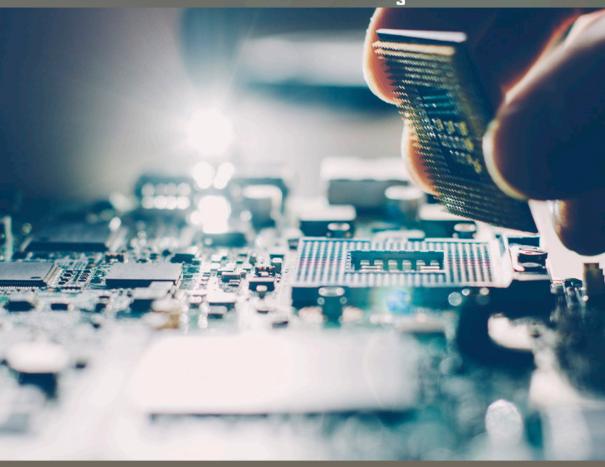
DESAFIOS DAS ENGENHARIAS:

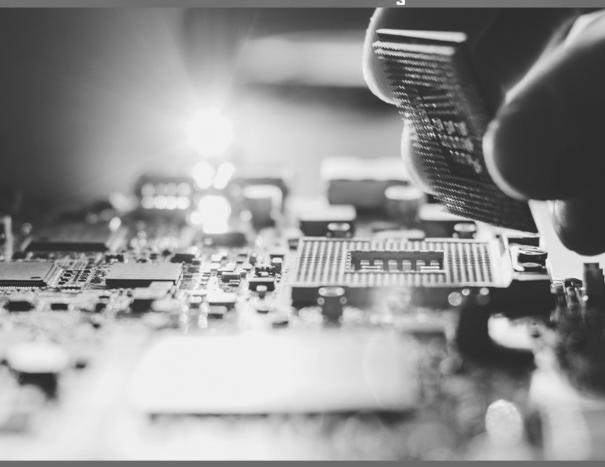
ENGENHARIA DE COMPUTAÇÃO 2



Ano 2021

DESAFIOS DAS ENGENHARIAS:

ENGENHARIA DE COMPUTAÇÃO 2



ERNANE ROSA MARTINS (ORGANIZADOR)



Editora chefe

Prof^a Dr^a Antonella Carvalho de Oliveira

Assistentes editoriais

Natalia Oliveira

Flávia Roberta Barão

Bibliotecária

Janaina Ramos

Proieto gráfico

Natália Sandrini de Azevedo

Camila Alves de Cremo

Luiza Alves Batista

Maria Alice Pinheiro 2021 by Atena Editora

Imagens da capa Copyright © Atena Editora

iStock Copyright do Texto © 2021 Os autores

iStock Copyright do Texto © 2021 Os autores

Edição de arte Copyright da Edição © 2021 Atena Editora

Luiza Alves Batista Direitos para esta edição cedidos à Atena

Revisão Editora pelos autores.

Os autores Open access publication by Atena Editora



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

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

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

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

Conselho Editorial

Ciências Humanas e Sociais Aplicadas

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

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

Prof^a Dr^a Andréa Cristina Marques de Araújo – Universidade Fernando Pessoa

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

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

Prof. Dr. Antonio Isidro-Filho - Universidade de Brasília



Prof. Dr. Arnaldo Oliveira Souza Júnior - Universidade Federal do Piauí

Prof. Dr. Carlos Antonio de Souza Moraes - Universidade Federal Fluminense

Prof. Dr. Crisóstomo Lima do Nascimento - Universidade Federal Fluminense

Profa Dra Cristina Gaio - Universidade de Lisboa

Prof. Dr. Daniel Richard Sant'Ana - Universidade de Brasília

Prof. Dr. Devvison de Lima Oliveira - Universidade Federal de Rondônia

Profa Dra Dilma Antunes Silva - Universidade Federal de São Paulo

Prof. Dr. Edvaldo Antunes de Farias - Universidade Estácio de Sá

Prof. Dr. Elson Ferreira Costa - Universidade do Estado do Pará

Prof. Dr. Eloi Martins Senhora - Universidade Federal de Roraima

Prof. Dr. Gustavo Henrique Cepolini Ferreira - Universidade Estadual de Montes Claros

Prof. Dr. Humberto Costa - Universidade Federal do Paraná

Prof^a Dr^a Ivone Goulart Lopes – Istituto Internazionele delle Figlie de Maria Ausiliatrice

Prof. Dr. Jadson Correia de Oliveira - Universidade Católica do Salvador

Prof. Dr. José Luis Montesillo-Cedillo - Universidad Autónoma del Estado de México

Prof. Dr. Julio Candido de Meirelles Junior - Universidade Federal Fluminense

Profa Dra Lina Maria Gonçalves - Universidade Federal do Tocantins

Prof. Dr. Luis Ricardo Fernandes da Costa - Universidade Estadual de Montes Claros

Profa Dra Natiéli Piovesan - Instituto Federal do Rio Grande do Norte

Prof. Dr. Marcelo Pereira da Silva - Pontifícia Universidade Católica de Campinas

Prof^a Dr^a Maria Luzia da Silva Santana - Universidade Federal de Mato Grosso do Sul

Prof. Dr. Miguel Rodrigues Netto – Universidade do Estado de Mato Grosso

Prof. Dr.Pablo Ricardo de Lima Falcão - Universidade de Pernambuco

Prof^a Dr^a Paola Andressa Scortegagna – Universidade Estadual de Ponta Grossa

Prof^a Dr^a Rita de Cássia da Silva Oliveira - Universidade Estadual de Ponta Grossa

Prof. Dr. Rui Maia Diamantino - Universidade Salvador

Prof. Dr. Saulo Cerqueira de Aguiar Soares – Universidade Federal do Piauí

Prof. Dr. Urandi João Rodrigues Junior - Universidade Federal do Oeste do Pará

Prof^a Dr^a Vanessa Bordin Viera – Universidade Federal de Campina Grande

Profa Dra Vanessa Ribeiro Simon Cavalcanti - Universidade Católica do Salvador

Prof. Dr. William Cleber Domingues Silva - Universidade Federal Rural do Rio de Janeiro

Prof. Dr. Willian Douglas Guilherme - Universidade Federal do Tocantins

Ciências Agrárias e Multidisciplinar

Prof. Dr. Alexandre Igor Azevedo Pereira - Instituto Federal Goiano

Prof. Dr. Arinaldo Pereira da Silva - Universidade Federal do Sul e Sudeste do Pará

Prof. Dr. Antonio Pasqualetto - Pontifícia Universidade Católica de Goiás

Profa Dra Carla Cristina Bauermann Brasil - Universidade Federal de Santa Maria

Prof. Dr. Cleberton Correia Santos - Universidade Federal da Grande Dourados

Prof^a Dr^a Diocléa Almeida Seabra Silva - Universidade Federal Rural da Amazônia

Prof. Dr. Écio Souza Diniz - Universidade Federal de Viçosa

Prof. Dr. Fábio Steiner - Universidade Estadual de Mato Grosso do Sul

Prof. Dr. Fágner Cavalcante Patrocínio dos Santos - Universidade Federal do Ceará

Profa Dra Girlene Santos de Souza - Universidade Federal do Recôncavo da Bahia

Prof. Dr. Jael Soares Batista – Universidade Federal Rural do Semi-Árido

Prof. Dr. Jayme Augusto Peres - Universidade Estadual do Centro-Oeste

Prof. Dr. Júlio César Ribeiro – Universidade Federal Rural do Rio de Janeiro

Profa Dra Lina Raquel Santos Araújo - Universidade Estadual do Ceará

Prof. Dr. Pedro Manuel Villa - Universidade Federal de Viçosa

Prof^a Dr^a Raissa Rachel Salustriano da Silva Matos - Universidade Federal do Maranhão

Prof. Dr. Ronilson Freitas de Souza - Universidade do Estado do Pará

Profa Dra Talita de Santos Matos - Universidade Federal Rural do Rio de Janeiro



Prof. Dr. Tiago da Silva Teófilo – Universidade Federal Rural do Semi-Árido Prof. Dr. Valdemar Antonio Paffaro Junior – Universidade Federal de Alfenas

Ciências Biológicas e da Saúde

Prof. Dr. André Ribeiro da Silva - Universidade de Brasília

Profa Dra Anelise Levay Murari - Universidade Federal de Pelotas

Prof. Dr. Benedito Rodrigues da Silva Neto - Universidade Federal de Goiás

Prof^a Dr^a Daniela Reis Joaquim de Freitas – Universidade Federal do Piauí

Profa Dra Débora Luana Ribeiro Pessoa - Universidade Federal do Maranhão

Prof. Dr. Douglas Siqueira de Almeida Chaves - Universidade Federal Rural do Rio de Janeiro

Prof. Dr. Edson da Silva - Universidade Federal dos Vales do Jequitinhonha e Mucuri

Prof^a Dr^a Elizabeth Cordeiro Fernandes – Faculdade Integrada Medicina

Profa Dra Eleuza Rodrigues Machado - Faculdade Anhanguera de Brasília

Profa Dra Elane Schwinden Prudêncio - Universidade Federal de Santa Catarina

Prof^a Dr^a Eysler Gonçalves Maia Brasil – Universidade da Integração Internacional da Lusofonia Afro-Brasileira

Prof. Dr. Ferlando Lima Santos - Universidade Federal do Recôncavo da Bahia

Prof^a Dr^a Fernanda Miguel de Andrade – Universidade Federal de Pernambuco

Prof. Dr. Fernando Mendes - Instituto Politécnico de Coimbra - Escola Superior de Saúde de Coimbra

