

Conceitos e Métodos de Inteligência Artificial para Análise de Dados

Artificial intelligence concepts and methods for data analysis

Gabriel Mendes Cirac¹, Tiago Cavalca², Luiz Fernando Nunes³, Leandro Duarte Pereira⁴, Valdinei de Paula Rodrigues⁵

1 – Doutor em Engenharia Elétrica – Unifei. E-mail.: gabriel_cirac@yahoo.com.br

2 – Bacharel em Sistemas de Informação pela FEPI. E-mail.: tiagocavalca47@gmail.com

3 – Mestre em Ciência e Tecnologia da Informação. Professor da FEPI. E-mail.: luiz.nunes@fepi.br

4 – Mestre em Engenharia de Produção e Professor da FEPI. E-mail.: leandro.pereira@fepi.br

5 – Mestre em Engenharia Eletrônica e Computação. Professor da FEPI. E-mail.: prof.valdinei.rodrigues@gmail.com

Recebido em: 25/09/2023

Revisado em: 17/10/2024

Aprovado em: 16/12/2024

Resumo: Inteligência Artificial (IA) é um conceito antigo, com debates especulativos datando de dois milênios, quando filósofos sonhavam em replicar a lógica humana por meio de dispositivos. Na origem, visavam entender processos de aprendizagem, memória, visão e raciocínio, visando a mecanização da inteligência. Com a evolução tecnológica, a IA ressurgiu como tópico relevante, introduzido sistematicamente na década de 1950 pelo artigo "Computing Machinery And Intelligence", prenunciando o potencial do aprendizado de máquina. Os avanços e diversificação da IA têm facilitado a vida humana, aplicando-se desde dispositivos eletrônicos, como Alexa, até análises de marketing e financeiras, auxiliando na tomada de decisões e automatização de processos. Este trabalho revisa a literatura de IA e realiza um estudo de caso com algoritmos de Classificação (Árvore de decisão, SVM e SGDClassifier), Agrupamento (K-Means), método Average Silhouette e decomposição PCA, dividindo os dados para treinamento, com foco na análise e exploração de dados de consumo. Os resultados são apresentados através de tabelas e gráficos comparativos, onde é possível verificar padrões de compras e perfis de consumo.

Palavras-chave: Inteligencia Artificial. Análise de Dados. Python. Aprendizado de Máquina.

Abstract: Artificial Intelligence (AI) is an ancient concept, with speculative debates dating back two millennia when philosophers dreamed of replicating human logic through devices. Originally, they aimed to understand processes of learning, memory, vision, and reasoning, seeking the mechanization of intelligence. With technological evolution, AI re-emerged as a relevant topic, systematically introduced in the 1950s by the article "Computing Machinery and Intelligence," foreshadowing the potential of machine learning. The advances and diversification of AI have facilitated human life, being applied from electronic devices like Alexa to marketing and financial analyses, assisting in decision-making and process automation. This work reviews the AI literature and conducts a case study with Classification algorithms (Decision Tree, SVM, and SGDClassifier), Clustering (K-Means), the Average Silhouette method, and PCA decomposition, splitting the data for training, focusing on the analysis and exploration of consumption data. The results are presented through comparative tables and graphs, where it is possible to identify purchasing patterns and consumption profiles.

Keywords: Artificial Intelligence, Data Analysis. Python. Machine Learning.

Introdução

Este trabalho aborda a inteligência artificial (IA), um campo fundamental da Ciência da Computação focado em equipar máquinas com capacidade de raciocínio semelhante aos humanos. A IA vem ganhando destaque em diversas áreas da computação, otimizando a interação entre máquinas e usuários (Mukhamediev et al, 2022).

Este estudo adota uma abordagem teórica e prática, incluindo uma revisão da literatura e uma análise de um conjunto de dados representando o comportamento do consumidor. O objetivo é desenvolver um algoritmo baseado em IA para entender os perfis de consumo e facilitar recomendações futuras.

A motivação deste trabalho reside na crescente importância da IA no cenário da computação e o impacto significativo que tem em diversas áreas. O estudo de caso busca agregar habilidades teóricas e práticas na linguagem Python, altamente demandada no mercado de IA.

