

ESTUDO INTRODUTÓRIO DA UTILIZAÇÃO DO CHATGPT POR DISCENTES DE CIÊNCIA DE DADOS

LEONARDO HENRIQUE DA SILVA MUNIZ¹
PAULO ROBERTO DA SILVA RUIZ¹

RESUMO

Atualmente, há um grande interesse no impacto que a inteligência artificial possui em diversos setores da sociedade, especialmente no ensino, onde ferramentas foram desenvolvidas para aliar rapidez na busca, praticidade e alta disponibilidade de conteúdo. Essas tecnologias são capazes de extrair dados e fornecer resultados com alto índice assertivo, proporcionando melhorias significativas nos estudos e aplicações no âmbito acadêmico. Este estudo combinou análises estatísticas, como contagens e médias, com análises qualitativas, explorando os sentimentos, motivações e desafios dos usuários. Essa abordagem integrada permitiu uma compreensão holística das necessidades dos usuários e fortaleceu a validade dos resultados, proporcionando insights valiosos para o desenvolvimento de produtos e serviços baseados no usuário. A pesquisa realizada com os estudantes do curso de Ciência de Dados revelou que todos utilizam o *ChatGPT* de alguma forma em seu dia a dia acadêmico. No contexto escolar, a ferramenta é utilizada principalmente para pesquisa sobre os temas abordados em sala de aula. Com base nos resultados obtidos, ficou evidente a importância dessa ferramenta no fortalecimento do aprendizado individual, tanto de forma educacional quanto prática.

Palavras-chave: Inteligência Artificial; Ciência cognitiva; Ciência de Dados.

ABSTRACT

There is currently significant interest in the impact of artificial intelligence across various sectors of society, particularly in education, where tools have been developed to combine speed in search, convenience, and high availability of content. These technologies can extract data and provide results with a high level of accuracy, leading to significant improvements in studies and applications within the academic realm. This study combined statistical analyses, such as counts and averages, with qualitative analyses, exploring users' feelings, motivations, and challenges. This integrated approach allowed for a holistic understanding of users' needs and strengthened the validity of the results, providing valuable insights for the development of user-based products and services. Research conducted with students in the Data Science course revealed that all of them use ChatGPT in some form in their academic daily lives. In the educational context, the tool is primarily used for research on topics covered in class. Based on the results obtained, the importance of this tool in strengthening both educational and practical individual learning became evident.

Keywords: Artificial Intelligence; Cognitive science; Data Science.

¹ Faculdade de Tecnologia de Adamantina (Fatec), Adamantina-SP, Brasil

1 INTRODUÇÃO

A utilização do *ChatGPT* no ensino superior e na área de Ciência de Dados é importante devido à sua capacidade de fornecer respostas precisas e rápidas, extrair informações de grandes volumes de dados, e facilitar a pesquisa e colaboração acadêmica. Essa tecnologia demonstra um avanço significativo, na qual promove melhorias nos estudos e aplicações realizadas na graduação, além de transmitir o potencial transformador da inteligência artificial nesse setor. (Ferreira; Garcia, 2023).

A Ciência de Dados emerge como um campo necessário na era da informação, desempenhando um papel fundamental na transformação digital e na resolução de problemas complexos. O curso de Ciência de Dados da Fatec Adamantina combina teoria e prática, capacitando os alunos com habilidades técnicas avançadas e preparando-os para enfrentar os desafios contemporâneos da sociedade (Rautenberg; Carmo, 2019).

No contexto atual, onde a informação é abundante e a interpretação precisa é essencial, os Cientistas de Dados se destacam como especialistas para extrair dados valiosos. Esses profissionais não apenas dominam técnicas estatísticas e habilidades de programação, mas também possuem uma compreensão profunda do contexto empresarial e das necessidades do mercado. Além disso, utilizam ferramentas avançadas para desenvolver modelos preditivos, otimizar processos e tomar decisões fundamentadas (Sandes, 2019).

Em um cenário onde a interação humana é cada vez mais mediada pela tecnologia, ferramentas como os chatbots desempenham um papel fundamental. Esses assistentes virtuais não só auxiliam no estudo e na pesquisa, mas também facilitam o trabalho profissional ao automatizar tarefas repetitivas e fornecer suporte personalizado em tempo real. Assim, a integração das ferramentas de chatbot com a Ciência de Dados não apenas melhora a eficiência e a produtividade, mas também promove uma abordagem mais acessível e interativa para o aprendizado e o trabalho profissional (Monteiro, 2023).

1.1 Problemática

Como o *ChatGPT* pode ajudar os estudantes de Ciência de Dados a inovarem

nos métodos de pesquisa, especialmente enfrentando os desafios na interpretação e compreensão dos resultados? Essa abordagem inclusiva e diversificada poderia melhorar o processo para o desenvolvimento de habilidades analíticas mais robustas e para uma compreensão mais profunda dos princípios subjacentes aos métodos de análise de dados, preparando assim os estudantes para lidar com a crescente complexidade e diversidade dos problemas enfrentados no campo da Ciência de Dados.