Prof^a Dr^a Gabriela Vieira do Amaral – Universidade de Vassouras

Prof. Dr. Gianfábio Pimentel Franco - Universidade Federal de Santa Maria

Prof. Dr. Helio Franklin Rodrigues de Almeida - Universidade Federal de Rondônia

Prof^a Dr^a Iara Lúcia Tescarollo – Universidade São Francisco

Prof. Dr. Igor Luiz Vieira de Lima Santos - Universidade Federal de Campina Grande

Prof. Dr. Jefferson Thiago Souza - Universidade Estadual do Ceará

Prof. Dr. Jesus Rodrigues Lemos - Universidade Federal do Piauí

Prof. Dr. Jônatas de França Barros – Universidade Federal do Rio Grande do Norte

Prof. Dr. José Max Barbosa de Oliveira Junior - Universidade Federal do Oeste do Pará

Prof. Dr. Luís Paulo Souza e Souza - Universidade Federal do Amazonas

Profa Dra Magnólia de Araújo Campos - Universidade Federal de Campina Grande

Prof. Dr. Marcus Fernando da Silva Praxedes - Universidade Federal do Recôncavo da Bahia

Prof^a Dr^a Maria Tatiane Gonçalves Sá - Universidade do Estado do Pará

Profa Dra Mylena Andréa Oliveira Torres - Universidade Ceuma

Profa Dra Natiéli Piovesan - Instituto Federacl do Rio Grande do Norte

Prof. Dr. Paulo Inada - Universidade Estadual de Maringá

Prof. Dr. Rafael Henrique Silva - Hospital Universitário da Universidade Federal da Grande Dourados

Profa Dra Regiane Luz Carvalho - Centro Universitário das Faculdades Associadas de Ensino

Profa Dra Renata Mendes de Freitas - Universidade Federal de Juiz de Fora

Profa Dra Vanessa da Fontoura Custódio Monteiro - Universidade do Vale do Sapucaí

Prof^a Dr^a Vanessa Lima Gonçalves – Universidade Estadual de Ponta Grossa

Profa Dra Vanessa Bordin Viera - Universidade Federal de Campina Grande

Prof^a Dr^a Welma Emidio da Silva – Universidade Federal Rural de Pernambuco

Ciências Exatas e da Terra e Engenharias

Prof. Dr. Adélio Alcino Sampaio Castro Machado - Universidade do Porto

ProF^a Dr^a Ana Grasielle Dionísio Corrêa – Universidade Presbiteriana Mackenzie

Prof. Dr. Carlos Eduardo Sanches de Andrade - Universidade Federal de Goiás

Profa Dra Carmen Lúcia Voigt - Universidade Norte do Paraná

Prof. Dr. Cleiseano Emanuel da Silva Paniagua – Instituto Federal de Educação, Ciência e Tecnologia de Goiás

Prof. Dr. Douglas Gonçalves da Silva - Universidade Estadual do Sudoeste da Bahia

Prof. Dr. Eloi Rufato Junior - Universidade Tecnológica Federal do Paraná

Profa Dra Érica de Melo Azevedo - Instituto Federal do Rio de Janeiro



Prof. Dr. Fabrício Menezes Ramos - Instituto Federal do Pará

Profa Dra. Jéssica Verger Nardeli – Universidade Estadual Paulista Júlio de Mesquita Filho

Prof. Dr. Juliano Carlo Rufino de Freitas - Universidade Federal de Campina Grande

Prof^a Dr^a Luciana do Nascimento Mendes – Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte

Prof. Dr. Marcelo Marques - Universidade Estadual de Maringá

Prof. Dr. Marco Aurélio Kistemann Junior - Universidade Federal de Juiz de Fora

Prof^a Dr^a Neiva Maria de Almeida - Universidade Federal da Paraíba

Profa Dra Natiéli Piovesan - Instituto Federal do Rio Grande do Norte

Profa Dra Priscila Tessmer Scaglioni - Universidade Federal de Pelotas

Prof. Dr. Sidney Gonçalo de Lima - Universidade Federal do Piauí

Prof. Dr. Takeshy Tachizawa - Faculdade de Campo Limpo Paulista

Linguística, Letras e Artes

Profa Dra Adriana Demite Stephani - Universidade Federal do Tocantins

Prof^a Dr^a Angeli Rose do Nascimento – Universidade Federal do Estado do Rio de Janeiro

Prof^a Dr^a Carolina Fernandes da Silva Mandaji - Universidade Tecnológica Federal do Paraná

Profa Dra Denise Rocha - Universidade Federal do Ceará

Profa Dra Edna Alencar da Silva Rivera - Instituto Federal de São Paulo

Prof^a Dr^aFernanda Tonelli - Instituto Federal de São Paulo,

Prof. Dr. Fabiano Tadeu Grazioli - Universidade Regional Integrada do Alto Uruguai e das Missões

Prof. Dr. Gilmei Fleck - Universidade Estadual do Oeste do Paraná

Prof^a Dr^a Keyla Christina Almeida Portela – Instituto Federal de Educação, Ciência e Tecnologia do Paraná

Prof^a Dr^a Miranilde Oliveira Neves - Instituto de Educação, Ciência e Tecnologia do Pará

Prof^a Dr^a Sandra Regina Gardacho Pietrobon - Universidade Estadual do Centro-Oeste

Prof^a Dr^a Sheila Marta Carregosa Rocha - Universidade do Estado da Bahia



Coleção desafios das engenharias: engenharia de computação 2

Diagramação: Maria Alice Pinheiro

Correção: Giovanna Sandrini de Azevedo **Indexação:** Gabriel Motomu Teshima

Revisão: Os autores

Organizador: Ernane Rosa Martins

Dados Internacionais de Catalogação na Publicação (CIP)

C691 Coleção desafios das engenharias: engenharia de computação 2 / Organizador Ernane Rosa Martins. – Ponta Grossa - PR: Atena. 2021.

Formato: PDF Requisitos de sistema: Adobe Acrobat Reader Modo de acesso: World Wide Web Inclui bibliografia ISBN 978-65-5983-384-9

DOI: https://doi.org/10.22533/at.ed.849211808

 Engenharia da computação. I. Martins, Ernane Rosa (Organizador). II. Título.

CDD 621.39

Elaborado por Bibliotecária Janaina Ramos - CRB-8/9166

Atena Editora

Ponta Grossa – Paraná – Brasil Telefone: +55 (42) 3323-5493 www.atenaeditora.com.br contato@atenaeditora.com.br



DECLARAÇÃO DOS AUTORES

Os autores desta obra: 1. Atestam não possuir qualquer interesse comercial que constitua um conflito de interesses em relação ao artigo científico publicado; 2. Declaram que participaram ativamente da construção dos respectivos manuscritos, preferencialmente na: a) Concepção do estudo, e/ou aquisição de dados, e/ou análise e interpretação de dados; b) Elaboração do artigo ou revisão com vistas a tornar o material intelectualmente relevante; c) Aprovação final do manuscrito para submissão.; 3. Certificam que os artigos científicos publicados estão completamente isentos de dados e/ou resultados fraudulentos; 4. Confirmam a citação e a referência correta de todos os dados e de interpretações de dados de outras pesquisas; 5. Reconhecem terem informado todas as fontes de financiamento recebidas para a consecução da pesquisa; 6. Autorizam a edição da obra, que incluem os registros de ficha catalográfica, ISBN, DOI e demais indexadores, projeto visual e criação de capa, diagramação de miolo, assim como lançamento e divulgação da mesma conforme critérios da Atena Editora.



DECLARAÇÃO DA EDITORA

A Atena Editora declara, para os devidos fins de direito, que: 1. A presente publicação constitui apenas transferência temporária dos direitos autorais, direito sobre a publicação, inclusive não constitui responsabilidade solidária na criação dos manuscritos publicados, nos termos previstos na Lei sobre direitos autorais (Lei 9610/98), no art. 184 do Código penal e no art. 927 do Código Civil; 2. Autoriza e incentiva os autores a assinarem contratos com repositórios institucionais, com fins exclusivos de divulgação da obra, desde que com o devido reconhecimento de autoria e edição e sem qualquer finalidade comercial; 3. Todos os e-book são open access, desta forma não os comercializa em seu site, sites parceiros, plataformas de e-commerce, ou qualquer outro meio virtual ou físico, portanto, está isenta de repasses de direitos autorais aos autores; 4. Todos os membros do conselho editorial são doutores e vinculados a instituições de ensino superior públicas, conforme recomendação da CAPES para obtenção do Qualis livro; 5. Não cede, comercializa ou autoriza a utilização dos nomes e e-mails dos autores, bem como nenhum outro dado dos mesmos, para qualquer finalidade que não o escopo da divulgação desta obra.



APRESENTAÇÃO

A Engenharia de Computação é a área que estuda as técnicas, métodos e ferramentas matemáticas, físicas e computacionais para o desenvolvimento de circuitos, dispositivos e sistemas. Esta área tem a matemática e a computação como seus principais pilares. O foco está no desenvolvimento de soluções que envolvam tanto aspectos relacionados ao software, quanto à elétrica/eletrônica. Os profissionais desta área são capazes de atuar principalmente na integração entre software e hardware, tais como: automação industrial e residencial, sistemas embarcados, sistemas paralelos e distribuídos, arquitetura de computadores, robótica, comunicação de dados e processamento digital de sinais.

