



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
UNIDADE ACADÊMICA DE BELO JARDIM
CURSO DE BACHARELADO EM ENGENHARIA DA COMPUTAÇÃO

ELIAS VITOR DA SILVA

**AVALIAÇÃO DO IMPACTO DO CHATGPT NA APRENDIZAGEM
DE HEAP E TABELAS HASH EM ALGORITMOS E ESTRUTURAS
DE DADOS**

Belo Jardim
2026

ELIAS VITOR DA SILVA

**AVALIAÇÃO DO IMPACTO DO CHATGPT NA APRENDIZAGEM
DE HEAP E TABELAS HASH EM ALGORITMOS E ESTRUTURAS
DE DADOS**

Trabalho de Conclusão de Curso apresentado
ao Curso de Engenharia da Computação da
Universidade Federal Rural de Pernambuco, como
requisito parcial para obtenção do título de
Bacharelado em Engenharia da Computação.

Orientador: Prof. Dr. Waldemar Neto,
Coorientador: Prof. Dr. Waldemar Neto

Belo Jardim
2026

Dados Internacionais de Catalogação na Publicação
Sistema Integrado de Bibliotecas da UFRPE
Bibliotecário(a): Nubia Matos – CRB-4 2234

S586a Silva, Elias Vitor da.
Avaliação do impacto do ChatGPT na aprendizagem de
Heap e Tabelas Hash em algoritmos e estruturas de dados
/ Elias Vitor da Silva. – Belo Jardim, 2026.
35 f.; il.

Orientador(a): Waldemar Pires Ferreira Neto.

Trabalho de Conclusão de Curso (Graduação) –
Universidade Federal Rural de Pernambuco, Unidade
Acadêmica de Belo Jardim - UABJ, Bacharelado em
Engenharia da Computação, Belo Jardim, BR-PE, 2026.

Inclui referências.

1. ChatGPT (Programa de computador). 2. Inteligência
artificial generativa. 3. Algoritmos e estruturas de dados.
4. Heap (estrutura de dados) 5. Tabelas Hash (estrutura de
dados). I. Ferreira Neto, Waldemar Pires, orient. II. Título

CDD 621.3

Elias Vitor da Silva

AVALIAÇÃO DO IMPACTO DO CHATGPT NA APRENDIZAGEM DE HEAP E TABELAS HASH EM ALGORITMOS E ESTRUTURAS DE DADOS

Trabalho de Conclusão de Curso apresentado
ao Curso de Engenharia da Computação da
Universidade Federal Rural de Pernambuco, como
requisito parcial para obtenção do título de
Bacharelado em Engenharia da Computação.

Aprovado em: 16 de Julho de 2026

BANCA EXAMINADORA

Prof. Dr. Waldemar Neto
Universidade Federal Rural de Pernambuco

Prof. Dr. Denini Gabriel (Examinador Interno)
Universidade Federal Rural de Pernambuco

Prof. Dr. Anderson Pinheiro (Examinador
Interno)
Universidade Federal Rural de Pernambuco

Dedico este trabalho à minha família, pelo apoio constante, pela confiança e pelo incentivo durante toda a minha trajetória acadêmica.

AGRADECIMENTOS

Agradeço, primeiramente, à minha família, pelo apoio, incentivo e compreensão ao longo de toda a minha trajetória acadêmica. Sua presença foi essencial para que eu pudesse superar os desafios encontrados durante a graduação e manter o compromisso com a conclusão deste trabalho.

Agradeço também ao meu orientador, aos professores e colegas que contribuíram direta ou indiretamente para minha formação, compartilhando conhecimentos, experiências e orientações importantes para o desenvolvimento desta pesquisa. Estendo meus agradecimentos à Universidade Federal Rural de Pernambuco, especialmente à Unidade Acadêmica de Belo Jardim, pela oportunidade de formação e pelo ambiente acadêmico que possibilitou a realização deste Trabalho de Conclusão de Curso.

Aprender não consiste apenas em obter respostas, mas em desenvolver o pensamento crítico, compreender problemas, avaliar caminhos e construir soluções.

RESUMO

Uma competência fundamental na formação em Computação é a capacidade de compreender, implementar e selecionar estruturas de dados adequadas para diferentes problemas. Nesse contexto, o avanço de ferramentas de inteligência artificial generativa, como o ChatGPT, tem suscitado discussões sobre seus possíveis efeitos no ensino de programação. Este trabalho investiga o impacto do uso do ChatGPT na compreensão dos estudantes da disciplina de Algoritmos e Estruturas de Dados, com foco nos conteúdos de Heap e de Tabelas Hash. O estudo adota um desenho experimental com grupos cruzados, no qual cada grupo utiliza a ferramenta em um tópico e não a utiliza no outro. A compreensão foi avaliada por meio de questionários objetivos aplicados ao final de cada atividade prática, sem apoio do ChatGPT. Participaram 25 estudantes, distribuídos em grupos, dos quais 20 concluíram todas as etapas do experimento. A análise estatística empregou os testes de Shapiro–Wilk, Mann–Whitney U e Cliff's Delta. Não foi identificada diferença estatisticamente significativa na compreensão entre as condições com e sem uso do ChatGPT ($p = 1,000$), e o tamanho de efeito foi classificado como negligível ($\delta = -0,003$). Os resultados convergem com estudos que também não observaram impacto significativo da ferramenta no desempenho de estudantes em disciplinas de programação e estruturas de dados, embora a alta familiaridade prévia dos participantes com IA generativa deva ser considerada na interpretação dos achados.

Palavras-chave: ChatGPT (programa de computador), inteligência artificial generativa, algoritmos e estruturas de dados, Heap (estrutura de dados), tabelas hash (estrutura de dados), compreensão.