1.2 Justificativa

As Redes Neurais Artificiais (RNAs) destacam-se como ferramentas poderosas no campo de Ciência de Dados, proporcionando amplo processamento e interpretação de dados complexos. Ao compreender o potencial transformador dessas tecnologias, é possível explorar novas abordagens para aprimorar o aprendizado dos alunos. O *ChatGPT*, desenvolvido com base em RNAs, apresenta uma abordagem inovadora para a interatividade aluno-dados. Sua capacidade de simular conversas humanas pode facilitar a compreensão de conceitos complexos, tornando o processo de aprendizado mais dinâmico e acessível. A pesquisa conduzida com os alunos do CST em Ciência de Dados da Fatec Adamantina – SP fornecerá uma percepção e aceitação do *ChatGPT* como uma ferramenta auxiliar nos estudos, contribuindo para avaliar seu impacto prático e sua relevância no contexto educacional. Dessa forma, a investigação sobre essa integração na educação de Ciência de Dados emerge como uma possibilidade em prol do aprendizado dos discentes.

1.3 Objetivo

Este artigo tem como objetivo compreender a importância do *ChatGPT* na formação de alunos que cursam Curso Superior de Tecnologia em Ciência de dados na Fatec Adamantina, explorando suas aplicações, benefícios e contribuições para o aprendizado.

Para atingir esse objetivo, este trabalho contemplará os seguintes objetivos específicos:

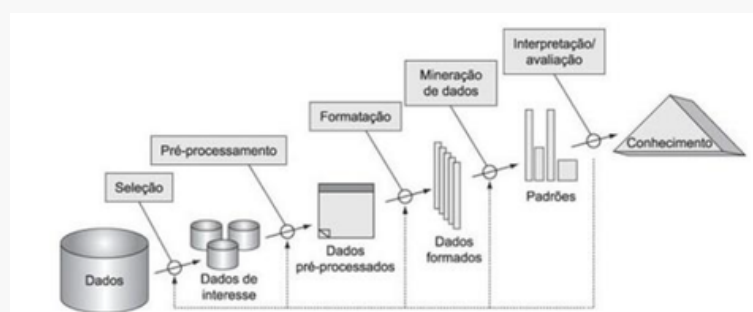
- Analisar a importância da RNAs e seus fundamentos. Para isso, será explorada a funcionalidade desse algoritmo para processar dados e conduzir aprendizados, evidenciando seus fundamentos.
- Compreender as funcionalidades do *ChatGPT* e seu papel fundamental no desenvolvimento de modelos de análise de dados utilizando Inteligência Artificial (IA).
- Realização de uma pesquisa junto aos alunos do CST em Ciência Dados da Fatec Adamantina – SP, a qual levantará a opinião dos alunos sobre o *ChatGPT* e sua utilização na ajuda aos estudos.

2 FUNDAMENTAÇÃO TEÓRICA

O processo conhecido como *knowledge-discovery in databases (KDD)*, equivalente em português à Descoberta de Conhecimento em Banco de Dados, abrange diversas etapas na identificação e remoção de informações em grandes conjuntos de dados para a extração de conhecimento.

A Figura 1 apresenta o fluxo de trabalho no contexto do KDD. A primeira etapa envolve a escolha de um agrupamento de dados. Isso permite que informações relevantes sejam extraídas por meio da mineração de dados, seja por meio de conjuntos de dados completos ou subconjuntos criteriosamente selecionados de variáveis. Por sua vez, a segunda etapa está concentrada na *Data Cleaning* (Limpeza de dados) e Pré-processamento, as quais consistem na organização dos dados de maneira que a qualidade seja uma característica intrínseca das operações de KDD, separando ruídos, também conhecidos como atributos nulos ou inconsistentes (Dantas et al., 2008). A terceira etapa começa com a transformação dos dados previamente selecionados, que já passaram pelo processo de limpeza e pré-processamento (Fayyad et al., 1996).

Figura 1: Fases do KDD



Fonte: Adaptado de Fayyad et al. (1996)

A etapa de seleção de dados determina com discernimento quais conjuntos de dados têm o potencial de alcançar as melhores informações. A seguir, na etapa de pré-processamento, realiza-se uma limpeza meticulosa e a seleção cuidadosa de atributos, buscando selecionar os melhores dados. Corrigindo informações ausentes, errôneas ou inconsistentes, garantindo que a pureza dos dados seja mantida, o que é essencial para manter a qualidade dos modelos de conhecimento que serão criados. Na etapa de transformação, os dados brutos são convertidos em uma forma que se torna mais fácil de entender. Uma cuidadosa seleção de método e algoritmo mais afinado com os objetivos da extração revelará padrões marcantes nos dados, esse é o contexto da mineração de dados. Esses padrões servirão como alicerces fundamentais na jornada da descoberta de informação e conhecimento oculto (Félix, 1998).