Dentro deste contexto, esta obra aborda diversos aspectos tecnológicos computacionais, tais como: implementação e modificações numéricas a serem feitas no algorítmo de Anderson (2010) para simular o escoamento sobre uma asa finita submetida a ângulos de ataque próximos ao estol; modelo distribuído para analisar a influência da formação e do adensamento de geada sobre o desempenho de evaporadores do tipo tubo-aletado, comumente usados em refrigeradores frost-free; um algoritmo de Redes Neurais Convolucionais(CNN) que identifica se a pessoa está ou não utilizando a máscara; potencialidades do M-Learning e Virtual Reality no curso técnico em Agropecuária; avaliação da qualidade da energia elétrica em um sistema de geração de energia fotovoltaica; uma abordagem para a segmentação de imagens cerebrais, utilizando o método baseado em algoritmos genéticos pelo método de múltiplos limiares; estudo numérico de uma âncora torpedo sem aletas cravada em solo isotrópico puramente coesivo, utilizando um modelo axissimétrico não-linear em elementos finitos; estudo acerca da análise numérica de placas retangulares por meio do método das diferencas finitas, obtendo soluções aproximadas para o campo de deslocamentos transversais bem como os correspondentes momentos fletores, para problemas envolvendo uma série de condições de contorno, utilizando-se o software Matlab® para simulação; desenvolvimento e aplicação da Realidade Virtual (RV) como Tecnologia de Informação e Comunicação (TIC) para auxiliar no processo de ensino-aprendizado de disciplinas do Ensino Médio; avaliação dos resultados obtidos em campanhas de medição de qualidade da energia elétrica (QEE) na rede básica em 500 kV; examinar o comportamento mecânico-estático de uma longarina compósita projetada para uma aeronave esportiva leve através de investigações numéricas, empreendidas em software (ANSYS Release 19.2) comercial de elementos finitos; construção de um sistema para monitoramento de ativos públicos; a relação da Sociedade 5.0 envolvida no contexto da Indústria 4.0 e a Transformação Digital; algoritmos de seleção e de classificação de atributos, identificando as vinte principais características que contribuem para o desempenho alto ou baixo dos estudantes; a Mask R-CNN, utilizada para a segmentação de produtos automotivos (parabrisas, faróis, lanternas, parachoques e retrovisores) em uma empresa do ramo de reposição automotiva; o nível de usabilidade do aplicativo protótipo

para dispositivo móvel na área da saúde voltado ao auxílio do monitoramento móvel no uso de medicamentos em seres humanos.

Sendo assim, está obra é significativa por ser composta por uma gama de trabalhos pertinentes, que permitem aos seus leitores, analisar e discutir diversos assuntos importantes desta área. Por fim, desejamos aos autores, nossos mais sinceros agradecimentos pelas significativas contribuições, e aos nossos leitores, desejamos uma proveitosa leitura, repleta de boas reflexões.

Ernane Rosa Martins

SUMÁRIO
CAPÍTULO 11
NONLINEAR LIFTING LINE IMPLEMENTATION AND VALIDATION FOR AERODYNAMICS AND STABILITY ANALYSIS André Rezende Dessimoni Carvalho Pedro Paulo de Carvalho Brito https://doi.org/10.22533/at.ed.8492118081
CAPÍTULO 211
INFLUÊNCIA DA FORMAÇÃO DE GEADA EM EVAPORADORES DE TUBO ALETADO USANDO UM MODELO DISTRIBUÍDO Caio Cezar Neves Pimenta André Luiz Seixlack
€ https://doi.org/10.22533/at.ed.8492118082
CAPÍTULO 324
INFLUÊNCIA DO NÚMERO DE SEÇÕES DE CONECTORES NA EFICIÊNCIA DA RUPTURA POR SEÇÃO LÍQUIDA EM CANTONEIRA DE CHAPA DOBRADA Jéssica Ferreira Borges Luciano Mendes Bezerra Francisco Evangelista Jr Valdeir Francisco de Paula https://doi.org/10.22533/at.ed.8492118083
CAPÍTULO 437
INFORMATION THEORY BASED STOCHASTIC HETEROGENEOS MULSTISCALE lanyqui Falcão Costa Liliane de Allan Fonseca Ézio da Rocha Araújo https://doi.org/10.22533/at.ed.8492118084
CAPÍTULO 559
INTELIGÊNCIA ARTIFICIAL PARA IDENTIFICAR O USO DE MÁSCARA NA PREVENÇÃO DA COVID-19 Roberson Carlos das Graças Edyene Cely Amaro Oliveira Guilherme Ribeiro Brandao Igor Siqueira da Silva Samara de Jesus Duarte Samara Lana da Rocha Hermes Francisco da Cruz Oliveira Guilherme Henrique Chaves Batista https://doi.org/10.22533/at.ed.8492118085

CAPÍTULO 6
ANÁLISE DE DESEMPENHO MECÂNICO DE PLACAS A PARTIR DE MÉTODOS APROXIMADOS Gabriel de Bessa Spínola Edmilson Lira Madureira Eduardo Morais de Medeiros
inttps://doi.org/10.22533/at.ed.8492118086
CAPÍTULO 785
M-LEARNING E VIRTUAL REALITY NO ENSINO TÉCNICO DE AGROPECUÁRIA Gabriel Pinheiro Compto Jeconias Ferreira dos Santos https://doi.org/10.22533/at.ed.8492118087
CAPÍTULO 895
MODELLING AND ANALYSIS OF AEROBOAT JAHU João B. de Aguiar Júlio C.S. Sousa José M. de Aguiar https://doi.org/10.22533/at.ed.8492118088
CAPÍTULO 9113
MONITORAMENTO DA QUALIDADE DE ENERGIA EM SISTEMA DE GERAÇÃO FOTOVOLTAICA - ANÁLISE DAS CAMPANHAS DE MEDIÇÃO DE TENSÃO E CORRENTE E CARACTERÍSTICAS DE INJEÇÃO DE HARMÔNICOS DOS SISTEMAS DE BAIXA, MÉDIA E ALTA TENSÃO Nelson Clodoaldo de Jesus João Roberto Cogo Luiz Marlus Duarte Jesus Daniel de Oliveira Luis Fernando Ribeiro Ferreira Éverson Júnior de Mendonça Leandro Martins Fernandes https://doi.org/10.22533/at.ed.8492118089
CAPÍTULO 10
OTIMIZAÇÃO MULTI-LIMIAR PARA SEGMENTAÇÃO DE MRI POR ALGORÍTIMO GENÉTICO Tiago Santos Ferreira Paulo Fernandes da Silva Júnior Ewaldo Eder Carvalho Santana Mauro Sérgio Silva Pinto Jayne Muniz Fernandes Ana Flávia Chaves Uchôa Jarbas Pinto Monteiro Guedes https://doi.org/10.22533/at.ed.84921180810

CAPITULO 11
ANÁLISE NUMÉRICA DA CAPACIDADE DE CARGA DE ÂNCORAS TORPEDO CONSIDERANDO EFEITOS DE SETUP Guilherme Kronemberger Lopes José Renato Mendes de Sousa Gilberto Bruno Ellwanger
€ https://doi.org/10.22533/at.ed.84921180811
CAPÍTULO 12156
ANÁLISE NUMÉRICA DE PLACAS EM ESTRUTURAS AEROESPACIAIS POR DIFERENÇAS FINITAS Júlio César Fiorin Reyolando Manoel Lopes Rebello da Fonseca Brasil
€ https://doi.org/10.22533/at.ed.84921180812
CAPÍTULO 13172
NUMERICAL SIMULATION OF LABYRINTH SEALS FOR PULSED COMPRESSION REACTORS (PCR) Hermann Enrique Alcázar Rojas Briam Rudy Velasquez Coila Arioston Araújo de Morais Júnior Leopoldo Oswaldo Alcázar Rojas
o https://doi.org/10.22533/at.ed.84921180813
CAPÍTULO 14183
PRÁTICAS E CONTROLE DA CORRUPÇÃO NO MERCADO SEGURADOR: UMA PROPOSTA DE DADOS PARA SISTEMAS DE CONTROLE E COMPLIANCE Lucas Cristiano Ferreira Alves Melissa Mourão Amaral Liza Dantas Noguchi https://doi.org/10.22533/at.ed.84921180814
CAPÍTULO 15198
PREDICTING EFFECTIVE CONSTITUTIVE CONSTANTS FOR WOVEN-FIBRE COMPOSITE MATERIALS Jonas Tieppo da Rocha Tales de Vargas Lisbôa Rogério José Marczak
o https://doi.org/10.22533/at.ed.84921180815
CAPÍTULO 16210
PREVENTING SPURIOUS ARTIFACTS WITH CONSISTENT INTERPOLATION OF PROPERTIES BETWEEN CELL CENTERS AND VERTICES IN TWO-DIMENSIONAL RECTILINEAR GRIDS Alexandre Antonio de Oliveira Lopes