Este estudo tem como objetivo investigar técnicas de IA e desenvolver uma solução de análise de dados usando aprendizado de máquina. Os objetivos específicos incluem uma revisão bibliográfica sobre IA, análise exploratória de dados, treinamento de algoritmos de classificação e agrupamento, entre outros.

O estudo está organizado em vários tópicos, que abrangem desde a história e desenvolvimento da IA, a aplicação prática, a ética em IA, até a metodologia usada, as ferramentas e técnicas aplicadas, os resultados alcançados e uma discussão sobre eles. A conclusão sintetiza as descobertas e observações da investigação.

Materiais e Métodos

Este estudo consiste em uma revisão bibliográfica e um estudo de caso envolvendo ferramentas de IA para contextualizar sua aplicabilidade geral, além de apresentar os métodos e algoritmos mais tradicionais, utilizando a linguagem Python e suas bibliotecas associadas ao tema. O objetivo é fornecer ao leitor uma visão geral sobre o assunto e ao mesmo tempo um guia prático que pode ser empregado em desafios reais da indústria e academia científica. Finalmente, os resultados do experimento prático do estudo de caso são avaliados e confrontados para permitir qualificar e quantificar os algoritmos empregados na investigação. Os softwares utilizados estão listados a seguir.

Google Colaboratory: O Google Colaboratory é um serviço gratuito fornecido pela Google que permite armazenar dados na nuvem para pesquisa de inteligência artificial e aprendizado de máquina. Ele permite o uso de blocos de código, textos com imagens e os

resultados dos códigos em um ambiente colaborativo (Johnston,2020).

Numpy ou Numerical Python, é uma biblioteca Python que oferece objetos de arrays (matrizes) multidimensionais e uma coleção de rotinas capazes de processar tais arrays. É utilizado em cálculos numéricos e tarefas matemáticas, sendo muito útil na execução de tarefas de Machine Learning e em diversas tarefas matemáticas (Junior,2018).

Scikit-Learn: Scikit-Learn é uma biblioteca Python para Machine Learning. Foi desenvolvida para suportar e possibilitar o treinamento de diversas técnicas de estatística e Machine Learning, tanto supervisionada quanto não supervisionada. Ela oferece um conjunto de ferramentas para todas as etapas de um projeto de Machine Learning (Cavalcanti, 2021).

Pandas Data Frame: Pandas é uma biblioteca Python usada para manipular e analisar dados. Ela oferece uma estrutura de dados bidimensional, potencialmente heterogênea e de tamanho variável, com rótulos em seus eixos (colunas e linhas). É especialmente útil para manipulação de dados e análise, com uma grande comunidade e muitos materiais de apoio disponíveis (Coutinho, 2021).

Base de dados: Para este estudo, um conjunto de dados sobre padrões de consumo foi escolhido. O conjunto de dados contém

2240 linhas e várias colunas contendo informações sobre os usuários e seus comportamentos de consumo. A base é referente a Marketing Digital. Nela se encontra dados pertinente ao usuário, que foram usados como base para o treinamento dos algoritmos selecionados.

Algoritmos: Diversos algoritmos foram utilizados para treinar na base de dados e gerar resultados para análise. Os algoritmos de classificação (Awari, 2020) usados foram a Árvore de Decisão (Campos, 2017), SVM (Santos, 2020) e SGDClassifier, e o algoritmo de agrupamento usado foi o Kmeans.

Resultados e Discussão

História da IA

A ideia de inteligência artificial teve origem a mais de dois mil anos (Gomes, 2010), com a busca por métodos ou dispositivos com a capacidade de reproduzir o raciocínio humano. Seu início aconteceu através de filósofos que procuravam entender os processos de aprendizagem, lembranças, visão e raciocínio; os quais fizeram várias tentativas de mecanizar a inteligência. Dada a limitação da tecnologia da época o conceito não se concretizou.

No ano de 1950 foi publicado um artigo chamado “Computing Machinery And Intelligence” (Turing, 1950) na revista Mind, onde o trabalho introduziu e sugeriu o

potencial das máquinas de pensar e resolver tarefas. O texto enviado não detalhava a técnica em si, ou tinha especificações para os profissionais da área de tecnologia da informação, podendo ser lido facilmente e entendido por pessoas de todas as áreas, era um trabalho introdutório. Alan Turing (conhecido como o pai da computação) seguiu a ideia do tema e indagou: “As máquinas podem pensar?”. Com este questionamento, iniciou-se aos primeiros esforços para criar técnicas capazes de endereçar esse desafio.

