

**NELSON DE SOUZA AMORIM
FERNANDO MANUEL ARAÚJO MOREIRA
CARLOS CÉLIO SOUSA DA CRUZ
UBIRAEALSON DE LIMA RUELA
PAULA RENATHA NUNES DA SILVA
JOSÉ ROBERTO BRANCO RAMOS FILHO
VICENTE MOREIRA RODRIGUES
THIAGO AUGUSTO DE SOUSA MOREIRA
GILSON FERNANDES BRAGA JUNIOR
ESTEFANY COUTO MILÉO
(ORGANIZADORES)**

ANAIS DO XIV SIMPÓSIO BRASILEIRO DE ENGENHARIA FÍSICA



Atena
Editora

Ano 2020

**NELSON DE SOUZA AMORIM
FERNANDO MANUEL ARAÚJO MOREIRA
CARLOS CÉLIO SOUSA DA CRUZ
UBIRAEALSON DE LIMA RUELA
PAULA RENATHA NUNES DA SILVA
JOSÉ ROBERTO BRANCO RAMOS FILHO
VICENTE MOREIRA RODRIGUES
THIAGO AUGUSTO DE SOUSA MOREIRA
GILSON FERNANDES BRAGA JUNIOR
ESTEFANY COUTO MILÉO
(ORGANIZADORES)**

ANAIS DO XIV SIMPÓSIO BRASILEIRO DE ENGENHARIA FÍSICA



Atena
Editora

Ano 2020

2020 by Atena Editora

Copyright © Atena Editora

Copyright do Texto © 2020 Os autores

Copyright da Edição © 2020 Atena Editora

Editora Chefe: Profª Drª Antonella Carvalho de Oliveira

Diagramação: Geraldo Alves

Edição de Arte: Lorena Prestes

Revisão: Os Autores



Todo o conteúdo deste livro está licenciado sob uma Licença de Atribuição *Creative Commons*. Atribuição 4.0 Internacional (CC BY 4.0).

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores. 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.

Conselho Editorial

Ciências Humanas e Sociais Aplicadas

Profª Drª Adriana Demite Stephani – Universidade Federal do Tocantins
Prof. Dr. Álvaro Augusto de Borba Barreto – Universidade Federal de Pelotas
Prof. Dr. Alexandre Jose Schumacher – Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso
Prof. Dr. Antonio Carlos Frasson – Universidade Tecnológica Federal do Paraná
Prof. Dr. Antonio Gasparetto Júnior – Instituto Federal do Sudeste de Minas Gerais
Prof. Dr. Antonio Isidro-Filho – Universidade de Brasília
Prof. Dr. Carlos Antonio de Souza Moraes – Universidade Federal Fluminense
Prof. Dr. Constantino Ribeiro de Oliveira Junior – Universidade Estadual de Ponta Grossa
Profª Drª Cristina Gaio – Universidade de Lisboa
Profª Drª Denise Rocha – Universidade Federal do Ceará
Prof. Dr. Deyvison de Lima Oliveira – Universidade Federal de Rondônia
Prof. Dr. Edvaldo Antunes de Farias – Universidade Estácio de Sá
Prof. Dr. Eloi Martins Senhora – Universidade Federal de Roraima
Prof. Dr. Fabiano Tadeu Grazioli – Universidade Regional Integrada do Alto Uruguai e das Missões
Prof. Dr. Gilmei Fleck – Universidade Estadual do Oeste do Paraná
Profª Drª Ivone Goulart Lopes – Istituto Internazionale delle Figlie di Maria Ausiliatrice
Prof. Dr. Julio Candido de Meirelles Junior – Universidade Federal Fluminense
Profª Drª Keyla Christina Almeida Portela – Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso
Profª Drª Lina Maria Gonçalves – Universidade Federal do Tocantins
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte
Prof. Dr. Marcelo Pereira da Silva – Universidade Federal do Maranhão
Profª Drª Miranilde Oliveira Neves – Instituto de Educação, Ciência e Tecnologia do Pará
Profª Drª Paola Andressa Scortegagna – Universidade Estadual de Ponta Grossa
Profª Drª Rita de Cássia da Silva Oliveira – Universidade Estadual de Ponta Grossa
Profª Drª Sandra Regina Gardacho Pietrobon – Universidade Estadual do Centro-Oeste
Profª Drª Sheila Marta Carregosa Rocha – Universidade do Estado da Bahia
Prof. Dr. Rui Maia Diamantino – Universidade Salvador
Prof. Dr. Urandi João Rodrigues Junior – Universidade Federal do Oeste do Pará
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande
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. Antonio Pasqualetto – Pontifícia Universidade Católica de Goiás
Profª Drª Daiane Garabeli Trojan – Universidade Norte do Paraná

Profª Drª Diocléa Almeida Seabra Silva – Universidade Federal Rural da Amazônia
Prof. Dr. Écio Souza Diniz – Universidade Federal de Viçosa
Prof. Dr. Fábio Steiner – Universidade Estadual de Mato Grosso do Sul
Prof. Dr. Fágner Cavalcante Patrocínio dos Santos – Universidade Federal do Ceará
Profª Drª Girlene Santos de Souza – Universidade Federal do Recôncavo da Bahia
Prof. Dr. Júlio César Ribeiro – Universidade Federal Rural do Rio de Janeiro
Profª Drª Lina Raquel Santos Araújo – Universidade Estadual do Ceará
Prof. Dr. Pedro Manuel Villa – Universidade Federal de Viçosa
Profª Drª Raissa Rachel Salustriano da Silva Matos – Universidade Federal do Maranhão
Prof. Dr. Ronilson Freitas de Souza – Universidade do Estado do Pará
Profª Drª Talita de Santos Matos – Universidade Federal Rural do Rio de Janeiro
Prof. Dr. Tiago da Silva Teófilo – Universidade Federal Rural do Semi-Árido
Prof. Dr. Valdemar Antonio Paffaro Junior – Universidade Federal de Alfenas