ABSTRACT

An essential competency in Computing education is the ability to understand, implement, and select suitable data structures for different problems. In this context, the advance of generative artificial intelligence tools, such as ChatGPT, has raised discussions about their possible effects on programming education. This work investigates the impact of ChatGPT usage on students' performance in the Data Structures and Algorithms course, focusing on Heap and Hash Table topics. The study adopts a crossover experimental design in which each group uses the tool in one topic and does not use it in the other. Performance was measured by scores on objective questionnaires applied at the end of each practical activity, without ChatGPT support. Twenty-five students were assigned to the groups, of whom 20 completed all experimental stages. Statistical analysis employed Shapiro–Wilk, Mann–Whitney U, and Cliff's Delta. No statistically significant difference in performance was identified between conditions with and without ChatGPT usage ($p = 1.000$), and the effect size was classified as negligible ($\delta = -0.003$). These results converge with studies that also did not observe a significant impact of the tool on student performance in programming and data structures courses, although the participants' high prior familiarity with generative AI should be considered when interpreting the findings.

Keywords: ChatGPT (software), generative artificial intelligence, data structures and algorithms, Heap (data structure), hash tables (data structure), performance.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo ilustrativo de tabela hash: chaves, função hash e baldes.	17
Figura 2 – Exemplo ilustrativo de Max-Heap: representação em árvore e em vetor. . .	19
Figura 3 – Visão geral do processo experimental.	25
Figura 4 – Distribuição das pontuações e estatística U do teste de Mann–Whitney. . .	30
Figura 5 – Comparações par a par e magnitude do Cliff's Delta.	31

LISTA DE TABELAS

Tabela 1 – Perfil dos participantes quanto ao uso de IA generativa ($n = 28$, formulário inicial)	26
Tabela 2 – Resumo da participação no experimento	26
Tabela 3 – Estatísticas descritivas do desempenho nos questionários	29
Tabela 4 – Desempenho médio por conteúdo e condição de uso do ChatGPT	29
Tabela 5 – Resultado do teste de Mann–Whitney U na comparação agregada	30
Tabela 6 – Coeficiente Cliff’s Delta e magnitude do efeito na comparação agregada . .	30

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
DSA	Data Structures and Algorithms
IA	Inteligência Artificial
TCC	Trabalho de Conclusão de Curso
UABJ	Unidade Acadêmica de Belo Jardim
UFRPE	Universidade Federal Rural de Pernambuco

LISTA DE SÍMBOLOS

H_0	Hipótese nula
H_1	Hipótese alternativa
n	Tamanho da amostra
$O(\cdot)$	Notação assintótica de complexidade
p	Valor de significância estatística
δ	Tamanho de efeito de Cliff

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Objetivos	14
1.1.1	Objetivo geral	14
1.1.2	Pergunta de pesquisa	14
1.1.3	Objetivos específicos	15
2	REFERENCIAL TEÓRICO	16
2.1	Estrutura de Dados	16
2.1.1	Hash	16
2.1.2	Heap	17
2.2	Inteligência Artificial Generativa e ChatGPT na Educação	19
2.3	ChatGPT no Ensino de Programação e em Estruturas de Dados	20
2.4	Instrumentos Estatísticos do Estudo	20
2.4.1	Teste de Shapiro–Wilk	21
2.4.2	Teste de Mann–Whitney U	21
2.4.3	Cliff’s Delta	21
2.5	Síntese do Referencial	22
3	METODOLOGIA	23
3.1	Caracterização da Pesquisa	23
3.2	Contexto e Participantes	23
3.3	Conteúdos Investigados	24
3.4	Hipótese da Pesquisa	24
3.5	Desenho Experimental	24
3.6	Instrumentos e Materiais	27
3.7	Procedimentos de Coleta de Dados	27
3.8	Procedimentos de Análise dos Dados	28
4	RESULTADOS E DISCUSSÕES	29
4.1	Organização dos Dados	29
4.2	Estatísticas Descritivas	29
4.3	Análise Estatística da Hipótese H_0	30
4.4	Discussão dos Resultados	31
4.5	Ameaças à Validade	32
5	CONCLUSÕES	33
	REFERÊNCIAS	34

1 INTRODUÇÃO

Algoritmos e Estruturas de Dados constituem uma área fundamental na formação em Computação, pois fornecem a base para organizar, processar e recuperar informações de maneira eficiente (LEBLANC et al., 2015; CORMEN et al., 2022). Entre os conteúdos dessa disciplina, Heap e Tabelas Hash são estruturas amplamente utilizadas e exigem articulação entre teoria, análise de complexidade e implementação prática (CORMEN et al., 2022).

Nos últimos anos, ferramentas de inteligência artificial generativa, como o ChatGPT, passaram a integrar o cotidiano dos estudantes de programação, auxiliando na explicação de conceitos, na geração de exemplos e na sugestão de soluções (BAIDOO-ANU; ANSAH, 2023; RAHMAN; WATANOBÉ, 2023; GARCIA, 2025). Ao mesmo tempo, seu uso suscita questionamentos sobre autonomia, pensamento crítico e possível dependência de respostas prontas (DWIVEDI et al., 2023; DELIKOURA; FUNG; HUI, 2025; QIN et al., 2023), especialmente em disciplinas que exigem raciocínio lógico e a implementação de algoritmos.

Diante desse cenário, o problema investigado neste trabalho foi formulado por meio da seguinte questão de pesquisa: *o uso do ChatGPT durante atividades práticas de Algoritmos e Estruturas de Dados altera o desempenho dos estudantes, medido pela pontuação nos questionários sobre Heap e Tabelas Hash, quando comparadas as condições com e sem a ferramenta?* Essa questão insere-se em investigações recentes sobre o impacto do ChatGPT em programação e em disciplinas de estruturas de dados (JAMIE; HAJIHASHEMI; ALIPOUR, 2024; KOSAR et al., 2024).

Este Trabalho de Conclusão de Curso foi motivado por um artigo recentemente aprovado no Congresso Brasileiro de Engenharia de Software (CBSOFT) (NETO, 2026), que investiga o impacto do ChatGPT na aprendizagem de Algoritmos e Estruturas de Dados. A presente pesquisa dá continuidade a essa linha de investigação, com foco nos conteúdos de Heap e Tabelas Hash.

1.1 Objetivos

1.1.1 Objetivo geral

O objetivo geral deste trabalho é analisar se o uso do ChatGPT durante atividades práticas influencia o desempenho dos estudantes, medido pela pontuação nos questionários sobre Heap e Tabelas Hash.