Seis anos depois, em 1956, logo após a segunda guerra mundial, em uma conferência no campus Dartmouth College foi fundado o campo de pesquisa em IA definindo como “a ciência e engenharia de produzir máquinas inteligentes” (Eliezer). Este acontecimento envolveu quatro investigadores da América do Norte (John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon) que organizaram e protagonizaram o evento (Sgarbosa et al., 2020), a conferência analisou tópicos sobre a computação, criatividade e processamento da linguagem. Por um mês o grupo debateu, trocando ideias sobre o assunto, onde originou-se o termo “Inteligência artificial”, que há mais de cinquenta anos tem como objetivo aproximar homens e máquinas (Isabel).

Desde a sua criação e com o avanço da tecnologia, várias melhorias foram agregadas

a IA. Em 1965 foi proposto o primeiro chat bot inteligente nos laboratórios do MIT (Massachusetts Institute of Technology). Outro grande avanço aconteceu em 1997 com a criação do Deep Blue pela IBM (Campbell et al., 2002), uma IA capaz de jogar xadrez. A partir de 2012 (Teixeira, 2013), com o incremento do hardware e avanço das GPU (Unidades de Processamento Gráfico) o conceito de Aprendizado Profundo (Deep Learning) evoluiu consideravelmente. Com toda evolução da IA e com o emprego massivo das técnicas inteligentes no mercado, o seu estudo tornou-se um atrativo muito grande no meio acadêmico. Por isso, o conhecimento das técnicas e empregos de IA é matéria fecunda atualmente.

Aplicações da IA no Contexto Atual

City Brains - Alibaba: Na China, duas aldeias chamadas City Brain e TaoBao iniciaram um projeto inovador que está sendo desenvolvido e implementado na cidade de Hangzhou. Esse projeto (Zhang et al., 2019) busca digitalizar atividades públicas e serviços que englobam todos os aspectos da vida em uma metrópole. Usando a nuvem da Alibaba, as ferramentas de IA desta gigante da tecnologia buscam analisar as demandas de serviços urbanos e os padrões de uso, a fim de otimizar sua eficiência. A Alibaba, devido ao sucesso dessa iniciativa, demonstrou um progresso significativo no pacote de soluções

para cidades inteligentes, expandindo seu processo para mais de vinte metrópoles na Ásia.

Alexa - Amazon: A IA trouxe transformações para a sociedade moderna, desde carros autônomos até assistentes de voz. Ao longo dos anos, os conceitos de computadores com a capacidade de conversação foram aprimorados. Tecnologias como a Internet das Coisas (IoT) e Interfaces de Voz com o Usuário (VUI) existem há muitos anos e são potencializadas pela IA, aprendizado de máquina, computação cognitiva e o rápido avanço na capacidade de miniaturização de componentes de hardware, aliado ao baixo custo dos mesmos. A Alexa, lançada pela Amazon em 2007 e introduzida no mercado em 2014 junto com o Echo (um alto-falante inteligente da mesma marca), funciona como assistente virtual de automação, focada em auxiliar a vida cotidiana de seus usuários. A Alexa pode realizar várias tarefas cotidianas, além de possuir funcionalidades que permitem realizar compras e comandar dispositivos de IoT em uma "Casa inteligente", tudo por comandos de voz.

DeepText – Facebook: A produção de sentenças de uma linguagem se baseia na combinação de palavras que geram uma mensagem adequada e significativa em um determinado contexto. A linguagem é uma das principais formas de inteligência que

permitiu aos seres humanos realizar uma série significativa de tarefas. Nesse sentido, o DeepText do Facebook (Zhong et al., 2016) utilizou redes neurais para ultrapassar o contexto limitado oferecido por outros processadores de texto, sendo capaz de escrever automaticamente blocos de texto completos ou palavras, baseado em uma entrada e de acordo com o cenário treinado.

Debater - IBM: O projeto Debater da IBM, criado em 2018 após dois anos de pesquisas na área de processamento de linguagem natural, oferece suas próprias APIs para o público em geral, fornecendo a equipes de ciências de dados a possibilidade de construir resumos e textos elaborados (Bar-Haim et al., 2021). Para que a IA pudesse debater com um humano, algumas habilidades tiveram que ser implementadas, como argumentar com relevância em relação ao tópico discutido e detectar a posição dos argumentos e avaliar sua qualidade.