Ciências Biológicas e da Saúde

Prof. Dr. André Ribeiro da Silva – Universidade de Brasília
Profª Drª Anelise Levay Murari – Universidade Federal de Pelotas
Prof. Dr. Benedito Rodrigues da Silva Neto – Universidade Federal de Goiás
Prof. Dr. Edson da Silva – Universidade Federal dos Vales do Jequitinhonha e Mucuri
Profª Drª Eleuza Rodrigues Machado – Faculdade Anhanguera de Brasília
Profª Drª Elane Schwinden Prudêncio – Universidade Federal de Santa Catarina
Prof. Dr. Ferlando Lima Santos – Universidade Federal do Recôncavo da Bahia
Prof. Dr. Gianfábio Pimentel Franco – Universidade Federal de Santa Maria
Prof. Dr. Igor Luiz Vieira de Lima Santos – Universidade Federal de Campina Grande
Prof. Dr. José Max Barbosa de Oliveira Junior – Universidade Federal do Oeste do Pará
Profª Drª Magnólia de Araújo Campos – Universidade Federal de Campina Grande
Profª Drª Mylena Andréa Oliveira Torres – Universidade Ceuma
Profª Drª Natiéli Piovesan – Instituto Federaci do Rio Grande do Norte
Prof. Dr. Paulo Inada – Universidade Estadual de Maringá
Profª Drª Vanessa Lima Gonçalves – Universidade Estadual de Ponta Grossa
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande

Ciências Exatas e da Terra e Engenharias

Prof. Dr. Adélio Alcino Sampaio Castro Machado – Universidade do Porto
Prof. Dr. Alexandre Leite dos Santos Silva – Universidade Federal do Piauí
Prof. Dr. Carlos Eduardo Sanches de Andrade – Universidade Federal de Goiás
Profª Drª Carmen Lúcia Voigt – Universidade Norte do Paraná
Prof. Dr. Eloi Rufato Junior – Universidade Tecnológica Federal do Paraná
Prof. Dr. Fabrício Menezes Ramos – Instituto Federal do Pará
Prof. Dr. Juliano Carlo Rufino de Freitas – Universidade Federal de Campina Grande
Prof. Dr. Marcelo Marques – Universidade Estadual de Maringá
Profª Drª Neiva Maria de Almeida – Universidade Federal da Paraíba
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte
Prof. Dr. Takeshy Tachizawa – Faculdade de Campo Limpo Paulista

Conselho Técnico Científico

Prof. Msc. Abrãao Carvalho Nogueira – Universidade Federal do Espírito Santo
Prof. Msc. Adalberto Zorzo – Centro Estadual de Educação Tecnológica Paula Souza
Prof. Dr. Adailson Wagner Sousa de Vasconcelos – Ordem dos Advogados do Brasil/Seccional Paraíba
Prof. Msc. André Flávio Gonçalves Silva – Universidade Federal do Maranhão
Profª Drª Andreza Lopes – Instituto de Pesquisa e Desenvolvimento Acadêmico
Profª Msc. Bianca Camargo Martins – UniCesumar
Prof. Msc. Carlos Antônio dos Santos – Universidade Federal Rural do Rio de Janeiro
Prof. Msc. Cláudia de Araújo Marques – Faculdade de Música do Espírito Santo
Prof. Msc. Daniel da Silva Miranda – Universidade Federal do Pará
Profª Msc. Dayane de Melo Barros – Universidade Federal de Pernambuco

Prof. Dr. Edwaldo Costa – Marinha do Brasil
 Prof. Msc. Eliel Constantino da Silva – Universidade Estadual Paulista Júlio de Mesquita
 Prof. Msc. Gevair Campos – Instituto Mineiro de Agropecuária
 Prof. Msc. Guilherme Renato Gomes – Universidade Norte do Paraná
 Prof^a Msc. Jaqueline Oliveira Rezende – Universidade Federal de Uberlândia
 Prof. Msc. José Messias Ribeiro Júnior – Instituto Federal de Educação Tecnológica de Pernambuco
 Prof. Msc. Leonardo Tullio – Universidade Estadual de Ponta Grossa
 Prof^a Msc. Lilian Coelho de Freitas – Instituto Federal do Pará
 Prof^a Msc. Liliani Aparecida Sereno Fontes de Medeiros – Consórcio CEDERJ
 Prof^a Dr^a Lívia do Carmo Silva – Universidade Federal de Goiás
 Prof. Msc. Luis Henrique Almeida Castro – Universidade Federal da Grande Dourados
 Prof. Msc. Luan Vinicius Bernardelli – Universidade Estadual de Maringá
 Prof. Msc. Rafael Henrique Silva – Hospital Universitário da Universidade Federal da Grande Dourados
 Prof^a Msc. Renata Luciane Polsaque Young Blood – UniSecal
 Prof^a Msc. Solange Aparecida de Souza Monteiro – Instituto Federal de São Paulo
 Prof. Dr. Welleson Feitosa Gazel – Universidade Paulista

Dados Internacionais de Catalogação na Publicação (CIP) (eDOC BRASIL, Belo Horizonte/MG)	
S612a	<p>Simpósio Brasileiro de Engenharia Física (14 : 2019 : Santarém) Anais [...] / XIV Simpósio Brasileiro de Engenharia Física, 23-25 outubro 2019, Santarém, PA; organizadores Nelson de Souza Amorim... [et al.]. – Ponta Grossa, PR: Atena, 2020.</p> <p>Formato: PDF Requisitos de sistema: Adobe Acrobat Reader Modo de acesso: World Wide Web ISBN 978-65-86002-15-7 DOI 10.22533/at.ed.157200203</p> <p>1. Engenharia física – Congressos. I. Título.</p> <p style="text-align: right;">CDD 573.724</p>
Elaborado por Maurício Amormino Júnior – CRB6/2422	

Atena Editora
 Ponta Grossa – Paraná - Brasil
www.atenaeditora.com.br
contato@atenaeditora.com.br

APRESENTAÇÃO

O curso de Engenharia Física da Universidade Federal de São Carlos foi criado no ano 1999 e foi pioneiro nesta área no Brasil. No ano de 2019, o curso de engenharia física no Brasil completou 20 anos. Nesse contexto, a Universidade Federal de São Carlos (UFSCar) e a Universidade Federal do Oeste do Pará (UFOPa) promoveram o XIV Simpósio Brasileiro de Engenharia Física que foi realizado na cidade de Santarém-PA no período de 23 a 25 de Outubro de 2019 na Universidade Federal do Oeste do Pará – Campus Tapajós com o tema “Jubileu de 20 anos da Engenharia Física no Brasil”.

