769 *milicores* de CPU e 4959 *megabytes* de memória.

No contexto da análise do fluxo de teste de imagens, é de suma importância ressaltar a eficácia da arquitetura de microserviços em lidar com essa demanda específica. Tal eficácia deriva da habilidade dessa arquitetura em limitar o uso de recursos de hardware em áreas específicas da aplicação.

Enquanto na aplicação monolítica todo o poder de processamento é direcionado à aplicação como um todo, resultando em sobrecarga quando múltiplas requisições de processamento de imagens são recebidas, na arquitetura baseada em microserviços,



cada instância foi alocada com uma quantidade específica de recursos, limitando o uso máximo a 3000 *milicores*, conforme detalhado na Tabela 2.

Essa abordagem evita a sobrecarga de toda a aplicação, reservando 2000 *milicores* para outras tarefas, evidenciando a capacidade da arquitetura de microsserviços em gerenciar eficientemente os recursos e assegurar a estabilidade do sistema, mesmo em situações de alta demanda.

CONSIDERAÇÕES FINAIS

Este estudo proporcionou uma análise comparativa entre arquiteturas monolítica e de microsserviços em uma aplicação de chat, destacando suas diferenças em relação à escalabilidade, utilização de recursos, latência e desempenho.

Os resultados revelaram que as arquiteturas divergem significativamente em diferentes funcionalidades, com vantagens e desafios distintos em cada abordagem. Enquanto a arquitetura monolítica demonstrou tempos de resposta mais rápidos em algumas áreas, como autenticação, a arquitetura de microsserviços se destacou na gestão eficiente de recursos, especialmente na persistência de imagens.

No entanto, o estudo também ressalta a importância de otimizar a arquitetura de microsserviços, aprimorar a segurança e a tolerância a falhas, além de considerar o impacto na experiência do usuário como direções para futuras pesquisas.

REFERÊNCIAS

AL-DEBAGY, O.; MARTINEK, P. **A comparative review of microservices and monolithic architectures.** *In*: 2018 IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI). Budapest, Hungary: IEEE, 2018. p. 149–154. Disponível em: <https://ieeexplore.ieee.org/abstract/document/8928192>. Acesso em: 6 set. 2023.

AMARAL, O.; CARVALHO, M. **Arquitetura de Micro Serviços: Uma comparação com sistemas monolíticos.** Universidade Federal da Paraíba, 2017. Disponível em: <https://repositorio.ufpb.br/jspui/handle/123456789/3235>. Acesso em: 6 set. 2023.

APACHE. **Apache JMeter.** 2023. Disponível em: <https://jmeter.apache.org/>. Acesso em: 19 set. 2023.

BOSCH, J. Design and use of industrial software architectures. *In*: **Proceedings Technology of Object-Oriented Languages and Systems. TOOLS 29 (Cat. No.PR00275).** [S.l.: s.n.], 1999. p. 404–404.

DOCKER. **Docker overview.** 2023. Disponível em: <https://docs.docker.com/get-started/overview/>. Acesso em: 19 set. 2023.

DRAGONI, N. *et al.* Microservices: Yesterday, today, and tomorrow. *In*: MAZZARA, M.; MEYER, B. **Present and Ulterior Software Engineering.** Cham: Springer International Publishing, 2017. p. 195–216. Disponível em: https://doi.org/10.1007/978-3-319-67425-4_12. Acesso em: 6 set. 2023.

DUARTE, J. P. L. **Análise comparativa entre arquitetura monolítica e de microsserviços.** Universidade Federal de Santa Catarina, 2017. Disponível em: <https://repositorio.ufsc.br/handle/123456789/182309>. Acesso em: 6 set. 2023.

FETTE, I.; MELNIKOV, A. **The websocket protocol.** 2011. Disponível em: <https://www.rfc-editor.org/rfc/rfc6455>. Acesso em: 7 set. 2023.

FOWLER, M.; LEWIS, J. **Microservices.** 2014. Disponível em: <http://martinfowler.com/articles/microservices.html>. Acesso em: 6 set. 2023.

KUBERNETES. **Kubernetes.** 2023. Disponível em: <https://kubernetes.io/>. Acesso em: 19 set. 2023.



ROSEN, L.; SHKLAR, L. Before the web: Tcp/ip. *In*: ROSEN, L.; SHKLAR, L. **Web application architecture: Principles, protocol and practices**. 1. ed. Chichester, West Sussex, Inglaterra: John Wiley & Sons Ltd, 2009. p. 11–28.

STATISTA. **Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025**. 2021. Disponível em: <https://www.statista.com/statistics/871513/worldwide-data-created/>. Acesso em: 19 set. 2023

WAZLAWICK, R. S. Preparação de um trabalho de pesquisa. *In*: WAZLAWICK, R. S. **Metodologia de pesquisa para ciência da computação**. Rio de Janeiro: Elsevier, 2009. v. 2.

ÍNDICE REMISSIVO

A

accuracy 35, 36, 37, 38, 83, 96
 aprendizagem ativa 20, 60, 69
 aprendizagem baseada em projetos 25, 58
 árvore de decisão 85, 86, 89, 90, 91, 92
 autonomia 57, 67, 74, 76

B

Brincar de Aprender 74, 75, 76

C

cibridismo 57, 59
 comunicação 22, 31, 43, 46, 47, 48, 49, 51, 128, 130, 132, 135
 confiabilidade 32, 33, 88, 89, 100, 102

D

detecção de spam 29, 31, 34, 37, 38

E

ecossistema tecnológico 16, 22, 24
 Educação Especial 75, 80
 estágio supervisionado 43, 44, 45, 46, 47, 48, 50
 experiência 74, 76, 137

F

F1-score 33, 83, 84, 86, 87, 89, 90, 91, 96, 97, 99, 100, 102, 103, 104

G

gamificação 58, 80

H

hard skills 42, 43, 44, 45, 48, 49, 50

J

jogos educativos 73, 75

K

K-Nearest Neighbors 28, 29, 35, 36, 37, 92, 105

M

metodologias ativas 56, 57, 58, 60, 61, 62, 64, 67, 68

N

Naive Bayes Multinomial 28, 29, 35, 36

O

onlife 57, 59, 67
 Orientação a Objetos 18, 19
 Oversampling 35, 37

P

precision 28, 35, 36, 83, 96
 Processamento de Linguagem Natural 29, 31, 33
 PROGNITIVE 10, 13, 15, 16, 17, 18, 21, 24
 Programação 10, 13, 17, 62, 64, 66, 68, 110, 111, 112, 113, 116, 117, 120, 127

R

raciocínio lógico 16, 20, 21, 24, 75, 78
 recall 28, 29, 31, 33, 35, 36, 37, 38, 83, 84, 86, 87, 89, 90, 96, 97, 99, 100, 102, 103
 Regressão Logística 29, 35, 36
 Road Runner 20, 21
 robótica 16, 17, 18, 19, 20, 21, 22, 24

S

Screen Programming 66, 67
 Síndrome de Down 11, 70, 73, 74
 soft skills 42, 43, 44, 45, 48, 49, 52
 STEAM 15, 16, 17, 20, 21, 24, 25
 Support Vector Classifier 28, 29, 35, 36

T

trabalho em equipe 16, 20, 22, 46, 51

U

undersampling 34, 37

www.PIMENTACULTURAL.com

anais
do **VI** CONGRESSO
SERTANEJO
DE COMPUTAÇÃO

AVANÇOS EM COMPUTAÇÃO APLICADA
Robótica, IA e Educação no Sertão