Covid-19: O artigo (Manoj et al., 2022) apresenta uma análise abrangente da detecção da COVID-19 usando técnicas de aprendizado profundo em imagens radiográficas. Os resultados indicam que o aprendizado profundo é uma modalidade confiável e acessível para o diagnóstico preciso do vírus. Além disso, são discutidas a relação custo-efetividade e implicações financeiras desses métodos em comparação com outras abordagens de detecção. Essas

descobertas têm relevância para o aprimoramento da caracterização de imagens e a antecipação de futuras pandemias.

ChatGPT: O ChatGPT é um modelo de inteligência artificial treinado pela OpenAI baseado na arquitetura GPT (Generative Pretrained Transformer) (Salvagno, 2023). Ele usa aprendizado de máquina para gerar respostas textuais em conversas, sendo capaz de entender e responder a várias entradas do usuário. O GPT foi treinado em uma variedade de fontes na internet, o que lhe permite simular a conversação humana de forma natural e coerente. Além disso, é atualizado periodicamente para aprimorar sua capacidade de compreensão e resposta.

A IA e a Ética

A IA (Inteligência Artificial) tem se destacado nos últimos anos como uma das principais tecnologias criadas pelo homem. Sua capacidade de processar grandes quantidades de informação e conhecimento tem impacto significativo na tomada de decisões em diversos setores, como educação e automobilismo. No entanto, o uso de algoritmos de machine learning também levanta desafios éticos (Hallamaa et al., 2022).

Com a IA assumindo tarefas anteriormente realizadas por humanos, surgem novos desafios éticos relacionados aos algoritmos, que herdam requisitos

sociais. É fundamental compreender o impacto da IA nas organizações e na vida das pessoas, o que exige uma reflexão crítica sobre nossas responsabilidades.

Há uma série de desafios éticos a serem enfrentados no contexto da IA, como a redução do controle humano, a remoção da responsabilidade humana, a desvalorização das competências humanas, a erosão da autodeterminação humana, a facilitação de condutas controversas ou maléficas, o preconceito e a injustiça, entre outros. No entanto, a IA também pode ser utilizada como solução para esses desafios, como na redução de preconceitos conscientes e inconscientes.

A IA tem o potencial de trazer tanto benefícios quanto preocupações graves. A ética desempenha um papel importante ao colaborar com as empresas, fornecendo orientações e seguindo as recomendações dos órgãos reguladores. Isso contribui para a utilização saudável e eficaz da IA, criando uma base de diálogo sustentável com os reguladores e estabelecendo regras por meio de códigos de conduta, instrumentos jurídicos e outros mecanismos de co-regulação na governança do mundo digital. A IA (Inteligência Artificial) tem se destacado nos últimos anos como uma das principais tecnologias criadas pelo homem. Sua capacidade de processar grandes quantidades de informação e conhecimento tem impacto significativo na tomada de decisões em

diversos setores, como educação e automobilismo. No entanto, o uso de algoritmos de machine learning também levanta desafios éticos.

Com a IA assumindo tarefas anteriormente realizadas por humanos, surgem novos desafios éticos relacionados aos algoritmos, que herdam requisitos sociais. É fundamental compreender o impacto da IA nas organizações e na vida das pessoas, o que exige uma reflexão crítica sobre nossas responsabilidades.

Há uma série de desafios éticos a serem enfrentados no contexto da IA, como a redução do controle humano, a remoção da responsabilidade humana, a desvalorização das competências humanas, a erosão da autodeterminação humana, a facilitação de condutas controversas ou maléficas, o preconceito e a injustiça, entre outros. No entanto, a IA também pode ser utilizada como solução para esses desafios, como na redução de preconceitos conscientes e inconscientes.

A IA tem o potencial de trazer tanto benefícios quanto preocupações graves. A ética desempenha um papel importante ao colaborar com as empresas, fornecendo orientações e seguindo as recomendações dos órgãos reguladores. Isso contribui para a utilização saudável e eficaz da IA, criando uma base de diálogo sustentável com os reguladores e estabelecendo regras por meio de códigos de conduta, instrumentos jurídicos

e outros mecanismos de co-regulação na governança do mundo digital.