Com a proposta de promover o conhecimento científico e inovação tecnológica bem como a integração entre especialistas, docentes e discentes da área, foram discutidos os 20 anos de existência do curso no Brasil e o intercâmbio de informações técnicas-científicas através de minicursos e palestras relacionados as diferentes temáticas da Engenharia Física e suas perspectivas futuras.

A coleção Anais do XIV Simpósio Brasileiro de Engenharia Física é uma obra que tem como objetivo divulgar os diversos trabalhos que participaram do evento através de trabalhos acadêmicos que abordaram diferentes temas, tais como: termodinâmica, propriedades dielétricas de materiais, ciência dos dados e machine learning, internet das coisas, deep learning, processos oxidativos avançados, energia solar, gerenciamento de projetos, física quântica e automação. Deste modo a obra contribui para disseminar os resultados obtidos pelos acadêmicos e fortalecer a diversidade científica no país, de forma multidisciplinar.

Comitê Organizador

SUMÁRIO

CAPÍTULO 1	1
A IMPORTÂNCIA DO REALIMENTADOR NA EFICIÊNCIA DE UM CICLO DE RANKINE UTILIZANDO O EES	
Muller Gabriel da Silva Chaves Carlos Eduardo Ribeiro Silva Vitor Azevedo Pinto Carlos Célio Sousa da Cruz	
DOI 10.22533/at.ed.1572002031	
CAPÍTULO 2	11
ANÁLISE TEÓRICA DAS PROPRIEDADES DIELÉTRICAS DA MACAÚBA (<i>Acrocomia acuelata</i>)	
Alex Torres da Silva Nelson de Souza Amorim	
DOI 10.22533/at.ed.1572002032	
CAPÍTULO 3	19
APLICAÇÃO DE TÉCNICAS DE DATA SCIENCE E MACHINE LEARNING EM UM PROBLEMA DE CLASSIFICAÇÃO DE UM DATASET DE MARKETING BANCÁRIO	
Yasmin Braga Teixeira João Vitor Rebelo Viana Josecley Fialho Góes Anderson Alvarenga de Moura Meneses	
DOI 10.22533/at.ed.1572002033	
CAPÍTULO 4	28
AQUISIÇÃO DE DADOS DO CONSUMO ELÉTRICO EM UMA EDIFICAÇÃO DA UFOPA UTILIZANDO CONCEITOS DE IOT	
Leonardo Paz Amoêdo Dalton Felipe Silva Varão João Elias Brasil Bentes Júnior Anderson Alvarenga de Moura Meneses	
DOI 10.22533/at.ed.1572002034	
CAPÍTULO 5	37
DEEP LEARNING PARA REGRESSÃO DE POTÊNCIA ELÉTRICA DE UMA USINA DE ENERGIA DE CICLO COMBINADO	
Mauro Sérgio dos Santos Moura Anderson Alvarenga de Moura Meneses	
DOI 10.22533/at.ed.1572002035	
CAPÍTULO 6	46
DEGRADAÇÃO DE CORANTES EM MEIO AQUOSO EMPREGANDO DIÓXIDO DE TITÂNIO NA FORMA DE FILMES FINOS PREPARADOS SOBRE SUBSTRATO CERÂMICO COMERCIAL	
Graziele Daiana Sena de Sousa Adriano Cesar Rabelo	
DOI 10.22533/at.ed.1572002036	

CAPÍTULO 7	59
FATORES CRÍTICOS QUE INTERFEREM NO GERENCIAMENTO DO TEMPO EM PROJETOS DE ENGENHARIA: ESTUDO DE CASO EM SANTARÉM – PA	
Raíssa Coelho Almeida Kevin de Matos Costa	
DOI 10.22533/at.ed.1572002037	
CAPÍTULO 8	70
OBTENÇÃO DA EQUAÇÃO DE KLEIN-GORDON-FOCK EM COORDENADAS DO CONE DE LUZ	
Jorge Kysnney Santos Kamassury Damião Pedro Meira Filho Sérgio Antônio de Souza Farias Natalie Von Paraski	
DOI 10.22533/at.ed.1572002038	
CAPÍTULO 9	83
RÁPIDO RECONHECIMENTO DE MODULAÇÕES ANALÓGICAS E DIGITAIS VIA REDES RESIDUAIS PROFUNDAS	
Jorge Kysnney Santos Kamassury Vinícius Felipe de Oliveira da Silva	
DOI 10.22533/at.ed.1572002039	
CAPÍTULO 10	98
REVISÃO DE LITERATURA SOBRE A AUTOMAÇÃO NO ENSINO DE FÍSICA NO BRASIL	
Davi Henrique da Silva Pedroso Gabriel Gonçalves da Silva Gilson Fernandes Braga Junior	
DOI 10.22533/at.ed.15720020310	
CAPÍTULO 11	109
DIAGNÓSTICO DOS MICROSSISTEMAS DE ENERGIA SOLAR FOTOVOLTAICO INSTALADOS NAS COMUNIDADES RURAIS DO MUNICÍPIO DE SANTARÉM	
Fabiane da Conceição Almeida Manoel Roberval Pimentel Santos	
DOI 10.22533/at.ed.15720020311	
SOBRE OS ORGANIZADORES.....	119
ÍNDICE REMISSIVO	121