1.1.2 Pergunta de pesquisa

A partir do objetivo geral, formula-se a seguinte pergunta de pesquisa, alinhada ao problema apresentado na seção anterior: *o uso do ChatGPT durante atividades práticas de*

Algoritmos e Estruturas de Dados altera o desempenho dos estudantes, medido pela pontuação nos questionários sobre Heap e Tabelas Hash, quando comparadas as condições com e sem a ferramenta?

1.1.3 Objetivos específicos

Os objetivos específicos são:

- caracterizar a participação e o perfil dos estudantes quanto ao uso prévio de IA generativa;
- comparar o desempenho nos questionários entre as condições com e sem uso do ChatGPT;
- discutir os resultados à luz da literatura e das limitações do desenho experimental adotado.

Para atingir esses objetivos, adota-se um desenho experimental com grupos cruzados, no qual cada grupo utiliza a ferramenta em um tópico e não a utiliza no outro. O desempenho é operacionalizado como a pontuação obtida em questionários objetivos aplicados ao final de cada atividade prática, sem apoio do ChatGPT.

2 REFERENCIAL TEÓRICO

Este capítulo apresenta os conceitos e estudos que fundamentam a pesquisa, com ênfase em Heap, Tabelas Hash, no uso do ChatGPT em disciplinas de programação e nos testes estatísticos empregados na análise dos dados.

2.1 Estrutura de Dados

As estruturas de dados constituem um dos fundamentos da Engenharia da Computação, sendo responsáveis pela organização, armazenamento e manipulação eficientes das informações utilizadas pelos algoritmos. A escolha da estrutura de dados adequada influencia diretamente o desempenho de uma aplicação, afetando aspectos como o tempo de execução, o consumo de memória e a escalabilidade. Dentre as diversas estruturas de dados, apresentaremos as duas que são cobertas nesta pesquisa: Heap e Tabelas Hash.

2.1.1 Hash

As estruturas de dados baseadas em *Hash* constituem uma das principais alternativas para o armazenamento e recuperação eficientes de informações em aplicações computacionais. Seu funcionamento baseia-se no uso de uma **função hash**, responsável por transformar uma chave de pesquisa em um índice correspondente a uma posição em uma tabela, denominada **tabela hash**. Esse mecanismo permite que operações como inserção, busca e remoção sejam realizadas, em média, em tempo constante, tornando as tabelas hash uma das estruturas mais eficientes para manipulação de grandes volumes de dados (CORMEN et al., 2022).

Uma função hash recebe uma chave de entrada e calcula um valor inteiro que determina a posição em que o elemento será armazenado. Idealmente, essa função deve distribuir uniformemente as chaves ao longo da tabela, minimizando a ocorrência de colisões. Quanto mais homogênea for essa distribuição, maior será o desempenho das operações realizadas sobre a estrutura.

Entretanto, diferentes chaves podem produzir o mesmo índice da tabela, fenômeno conhecido como **colisão**. Como as colisões são inevitáveis quando o número de possíveis chaves é maior que o de posições disponíveis na tabela, diversos métodos foram desenvolvidos para tratá-las. Entre as abordagens mais utilizadas estão o **encadeamento separado** (*separate chaining*), no qual cada posição da tabela mantém uma lista contendo todos os elementos associados ao mesmo índice, e o **endereçoamento aberto** (*open addressing*), em que novos elementos são armazenados em outras posições livres da própria tabela por meio de estratégias de sondagem, como sondagem linear (CORMEN et al., 2022).

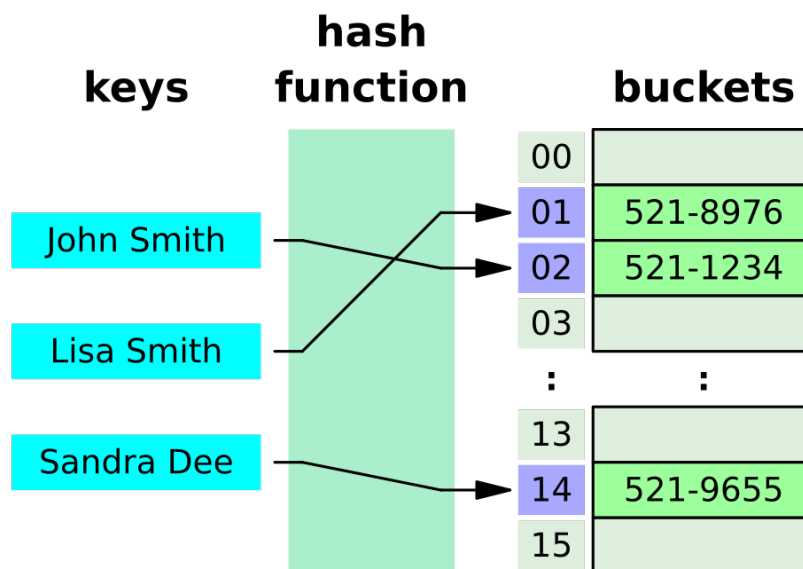
Na linguagem Java, as estruturas baseadas em hash são disponibilizadas principalmente por meio das classes *HashMap* e *HashSet*, integrantes da **Java Collections Framework**. A

classe *HashMap* permite armazenar pares chave-valor, oferecendo acesso rápido aos elementos por meio da chave de busca. Já a classe *HashSet* utiliza internamente uma tabela hash para armazenar apenas elementos únicos, impedindo a inserção de valores duplicados. Em ambas as estruturas, as operações de inserção, remoção e consulta apresentam complexidade média de **O(1)**, embora, no pior caso, possam atingir **O(n)** quando ocorrem muitas colisões.

As tabelas hash são amplamente empregadas em aplicações que exigem acesso rápido às informações, como mecanismos de autenticação, sistemas de cache, bancos de dados, compiladores, indexação de arquivos e tabelas de símbolos. Sua eficiência faz com que sejam frequentemente utilizadas quando a ordem dos elementos não constitui um requisito importante, mas a velocidade de acesso é um fator crítico para o desempenho da aplicação.