Conceitos Chave

O processo de aprendizado envolve a construção de algoritmos que "aprendem" a partir dos dados (Abu-Mostafa et al, 2012), reconhecendo padrões e fazendo previsões. Existem diferentes áreas do aprendizado de máquina, incluindo o aprendizado supervisionado, o aprendizado não supervisionado e o aprendizado por reforço. O aprendizado supervisionado utiliza dados rotulados para fazer previsões, enquanto o aprendizado não supervisionado busca identificar padrões e agrupar os dados sem informações de rótulos. Já o aprendizado por reforço envolve a interação de um agente com um ambiente, aprendendo com as ações e recompensas.

Dentro do aprendizado supervisionado, existem as subáreas de classificação e regressão (Castillo-Botón et al., 2022). A classificação envolve a previsão de uma observação em categorias, enquanto a regressão (Almeida et al., 2022) visa prever valores contínuos. O agrupamento é uma técnica de aprendizado não supervisionado que busca agrupar observações em clusters com base em suas similaridades (Probierz, Barbara, et al., 2022).

Os algoritmos de classificação mais comuns incluem as Máquinas de Vetores de

Suporte (SVM), Árvores de Decisão e o K-Nearest Neighbours (KNN). As Máquinas de Vetores de Suporte utilizam um hiperplano para separar os dados em classes (Hmeidi et al., 2007), enquanto as Árvores de Decisão utilizam estruturas de decisão para classificar os dados. O KNN classifica os dados com base na similaridade entre os atributos.

No caso dos algoritmos de regressão, eles são usados para prever valores contínuos. Esses algoritmos também fazem parte do aprendizado supervisionado e podem ser aplicados em problemas como previsão de preços ou valores.

A engenharia de atributos é uma etapa importante no processo de aprendizado de máquina, onde as características relevantes dos dados são identificadas e selecionadas para melhor descrever o problema.

O pré-processamento dos dados envolve técnicas de limpeza, transformação e redução de dados. A limpeza inclui lidar com dados ausentes, redução de ruído e resolução de inconsistências. A transformação envolve a normalização dos dados e a seleção de atributos relevantes. A redução de dados visa melhorar a eficiência e reduzir o custo computacional, através de técnicas como agregação de cubo de dados, seleção de subconjunto de atributos e redução de dimensionalidade. A padronização e normalização são técnicas usadas no pré-processamento de dados. A normalização

coloca os dados em um intervalo específico, enquanto a padronização os coloca na mesma escala, com média 0 e desvio padrão 1.

Resultados do Estudo de Caso

Os resultados obtidos neste trabalho serão apresentados por meio de imagens geradas a partir dos algoritmos de classificação, agrupamento e das comparações feitas por *plots*, além de serem anexados ao trabalho e à documentação do Google Colaboratory.

Os resultados avaliando os algoritmos de classificação são apresentados na Tabela 1, onde o leitor pode verificar a acurácia (precisão em determinar o tipo do perfil do consumidor de acordo com o padrão de consumo).

Tabela 1: Tabela de Resultados

Algoritmo	Treino	Teste
Árvore de Decisão	1	0.82
SVM	0.84	0.85
SGDClassifier	0.79	0.80

Os três algoritmos apresentaram resultados significativos, mantendo suas acurácias com um valor mínimo de 79.9%. A acurácia deve ser avaliada em um projeto de acordo com o cenário do mesmo. Por exemplo, um cenário bancário requer uma acurácia muito alta, idealmente acima de 99%, com a finalidade evitar todos os tipos de fraudes, enquanto um cenário de jogos online pode ter uma acurácia menor, pois o sistema

não é crítico. Pode-se dizer que nos três casos apresentados, os resultados foram interessantes. A árvore de decisão obteve 100% no conjunto de treino e 82% nos dados de teste, o que indica que seu treinamento foi satisfatório. Por outro lado, tanto o SVM quanto o SGDClassifier tiveram desempenho melhor em seus respectivos testes em comparação com os dados de treino, alcançando 85% e 84%, e 80% e 79%, respectivamente.