A mineração de dados pode contemplar diferentes metodologias. O aprendizado supervisionado é um tipo de modelo que aprende com exemplos rotulados, ou seja, dados que já têm uma classificação conhecida. Ele é muito útil em tarefas como classificação e regressão, onde pode entender diferentes padrões e contextos. Por exemplo, ao analisar avaliações, o aprendizado supervisionado pode ser treinado para automaticamente classificar sentimentos, facilitando a organização e interpretação de grandes conjuntos de dados. No aprendizado não supervisionado, os algoritmos tentam encontrar estruturas ou padrões que consigam analisar os dados sem orientação de rótulos, ou classes. Esse aprendizado normalmente é utilizado em atividades de clusterização, onde irá agrupar os dados com base em características semelhantes (Sanches, 2003).

A última etapa contempla a interpretação de resultados, na qual são avaliados o desempenho dos modelos de classificação. As regras utilizadas para minerar os dados são cuidadosamente decifradas e avaliadas. Nesse momento, surgem diversos padrões e relacionamentos, os quais lançam luz sobre novas descobertas que podem ser usadas em pesquisas, otimização e outras aplicações (Fayyad et al., 1996).

2.1 Rede Neural Artificial

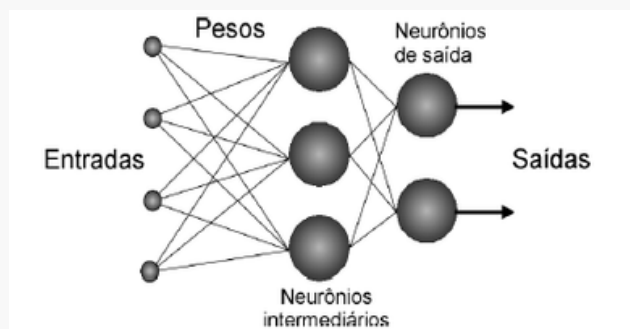
As RNAs têm uma história que remete à década de 1940, McCulloch e Pitts (1943) apresentaram um modelo matemático inspirado no sistema nervoso biológico, especialmente no cérebro humano. No entanto, naquela época, o avanço prático das RNAs foi limitado devido à falta de poder computacional e

dados necessários para treiná-las. Foi apenas nas décadas de 1980 e 1990 que as RNAs ressurgiram com força, graças à popularização de algoritmos de treinamento eficientes, destacando-se o famoso algoritmo de retropropagação. Nesse período, modelos como o Perceptron Multicamadas (Multilayer Perceptron - MLP) que se fundamenta em um tipo de arquitetura de rede neural, na qual se baseia em múltiplas camadas de neurônios, incluindo uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída (Bonini Neto; Bonini, 2010). Nos últimos dez anos, o campo de *Deep Learning* (Aprendizado Profundo, a qual presume-se muitas camadas ocultas) assumiu o protagonismo. Modelos como as Redes Neurais Convolucionais (*Convolutional Neural Network* - *CNNs*), utilizadas em processamento de imagens, e as Redes Neurais Recorrentes (*Recurrent Neural Network* - *RNNs*), aplicadas em sequências temporais, tendo impacto em áreas como reconhecimento de imagem, processamento de linguagem natural e jogos. Esses avanços têm transformado significativamente a forma como lidamos com informações e dados em várias disciplinas. (Lopez; Mattos, 2021).

Uma vez ajustadas, tornam-se poderosas ferramentas no campo da ciência de dados e da inteligência artificial, capacitando a rápida classificação e agrupamento de dados. Tarefas como o reconhecimento de fala ou de imagens podem ser realizadas em questão de minutos, em vez de horas, quando comparadas com a identificação manual realizada por especialistas. (Walter, 2005).

A disseminação e implementação prática de modelos de RNAs foram facilitadas por plataformas como *TensorFlow* e *PyTorch*, ampliando sua aplicabilidade em uma variedade de domínios e setores. Esse desenvolvimento contínuo destaca esses algoritmos como ferramenta fundamental na vanguarda da pesquisa e aplicação em Inteligência Artificial. A Figura 2 apresente uma arquitetura típica de uma RNA com 4 neurônios de entrada, 3 na camada oculta e 2 neurônios na de saída.