Segundo Cormen et al. (CORMEN et al., 2022), as tabelas hash representam uma das soluções mais eficientes para problemas de pesquisa dinâmica, desde que a função hash seja adequadamente projetada e o fator de carga seja mantido sob controle. Dessa forma, compreender os princípios de funcionamento das funções hash, os mecanismos de resolução de colisões e os fatores que influenciam seu desempenho é essencial para a implementação de sistemas computacionais eficientes e escaláveis.

Figura 1 – Exemplo ilustrativo de tabela hash: chaves, função hash e baldes.



Fonte: Jorge Stolfi / Wikimedia Commons (domínio público).

2.1.2 Heap

A Heap é uma estrutura de dados hierárquica amplamente empregada em algoritmos que necessitam de acesso rápido ao maior ou ao menor elemento de um conjunto de dados. Diferentemente das árvores binárias de busca, nas quais os elementos são organizados de acordo com relações de ordem entre subárvores, a *Heap* obedece a uma propriedade específica

denominada propriedade de heap, que estabelece uma relação entre cada nó e seus descendentes imediatos. Essa característica permite identificar rapidamente o elemento de maior ou menor prioridade, dependendo do tipo de heap utilizada (CORMEN et al., 2022).

Existem dois tipos principais de *Heap*: a *Max-Heap* e a *Min-Heap*. Em uma *Max-Heap*, o valor armazenado em cada nó é sempre maior ou igual aos valores de seus filhos, garantindo que o maior elemento esteja na raiz da árvore. Já em uma *Min-Heap* ocorre o comportamento inverso: cada nó possui valor menor ou igual ao de seus descendentes diretos, fazendo com que o menor elemento permaneça na raiz. É importante destacar que essa propriedade não implica que todos os elementos estejam completamente ordenados, mas apenas que a relação de prioridade entre pais e filhos seja preservada. Nosso curso adotou o *Max-Heap*.

Embora seja representada conceitualmente como uma árvore binária, a Heap normalmente é implementada com um vetor. Nessa representação, a posição dos elementos no vetor corresponde à estrutura hierárquica da árvore, permitindo localizar facilmente o pai e os filhos de qualquer nó por meio de operações aritméticas simples sobre os respectivos índices. Essa implementação elimina a necessidade de ponteiros entre os nós, reduzindo o consumo de memória e tornando as operações mais eficientes.

As principais operações realizadas em uma Heap incluem a inserção de novos elementos, a remoção da raiz e a construção da estrutura a partir de um conjunto de dados. Após cada inserção ou remoção, pode ser necessário reorganizar a árvore para restaurar a propriedade de heap. Esse processo é realizado por algoritmos conhecidos como heapify-up (ou sift-up) e heapify-down (ou sift-down), responsáveis por movimentar os elementos até que a estrutura volte a satisfazer suas propriedades.

As operações de inserção e remoção têm complexidade temporal de $O(\log n)$, uma vez que, no pior caso, o elemento precisa percorrer a altura da árvore para alcançar sua posição correta. Por outro lado, a consulta ao elemento de maior prioridade, localizado na raiz, tem complexidade $O(1)$. Além disso, a construção de uma Heap a partir de um conjunto de elementos pode ser realizada em tempo $O(n)$, tornando essa estrutura bastante eficiente para diversas aplicações (CORMEN et al., 2022).

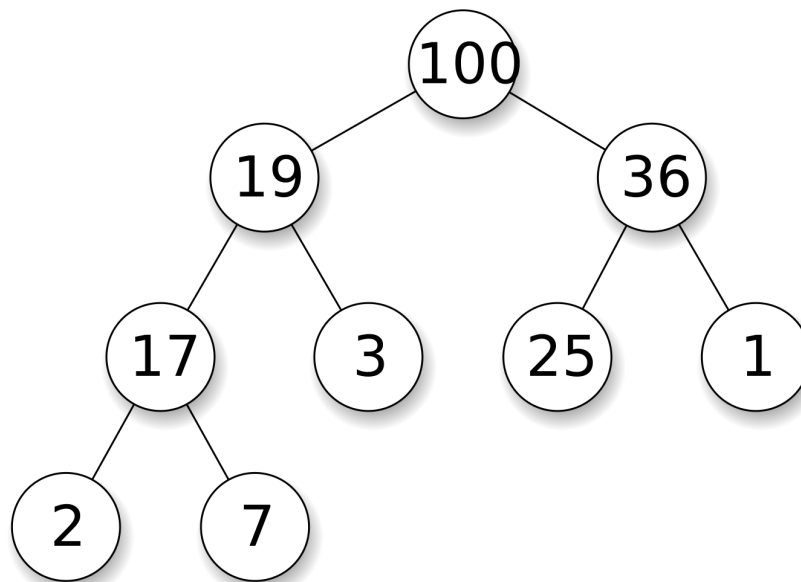
Apesar de seu elevado desempenho para operações de prioridade, a Heap não constitui uma estrutura adequada para buscas arbitrárias, uma vez que apenas o elemento localizado na raiz possui posição garantida. A localização de qualquer outro elemento pode exigir a inspeção de grande parte da árvore, o que resulta em complexidade linear. Dessa forma, sua principal finalidade não é armazenar dados ordenados, mas oferecer acesso eficiente ao elemento de maior ou menor prioridade.

Segundo Cormen et al. (CORMEN et al., 2022), as Heaps desempenham um papel fundamental no desenvolvimento de algoritmos eficientes, especialmente em problemas relacionados à ordenação, escalonamento e gerenciamento de prioridades. Assim, compreender seu funcionamento, suas propriedades e suas aplicações representa um conhecimento essencial para estudantes e profissionais da área de Computação, contribuindo para a escolha da estrutura de

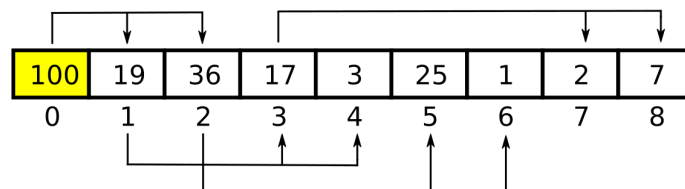
dados mais adequada em diferentes contextos de desenvolvimento de software.