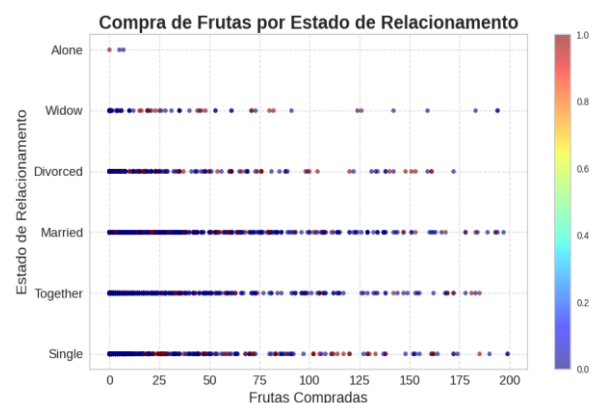
Resultados do Algoritmo de Agrupamento

Considerando o algoritmo K-Means (Alves, 2018), o resultado gerado foi um gráfico identificando os grupos e seus centróides, conforme ilustrado na Figura 2. O resultado mostrou-se inconclusivo, o que é natural em problemas de agrupamento, pois os centróides dos grupos 0 e 1 estão relativamente próximos um do outro, enquanto o centróide do grupo 2 está totalmente isolado. Por isso, foi necessário utilizar outro método para obter os resultados, que foi feito por meio de comparações usando plots de dispersão. Assim, foi possível analisar informações sobre a base de dados agrupando essas informações em pares e comparando com a resposta. A escolha dos pares considerou diversos cenários hipotéticos para testar diferentes refalações, como a relação entre número de visitas ao website e quantidade de compras online, a

ideia é verificar como as variáveis se influenciam.

A Figura 1 mostra um cenário que relaciona a compra de frutas e o estado de relacionamento. Pode-se observar que entre as possíveis compras, são de usuários dos grupos 4 e 6, casados e solteiros, o que indica maior aquisição de frutas por estes clientes dos quais tiveram suas informações de compras inseridas na base de dados para a realização deste comparativo e obtenção deste resultado.

Figura 1: Relação entre compra de frutas e estado de relacionamento.



A Figura 2 mostra um cenário de análise das compras de carne por pessoas que possuem crianças em casa. Nesta condição, é possível concluir que há um maior consumo do grupo de pessoas que não possuem crianças, seguidas pelo grupo de pessoas que possuem uma criança apenas e, logo após, as que têm duas crianças em casa.

A Figura 3 mostra um cenário de compras relacionado à renda versus quantidade de vinho adquirida. Nota-se que

as aquisições dos produtos são realizadas por pessoas com renda de até 100.000 dólares, por ano, sendo que a quantidade aumenta com o salário.

Figura 2: Relação entre compra de carne e presença de crianças em casa.

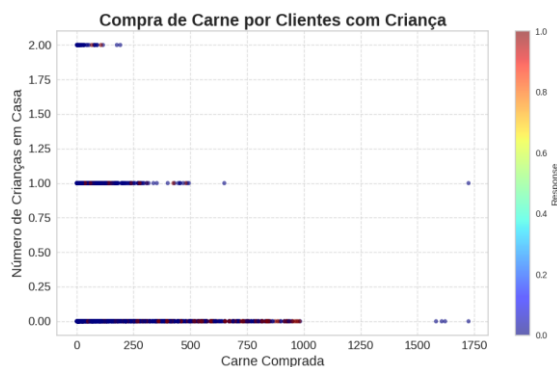


Figura 3: Relação entre renda e quantidade de vinho adquirida.

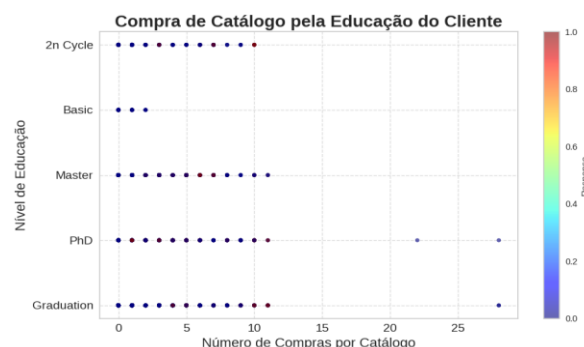


A Figura 4 mostra um cenário da quantidade de compras no catálogo em relação ao nível educacional dos usuários. O ciclo básico foi o grupo com menor probabilidade de compra, já os níveis de graduação e PhD apresentam uma maior probabilidade de compra do catálogo da loja.