APLICAÇÃO DE TÉCNICAS DE DATA SCIENCE E MACHINE LEARNING EM UM PROBLEMA DE CLASSIFICAÇÃO DE UM DATASET DE MARKETING BANCÁRIO

Data de aceite: 27/01/2020

Yasmin Braga Teixeira

UFOPA, Laboratório de Inteligência Computacional (LabIC),
Santarém, Pará

<http://lattes.cnpq.br/3762230290151180>

João Vitor Rebelo Viana

UFOPA, Laboratório de Modelagem Computacional (LabMC)
Santarém – Pará

<http://lattes.cnpq.br/6370552438773776>

Josecley Fialho Góes

UFOPA, Laboratório de Modelagem Computacional (LabMC)
Santarém – Pará

<http://lattes.cnpq.br/6689337073298410>

Anderson Alvarenga de Moura Meneses

UFOPA, Laboratório de Inteligência Computacional (LabIC),
Santarém, Pará

<http://lattes.cnpq.br/6272872215125680>

RESUMO: A pesquisa explora o conceito de análise de dados com Ciência dos Dados (*Data Science*) ao utilizar Aprendizado de Máquina (*Machine Learning*) para resolver o problema de classificação de uma proposta de marketing Bancário. O *dataset* utilizado consiste no resultado de uma campanha de *marketig* em um

banco de Portugal. Ele foi obtido no repositório da Universidade da Califórnia (UCI) e chama-se *Bank Marketing Data Set*. A composição do *dataset* está dividida em 45211 observações e 17 atributos incluindo os dados sobre a aceitação ou não da proposta pelo cliente do banco. Com isso, a pesquisa foca em desenvolver algoritmos de Aprendizado de Máquina que resolvam o problema de classificação para prever se o cliente aceitou ou não a proposta da propaganda. Os processos da metodologia incluem analisar o *dataset*, fazer análise exploratória de dados e plotagem de gráficos para conhecimento de quais variáveis são necessários manter ou retirar para aplicação do algoritmo, implementar a algoritmo KNN e Redes Neurais para classificação de valores. Como resultados a acurácia foi de 88% para a Rede Neural MLP e para o KNN. Como trabalhos futuros, é necessário mais testes para uma melhor comparação, porém foi possível prever a proposta do cliente e obter informações sobre perfis de consumo de clientes de um banco.

PALAVRAS-CHAVE: Ciência de Dados, Aprendizado de Máquina, Redes Neurais, K-Nearest Neighbors, Marketing Bancário.

APPLICATION OF DATASCIENCE AND
MACHINE LEARNING TECHNIQUES IN
A PROBLEM OF CLASSIFICATION OF A
BANKING MARKETING DATASET

ABSTRACT: The research explores the concept of Data Science by using Machine Learning to solve the classification problem of a Banking marketing proposal. The dataset used is the result of a marketing campaign in a bank of Portugal. It was obtained from the University of California (UCI) repository and is called Bank Marketing Data Set. The composition of the dataset is divided into 45211 observations and 17 attributes including data on whether or not the bank's client accepts the proposal. Thus, the research focuses on developing Machine Learning algorithms that solve the classification problem to predict whether or not the customer accepted the advertising proposal. The methodology processes include analyzing the dataset, doing exploratory data analysis with graph plotting to know which variables are necessary to maintain or remove for application of the algorithm, implementing the KNN algorithm and Neural Networks for comparison of values. As a result the accuracy was 88% for the Neural Network and the KNN. In future researches, further testing is needed for a better comparison, but it was possible to predict the customer's proposal and obtain information on customer consumption profiles from a bank.

KEYWORDS: Data Science, Machine Learning, Neural Network, K-Nearest Neighbors, Bank Marketing.

1 | INTRODUÇÃO

Nos dias atuais, uma imensa quantidade de dados é coletada de diversos âmbitos. Esses dados são analisados e podem ser utilizados para inferir na tomada de decisões que contribuem para a evolução dependendo do contexto ou local onde será aplicado. Com isso, diversas indústrias estão investindo para obter vantagens com a utilização da análise de dados (WISAENG, 2013). Incluídas nessas empresas estão as de finanças ou bancárias. Segundo (PROVOST e FAWCETT, 2013) a análise de dados de propagandas em bancos é utilizada para o desenvolvimento de perfis de clientes. Com isso, favorecem a indicação de produtos ou serviços específicos para cada perfil (ELSALAMONY e ELSAYAD, 2013).

Essa análise de dados é feita através de Ciência de Dados (Data Science) que segundo (CIELEN, MAYSMAN e ALI, 2016) é o processo de extrair conhecimento a partir da análise de uma série de dados e fazer previsões a partir disso.

Nesse sentido, para a previsão é utilizado o conceito de Aprendizado de Máquina (*Machine Learning*) (GRUS, 2014) em que sua definição é a criação de modelos que aprendem a partir de dados analisados.

No contexto de *Machine Learning* existem dois tipos de problemas a serem resolvidos, que consistem nos problemas de Regressão, com variáveis-alvo quantitativas, e Classificação que por sua vez as variáveis-alvo são qualitativas (CAETANO, 2015).

Dentre os diversos algoritmos de aprendizado de máquina, este trabalho foca

na classificação com *K-Nearest Neighbors* (KNN) (ABDELMOULA, 2015) e Redes Neurais Artificiais (*Artificial Neural Network*) (KIRK, 2017). O algoritmo KNN calcula a distância Euclidiana dos pontos e os classifica a partir de suas proximidades. Já as Redes Neurais simulam neurônios artificiais com conjuntos de entradas e saídas em que cada conexão (neurônio) tem um peso associado (ELSAMONY, 2014).

O *dataset* explorado chama-se Bank Marketing Data Set e consiste em uma reunião de dados de uma campanha de marketing feita com os clientes de um banco de Portugal em que tinha como proposta a venda de uma aplicação bancária de longo prazo. Ele é composto por 45211 observações e 17 atributos onde é incluído os dados sobre se o cliente aceitou ou não a proposta além de características dos clientes, da campanha de propaganda e do próprio banco.