Figura 2 – Exemplo ilustrativo de Max-Heap: representação em árvore e em vetor.

Tree representation



Array representation



Fonte: Wikimedia Commons (domínio público).

2.2 Inteligência Artificial Generativa e ChatGPT na Educação

A inteligência artificial generativa ganhou destaque na educação superior com sistemas capazes de gerar texto, código e explicações a partir de instruções em linguagem natural. Baidoo-Anu e Ansah (BAIDOO-ANU; ANSAH, 2023) discutem oportunidades e desafios do ChatGPT no ensino, destacando potencial de apoio à aprendizagem e, ao mesmo tempo, riscos à integridade acadêmica e à avaliação tradicional. Em direção semelhante, Rahman e Watanobe (RAHMAN; WATANOBÉ, 2023) analisam o ChatGPT na educação em programação, apontando benefícios no esclarecimento de dúvidas e na geração de exemplos, mas também a necessidade de uso crítico.

No apoio à escrita e à compreensão de código, MacNeil et al. (MACNEIL et al., 2022) investigam explicações geradas por modelos de linguagem para trechos de programa, evidenciando o papel da ferramenta como recurso complementar — e não substituto — do processo de estudo. Outros trabalhos relatam efeitos sobre motivação e engajamento discente no uso de ChatGPT (MUÑOZ et al., 2023).

Em contrapartida, a literatura alerta para respostas incorretas ou genéricas, aprendizagem superficial e possível redução do engajamento cognitivo quando o uso é recorrente e pouco orientado (DWIVEDI et al., 2023; DELIKOURA; FUNG; HUI, 2025; QIN et al., 2023; KOSMYNA et al., 2025). Esses pontos reforçam a importância de investigar empiricamente o impacto do ChatGPT em disciplinas que exigem raciocínio e produção autônoma, como Algoritmos e Estruturas de Dados.

2.3 ChatGPT no Ensino de Programação e em Estruturas de Dados

Garcia (GARCIA, 2025) revisa usos do ChatGPT no ensino de programação — tutoria, geração de materiais, produção de código e feedback — e identifica desafios recorrentes, como dependência excessiva da ferramenta e limitações técnicas das respostas. A revisão sugere que o efeito pedagógico depende fortemente do desenho da atividade e do modo de integração da IA à disciplina.

Em estudo empírico com programação orientada a objetos, Kosar et al. (KOSAR et al., 2024) não observaram impacto estatisticamente significativo no desempenho dos estudantes quando o ChatGPT estava disponível, embora os participantes tenham relatado utilidade no apoio à leitura e à compreensão do código. Já Jamie et al. (JAMIE; HAJIHASHEMI; ALIPOUR, 2024) investigaram o ChatGPT como complemento a monitores em Algoritmos e Estruturas de Dados, com *prompts* estruturados e mediação humana, relatando ganhos quando o uso foi supervisionado.

Esses achados indicam que evidências obtidas em um tópico ou condição de uso não se generalizam automaticamente a outros contextos. Em particular, permanece aberta a questão do impacto do ChatGPT em conteúdos específicos como Heap e Tabelas Hash, avaliados por pontuação em questionários objetivos após atividades práticas — foco deste TCC.

2.4 Instrumentos Estatísticos do Estudo

A análise dos dados deste trabalho combina verificação de premissas, comparação entre condições e estimativa da magnitude do efeito. Os três instrumentos descritos a seguir são os efetivamente empregados na Metodologia e nos Resultados.

2.4.1 Teste de Shapiro–Wilk

O teste de Shapiro–Wilk avalia se uma amostra pode ser considerada proveniente de uma distribuição normal (SHESKIN, 2003). A estatística do teste é dada por:

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (1)$$

em que $x_{(i)}$ são as observações ordenadas, $\bar{x}$ é a média amostral e a_i são coeficientes tabulados em função do tamanho amostral n . Valores de W próximos de 1 indicam compatibilidade com a normalidade. No presente estudo, o teste é aplicado às pontuações de cada condição experimental (com e sem ChatGPT). A decisão orienta a escolha de procedimentos posteriores: violações relevantes da normalidade, em conjunto com amostra reduzida, favorecem o uso de métodos não paramétricos.

2.4.2 Teste de Mann–Whitney U

O teste de Mann–Whitney U compara duas amostras quanto à tendência central, com base em postos, sem exigir normalidade (SHESKIN, 2003; WOHLIN et al., 2012). Uma forma usual de calcular a estatística U é:

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1, \quad (2)$$

em que n_1 e n_2 são os tamanhos das duas amostras e R_1 é a soma dos postos da primeira amostra. Neste TCC, após a reorganização das pontuações por condição (com versus sem ChatGPT), o Mann–Whitney U constitui o procedimento principal para avaliar a hipótese nula de ausência de diferença mensurável entre as condições.

2.4.3 Cliff's Delta

Além da significância estatística, estima-se o tamanho do efeito por meio do Cliff's Delta (δ), medida não paramétrica que expressa a probabilidade de uma observação de um grupo superar a de outro (KITCHENHAM; MADEYSKI; BRERETON, 2019):

$$\delta = \frac{\#(x_i > y_j) - \#(x_i < y_j)}{n_1 \cdot n_2}, \quad (3)$$

em que x_i e y_j são as observações das duas condições e $\#$ denota a contagem dos pares satisfazendo a desigualdade correspondente. O coeficiente complementa o valor- p : mesmo quando não se rejeita H_0 , δ informa se a diferença observada é, na prática, negligível, pequena, média ou grande, segundo limiares consolidados na literatura de experimentos em Engenharia de Software (KITCHENHAM; MADEYSKI; BRERETON, 2019).

2.5 Síntese do Referencial

O referencial evidencia que Heap e Tabelas Hash exigem articulação entre conceito, complexidade e implementação, e que o ChatGPT pode apoiar o estudo de programação, embora seu impacto dependa do contexto e do modo de uso. Ainda há pouca evidência empírica sobre o efeito da ferramenta nesses tópicos específicos quando o desempenho é medido por pontuação em questionários objetivos. Este TCC investiga esse impacto por meio de um experimento com grupos cruzados, comparando as condições com e sem ChatGPT com o suporte dos testes de Shapiro–Wilk, Mann–Whitney U e Cliff’s Delta.