Flávio Pereira Nascimento

Tulio Ligneui Santos
Alberto Barbosa Júnior Luca Pallozzi Lavorante
di https://doi.org/10.22533/at.ed.84921180816
CAPÍTULO 17230
REALIDADE VIRTUAL APLICADA COMO FERRAMENTA DE AUXÍLIO AO ENSINO Simone Silva Frutuoso de Souza Everton Welter Correia Gabrielly Chiquezi Falcão Leonardo Plaster Silva Érica Baleroni Pacheco Fábio Roberto Chavarette Fernando Parra dos Anjos Lima
€ https://doi.org/10.22533/at.ed.84921180817
CAPÍTULO 18245
RESULTADOS DE CAMPANHAS DE MEDIÇÃO DE QUALIDADE DA ENERGIA EM SISTEMAS COM COMPENSADORES ESTÁTICOS DE REATIVOS - ANÁLISE DO IMPACTO DE OUTROS AGENTES NA AMPLIFICAÇÃO DE HARMÔNICOS EM SISTEMA DE 500 kV Nelson Clodoaldo de Jesus João Roberto Cogo Luis Fernando Ribeiro Ferreira Luiz Marlus Duarte Éverson Júnior de Mendonça Leandro Martins Fernandes Jesus Daniel de Oliveira https://doi.org/10.22533/at.ed.84921180818
CAPÍTULO 19258
SIMPLIFIED NUMERICAL MODEL FOR ANALYSIS OF STEEL-CONCRETE COMPOSITE BEAMS WITH PARTIAL INTERACTION Samuel Louzada Simões Tawany Aparecida de Carvalho Ígor José Mendes Lemes Rafael Cesário Barros Ricardo Azoubel da Mota Silveira https://doi.org/10.22533/at.ed.84921180819
CAPÍTULO 20266
SIMULAÇÃO DE UMA LONGARINA COMPÓSITA DE UMA AERONAVE ESPORTIVA LEVE Felipe Silva Lima
Álvaro Barbosa da Rocha Daniel Sarmento dos Santos

Francisco Ismael Pinillos Nieto

Wanderley Ferreira de Amorim Júnior
ohttps://doi.org/10.22533/at.ed.84921180820
CAPÍTULO 21279
SISTEMA RFID PARA CONTROLE DE ATIVOS PÚBLICOS João Felipe Fonseca Nascimento Jislane Silva Santos de Menezes Jean Louis Silva Santos Jennysson D. dos Santos Júnior Luccas Ribeiro Cruz Jean Carlos Menezes Oliveira João Marcos Andrade Santos
di https://doi.org/10.22533/at.ed.84921180821
CAPÍTULO 22
Carlos Henrique Leal Viana Sávio Torres Melo Rebeka Manuela Lobo Sousa Tiago Monteiro de Carvalho Thiago Rodrigues Piauilino Ribeiro doi https://doi.org/10.22533/at.ed.84921180822
CAPÍTULO 23303
SOCIEDADE 5.0 CORRELACIONADA COM A INDÚSTRIA 4.0 E A TRANSFORMAÇÃO DIGITAL Pablo Fernando Lopes Thiago Silva Souza Fernando Hadad Zaidan https://doi.org/10.22533/at.ed.84921180823
CAPÍTULO 24313
TÉCNICA DE DIAGNÓSTICO DE BARRAS QUEBRADAS EM MOTOR DE INDUÇÃO TRIFÁSICO SEM CARGA POR MEIO DA TRANSFORMADA WAVELET Carlos Eduardo Nascimento Cesar da Costa to https://doi.org/10.22533/at.ed.84921180824
CAPÍTULO 25332
UNCERTAINTY QUANTIFICATION OF FRACTURE POTENTIAL AT CONCRETE-ROCK INTERFACE Mariana de Alvarenga Silva
Francisco Evangelista Junior
ohttps://doi.org/10.22533/at.ed.84921180825

CAPÍTULO 26342
USANDO MINERAÇÃO DE DADOS PARA IDENTIFICAR FATORES MAIS IMPORTANTES DO ENEM DOS ÚLTIMOS 22 ANOS Jacinto José Franco Fernanda Luzia de Almeida Miranda Davi Stiegler Felipe Rodrigues Dantas Jacques Duílio Brancher Tiago do Carmo Nogueira https://doi.org/10.22533/at.ed.84921180826
CAPÍTULO 27355
ARTIFICIAL INTELLIGENCE USAGE FOR IDENTIFYING AUTOMOTIVE PRODUCTS Leandro Moreira Gonzaga Gustavo Maia de Almeida https://doi.org/10.22533/at.ed.84921180827
CAPÍTULO 28366
UTILIZAÇÃO DE APLICATIVO PARA DISPOSITIVO MÓVEL PARA ADMINISTRAÇÃO DE MEDICAMENTOS Luísa de Castro Guterres Allan Rafael da Silva Lima Wender Antônio da Silva
https://doi.org/10.22533/at.ed.84921180828
CAPÍTULO 29399
VIBRATIONS ANALYSIS UNCOUPLED AND COUPLED FLUID-STRUCTURE BETWEEN SHELL AND ACOUSTIC CAVITY CYLINDRICAL FOR VARIOUS BOUNDARY CONDITIONS Davidson de Oliveira França Júnior Lineu José Pedroso
dinttps://doi.org/10.22533/at.ed.84921180829
SOBRE O ORGANIZADOR410
ÍNDICE REMISSIVO411

CAPÍTULO 26

USANDO MINERAÇÃO DE DADOS PARA IDENTIFICAR FATORES MAIS IMPORTANTES DO ENEM DOS ÚLTIMOS 22 ANOS

Data de aceite: 02/08/2021

Jacinto José Franco

Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso (IFMT) Barra do Garças, MT – Brazil

Fernanda Luzia de Almeida Miranda

Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso (IFMT) Barra do Garças, MT – Brazil

Davi Stiegler

Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso (IFMT) Barra do Garças, MT – Brazil

Felipe Rodrigues Dantas

Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso (IFMT) Barra do Garças, MT – Brazil

Jacques Duílio Brancher

Departamento de Ciências da Computação Universidade Estadual de Londrina (UEL) -Londrina, PR – Brasil

Tiago do Carmo Nogueira

Instituto Federal de Educação, Ciência e Tecnologia Baiano (IFBaiano) Guanambi, BA – Brasil

RESUMO: O Exame Nacional do Ensino Médio (Enem) é uma prova requerida pela maioria das universidades brasileiras para seleção de estudantes. Neste exame, são coletadas

várias características sobre os candidatos. Algumas dessas características não contribuem de forma significativa para a predição de desempenho, representando assim, um conjunto demasiadamente grande de dados, o que requer recursos computacionais exponenciais para identificá-las. Para solucionar tal problema, este trabalho aplica algoritmos de seleção e de classificação de atributos, identificando as vinte principais características que contribuem para o desempenho alto ou baixo dos estudantes neste exame, nos últimos vinte e dois anos (1998-2019).

PALAVRAS - CHAVE: Enem, Mineração de Dados Educacionais, Experimentos, Algoritmos de selecão

ABSTRACT: The national high school exam (Enem) is a test required by most Brazilian universities to select students. In this exam, several characteristics about the candidates are collected. Some of these characteristics do not contribute significantly to the performance prediction, thus representing an excessively large dataset, requiring exponential computational resources to identify them. To solve this problem, this work applies algorithms for the selection and classification of attributes, identifying twenty main characteristics that contribute to the high or low performance of students in this exam, in the last twenty-two years (1998-2019).

KEYWORDS: Enem, Educational Data Mining, Experiments, Feature selection Algorithms.

1 I INTRODUÇÃO

No Brasil, os estudantes de nível médio que almejam ingressar em algum curso de nível superior precisam realizar testes para avaliar seus conhecimentos. Essa avaliação pode ocorrer de duas formas, via vestibular ou pelo Exame Nacional do Ensino Médio, o Enem. Este é aplicado anualmente pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) (Lima et al., 2019).

Nos anos iniciais de aplicação do Enem, o teste contou com um número pequeno de inscritos; contudo, esse número foi crescendo ao longo dos anos ao longo dos anos à medida que as instituições brasileiras de ensino superior foram adotando-o como requisito para o ingresso em seus cursos. Atualmente, essa prova é largamente utilizada por grande parte das instituições públicas e privadas do país, pois provê uma métrica confiável para avaliar os conhecimentos do público estudantil.

O resultado de cada edição do Enem realizada nos últimos anos está disponível via microdados anonimizados pelo Inep¹, o que viabiliza análises quanto a relevância de cada fator. Uma dificuldade que muitos pesquisadores podem ter ao analisar o Enem diz respeito ao número de fatores coletados pelo Inep em cada prova, que pode chegar a 327 variáveis coletadas em um único ano. Para resolver esse tipo de problema, em mineração de dados, vários métodos foram desenvolvidos com o intuito de reduzir a dimensionalidade, permitindo dizer quais fatores são mais importantes e eliminar os que são redundantes ou irrelevantes (Stańczyk & Jain, 2015).

Considerando a mencionada constatação, este trabalho fez uso de métodos de seleção de atributos e de classificação para identificar os fatores que mais contribuem para o desempenho do estudante como sendo baixo ou alto. Nos datasets avaliados foram considerados todos os fatores como entrada dos algoritmos de seleção de atributos, com exceção dos que se referiam às notas, às informações do gabarito e à presença dos candidatos.. Ao final, escolheram-se somente os 10 melhores atributos retornados por cada algoritmo de acordo com o *ranking*. As 10 características de cada ano foram aglutinadas em um *ranking* global após a conclusão dos experimentos de cada ano.

Para mensurar a qualidade dos fatores retornados pelos métodos de seleção, este trabalho fez uso de algoritmos de classificação e da média das dez execuções, estabelecendo-se a melhor combinação de fatores cuja taxa de acertos foi maximizada.