A Figura 5 mostra um cenário onde é possível analisar a quantidade de visitas ao website por mês em relação à quantidade de compras online. A análise demonstra que existem casos de compras de produtos a partir

da segunda visita ao website, porém o número de compras é consolidado a partir da terceira visita, onde a maioria das compras são realizadas. A partir da terceira visita, há uma queda no número de compras por visita.

Figura 4: Relação entre quantidade de compras no catálogo e nível educacional dos usuários.



Com relação aos resultados, pode-se afirmar que os algoritmos de classificação apresentaram uma acurácia satisfatória. Foi surpreendente observar que a Árvore de Decisão alcançou uma acurácia de 100% nos dados de treino, indicando um bom desempenho após o preparo adequado da base de dados. No entanto, é importante ressaltar que a avaliação principal deve ser realizada nos dados de teste para uma análise mais precisa.

Figura 5: Relação entre número de visitas ao website e quantidade de compras online.



Quanto ao algoritmo de agrupamento, os primeiros resultados obtidos não foram conclusivos, devido à divisão dos centroides e à dificuldade de interpretação. Isso evidenciou a necessidade de uma exploração mais aprofundada, utilizando o método de comparação das características por meio de gráficos de dispersão e relacionando-as com a variável de resposta. Essa abordagem permitiu obter os resultados desejados para concluir a Análise e Exploração. Foi possível estabelecer critérios de comparação utilizando as features predefinidas para investigar combinações específicas. Além disso, uma vez que o método está pronto, é possível combinar diferentes características para análises mais abrangentes.

Conclusão

Este trabalho consistiu em um estudo de caso de Inteligência Artificial (IA) utilizando uma base de dados representando padrões de consumo, e os resultados obtidos sugerem um grande potencial de aplicações nessa área. A revisão bibliográfica e teórica desempenharam um papel fundamental no entendimento abrangente dos tópicos abordados. Os métodos de IA propostos foram aplicados com sucesso nas etapas de desenvolvimento, treinamento e teste. Foi necessário realizar uma revisão bibliográfica sobre o tema, e o embasamento teórico auxiliou na estruturação dos métodos

algorítmicos, com base na proposta do estudo de caso. As metas estabelecidas foram alcançadas e podem ser resumidas da seguinte forma:

- Realização de pesquisas bibliográficas sobre IA, incluindo sua história, aplicações no contexto atual e a ética envolvida nesse campo;
- Exploração de pesquisas relacionadas a Aprendizado de Máquina, Algoritmos (Classificação, Regressão, Agrupamento), Engenharia de Atributos (Features) e Técnicas de Pré-processamento;
- Busca por uma base de dados adequada ao estudo, que permitisse a aplicação dos métodos;
- Definição dos métodos e materiais utilizados durante o desenvolvimento do cenário, para realização dos testes e treinamento da IA;
- Aplicação dos métodos para gerar resultados, seguida da análise e exploração dos mesmos por meio da criação de gráficos e comparação das variáveis de interesse com a *feature "response"*.

Através da análise dos resultados, podemos afirmar que a IA pode ser empregada com sucesso na análise de padrões de consumo e trazer valor para diversos tipos de negócios, além de facilitar a compreensão das necessidades e demandas dos clientes. Por fim, é importante ressaltar que, devido à limitação dos dados utilizados, que são teóricos, uma base de dados real e

mais abrangente permitiria uma análise mais aprofundada e condizente com a realidade.