3 METODOLOGIA

Este capítulo descreve os procedimentos metodológicos adotados para investigar o impacto do uso do ChatGPT no desempenho dos estudantes em Algoritmos e Estruturas de Dados, com foco em Heap e Tabelas Hash. O estudo adota um desenho experimental com grupos cruzados, no qual as condições de uso da ferramenta são invertidas entre os dois tópicos investigados.

3.1 Caracterização da Pesquisa

A pesquisa caracteriza-se como um estudo empírico em contexto educacional, com abordagem quantitativa e desenho comparativo. O objetivo é observar se o uso do ChatGPT em atividades práticas de programação influencia o desempenho dos estudantes nos conteúdos trabalhados. Para isso, adota-se um desenho experimental com dois grupos e condições invertidas de uso da ferramenta entre os tópicos Heap e Tabelas Hash (WOHLIN et al., 2012).

Neste estudo, o **desempenho** é operacionalizado como a pontuação obtida por cada estudante em questionários objetivos de múltipla escolha, aplicados ao final de cada atividade prática, sem apoio do ChatGPT. O questionário de Heap continha 20 questões e o de Tabelas Hash, 21. A pontuação corresponde ao número de acertos, com o mesmo peso por questão, resultando em escalas de 0 a 20 e de 0 a 21, respectivamente. Essa medida permite comparar o desempenho nas condições com e sem uso da ferramenta durante as implementações.

3.2 Contexto e Participantes

O estudo foi conduzido na disciplina de Algoritmos e Estruturas de Dados, vinculada ao curso de Engenharia da Computação da Universidade Federal Rural de Pernambuco, na Unidade Acadêmica de Belo Jardim, no primeiro semestre letivo de 2026 (2026.1). A disciplina integra tipicamente o terceiro período da grade curricular e tem, como formação prévia esperada, conteúdos de Programação Orientada a Objetos (ou equivalente). O experimento ocorreu no final da disciplina. As atividades práticas foram implementadas em Java, linguagem adotada no componente curricular.

Os participantes foram estudantes regularmente matriculados no período de aplicação do estudo. A participação foi voluntária, precedida da apresentação dos objetivos da pesquisa e do aceite por meio do termo de consentimento livre e esclarecido (TCLE). Como incentivo acadêmico, a participação nas atividades experimentais valeu 1 ponto na disciplina. Os dados analisados foram tratados de forma agregada ou anonimizada, preservando a confidencialidade dos participantes. Ao todo, 28 estudantes responderam ao formulário inicial vinculado ao TCLE; desses, 25 integraram a distribuição experimental nos grupos — três estudantes que

responderam ao formulário não compareceram no dia da alocação aos grupos e, portanto, não participaram das atividades experimentais.

3.3 Conteúdos Investigados

Este TCC concentra-se nos conteúdos de Heap e Tabelas Hash, escolhidos por combinarem fundamentos conceituais, análise de complexidade e desafios de implementação relevantes na formação em Computação (CORMEN et al., 2022). Heap foi o primeiro conteúdo trabalhado no experimento; Tabelas Hash, o segundo, com inversão das condições de uso do ChatGPT entre os grupos.

3.4 Hipótese da Pesquisa

A investigação considera uma única hipótese nula, associada ao desempenho nos questionários após as atividades práticas:

- H_0 : não há diferença significativa no desempenho dos estudantes — medido pela pontuação nos questionários — quando comparamos condições com e sem uso do ChatGPT nas atividades práticas de Algoritmos e Estruturas de Dados.

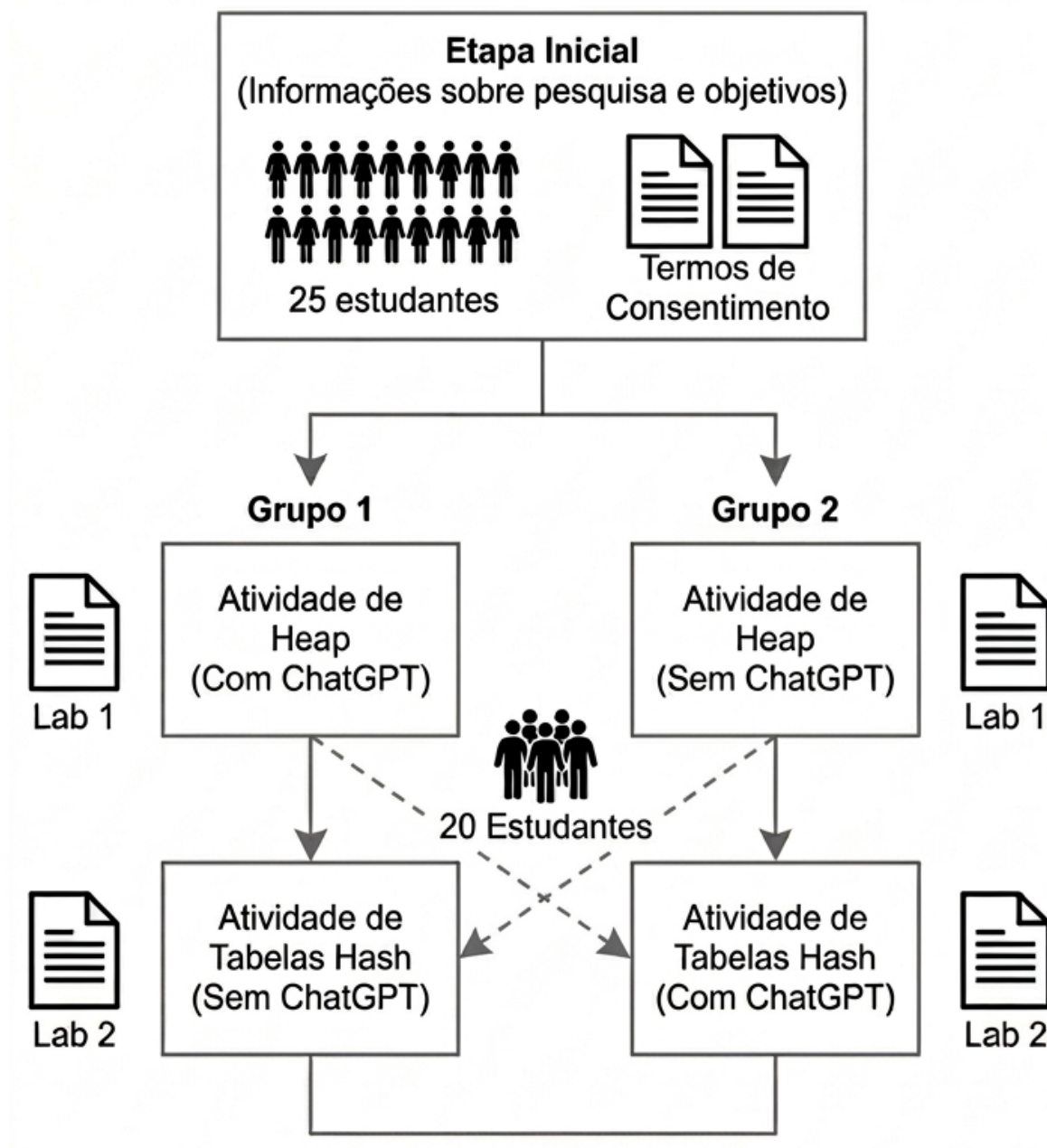
A hipótese H_0 é avaliada por meio das pontuações obtidas nos questionários aplicados após cada atividade prática, comparando o desempenho dos estudantes em Heap e Tabelas Hash sob cada condição experimental. A análise estatística considera o tamanho da amostra e a distribuição dos dados, empregando testes não paramétricos quando necessário (KITCHENHAM; MADEYSKI; BRERETON, 2019; SHESKIN, 2003).