O objetivo da pesquisa é desenvolver algoritmos em Python que resolvam com Aprendizado de Máquina o problema de Classificação para prever se o cliente aceitará a proposta de investimento da propaganda ou não.

2 | METODOLOGIA

A base de dados chamada Bank Marketing Data set foi baixada do repositório de dados da UCI. Os atributos “*age*”, “*job*”, “*education*”, “*marital*”, “*default*”, “*balance*”, “*housing*” e “*loan*” contidos nela são relacionados ao cliente; “*contact*”, “*day*”, “*contact*”, “*month*”, “*duration*” e “*campaign*” referem-se a campanha em si e, por fim, “*pdays*”, “*previous*” e “*poutcome*” fazem parte do histórico bancário.

Toda a solução do problema foi implementada na linguagem Python no ambiente de desenvolvimento *Spyder*. Para o tratamento dos dados foi utilizada a biblioteca *Pandas*, com ela a análise dos dados se torna rápida e poderosa, junto com as bibliotecas de visualização gráfica *Matplotlib* e *Seaborn* onde foi feita a Análise Exploratória de Dados com a plotagem dos gráficos relação entre os clientes contatados e seu estado de empréstimo habitacional e pessoal, gráfico entre os dados de inadimplência dos clientes e a aceitação do investimento, gráfico de relação entre o número de ligações e a duração das chamadas, gráfico da taxa de inscrição na proposta pelo mês de contato, gráfico da taxa de inscrição pelo emprego do cliente e gráfico da taxa de inscrição por número de contato por mês. Na preparação dos dados para os testes, as variáveis “*job*”, “*education*”, “*housing*”, “*default*”, “*loan*” e “*y*” do *dataset* foram transformadas para variáveis do tipo *dummy* para evitar tendências na rede.

O data set foi dividido em variável alvo (*y*) e variáveis preditoras (todas as outras 16 variáveis). Para o cálculo do desempenho do classificador foi tirada a média das acurácias dos *Cross Validation* (CVs) de dez testes. O modelo utilizado para classificação foi o *K-Nearest Neighbors* (KNN) e Redes Neurais Artificiais Perceptron

Multicamadas (*Multilayer Perceptron - MLP*). Para o KNN foi realizado o modelo de seleção “*train_test_split*” que separa o *dataset* em conjunto de teste e conjunto de treinamento em uma proporção que evita tendências no modelo.

Os algoritmos foram implementados com auxílio da biblioteca de aprendizado de máquina *Scikit-Learn*. O algoritmo KNN foi treinado com o número de vizinhos igual a 6. A rede neural foi treinada com *cross-validation* igual a 10, com 3 camadas ocultas contendo, respectivamente, 100, 90 e 30 neurônios, todas com função de ativação *ReLU* e 1000 épocas.

3 | RESULTADOS

Foi realizado o processo de Análise Exploratória de Dados e a construção de gráficos que correlacionam as variáveis do problema para maior conhecimento do *dataset* a ser estudado.

As Fig. 1 e Fig. 2 relacionam o estado de empréstimo bancário dos clientes do banco que foram contatados na campanha. Na Fig. 1 os empréstimos habitacionais e em Fig. 2 empréstimos pessoais. Em ambas a taxa de aceitação da proposta é maior quando os clientes não possuem nenhum empréstimo.

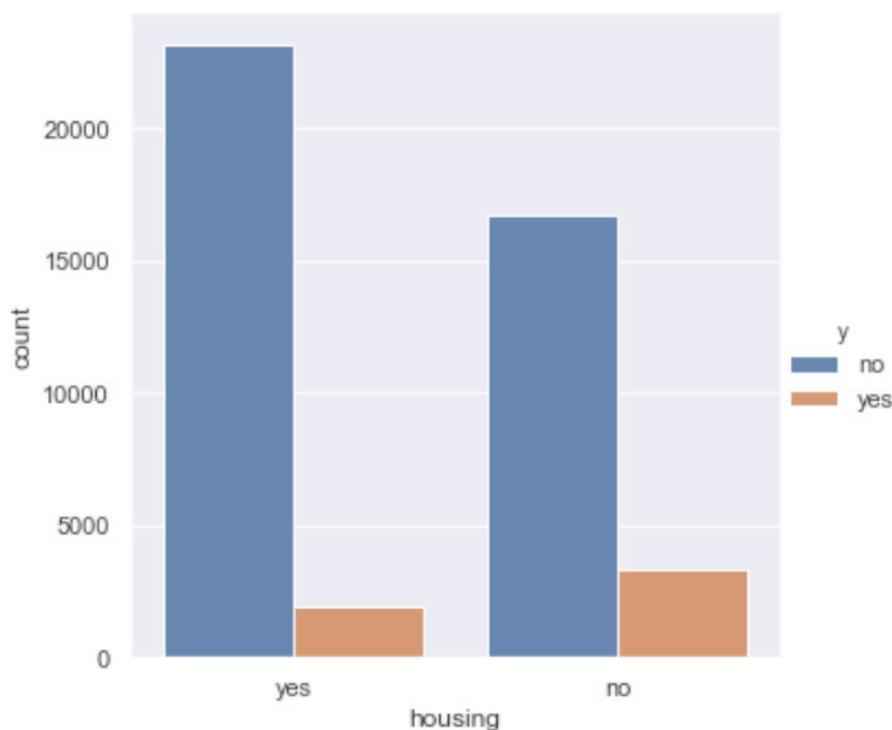


Figura 1: Relação entre os clientes contatados que possuem empréstimo habitacional e resposta da campanha.

Fonte: Autores.