Com a finalidade de avaliar os dados entre 1998 e 1999, este trabalho aplicou uma abordagem exploratória para identificar algoritmos que possuem melhores desempenhos em relação às informações contidas nos *datasets* do Enem. A partir dos resultados obtidos, os quatro melhores métodos de seleção e os dois melhores algoritmos de classificação foram selecionados para evidenciar os fatores mais relevantes no contexto educacional para a performance dos estudantes.

¹ http://inep.gov.br/microdados

Quanto à sua estrutura, além desta introdução (seção 1), este artigo apresenta: a seção 2, que traz a revisão da literatura, com intuito de identificar os principais trabalhos sobre mineração de dados educacionais no Enem; a seção 3, na qual se realiza o delineamento experimental demonstrando as fases de pré-processamento, seleção e classificação das melhores características aplicadas no Enem; a seção 4, que apresenta os resultados e discussões; e a secão 5, que traz as conclusões deste estudo.

2 I REVISÃO DA LITERATURA

Para a identificação de possíveis lacunas na literatura, buscou-se por trabalhos que aplicassem a mineração de dados educacionais no conjunto de dados do Enem. Essa procura ocorreu de maneira exploratória, utilizando-se as engines de buscas *Google Scholar, ISI Web of Science, ACM Digital Library, IEEE Digital Library, Scielo, Science Direct, Scopus* e *Springer*.

Dessa forma, definiu-se uma *string* de busca genérica a fim de detectar o maior número de estudos possíveis sobre a temática, elencando-se, como a principal *string*, a palavra-chave 'Enem'. Todos os resultados advindos das *engines* acima elencadas foram considerados, desde que a quantidade de entradas fosse inferior a 1.000 registros. Não obstante, por ser considerado uma *string* de busca genérica, que gera muitos registros no Google Scholar passíveis de descarte por estarem fora do escopo deste trabalho, acrescentou-se ao protocolo mais uma palavra-chave. Assim, utilizando-se dos conectores lógicos, a *string* de busca final foi ('Enem' and 'mineração de dados'), limitando as buscas em 500 registros. Ao final da aplicação deste protocolo, obteve-se o quantitativo de 11 estudos sobre mineração de dados e Enem, em todas as engines utilizadas nesta pesquisa.

Para realizar a tarefa de filtrar os registros, foi considerado apenas artigos publicados em periódicos e eventos, sendo sua seleção feita pelo título, resumo, palavras chaves e conclusões, conforme recomenda (Rf & Mc, 2007) para revisões e o software usado nessa atividade foi o Parsifal. Com essa filtragem, constatou-se que somente 11 artigos de fato usavam algum método de mineração de dados aplicado aos datasets do Enem.

A grande parte dos trabalhos encontrados na literatura não levam em consideração o uso dos algoritmos para a redução da dimensionalidade, com exceção de (Santos et al., 2020) que aplica o *Principal Component Analysis* (PCA) no *dataset* de 2016 do Enem em conjunto com Redes Bayesianas, o que permite gerar redes reduzidas, facilitando assim a interpretação dos resultados. No estudo considerou-se 43 variáveis e após a aplicação do PCA o número foi reduzido para apenas 23. Segundo os autores, a renda, a escolaridade dos pais e o tipo de escola são fatores de grande influência no resultado.

A maioria dos trabalhos quando faz a seleção dos atributos, o faz sem o auxílio de algoritmos, apenas avaliando o conjunto de variáveis alvo do estudo. Isso é realizado por (Gomes et al., 2017) e (COSTA et al., 2017), para estabelecer a composição das variáveis

344

da pesquisa. Essa abordagem também é encontrada em trabalhos que realizam análises estatísticas da secão seguinte.

Os artigos identificados nas análises focam-se majoritariamente em classificadores, sendo esse o caso de (Alves et al., 2018), (Braga & Drummond, 2016), (COSTA et al., 2017), (Adeodato et al., 2014), (Silva & Bernardino, 2019), (Stearns et al., 2017), (Simon & Cazella, 2017), e (Gomes et al., 2017). Os algoritmos mais referenciados são: KNN, Naive Bayes, J48, SVM, Redes Neurais, Redes Bayesianas, XGBoost e AdaBoost.

Embora este trabalho se foque somente em algoritmos de classificação, no processo, identifica-se que os trabalhos de (Gomes et al., 2017), (Braga & Drummond, 2016), (Adeodato et al., 2014) e (Alves et al., 2018) fazem uso de regras de associação. Identifica-se também que (Ideas, 2019) e (Albertini & Backes, 2017) utilizam agrupamento e (Alves et al., 2018) de regressão alucidar conhecimentos dos *datasets*.

Nota-se que a maior parte dos artigos lidam com apenas um único *dataset*, com exceção de (Ideas, 2019) que explorou os *datasets* de 2009 à 2015 e (Albertini & Backes, 2017) com dados de 2010 e 2011, os demais artigos focaram-se em apenas um ano do Enem. Destes trabalhos, é possível notar que nenhum dos autores fez uso de métodos da mineração de dados para prover uma visão geral acerca dos fatores de todos os *datasets* do Enem.

Em decorrência dessa constatação, este trabalho utilizou métodos de seleção de atributos para identificar os 10 fatores mais importantes de todos os anos do Enem, tendo sua validação provida por meio de algoritmos de classificação.

A partir do que se verificou na literatura, observou-se que muito do que se coletou em todos os anos do Enem ainda não foi analisado utilizando métodos de mineração de dados. Sendo assim, o presente trabalho preenche essa lacuna com um desenho experimental replicado para todos os anos, provendo uma visão de alto nível acerca dos 20 fatores considerados mais importantes na classificação.

3 I DELINEAMENTO EXPERIMENTAL

Nesta seção são apresentados os três passos usados para obter as 10 melhores características do Enem, que consistem no pré-processamento, na seleção de 10 atributos mais relevantes para cada *dataset* desde 1998 e na validação pela classificação.

3.1 Pré-processamento

O primeiro passo efetuado consistiu em calcular as médias aritméticas de todas as notas para cada prova. O limite de 600 pontos determina se o aluno tem desempenho baixo ou alto, sendo considerado alto quando o aluno alcança a média cujo valor é igual ou superior a 600 pontos e baixo se o valor for inferior ao mencionado limite.

Essa abordagem foi aplicada somente para o ano de 2018; para os demais,

considerou-se uma constante, que é 1,13, multiplicada pela média para estabelecer o limite de notas altas. É válido esclarecer que se chegou ao valor da constante por meio da divisão de 600 pela média obtida em 2018. Isso se fez necessário para normalizar entre os anos o conceito de aluno com desempenho alto ou baixo, pois a média de cada ano foi distinta, em virtude, dentre outros motivos, de possíveis variações quanto ao grau de dificuldade das provas entre as diferentes edições do exame.

Saliente-se, ainda, que as informações de todos os estudantes foram consideradas no experimento, com exceção às que se referiam aos alunos que não preencheram os questionários, que eram treineiros, que zeraram na(s) prova(s) e aos que faltaram às etapas do exame ou a alguma delas.

Após a definição do filtro inicial e do perfil de aluno com alto desempenho, o próximo passo consistiu em converter o *dataset* de cada ano em números inteiros com o comando *apply* da biblioteca Pandas em todas as colunas dos *datasets*. Cada nível das características coletadas tem um valor inteiro único, mesmo os valores nulos. Para estes, o mesmo resultado pode ser obtido usando a *SimpleImputer*² disponível no *Scikit-learn*.

Um problema observado quando se utilizou de classificação foi o desbalanceamento das classes. Nessa situação, ao se aplicar os algoritmos de classificação, percebeu-se que enquanto houve grande índice de acerto na apresentação da classe de alunos com baixo desempenho, o mesmo não ocorreu em relação à categoria de alunos com alto desempenho, o que invalidou o resultado. Uma abordagem muito comum em ciência de dados é usar algum método sintético para balancear as classes. Essa tarefa foi executada usando o SMOTE (*Synthetic Minority Oversampling Technique*) implementado na biblioteca imblearn³.

Com o pré-processamento concluído, o *dataset* ficou pronto para a próxima etapa, a seleção de atributos, podendo, na sequência, ser realizada a execução dos classificadores. Um aspecto relevante a ser observado sobre os *datasets* em questão referiu-se ao número de fatores coletados. Nota-se que esse número variou entre 100 e 327, conforme a Figura 1, e com apenas 10 características foi possível obter taxas de acerto de 69.2% a 89%, o que pode ser verificado na Figura 2 da seção 4.

² https://scikit-learn.org/stable/modules/impute.html

³ https://imbalanced-learn.readthedocs.io/en/stable/index.html

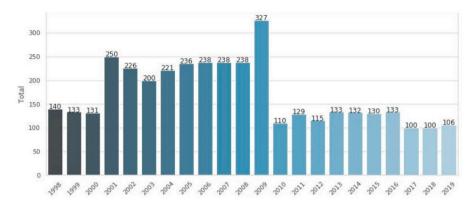


Figura 1: Total de características consideradas de todos os anos do Enem.

4 I SELEÇÃO DOS ATRIBUTOS MAIS RELEVANTES

Inicialmente o experimento partiu de uma abordagem exploratória no sentido de testar vários algoritmos de seleção de atributos e de classificação, não sendo viável executar em *datasets* muito grandes, devido ao tempo requerido para isso. Portanto, escolheram-se os dados de 1999 para apontar os algoritmos mais eficientes em identificar os fatores que propiciam melhor resultado na classificação.