Figura 2: Exemplo de arquitetura de RNA

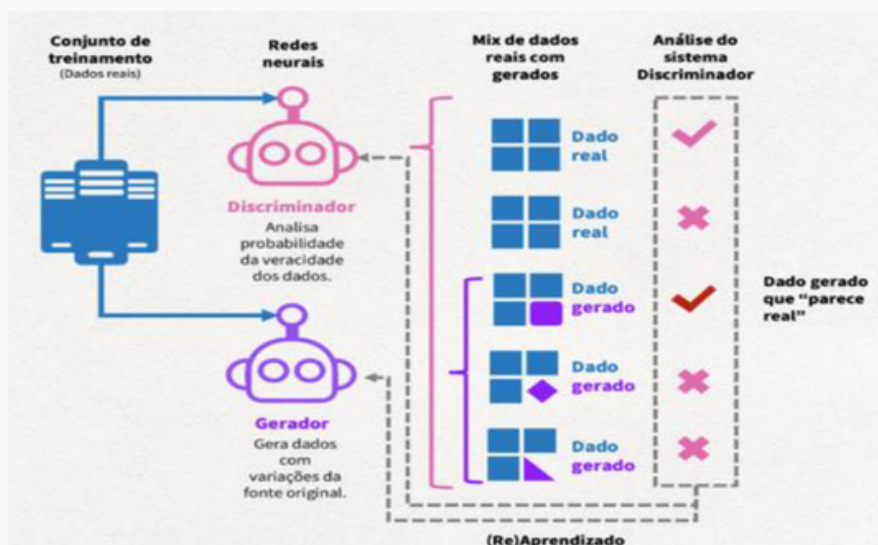


Fonte: Adaptado por Tafner (1998)

2.2 Rede Neural Generativa

As Redes Neurais Generativas (RNGs) representam uma classe significativa de modelos no campo do aprendizado de máquina, destacando-se pela capacidade de gerar dados que se assemelham a um conjunto original, sem depender de uma estrutura específica dos dados de entrada. Conforme a Figura 3, consistem em dois componentes principais: o Gerador e o Discriminador. Onde o primeiro examina um conjunto de dados de treinamento para entender seus principais atributos e padrões. Enquanto isso, o segundo também estuda esses mesmos dados, aprendendo a diferenciar entre exemplos reais e falsos. Depois dessa análise inicial, o Gerador cria amostras de dados, adicionando algumas mudanças aleatórias para introduzir variação (mix de dados reais com gerados). Essas novas amostras são então enviadas ao Discriminador, que verifica a probabilidade de cada amostra ser parte do conjunto de dados original, dando um feedback ao Gerador, dizendo o quão realistas são as amostras geradas. Com isso, o Gerador ajusta suas estratégias para fazer amostras ainda mais convincentes na próxima rodada. Esse processo continua de forma iterativa. De certa forma, o Gerador tenta enganar o Discriminador com amostras cada vez mais realistas, sendo que esse fica cada vez melhor em distinguir entre dados reais e gerados. Com o tempo, o Gerador se torna tão bom em criar dados que são quase idênticos aos dados reais, resultando em uma IA muito eficiente. Esse processo é formulado como um jogo adversarial, onde ambos componentes estão em competição, buscando atingir um equilíbrio dinâmico (Paula, 2023).

Figura 3: Visão simplificada do funcionamento da IA Generativa



Fonte: Adaptado Paula (2023)

As RNGs têm se destacado como uma ferramenta poderosa para a criação de dados sintéticos, proporcionando uma contribuição valiosa em áreas onde a disponibilidade de dados autênticos é restrita ou quando se busca inovação na geração de conteúdo visual e textual (Rodrigues, 2021).

2.3 ChatGPT e RNAs

Nos últimos anos, o avanço da Inteligência Artificial revolucionou a forma como as pessoas interagem com a tecnologia em várias áreas, desde desenvolvimento até pesquisa e análise (Silva et al., 2022). Uma dessas criações é o *ChatGPT*, uma aplicação que utiliza Inteligência Artificial para simular conversas humanas e fornecer respostas automatizadas a perguntas.

O *ChatGPT* (*Chat Generative Pre-trained Transformer*), em português "Transformador generativo Pré-treinado" é uma conquista notável no campo da Inteligência Artificial. A ferramenta permite a criação de conversas com os usuários, analisando vastos volumes de dados. Desenvolvida pela OpenAI, empresa de pesquisa com sede nos Estados Unidos da América, essa ferramenta representa um marco na capacidade das máquinas em compreender e interagir com a linguagem humana de maneira envolvente e precisa, proporcionando um novo horizonte de possibilidades na era da informação (Silva, 2023).

O *ChatGPT* é construído com base em RNAs. O funcionamento consiste em algoritmos com camadas de neurônios artificiais que processam dados por meio de operações matemáticas complexas. O poder das RNAs permite que programas interpretem e compreendam as perguntas formuladas pelos usuários, aprendendo contextos específicos e adaptando-se para responder de forma precisa (Anjos; Bastos, 2023).