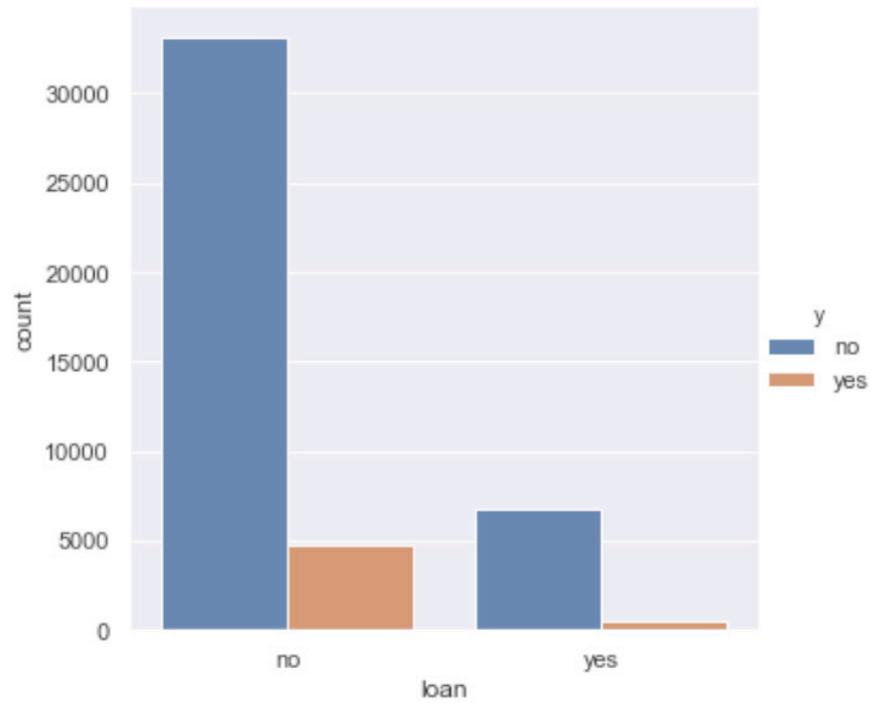


Figura 2: Relação entre o estado de empréstimo pessoal dos clientes e a resposta da proposta.

Fonte: Autores.

Na Fig. 3 a relação é entre os dados de inadimplência dos clientes no banco. Nela pode-se perceber que os clientes que se encontram no estado de inadimplência quase não aceitaram a proposta de investimento.

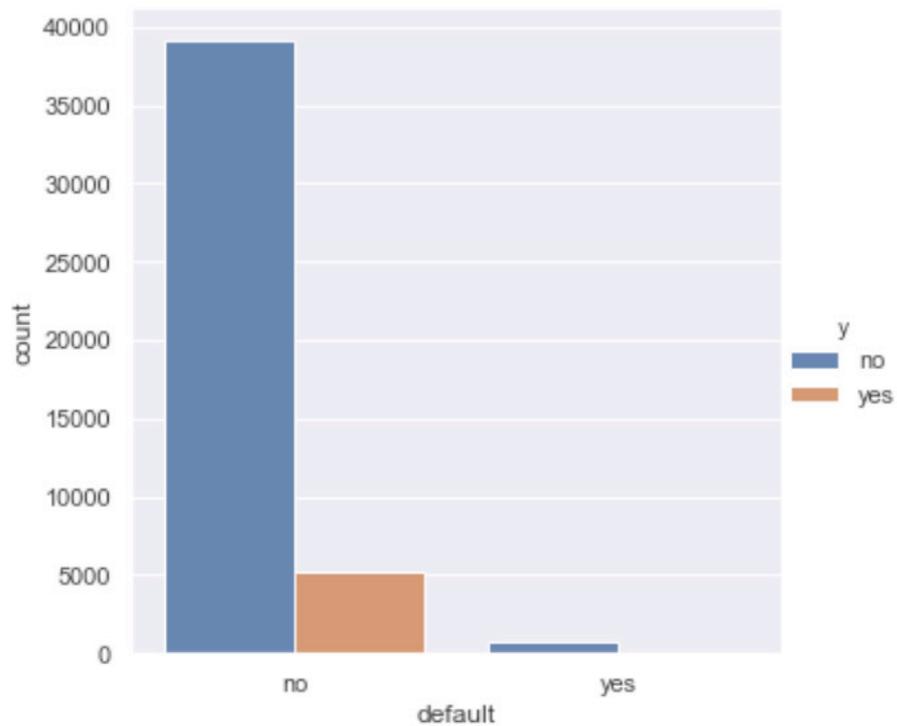


Figura 3: Relação entre os clientes contatados pela campanha que estão inadimplentes ou não e a resposta da campanha.

Fonte: Autores.

A Fig 4 configura-se um histograma com as contratações mensais feitas pela campanha. Com isso, é possível perceber que o mês em que a campanha mais divulgou o investimento para os clientes foi o mês de maio. Na Fig 5 por outro lado, que mostra a porcentagem de aceite da proposta por mês, ou seja, o mês em que as pessoas mais aceitam a proposta foi o mês de março.

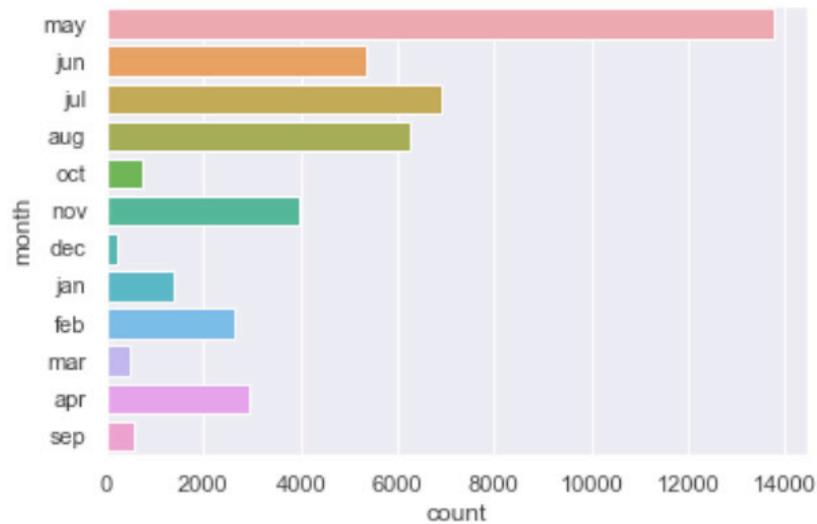


Figura 4: Número de contato na campanha por mês.