Considerando haver na literatura muitos algoritmos de seleção de atributos e que a execução de todos iria requerer muito tempo, neste trabalho foram escolhidos alguns algoritmos disponíveis no software Weka na versão 3.9.4, Scikit-learn na versão 0.23.1 e na biblioteca MLxtend⁴ na versão 0.17.2. Outra abordagem usada para evidenciar os atributos mais significativos consistiu em executar algoritmos de classificação e extrair o feature_importances_ após o treinamento do modelo. Somado aos algoritmos, também foi testado os atributos cujo fator de correlação e Predictive Power Score (PPS) é maior. O fator de correlação foi obtido pela biblioteca pandas na versão 1.0.1 e o PPS pela biblioteca ppscore 0.0.2.

Foram considerados inicialmente os seguintes algoritmos de seleção de atributos:

- Weka InfoGainAttributeEval, SymmetricalUncertAttributeSetEval e GainRatioAttributeEval.
- Scikit-learn SelectKBest, RFE, GenericUnivariateFeatures e PCA.
- MLxtend SequentialFeatureSelector (SFS) combinado com o classificador LinearRegression do Scikit-learn.

Os classificadores abaixo foram usados para extrair as características mais importantes providas pelo atributo feature_importances_:

XGBoost provido pela biblioteca xgboost na versão 1.0.2.

⁴ http://rasbt.github.io/mlxtend/

LinearSVC e ExtraTreesClassifier disponíveis no Scikit-learn.

5 I CLASSIFICAÇÃO

No experimento, os algoritmos de classificação serviram para dizer o quanto é possível chegar em termos de classificação e definir a melhor combinação de fatores. Quanto ao número de classificadores, o experimento não explorou tantas opções devido à quantidade de dados que precisaram ser processados.

Os seguintes classificadores foram considerados: XGBoost, LightGBM, redes neurais e árvores de decisão implementados nas classes MLPClassifier e DecisionTreeClassifier do Scikit-learn. Nos algoritmos XGBoost e LightGBM não foi aplicada nenhuma personalização, no DecisionTreeClassifier alterou-se somente o parâmetro n_jobs com valor -1 para permitir o uso pleno de todos os núcleos. Como entrada, os algoritmos tiveram 30% do dataset para teste e 70% para treinamento e cada classificador foi repetido 10 vezes com conjuntos de teste e de treinamento distintos gerados pelo train_test_split do Scikit-learn, alterando-se apenas o random_state de cada execução.

6 I RESULTADOS E DISCUSSÃO

Nesta seção são apresentados todos os resultados obtidos experimentalmente combinando fatores, algoritmos de seleção e de classificação que proporcionam a melhor taxa de acerto. Inicialmente, parte-se de uma abordagem exploratória quanto ao número de algoritmos testados e, depois, passa-se a discorrer sobre a escolha dos quatro melhores algoritmos para a seleção de atributos e dos dois melhores para a classificação.

6.1 Dataset de 1998 e 1999

Os dois primeiros anos do Enem foram testados com o intuito de identificar os algoritmos de seleção e classificações que seriam mais confiáveis quanto à qualidade dos resultados possíveis de serem obtidos nos demais anos.

Ranking	Método	Algoritmo de classificação	Quantidade de características	Taxa de acerto
1	XGBClassifier	LightGBM	10	76.31567
2	ExtraTreesClassifier	XGBClassifier	10	76.31279
3	SFS	XGBClassifier	10	76.24993
4	ExtraTreesClassifier	NeuralNet	10	76.24670

Tabela 1: 4 melhores resultados combinando todos os algoritmos no dataset de 1999.

Ranking	Método de seleção	Algoritmo de classificação	Quantidade de características	Taxa de acerto
1	SFS	LightGBM	10	80.46193
2	PCA	XGBoost	10	80.36647

Tabela 2: Duas melhores combinações com o dataset de 1998.

A partir da execução, repetida por 10 vezes, de todos os algoritmos com 10 características, foram escolhidos para a tarefa de classificação os algoritmos *XGBoost* e *LightGBM*; já para a tarefa de seleção, os *XGBoost, ExtraTreesClassifier, PCA* e *SFS*. A partir do ano 2000 somente esses algoritmos foram usados, pois executar todos os algoritmos disponíveis geraria um alto custo computacional.

7 I RESULTADO COMBINADO DE TODOS OS ANOS

Conforme a Figura 2, os resultados obtidos experimentalmente indicaram que, no período de 1998 a 2019 a taxa de acerto possível variou entre 69.2\% e 89.0\% com 10 características.

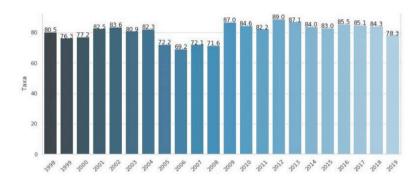


Figura 2: Taxa de acerto do experimento dos últimos 22 anos.

Em razão do grande número de características gerado pela busca efetuada nos 22 anos de exame, para elencar as mais relevantes em cada ano, optou-se por usar o atributo feature_importances_ obtido a partir de uma execução em cada *dataset* com o *XGBoost* para indicar os 20 atributos considerados mais importantes. A partir da média do valor feature_importances_, permitiu-se indicar os atributos que em média foram mais relevantes para a atividade de classificação.

Ranking	Descrição	Anos	Score Médio
1	Língua Estrangeira	2016, 2018, 2019, 2011, 2012, 2013, 2014, 2010	0,302309875
2	Grau de importância quanto aos motivos que levaram a participar do ENEM Para Conseguir uma bolsa de estudos (ProUni, outras)	2010	0,280232
3	O quanto você se interessa e acompanha: a política internacional	1999, 2000,	0,248994
4	Se indicou indígena, qual(is) língua(s) você domina	2005, 2006, 2007, 2008	0,2179125
5	Indique os motivos que levaram você a participar do ENEM: Conseguir uma bolsa de estudos (ProUni, outras).	2015,2016, 2012, 2013, 2014	0,1878978
6	Indique o que levou você a participar do ENEM: Conseguir uma bolsa de estudos (ProUni, outras)	2011	0,16978
7	Conseguir uma boisa de extudos (PTOUTIL, outras) Você tem em sua casa? Microcomputador		0,160423333333333
8	Fez curso de língua estrangeira	2009, 2000, 2001, 2002, 2003, 2004, 2013, 2010	0,155758625
9	Indique os cursos que você frequenta ou frequentou: Curso superior	2013, 2014, 2011, 2010	0,153418
10	Em que tipo escola cursou ou está cursando o ensino médio (2º grau)	2000, 2001, 2002, 2003, 2004, 2005, 2008, 2015, 2016, 2017, 2018, 1998	0,144833846153846
11	Você considera que conhece suficientemente a atividade de trabalho Que você escolheu.	1999	0,129628
12	O quanto você se interessa pela política dos outros países	2001, 2002, 2003, 2004, 2005, 2006	0,1238565
13	Qual é a renda mensal de sua família? (Some a sua renda com a dos Seus familiares.)	2015, 2016, 2017, 2018, 2019, 2005, 2006, 2007, 2008, 2012, 2010	0,1196753
14	Sexo	2011, 2013, 2014, 2017, 2003, 2004, 2005, 2006, 2007, 2008	0,11832
15	Tem Acesso à Internet e quantos	2003, 2004, 2009, 2011	0,11703875
16	Em que tipo de escola cursou o ensino fundamental (1º grau)		0,10934
17	Durante o Ensino Fundamental, você abandonou os estudos e/ou Foi reprovado?	2010, 2012 2015,2016	0,103279
18	Em que turno cursou ou esta cursando o ensino médio (2º grau)		0,093806428571429
		2005	
19	Ao concluir o ensino médio (2º grau) o candidato do ENEM pretende fazer curso(s) profissionalizante(s) e me preparar para o trabalho.	1999	0,087494

Tabela 3: Ranking dos melhores atributos dos últimos 22 anos do Enem.

A Tabela 3 elencou os fatores mais importantes para os classificadores. O resultado seria diferente se fosse utilizado o fator de correlação como referência. Embora a correlação dos fatores com o resultado não seja o foco principal deste trabalho, é importante dizer que a renda familiar se tornou o fator de maior correlação com a definição do desempenho como baixo ou alto somente a partir de 2011. Na Tabela 4, considerou-se o maior fator de correlação que mais se aproxima de 1 positivo ou negativo.

Quanto ao algoritmo que permitiu a melhor classificação, o *XGBoost* destacou-se em quase todos os anos. Em média, foi possível obter a taxa de 80,85% de acerto, conforme a Tabela 5, o que proporcionou maior segurança para dizer que os fatores elencados na

Tabela 3 poderiam ser considerados relevantes. O melhor resultado foi o obtido com o *XGBoost* para o ano de 2012, tendo assertividade de 89%.

Descrição	Fator	Ano
Tipo de escola que cursou o ensino médio (2ºgrau)	0,422846794455483	1998
O quanto você se interessa e acompanha: a politica internacional	-0,32451507473847	1999
O quanto voce se interessa e acompanna: a pontica internacional	-0,346604306461826	2000
Em que tipo escola cursou ou está cursando o ensino médio (2º grau)	0,380182101996118	2001
Tem Microcomputador e quantos	-0,290258678805769	2002
Está frequentando um curso profissionalizante	0,256403650140639	2003
Até quando a mãe estudou	0,28581360289315	2004
	0,249277195498918	2005
So indicou indicono, qual(is) língua(s) você domina	0,216173206398373	2006
Se indicou indígena, qual(is) língua(s) você domina	0,219766219644129	2007
	0,257849686576848	2008
Até quando sua mãe estudou	0,218394601633761	2009
Grau de importância quanto aos motivos que levaram a participar do ENEM para Conseguir uma bolsa de estudos (ProUni, outras)	-0,329040009597119	2010
	0,397946320171181	2011
	0,410388728724133	2012
	0,39844400990799	2013
Qual é a renda mensal de sua família?	0,401047691589809	2014
	0,248671127151669	2015
(Some a sua renda com a dos seus familiares.)	0,404529502005125	2016
	0,407309507632146	2017
	0,410839404783366	2018
	0,439258036363207	2019

Tabela 4: Maior fator de correlação de todos os anos do Enem.