O *ChatGPT* possui uma ampla gama de aplicações, desde assistentes virtuais em sites até *chatbots* em plataformas de mensagens, e é capaz de responder a perguntas comuns, gerar texto para suporte ao cliente, traduzir idiomas, entre outras tarefas. Ele está disponível em várias versões, como o GPT-1, GPT-2 e GPT-3, cada uma treinada em conjuntos de dados cada vez maiores, o que resulta em modelos mais capazes de entender e gerar texto com precisão.

Os principais objetivos do *ChatGPT* incluem aprimorar a interação humano-computador, fornecer suporte automatizado para várias tarefas na compreensão e geração de linguagem natural. Em suma, representa um avanço significativo na capacidade das máquinas de entender e gerar texto de forma semelhante aos humanos, com aplicações em diferentes áreas, desde assistência virtual até pesquisa e educação.

3 METODOLOGIA

Para contemplar os objetivos propostos, este trabalho utiliza uma metodologia quali-quantitativa. Para compreender a utilização do *ChatGPT* pelos discentes de Ciência de Dados, foi necessário a aplicação de um questionário para, posteriormente realizar análises das respostas.

Em Ciência de Dados, a combinação de métodos quantitativos e qualitativos em pesquisas por meio de entrevistas com usuários oferece uma abordagem para compreender as complexidades dos dados e as reais necessidades dos indivíduos. Como afirmado por Creswell (2014) a utilização de múltiplos métodos aumenta a validade e a confiabilidade dos resultados, proporcionando uma compreensão mais profunda do fenômeno em estudo. No contexto das entrevistas com usuários, os métodos quantitativos podem ser empregados para coletar dados estruturados e objetivos, como a frequência de uso de determinados recursos ou a classificação de preferências. Esses dados quantitativos fornecem uma base sólida para análise estatística e identificação de tendências (Gunther, 2006).

Por outro lado, os métodos qualitativos enriquecem a compreensão dos dados ao explorar os aspectos subjetivos e contextuais das experiências dos usuários. Conforme defendido por Patton (2002) a abordagem qualitativa permite uma exploração mais profunda dos significados e das perspectivas dos participantes, revelando que não podem ser capturados apenas por números. Nas entrevistas com usuários, os métodos qualitativos são úteis para capturar sentimentos, motivações e desafios enfrentados pelos indivíduos, fornecendo um importante contexto para interpretar os resultados quantitativos (Lüdke et al., 2013).

Integrando métodos quantitativos e qualitativos em pesquisas por meio de entrevistas com usuários, é possível obter uma visão geral das necessidades e comportamentos dos usuários. A triangulação de métodos fortalece a validade dos resultados, permitindo uma melhor compreensão do fenômeno em estudo (Denzin; Lincoln, 2011). Essa abordagem permite que os pesquisadores identifiquem padrões quantitativos, enquanto também exploram os contextos e significados, resultando em informações para a tomada de decisão.

O tipo de estudo deste trabalho foi o descritivo e exploratório por meio de questionário estruturado, o qual foi aplicado aos alunos matriculados no Curso Superior de Tecnologia em Ciência de Dados da Fatec Adamantina – SP, onde foram obtidas respostas de 36 alunos. O questionário foi distribuído e respondido de forma online na plataforma Google Forms, contemplando os seguintes temas:

- Conhecimento sobre a ferramenta *ChatGPT*.
- Experiência de Uso.
- Atividades de Uso.
- Utilização do *ChatGPT*.
- Percepção sobre o *ChatGPT*.
- Consideração sobre a Inovação do *ChatGPT*.
- Conhecimento sobre o Funcionamento.
- Avaliação do uso pessoal.
- Avaliação do *ChatGPT*.
- Confiança nas Informações do *ChatGPT*.

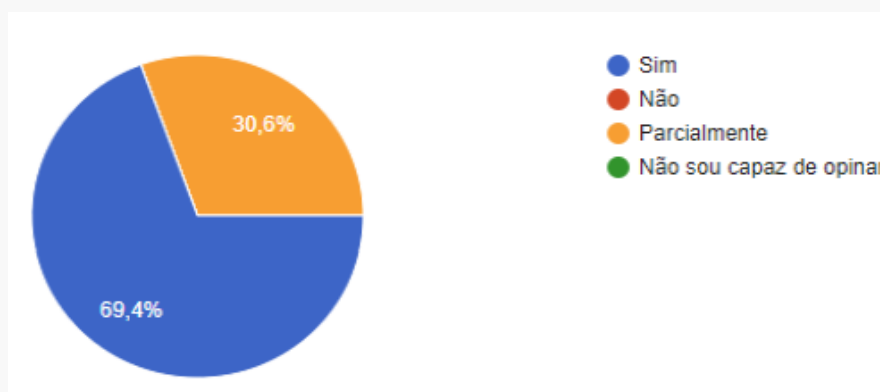
A partir das respostas foram realizadas análises estatísticas, como contagem e médias. A seguir foram realizadas análises qualitativas dos dados.