Fonte: Autores.

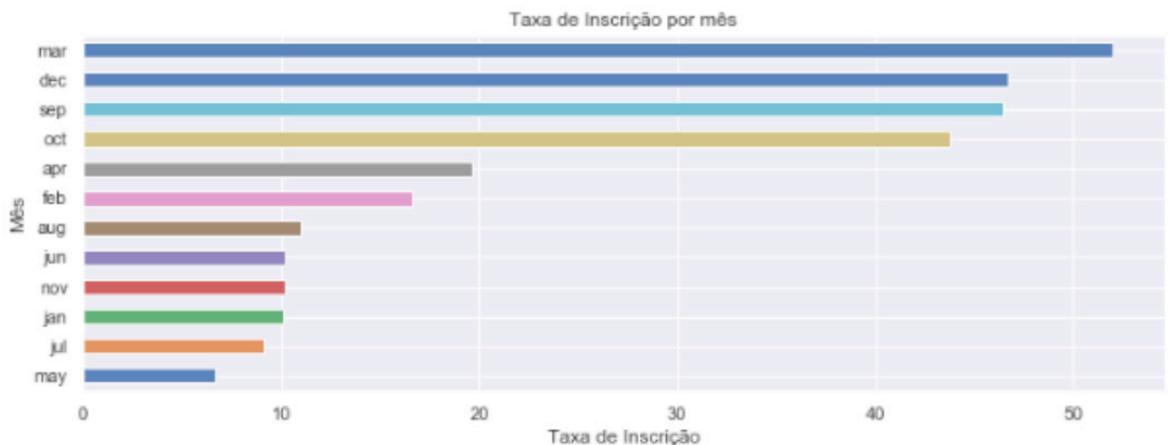


Figura 5: Relação da porcentagem de aceitação na proposta por mês.

Fonte: Autores.

A Fig 6 mostra a correlação entre duas colunas do *dataset* relacionadas as ligações feitas para os clientes que são “*campaign*” com o número de ligações feitas para cada cliente e “*duration*” com a duração das chamadas em segundos. Além disso, possui a resposta do cliente para cada caso o que informa uma correlação muito grande em ligações não muito longas com um baixo número de ligações.

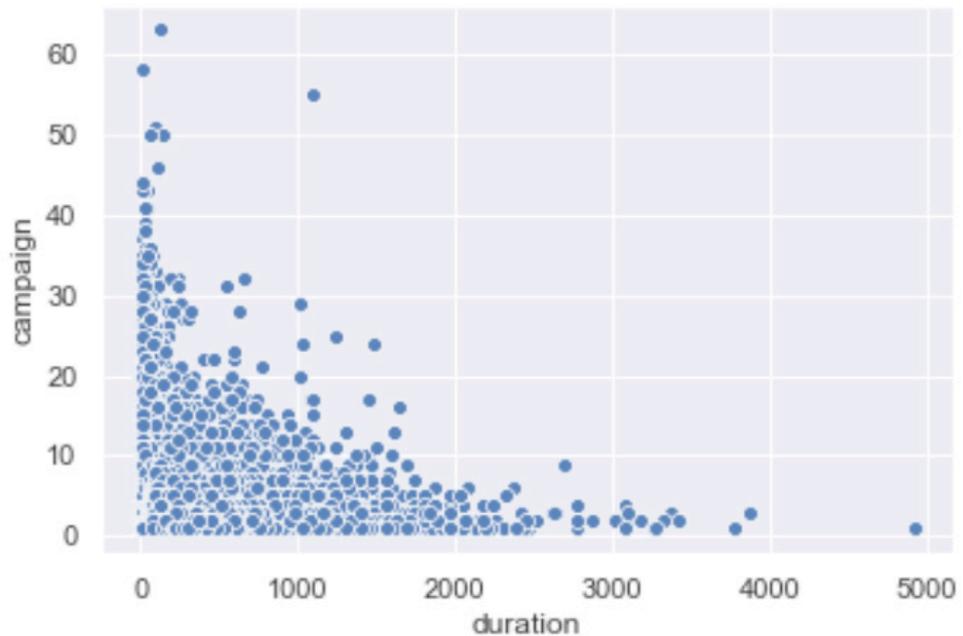


Figura 6: Relação entre duração da chamada e número de ligações com a resposta do cliente na campanha.

Fonte: Autores.

Com base na Análise Exploratória de Dados do problema de classificação as informações mais relevantes foram utilizadas, ou seja, as colunas diretamente relacionadas com a resposta na proposta de investimento. Sendo assim, as colunas “*marital*” que corresponde ao estado civil dos clientes, “*contact*” que correspondente ao tipo do meio de comunicação do cliente, “*day*” correspondente ao dia do mês do último contato, “*pdays*” número de dias que passaram depois do último contato na última campanha, “*previous*” correspondente a taxa de contato ao cliente antes da campanha e “*poutcome*” que corresponde ao resultado da campanha anterior para cada cliente, foram removidas do *dataframe*. Foram utilizadas, então, a coluna “*job*” com o tipo de emprego dos clientes, “*education*” com o tipo de estudo do cliente, “*age*” com a idade dos clientes, “*balance*” com o balanço bancário dos clientes, “*default*” com dados se o cliente está inadimplente ou não, “*housing*” com dados de empréstimo habitacional do cliente, “*loan*” com dados de empréstimo pessoal do cliente, “*month*” correspondente ao mês do último contato e “*campaign*” correspondente ao número de ligações realizados durante esta campanha e para cada cliente.