Referências

- Abu-Mostafa, Yaser S., et al. Learning from Data: A Short Course. AMLbook.com, 2012.
- Bar-Haim, Roy, et al. "Project Debater APIs: Decomposing the AI Grand Challenge". Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Association for Computational Linguistics, 2021, p. 267–74. DOI.org (Crossref), <https://doi.org/10.18653/v1/2021.emnlp-demo.31>.
- Campbell, Murray, et al. "Deep Blue". Artificial Intelligence, vol. 134, no 1–2, janeiro de 2002, p. 57–83. DOI.org (Crossref), [https://doi.org/10.1016/S0004-3702\(01\)00129-1](https://doi.org/10.1016/S0004-3702(01)00129-1).
- Castillo-Botón, C., et al. "Machine Learning Regression and Classification Methods for Fog Events Prediction". Atmospheric Research, vol. 272, julho de 2022, p. 106157. DOI.org (Crossref), <https://doi.org/10.1016/j.atmosres.2022.106157>.
- Cavalcanti, Toti. "Aula 08 – Scikit-Learn – Support Vector Machine ou máquina de vetores de suporte". 2021. Disponível em: <https://www.codigofluente.com.br/aula-08-scikit-learn-maquina-de-vetores-de-suporte/>
- Coutinho Thiago. "Quais são as vantagens e funcionalidades da biblioteca Pandas data frame?". 2021. Disponível em: <https://www.voitto.com.br/blog/artigo/dataframe/amp?gclid=Cj0KCQiAsdKbBhDHARIsANJ6-jcnbY0V1aB0Zw9b9QCLnlm8I5GWV3Du6tADQMVGiHjRDIFGjSZ9NpwaAsZAEALwwcB>
- Johnston Ben. "Educational Classroom Technologies". 2020. Disponível em: <https://mcgrawect.princeton.edu/tool/google-colab/>
- Junior S Luiz. "Entendendo a biblioteca NumPy". 2018. Disponível em: <https://medium.com/ensina-ai/entendendo-a-biblioteca-numpy-4858fde63355>
- Gomes, Dennis dos Santos. "Inteligência Artificial: conceitos e aplicações." *Revista Olhar Científico* 1.2 (2010): 234-246.
- Hallamaa, Jaana, e Taina Kalliokoski. "AI Ethics as Applied Ethics". *Frontiers in Computer Science*, vol. 4, abril de 2022, p. 776837. DOI.org (Crossref), <https://doi.org/10.3389/fcomp.2022.776837>.
- Hmeidi, Ismail, et al. "Performance of KNN and SVM Classifiers on Full Word Arabic Articles". *Advanced Engineering Informatics*, vol. 22, no 1, janeiro de 2008, p. 106–11. DOI.org (Crossref), <https://doi.org/10.1016/j.aei.2007.12.001>.
- Mukhamediev, Ravil I., et al. "Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges". *Mathematics*, vol. 10, no 15, julho de 2022, p. 2552. DOI.org (Crossref), <https://doi.org/10.3390/math10152552>.
- Manoj Kumar, et al. "Detection of COVID-19 Using Deep Learning Techniques and Cost Effectiveness Evaluation: A Survey". *Frontiers in Artificial Intelligence*, vol. 5, maio de 2022, p. 912022. DOI.org (Crossref), <https://doi.org/10.3389/frai.2022.912022>.
- Probierz, Barbara, et al. "Clustering of Scientific Articles Using Natural Language Processing". *Procedia Computer Science*, vol. 207, 2022, p. 3449–58. DOI.org (Crossref), <https://doi.org/10.1016/j.procs.2022.09.403>.
- Salvagno, M., Taccone, F.S. & Gerli, A.G. Can artificial intelligence help for scientific writing?. *Crit Care* 27, 75 (2023). <https://doi.org/10.1186/s13054-023-04380-2>
- Sgarbosa, Pietro, e Gustavo Henrique Del Vecchio. "INTELIGÊNCIA ARTIFICIAL E SUAS IMPLICAÇÕES: como os dispositivos inteligentes e assistentes virtuais influenciam o cotidiano das pessoas". *Revista Interface Tecnológica*, vol. 17, no 2, dezembro de 2020, p. 193–205. DOI.org (Crossref), <https://doi.org/10.31510/inf.v17i2.936>.
- Teixeira, Marcelo Andrade. "Sistemas Inteligentes para Tratamento das Estratégias e Táticas do Go." *10ª Semana de Pesquisa Extensão e Pós-Graduação*. 2013.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59, 433–460.
- Zhang, Jianfeng & Hua, Xian-Sheng & Huang, Jianqiang & Shen, Xu & Chen, Jingyuan & Zhou, Qin & Fu, Zhihang & Zhao, Yiru. (2019). The City Brain: Practice of Large-Scale Artificial Intelligence in the Real World. *IET Smart Cities*. 1. 10.1049/iet-smc.2019.0034.
- Zhong, Zhuoyao, et al. DeepText: A Unified Framework for Text Proposal Generation and Text Detection in Natural Images. *arXiv*, 24 de maio de 2016. [arXiv.org, http://arxiv.org/abs/1605.07314](http://arxiv.org/abs/1605.07314).