4 RESULTADOS

A pesquisa revelou que todos os alunos pesquisados conhecem a ferramenta, mas 94,4% realmente a utilizou. As atividades que os alunos mais utilizam são: trabalhos escolares, atividades no trabalho e em atividades de lazer. É interessante destacar que os entrevistados utilizam no contexto escolar para: pesquisar sobre o assunto abordado em aulas e trabalhos, para elaborar textos diversos, como ferramenta de dicas ou ideias de como fazer trabalho. Além disso, a utilização em atividades de codificação também foi lembrada, sendo utilizado para criar ou identificar erros em programas.

A Figura 4 apresenta o resultado da percepção positiva na aprendizagem e nos estudos. Sua análise revela que a imensa maioria acredita ser positivo o seu uso.

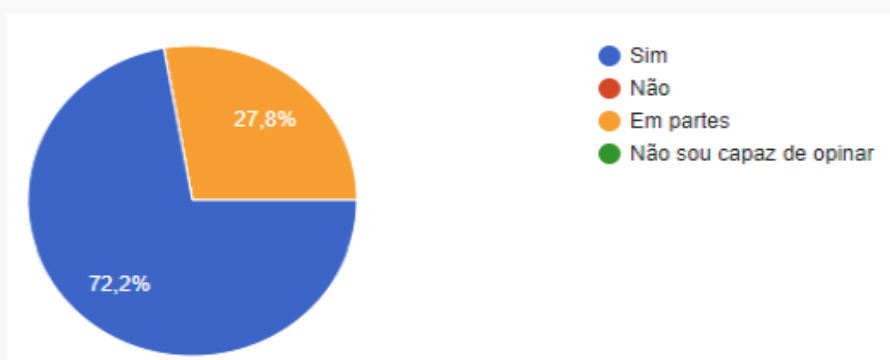
Figura 4: Percepção de uso da ferramenta



Fonte: Autoria própria (2024)

A Figura 5 apresenta a resposta sobre a percepção dos alunos em relação ao *ChatGPT* como uma ferramenta inovadora que facilita a vida do estudante, os resultados foram que 72,2% dos alunos responderam positivamente, concordando que o *ChatGPT* é uma ferramenta inovadora e útil para facilitar sua vida acadêmica. Os 27,8% restantes expressaram uma visão diferente, concordando apenas parcialmente com a afirmação.

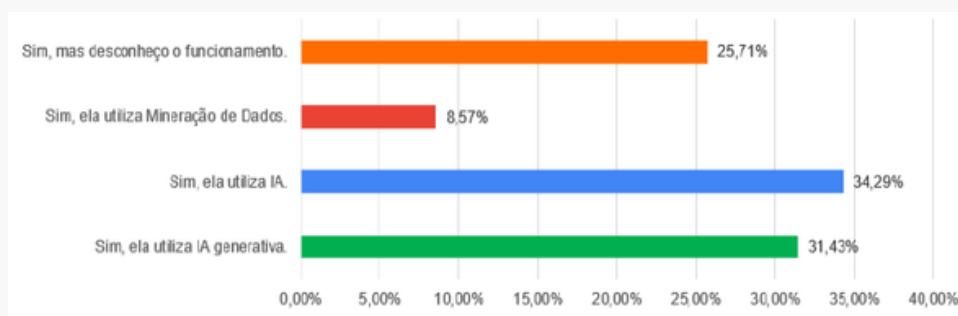
Figura 5: O ChatGPT é uma ferramenta inovadora?



Fonte: Autoria própria (2024)

Analizando a Figura 6 foi possível verificar que todos os alunos conhecem o ChatGPT, mas apenas 31,43% conhecem o funcionamento, assim acertando a resposta correta.

Figura 6. Você conhece o funcionamento da ferramenta?



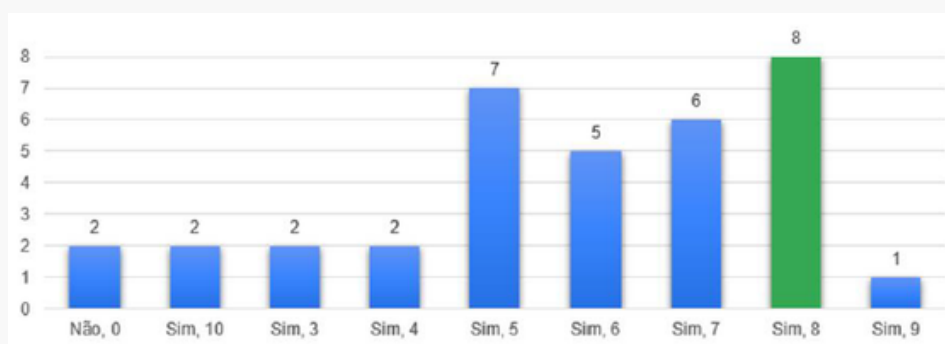
Fonte: Autoria própria (2024)

Conforme demonstrado na Figura 7, a maioria dos alunos respondeu que já utilizaram a ferramenta *ChatGPT*. A Oito alunos atribuíram a nota 8 para o nível de confiança nas informações que foram fornecidas.