Observa-se que a melhor combinação de algoritmos obtida no experimento foi o do *ExtraTreesClassifier* com o *XGBoost* para a classificação, em 8 anos, seguido do *XGBoost* para a classificação e seleção, em 7 anos, e do *SFS* com o *XGBoost*, em 4 anos. Entre os quatro métodos escolhidos para a seleção de atributos, o *PCA* é o seletor de atributos menos eficiente.

Utilizando o método de classificação supracitado, foi possível constatar que o que é perguntado aos candidatos no questionário de cada Enem influencia diretamente o resultado da classificação. Partindo dessa premissa, o questionário do exame de 2012 foi considerado o melhor estruturado de todas as edições, pois, no referido ano, foi possível obter a taxa de acerto de 88,97% executando o *XGBoost*.

Segundo a Tabela 5, corroborada pela Figura 3, o fator de classificação considerado mais importante para 2012 foi o tipo de escola que o estudante cursou o ensino fundamental. Uma possível justificativa para isso seria o fato de, na educação escolar brasileira, esse nível de ensino representar 3/4 dos estudos anteriores ao ingresso da maioria dos candidatos em algum curso de nível superior. Infelizmente tal informação não pôde ser coletada nas últimas edições do Enem, em virtude de seus respectivos questionários não contemplarem mais essa pergunta

Importância	Fator	Descrição		
0.231240	Q32	Em que tipo de escola você cursou o Ensino Fundamental?		
0.191640	TP_LINGUA	Tipo de Língua Estrangeira		
0.104269	Q3	Qual é a renda mensal de sua família? (Some a sua renda com a		
		Dos seus familiares.)		
0.087090	Q21	Você tem em sua casa? Banheiro		
		Caso você ingresse no Ensino Superior privado pretende		
0.082882	Q38	recorrer aos auxílios abaixo para custeio das mensalidades?		
		Bolsa de estudos da empresa onde trabalho.		
0.076745	Q10	Você tem em sua casa? Microcomputador		
0.064528	Q30	Quantos anos você levou para concluir o Ensino Fundamental?		
0.061394	O16	Você tem em sua casa?		
0.001394	Q16	Telefone celular		
0.050122	Q28	Indique os motivos que levaram você a participar do ENEM:		
0.030122	Q28	Conseguir uma bolsa de estudos (ProUni, outras)		
0.050088	Q4	Quantas pessoas moram em sua casa (incluindo você)?		

Tabela 5: Melhores características para a classificação de 2012.

Em 2012, o desempenho do aluno foi considerado alto, se obteve média maior ou igual a 570,30 pontos. A partir das porcentagens da Tabela6, notou-se que somente 7,71% dos alunos que cursaram o ensino fundamental na rede pública conseguiram atingir notas acima de 570,30 pontos, ao passo que as alcançaram 45,92% dos que cursaram esse nível escolar na rede particular de ensino. Essa constatação provavelmente se aplica aos demais anos, pois o Brasil não avançou muito em termos de qualidade do ensino fundamental, que, conforme se averiguou neste estudo, foi o nível que fez toda a diferença para classificar o desempenho dos estudantes em alto ou baixo.

Descrição	Conceito	Quantidade	Porcentagem
Somente em escola pública	Alto	336.782	7,71%
Somente em escola publica	Baixo	4.029.581	92,29%
Maior parte em escola pública	Alto	78.548	17,75%
Maior parte em escola publica	Baixo	364.073	82,25%
Somente em escola particular	Alto	338.140	45,92%
Somenie em escola particular	Baixo	398.168	54,08%
Maior parte em escola particular	Alto	68.270	28,17%
Maior parte em escola particular	Baixo	174.068	71,83%
Somente em escola indígena	Alto	99	6,86%
Somente em escola margena	Baixo	1.345	93,14%
Maior parte em escola indígena	Alto	55	7,97%
Maior parte em escola mulgena	Baixo	635	92,03%
Somente em escola situada em comunidade	Alto	53	8,33%
quilombola.	Baixo	583	91,67%
Maior parte em escola situada em comunidade	Alto	74	11,13%
Quilombola	Baixo	591	88,87%

Tabela 6: Porcentagem de alunos cuja nota é considerada é baixo ou alto para 2012.

352

81 CONCLUSÕES

A partir dos dados evidenciados nesta pesquisa, foi possível inferir que a redução da dimensionalidade do conjunto de características simplificou o estudo, pois diminuiu seu escopo, permitindo que se focasse no que é de fato relevante para avaliar o desempenho da maior parte dos estudantes brasileiros no ENEM.

Cumpre ressaltar que a consideração de um maior número de atributos não pressupõe necessariamente uma melhor classificação, pois alguns deles são redundantes e/ou irrelevantes para os resultados, conforme verificou-se em alguns classificadores pela variável *feature_importance* e pelo ranking fornecido pelos algoritmos de seleção de atributos.

Foi possível perceber que alguns atributos considerados importantes nos experimentos foram removidos do questionário do Enem dos últimos anos, a exemplo dos dados relacionados ao ensino fundamental do estudante. Do ponto de vista educacional, observou-se que algumas informações retiradas eram mais relevantes que as que permaneceram no questionário.

Notou-se, ainda, que embora sejam computacionalmente menos onerosos, os métodos *filters* selecionam características menos eficientes para classificar os alunos quanto aos seus desempenhos no referido Exame.

Pelos resultados obtidos, mesmo que este estudo não tenha explorado muitos fatores para todos os anos e tenha executado os ensembles sem qualquer otimização, mostrou-se mais eficiente que os trabalhos analisados na literatura.

Verificou-se que a importância e a correlação dos atributos entre os anos analisados foram distintas. Após análise das Tabelas 3, 4 e 5, percebeu-se também que os fatores nelas descritos foram os que permitiram classificar melhor os alunos brasileiros de nível médio. Além disso, constatou-se a necessidade de uma reformulação no questionário a ser preenchido pelos estudantes, a fim de melhorar a acurácia do modelo de aprendizagem de máquina, pois, em 2019, foi possível obter apenas 78.3% de acerto contra 88.97% de 2012.

Diante de todo o exposto, o presente estudo pode servir de base para um trabalho futuro, pois provê subsídios para a realização de análises estatísticas acerca dos fatores evidenciados experimentalmente, bem como fornece meios para averiguar quais interações os fatores podem ter entre si, permitindo, assim, uma melhor compreensão dos atributos e arranjos que mais influenciam o desempenho dos estudantes no ENEM.

REFERÊNCIAS

Adeodato, P. J. L., Santos Filho, M. M., & Rodrigues, R. L. (2014). Predição de desempenho de escolas privadas usando o ENEM como indicador de qualidade escolar. *Anais Do XXV Simpósio Brasileiro de Informática Na Educação (SBIE 2014)*, 1(1), 891. https://doi.org/10.5753/cbie.sbie.2014.891

- Albertini, M. K., & Backes, A. R. (2017). Visualization of Clusters in an Educational Data Set Based on Convex-Hull Shape Preservation Algorithm. *International Journal of Pattern Recognition and Artificial Intelligence*. *31*(02), 1750004.
- Alves, R. D., Cechinel, C., & Queiroga, E. (2018). Predição do desempenho de Matemática e Suas Tecnologias do ENEM utilizando técnicas de Mineração De Dados. *Anais Dos Workshops Do Congresso Brasileiro de Informática Na Educação*, **7**(1), 469.
- Braga, L. C. C., & Drummond, I. N. (2016). Uma abordagem de mineração descritiva aplicada a dados abertos governamentais empregando a ferramenta R. *Anais Do Computer on the Beach (2016)*, 51–60. https://siaiap32.univali.br/seer/index.php/acotb/article/view/10709
- COSTA, J. A. R., REIS, ; André L., SOUZA, ; Daniel C. L., CRISTINO, ; Kaessa G. S., M., ; Marcelo, AURELIANO, SOARES, ; Salles R., SANTOS, ; Thiago E., & SILVA, ; Yasmin V. S. (2017). *Técnicas de mineração de dados aplicadas em dados do enem 2015*. 2–5. https://jornada.ifsuldeminas.edu.br/index.php/jcmch4/jcmch4/paper/viewFile/3406/2282
- Gomes, T., Gouveia, R., & Batista, M. (2017). Dados Educacionais Abertos: associações em dados dos inscritos do Exame Nacional do Ensino Médio. *Anais Do Workshop de Informática Na Escola, 23*(1), 895.
- Ideas, N. (2019). Desempenho das Instituições Brasileiras no ENEM: uma Abordagem Usando Mineração de Dados. 106–113.
- Lima, P. da S. N., Ambrósio, A. P. L., Ferreira, D. J., & Brancher, J. D. (2019). Análise de dados do Enade e Enem: uma revisão sistemática da literatura. *Avaliação: Revista Da Avaliação Da Educação Superior (Campinas)*, *24*(1), 89–107.
- Rf, S. A., & Mc, E. M. A. (2007). Estudos De Revisão Sistemática: Um Guia Para Síntese. *Revista Brasileira de Fisioterapia*, *11*(1), 83–89. https://doi.org/10.1590/S1413-35552007000100013
- Santos, A. T. B., Paulino, J., Silva, M. S., & Rego, L. (2020). Educational Data Mining: A Study on Socioeconomic Indicators in Education in INEP Database. In S. Borah, V. Emilia Balas, & Z. Polkowski (Eds.), *Advances in Data Science and Management* (pp. 51–65). Springer Singapore.
- Silva, J. E. H., & Bernardino, H. S. (2019). *A Data Mining Approach Applied to the High School National Examination: Analysis of Aspects of Candidates to Brazilian Universities* (Vol. 1). https://doi.org/10.1007/978-3-030-30241-2
- Simon, A., & Cazella, S. (2017). Mineração de Dados Educacionais nos Resultados do ENEM de 2015. *Anais Dos Workshops Do VI Congresso Brasileiro de Informática Na Educação (CBIE 2017)*, 1(Cbie), 754. https://doi.org/10.5753/cbie.wcbie.2017.754
- Stańczyk, U., & Jain, L. C. (2015). Feature selection for data and pattern recognition. Springer.
- Stearns, B., Rangel, F., Firmino, F., Rangel, F., & Oliveira, J. (2017). Prevendo Desempenho dos Candidatos do ENEM Através de Dados Socioeconômicos. *Anais Do XXXVI Concurso de Trabalhos de Iniciação Cient{\^i\}ifica Da SBC*.