Com isso, como resultados dos testes de classificação encontram-se na Tabela 1. A média das acurácias com o algoritmo KNN foi em torno de 88,3% e com Redes Neurais MLP, a média foi de 88,2%, ou seja, aproximadamente iguais.

	MLP	KNN
	0.896	0.883
	0.911	0.882
	0.894	0.882
	0.873	0.883
	0.888	0.881
	0.873	0.884
	0.864	0.882
	0.883	0.882
	0.878	0.883
	0.861	0.885
Média	0.882	0.883
Desvio Padrão	0.0014	0.0014

Tabela 1: Resultados dos Testes e média dos valores.

Fonte: Autores.

4 | CONCLUSÃO

Em comparação, pode-se concluir que a classificação feita sobre a aceitação da proposta de marketing em uma aplicação bancária obteve um valor próximo de acurácia para o caso de uma Rede Neural Multi-layer Perceptron e o KNN, porém com uma diferença muito pequena, o que necessita, como trabalhos futuros, de mais testes para uma melhor comparação dos dois parâmetros. Por fim, foi possível prever a resposta de um cliente em uma campanha de marketing com base em dados coletados anteriormente durante a campanha. Além disso, foi possível obter informações sobre perfis de consumidores para possíveis melhorias nas próximas campanhas de marketing a serem realizadas.

AGRADECIMENTOS

Esse trabalho foi parcialmente financiado pelo Banco da Amazônia S.A. (projeto “Sistema para monitoramento de energia usando Internet das Coisas, Big Data e Machine Learning). Y.B. agradece o apoio do CNPq pelo apoio financeiro.

REFERÊNCIAS

ABDELMOULA, Aida Krichene. **Bank credit risk analysis with k-nearest-neighbor classifier: Case of Tunisian banks**. Accounting and Management Information Systems, v. 14, n. 1, p. 79, 2015.

CAETANO, Mateus et al. **Modelos de classificação: aplicações no setor bancário**. 2015.

CIELEN, Davy; MEYSMAN, Arno; ALI, Mohamed. **Introducing data science: big data, machine learning, and more, using Python tools**. Manning Publications Co., 2016.

ELSALAMONY, Hany A. **Bank direct marketing analysis of data mining techniques**. International Journal of Computer Applications, v. 85, n. 7, p. 12-22, 2014.

ELSALAMONY, Hany A.; ELSAYAD, Alaa M. **Bank direct marketing based on neural network and C5. 0 Models**. International Journal of Engineering and Advanced Technology (IJEAT), v. 2, n. 6, 2013.

GRUS, Joel. **Data science from scratch: first principles with python**. O'Reilly Media, 2019.

KIRK, Matthew. **Thoughtful Machine Learning with Python: A Test-driven Approach**. "O'Reilly Media, Inc.", 2017.

PROVOST, Foster; FAWCETT, Tom. **Data science and its relationship to big data and data-driven decision making**. Big data, v. 1, n. 1, p. 51-59, 2013.

WISAENG, K. **A comparison of different classification techniques for bank direct marketing**. International Journal of Soft Computing and Engineering (IJSCE), v. 3, n. 4, p. 116-119, 2013.

ÍNDICE REMISSIVO

A

Acompanhamento do consumo 28, 30
Aprendizado de máquina 19, 20, 21, 22, 37, 38, 83
Aprendizagem profunda 37

C

Campo eletromagnético clássico 70, 72, 79, 81
Ciclo de rankine 1, 4
Ciência de dados 19, 20
Comunidades rurais 109, 110
Constante dielétrica 11, 12, 13, 14, 15, 17
Coordenadas do cone de luz 70, 71, 72, 74, 75, 76, 77, 78, 79, 81
Cronogramas 59, 60, 61, 62, 67

E

Ees 1, 2, 4, 5, 9
Eficiência 1, 2, 4, 6, 7, 8, 9, 28, 29, 30, 31, 33, 35, 36, 37, 46, 50, 56, 84, 111, 115, 119
Eficiência energética 28, 29, 30, 31, 33, 35, 36, 119
Energia 2, 3, 4, 5, 6, 8, 11, 12, 26, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 41, 44, 49, 50, 75, 109, 110, 111, 112, 116, 117, 118, 119, 120
Energia elétrica 2, 3, 28, 29, 36, 37, 41, 44, 109, 110, 111, 117, 120
Energia solar 109, 110, 111, 117
Engenharia 7, 10, 11, 18, 36, 37, 44, 46, 58, 59, 62, 69, 103, 107, 108, 117, 119, 120
Ensino-aprendizado 98
Equação de klein-gordon-fock 70
Experimento 98, 100, 102, 104, 107

F

Falhas 61, 109, 115, 116, 117, 119
Fator de perda 11, 12, 13, 15, 16, 17
Filmes finos 46, 50, 51
Fotocatálise heterogênea 46, 47, 48

G

Gerenciamento 28, 29, 33, 59, 60, 61, 62, 63, 68, 69

I

Interdisciplinaridade 98, 102, 103, 106
Internet das coisas 26, 28, 30, 36, 44

K

K-nearest neighbors 19, 20, 21

M

Macaúba 11, 12, 13, 17, 18

Marketing bancário 19

Ms project 63

P

Potência elétrica 37

Processos oxidativos avançados 46, 47, 48, 57, 58

Projetos 30, 59, 60, 61, 62, 63, 67, 68, 69, 100, 103, 106, 120

Propriedades dielétricas 11, 12, 18

R

Realimentador 1, 2, 3, 4, 6, 7, 8, 9

Reconhecimento automático de modulações 83

Redes neurais 19, 21, 25, 38, 83, 84, 85, 87

Regressão 20, 37, 38, 43, 44, 89

Resnet 89, 90, 91

S

Substrato cerâmico 46

T

Tecnologia 1, 29, 36, 56, 57, 69, 70, 84, 98, 99, 100, 104, 105, 106, 107, 117, 120

Tempo de treinamento 83, 85, 92, 94

Tratamento de águas residuais 46, 57

 **Atena**
Editora

2 0 2 0