3.5 Desenho Experimental

O desenho experimental é organizado em três momentos principais: etapa inicial, atividade de Heap e atividade de Tabelas Hash. Em cada atividade prática, a avaliação do desempenho foi realizada por meio de um questionário aplicado ao final da implementação, sem uso do ChatGPT. A Figura 3 apresenta uma visão geral do processo. Nela, a etapa experimental inicia com 25 estudantes distribuídos nos grupos; a análise comparativa final considera apenas os 20 participantes com participação completa, isto é, que cumpriram todas as etapas da pesquisa — do formulário de perfil até o último questionário.

Na etapa inicial, os estudantes foram informados sobre a pesquisa, seus objetivos, as atividades previstas e as condições de participação. Antes do início das atividades experimentais, os participantes assinaram o termo de consentimento livre e esclarecido (TCLE) e responderam a perguntas sobre familiaridade prévia e frequência de uso de ferramentas de inteligência artificial generativa, especialmente em atividades de programação. Ao todo, 28 estudantes responderam a esse formulário, autorizando o uso dos dados exclusivamente para fins acadêmicos. A Tabela 1 resume o perfil desses 28 respondentes quanto ao conhecimento e ao uso dessas ferramentas.

Figura 3 – Visão geral do processo experimental.



Fonte: Elaboração própria.

Conforme a Tabela 1, todos os respondentes já possuíam conhecimento prévio sobre ferramentas de IA generativa e a maioria declarou utilizá-las com frequência em atividades de programação. Em seguida, os 25 participantes que integraram o experimento foram distribuídos aleatoriamente em dois grupos de tamanho semelhante — Grupo 1 com 12 estudantes e Grupo 2 com 13 — com condições invertidas de uso do ChatGPT:

- **Grupo 1:** utilizou o ChatGPT na atividade prática de Heap e não utilizou a ferramenta na atividade prática de Tabelas Hash.
- **Grupo 2:** não utilizou o ChatGPT na atividade prática de Heap e utilizou a ferramenta

Tabela 1 – Perfil dos participantes quanto ao uso de IA generativa ($n = 28$, formulário inicial)

Indicador	n	Percentual
Já conhecia IA generativa	28	100,0%
<i>Frequência de uso geral</i>		
Mais de uma vez ao dia	10	35,7%
Uma vez ao dia	2	7,1%
Várias vezes na semana	9	32,1%
Algumas vezes na semana	7	25,0%
<i>Uso em atividades de programação</i>		
Sempre	1	3,6%
Quase sempre	25	89,3%
Raramente	2	7,1%

Fonte: Elaboração própria, com base no termo de consentimento.

na atividade prática de Tabelas Hash.

Esse arranjo, conhecido como desenho com grupos cruzados, permite que cada tópico seja estudado sob as duas condições experimentais, reduzindo o efeito de diferenças sistemáticas entre os conteúdos. Cada participante vivenciou uma atividade com a ferramenta e outra sem ela. A Tabela 2 resume a participação e as perdas amostrais ao longo do experimento.

Tabela 2 – Resumo da participação no experimento

Indicador	Quantidade
Estudantes distribuídos nos grupos	25
Participação completa (Heap e Tabelas Hash)	20
Desistentes	5
Grupo 1 com participação completa	9
Grupo 2 com participação completa	11

Fonte: Elaboração própria, com base na planilha de coleta do experimento.

Dos 25 estudantes distribuídos, cinco não concluíram todas as etapas e foram excluídos da análise comparativa (três no Grupo 1 e duas no Grupo 2), resultando em 20 participantes com participação completa.

O experimento foi realizado em dois encontros distintos, cada um com duração aproximada de duas horas, abrangendo aula teórica, atividade prática e aplicação do questionário. No primeiro encontro, os participantes assistiram à aula teórica de Heap e, na sequência, realizaram a implementação prática em ambiente de laboratório e responderam ao questionário objetivo correspondente. As condições de uso do ChatGPT seguiram a divisão dos grupos descrita anteriormente. No grupo autorizado a usar a ferramenta, foi permitido o acesso à interface web do ChatGPT na versão gratuita, com seleção automática de modelo e uso livre, sem roteiro prévio de *prompts*; no grupo sem ChatGPT, os estudantes puderam consultar materiais didáticos, anotações e documentação, mas não ferramentas de IA generativa. O docente e monitores acompanharam o laboratório para reforçar o cumprimento das condições. Ao término da atividade, todos responderam ao questionário sobre Heap, sem apoio do ChatGPT.