Em uma escala de 0 a 10, os alunos foram solicitados a avaliar o quanto utilizam o *ChatGPT*. Os resultados revelaram que 22,2% dos alunos atribuíram a nota máxima, indicando um uso frequente e significativo da plataforma,

72,2% dos alunos deram notas positivas, destacando a utilidade percebida em suas atividades acadêmicas. Entretanto, é interessante observar que apenas 25% dos alunos demonstraram uma confiança elevada nas respostas fornecidas pela ferramenta. A maioria dos alunos, 75%, expressou certa desconfiança em relação às respostas geradas pelo sistema.

Figura 7. Você já utilizou o ChatGPT? Em caso positivo, qual seu nível de confiança nas informações fornecidas?



Fonte: Autoria própria (2024)

5 CONSIDERAÇÕES FINAIS

Os resultados desta pesquisa demonstram que o *ChatGPT* desempenha um papel importante na formação e no trabalho dos alunos de Ciência de Dados. A maioria dos participantes relatou uma familiaridade significativa com a ferramenta, utilizando-a não apenas para suporte em projetos complexos, mas também para explorar novas técnicas e aprimorar suas habilidades. Isso sugere que, como uma ferramenta educacional e prática, não só fortalece o aprendizado individual, mas também amplia o alcance do conhecimento, preparando os alunos para os desafios do mundo real.

Para otimizar ainda mais o uso do *ChatGPT* na vida dos alunos, recomenda-se integrar a ferramenta em currículos acadêmicos e programas de treinamento profissional. Além disso, promover workshops regulares e tutoriais específicos sobre como utilizá-lo para resolver problemas complexos, podendo ajudar os alunos a explorarem melhor as capacidades da ferramenta. Implementar casos de uso práticos, onde os alunos possam aplicar o *ChatGPT* em projetos reais.

Além das iniciativas mencionadas, é fundamental que instituições acadêmicas e profissionais incentivem a colaboração entre alunos e profissionais experientes na área, permitindo que os estudantes não só

utilizem o *ChatGPT* como uma ferramenta individual, mas também como uma plataforma para troca de conhecimentos e experiências. Isso não apenas enriquece o aprendizado coletivo, mas também prepara os alunos para enfrentar os desafios dinâmicos e interdisciplinares que caracterizam o campo da Ciência de Dados.

REFERÊNCIAS

ANJOS, R. M. dos; BASTOS, R. C. **Inteligência artificial e a engenharia: dos oráculos ao ChatGPT: Artificial Intelligence and Engineering: From Oracles to ChatGPT.** Revista UNICREA - Revista Técnico Científica da Universidade Corporativa do CREA/SC, [S. l.], v. 1, n. 2, p. 94–116, 2023.

BONINI NETO, A.; BONINI, C. dos S. B. Redes neurais artificiais: introdução e uso do algoritmo Perceptron em biossistemas. **Revista Brasileira de Engenharia de Biossistemas**, Tupã, v. 2, pág. 87–95, 2010. DOI: 10.18011/bioeng2010v4n2p87-95.

DANTAS, Eric Rommel G. et al. **O uso da descoberta de conhecimento em base de dados para apoiar a tomada de decisões.** V Simpósio de Excelencia em Gestão e Tecnologia, 2008.

DENZIN, Norman K; LINCOLN, Yvonna S. **The Sage handbook of qualitative research.** Academia.edu, 2011. Disponível em: <https://library.wur.nl/WebQuery/titel/2070514>. Acesso em: 06 jun. 2024.

FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. **From data mining to knowledge discovery in databases.** AI magazine, v. 17, n. 3, p. 37-37, 1996.

FÉLIX, Luis Carlos Molina. **Data Mining no processo de extração de conhecimento de base de dados.** 1998. Tese de Doutorado (Ciências de Computação e Matemática Computacional). Instituto de Ciências Matemáticas em Ciências /USP. São Carlos, 1998.

FERREIRA, R. C. V.; GARCIA, G. H. M.; BRASIL, D. R. **O surgimento do Chat GPT e a insegurança sobre o futuro dos trabalhos acadêmicos.** Cadernos de Direito Actual, [S. l.], n. 21, p. 130–143, 2023. Disponível em: <https://www.cadernosdedereitoactual.es/ojs/index.php/cadernos/article/view/917>. Acesso em: 7 jun. 2024.