354

ÍNDICE REMISSIVO

Α

Algoritmo 9, 59, 60, 62, 63, 64, 65, 66, 127, 172, 211, 320, 323, 324, 343, 350, 355, 370 Algoritmos de seleção 9, 342, 343, 347, 348, 353

ANSYS 9, 172, 173, 176, 177, 178, 180, 181, 204, 208, 266, 267, 272, 273, 399, 401

Aplicativo 9, 16, 65, 88, 89, 90, 92, 93, 273, 366, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 381, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395

Aprendizado 9, 59, 60, 63, 64, 65, 66, 87, 230, 232, 233, 235, 240, 242, 244, 281, 290 Artificial Intelligence 16, 60, 354, 355

В

Blender 231, 236, 237

C

Classificação 9, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 384

Computational Vision 355, 356

Comunicação 9, 85, 94, 95, 194, 230, 231, 232, 242, 243, 281, 283, 286, 304, 306, 307, 367, 384, 395

Coronavírus 59, 60, 65

Covid-19 11, 59, 60, 62, 65

D

Desempenho 9, 12, 11, 12, 13, 14, 19, 23, 62, 67, 113, 114, 173, 186, 257, 267, 310, 342, 343, 345, 346, 350, 352, 353, 354, 367, 370, 373, 374, 389

Diagnóstico 15, 127, 313, 314, 316, 317, 318, 328, 329, 371

Diagramas 115, 283, 284, 371, 372

Dispositivo Móvel 10, 16, 366, 368, 370, 371

Е

Educação 24, 85, 86, 87, 88, 93, 94, 230, 232, 233, 235, 240, 241, 242, 243, 244, 279, 292, 303, 313, 342, 351, 353, 354, 369, 410

Enem 16, 342, 343, 344, 345, 347, 348, 350, 351, 353, 354

Energia Elétrica 9, 113, 114, 116, 126, 245, 257, 314

Ensino 9, 12, 14, 85, 86, 87, 89, 90, 92, 93, 95, 230, 231, 232, 233, 235, 236, 239, 240, 241, 242, 243, 244, 281, 292, 342, 343, 351, 352, 353, 354

Equações 11, 13, 14, 15, 17, 18, 19, 22, 24, 25, 26, 27, 29, 33, 34, 37, 95, 399

Estruturação de dados 194

F

Finite Differences 38, 156, 157, 158, 159, 160, 162, 163, 165, 169, 170, 171 Fracture Mechanics 332, 334, 341

G

Genetic Algorithm 128, 129, 130, 132, 133, 136, 137, 172, 180 Geração Fotovoltaica 12, 113, 115, 124, 125

Image Processing 128, 130, 136, 356, 364

Indústria 4.0 9, 15, 303, 304, 305, 306, 308, 309, 310, 312

Informação 9, 37, 85, 86, 92, 94, 188, 195, 196, 230, 231, 232, 233, 242, 243, 280, 281, 282, 283, 304, 308, 319, 351, 366, 367, 368, 371, 395, 396, 410

Inteligência Artificial 11, 59, 304, 307, 308, 355, 356

Interface 51, 144, 146, 150, 152, 232, 235, 236, 239, 283, 284, 286, 332, 333, 334, 341, 369, 372, 376, 384, 385, 386, 397

Interpolation 13, 1, 4, 101, 102, 103, 178, 210, 215, 216, 217, 218, 221, 227

L

Labyrinth Seals 13, 172, 174, 176, 179, 181, 182

M

Máscara 9, 11, 59, 61, 62, 63, 64, 65, 66

MASK R-CNN 9, 355, 356, 359, 360, 361, 362, 364, 365

Method 1, 2, 5, 6, 7, 8, 9, 10, 38, 44, 55, 57, 67, 68, 73, 74, 75, 76, 77, 78, 82, 83, 107, 112, 128, 129, 130, 131, 136, 141, 145, 156, 157, 158, 163, 169, 170, 171, 174, 175, 177, 178, 180, 181, 198, 199, 208, 210, 211, 215, 216, 217, 226, 227, 228, 229, 258, 259, 260, 264, 313, 336, 357, 399, 401, 409

Metodologias Ativas 231, 232, 244

Mineração de dados 343, 344, 345, 354

M-Learning 9, 12, 85, 86, 87, 88, 89, 92, 93, 94

Modelagem 17, 18, 211, 236, 237, 271, 284, 312, 371, 372, 374, 375

Modelo distribuído 9, 11, 11, 14, 22

Modelo Numérico 259, 271

Monitoramento 9, 10, 12, 60, 66, 113, 114, 115, 116, 118, 120, 122, 124, 125, 246, 248, 253, 279, 280, 283, 285, 290, 313, 314, 328, 366, 367, 368, 395

Motor de Indução 15, 313, 314, 316, 318, 319, 321

P

Probabilidade 24, 31, 32, 34, 185, 332, 375

Protótipo 9, 234, 240, 241, 242, 283, 285, 286, 289, 366, 368, 371, 372, 374, 394

Pulsed compression reactor 172, 173, 175, 181, 182

R

Realidade Virtual 9, 14, 94, 230, 231, 232, 233, 234, 235, 239, 240, 241, 242, 243, 244

Rectilinear grids 13, 210, 212, 218, 227

Redes Neurais Artificiais 60, 62, 355, 364

RFID 15, 279, 280, 282, 283, 285, 286, 287, 288, 290, 291

S

Setup 13, 138, 139, 140, 146, 147, 148, 149, 150, 152, 153, 154, 155

Sistema 9, 12, 14, 15, 11, 15, 18, 64, 88, 90, 91, 113, 114, 115, 116, 117, 118, 120, 123, 124, 125, 126, 172, 184, 185, 186, 194, 195, 196, 231, 233, 234, 245, 246, 247, 248, 250, 251, 252, 253, 254, 255, 256, 257, 272, 279, 280, 283, 284, 285, 286, 287, 289, 290, 291, 292, 293, 297, 299, 300, 306, 307, 312, 356, 366, 367, 368, 369, 370, 371, 374, 375, 376, 381, 382, 384, 385, 386

Sistema de controle 194, 290

Sistema Estrutural 272, 292, 293, 297, 299, 300

Smartphone 90, 91, 94, 376

Sociedade 5.0 9, 15, 303, 304, 305, 306, 308, 309, 310

Sociedade Criativa 303, 304, 306, 308, 309

Software 9, 28, 67, 74, 137, 138, 139, 156, 157, 163, 176, 177, 200, 209, 231, 236, 266, 267, 282, 284, 287, 291, 292, 293, 298, 321, 323, 324, 325, 328, 344, 347, 371, 372, 375, 376, 386, 396, 397, 398, 399, 401

Т

Tecnologia 9, 24, 85, 86, 87, 91, 93, 94, 114, 230, 231, 232, 239, 240, 241, 242, 244, 267, 279, 280, 281, 282, 283, 290, 292, 301, 302, 304, 306, 307, 308, 309, 310, 311, 313, 332, 342, 366, 367, 368, 396, 410

TICs na Educação 85, 93

Torpedo anchors 138, 139, 140, 148, 150, 152, 155

Transformação Digital 9, 15, 303, 304, 305, 307, 308, 309, 310, 311

U

Uncertainty Quantification 15, 332, 336, 341

Usabilidade 9, 234, 366, 368, 372, 374, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393,

394, 395, 396, 397, 398

V

Virtual 9, 12, 14, 85, 86, 87, 88, 89, 93, 94, 100, 101, 209, 230, 231, 232, 233, 234, 235, 239, 240, 241, 242, 243, 244, 309, 402

Virtual Reality 9, 12, 85, 86, 87, 88, 231, 243, 244

W

Web 10, 35, 279, 280, 283, 286, 287, 290, 304, 344, 386, 396

DESAFIOS DAS ENGENHARIAS:

ENGENHARIA DE COMPUTAÇÃO 2





DESAFIOS DAS ENGENHARIAS:

ENGENHARIA DE COMPUTAÇÃO 2