RODRIGUES, Mateus Franco. **Geração de dados sintéticos utilizando redes neurais artificiais.** 34 f. Trabalho de Conclusão de Curso (Graduação em Engenharia de Software) - Universidade Federal do Ceará, Campus de Russas, Russas, 2021.

GÜNTHER, H. **Pesquisa qualitativa versus pesquisa quantitativa:** esta é a questão? Psicologia: Teoria e Pesquisa, v. 22, n. 2, p. 201–209, maio 2006.

LOPEZ, Martin Andreoni; MATTOS, Diogo M. F. **Resumo de grandes volumes de dados com filtro de bloom:** uma abordagem eficiente para aprendizado profundo com redes neurais convolucionais em fluxos de rede. In: Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC), 39. , 2021, Uberlândia. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2021 . p. 532-545. ISSN 2177-9384. DOI: <https://doi.org/10.5753/sbrc.2021.16745>.

LUDKE, Menga; ANDRÉ, Marli. **Pesquisa em educação:** abordagens qualitativas. Em Aberto, v. 5, n. 31, 1986.

MCCULLOCH, WS, PITTS, W. **Um cálculo lógico das ideias imanentes na atividade nervosa.** Boletim de Biofísica Matemática 5, 115–133, 1943. <https://doi.org/10.1007/BF02478259>.

MONTEIRO, Jean Carlos da Silva. Assistente ChatGPT na educação: possibilidades e desafios. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, [S. l.], v. 9, n. 6, p. 2899–2906, 2023. DOI: 10.51891/rease.v9i6.10482. Disponível em: <https://periodicorease.pro.br/rease/article/view/10482>. Acesso em: 7 jun. 2024.

PATTON, Michael Quinn. **Qualitative research & evaluation methods.** 3. Ed. London: Sage Publications, 2002. 56 p.

PAULA, Heller. **GenAI:** o que é e como funciona a inteligência artificial generativa. 2023. Disponível em: < <https://faberhaus.com.br/genai-ia-generativa/> >. Acesso em: 20 jun. 2024.

RAUBER, Thomas Walter. **Redes neurais artificiais**. Universidade Federal do Espírito Santo, v. 29, 2005. Disponível em:
https://www.researchgate.net/profile/Thomas-Rauber-2/publication/228686464_Redес_neurais_artificiais/links/02e7e521381602f2bd000000/Redes-neurais-artificiais.pdf > Acessado em: 06 jun. 2024.

RAUTENBERG, Sandro; DO CARMO, Paulo Ricardo Viviurka. Big data e ciência de dados: complementariedade conceitual no processo de tomada de decisão. **Brazilian Journal of Information Science**, v. 13, n. 1, p. 56-67, 2019.

SANCHES, Marcelo Kaminski. **Aprendizado de máquina semi-supervisionado**: proposta de um algoritmo para rotular exemplos a partir de poucos exemplos rotulados. 2003. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação, University of São Paulo, São Carlos, 2003.
doi:10.11606/D.55.2003.tde-12102003-140536. Acesso em: 2024-06-06.

SANDES, Jorge Reis de Vasconcellos. O perfil e o papel do cientista de dados = The profile and the role of the data scientist. **Revista do BNDES**, Rio de Janeiro, v. 26, n. 52, p. 275-319, dez. 2019.

SILVA FILHO, Abel Guilhermino da. **Classificação de padrões de imagens**: função objetivo para perceptron multicamada e máquina de aprendizado extremo convolucional. 2019. Tese (Doutorado em Ciência da Computação) – Universidade Federal de Pernambuco, Recife, 2019.

SILVA, Vinicius Lopes da. **Ética e responsabilidade na era da inteligência artificial**: aprendizagem digital no chat GPT. 2023. 27 p. Monografia (especialização) - Pós-graduação Lato Sensu em Mídia e Educação - Universidade Federal do Pampa/UAB - Universidade Aberta do Brasil, Campus São Borja, São Borja, 2023.

SILVA ALBUQUERQUE, C.L. **Análise de sentimento sobre comentários em sites de e-commerce no idioma português/BR**. Trabalho de Conclusão de Curso (Sistemas de Informação) - Centro Universitário Christus, Fortaleza, 2022.

TAFNER, Malcon. La Plasticidad Neuronal Artificial. RAN – **Revista Argentina de Neurociencias**. Argentina – Buenos Aires, v.2, n.4, p50-53, ago. 1998.

THIAGO SOARES LAITZ; LEVY BOCCATO. **Estudo e aplicação de redes neurais generativas**. In: XXIX CONGRESSO DE INICIAÇÃO CIENTÍFICA DA UNICAMP, 2021, Campinas. Anais eletrônicos... Campinas, Galoá, 2021. Disponível em: <<https://proceedings.science/unicamp-pibic/pibic-2021/trabalhos/estudo-e-aplicacao-de-redes-neurais-generativas?lang=pt-br>> Acesso em: 06 jun. 2024.