Aproximadamente 15 dias depois, ocorreu o segundo encontro, dedicado a Tabelas Hash, com a mesma estrutura (aula teórica, implementação prática e questionário), porém com as condições de uso do ChatGPT invertidas entre os grupos.

3.6 Instrumentos e Materiais

As atividades práticas foram disponibilizadas por meio de repositórios que contêm código-base e métodos a serem implementados. Essa estrutura padronizou as tarefas e garantiu que todos os estudantes trabalhassem sobre os mesmos requisitos. O código-base continha apenas a estrutura necessária para orientar a implementação, sem fornecer diretamente a solução dos métodos avaliados.

O instrumento de avaliação foi composto por questionários objetivos de múltipla escolha (itens fechados). As questões foram selecionadas aleatoriamente a partir de bancos públicos de questões de entrevista técnica da plataforma Sanfoundry, alinhadas aos conteúdos de Heap e Tabelas Hash¹. Cada acerto vale um ponto; a pontuação final de cada questionário é a soma dos acertos (0 a 20 em Heap; 0 a 21 em Tabelas Hash). Em ambos os casos, o uso do ChatGPT não foi permitido durante a aplicação do instrumento. Os questionários medem desempenho por reconhecimento e resposta a itens objetivos, e não pretendem capturar, por si só, toda a profundidade da aprendizagem prática.

Os materiais do experimento, os enunciados das atividades, os instrumentos e os scripts de análise estatística estão disponíveis em repositório público no GitHub².

3.7 Procedimentos de Coleta de Dados

A coleta de dados seguiu a sequência definida no desenho experimental. Inicialmente, os participantes receberam orientações sobre a pesquisa e foram distribuídos nos dois grupos. Em seguida, para o conteúdo de Heap, assistiram à aula teórica, realizaram a atividade prática sob a condição de uso do ChatGPT definida para seu grupo e responderam ao questionário correspondente, sem apoio da ferramenta.

Depois da conclusão dessa etapa, os estudantes participaram da aula teórica de Tabelas Hash, realizaram a atividade prática sob a condição invertida de uso do ChatGPT e responderam ao questionário sobre esse conteúdo, também sem apoio da ferramenta.

Os dados utilizados na análise comparativa foram as pontuações dos questionários de Heap e Tabelas Hash dos participantes com participação completa — aqueles que cumpriram todas as etapas da pesquisa, desde o formulário de perfil até o último questionário. Foram excluídos da análise comparativa os cinco estudantes que não concluíram todas as etapas, embora parte deles tenha participado da etapa de Heap.

¹ Heapsort: <https://www.sanfoundry.com/heapsort-multiple-choice-questions-answers-mcqs/>. Hash tables (chaining): <https://www.sanfoundry.com/hash-tables-chaining-binary-trees-multiple-choice-questions-answers-mcqs/>.

² <<https://github.com/hellyaxs/MW-CL-pesquisa-tcc>>

3.8 Procedimentos de Análise dos Dados

A análise dos dados foi organizada para avaliar a hipótese H_0 . As pontuações de Heap e de Tabelas Hash foram reorganizadas por condição experimental — com ChatGPT e sem ChatGPT — e não por grupo de alocação inicial, conforme o desenho com grupos cruzados. A comparação inferencial principal considerou todas as pontuações de cada condição, totalizando 20 observações por condição na amostra completa. As pontuações foram analisadas na escala original de número de acertos, sem conversão para percentual ou outra normalização entre os questionários; essa diferença de escala (máximos 20 e 21) deve ser considerada na interpretação da comparação agregada.

Foram calculadas estatísticas descritivas das pontuações, incluindo média, mediana, desvio padrão, valores mínimo e máximo, tanto na comparação agregada entre condições quanto, de forma exploratória, separadas por conteúdo. Em seguida, verificou-se a distribuição dos dados por meio do teste de Shapiro–Wilk, como forma de avaliar a normalidade das pontuações em cada condição experimental.

Embora os resultados do teste de Shapiro–Wilk tenham indicado desvio significativo da normalidade em uma das condições, optou-se pela utilização do teste não paramétrico de Mann–Whitney U para a comparação agregada entre as pontuações com e sem uso do ChatGPT. Essa escolha foi motivada pelo tamanho reduzido da amostra, uma vez que testes não paramétricos tendem a apresentar maior robustez em estudos com número limitado de participantes (KITCHENHAM; MADEYSKI; BRERETON, 2019; SHESKIN, 2003). A análise principal trata as pontuações reorganizadas por condição como duas amostras a serem comparadas quanto ao desempenho sob cada regime de uso da ferramenta. Reconhece-se, contudo, que o desenho com grupos cruzados produz, no nível do participante, um par de observações (uma com ChatGPT e outra sem); a opção pelo Mann–Whitney U agregado privilegiou a comparação entre condições na forma em que a hipótese H_0 foi formulada, em detrimento de um teste estritamente pareado, o que constitui uma limitação metodológica a ser considerada na interpretação dos resultados.

Os procedimentos descritivos e inferenciais foram executados em Python, com apoio da biblioteca SciPy para os testes de Shapiro–Wilk e de Mann–Whitney U, além do cálculo do Cliff’s Delta implementado a partir das pontuações reorganizadas por condição.

Além da verificação da significância estatística, também foi calculado o coeficiente Cliff’s Delta sobre a comparação agregada entre condições, com o objetivo de estimar a magnitude da diferença observada. Enquanto o teste de Mann–Whitney U permite verificar a existência de diferenças estatisticamente significativas, o Cliff’s Delta complementa essa análise ao quantificar a intensidade do efeito encontrado. Para interpretação, adotaram-se os limiares propostos por Romano et al., conforme reportado por Kitchenham et al.: $|\delta| < 0,147$ indica efeito negligível; entre 0,147 e 0,33, pequeno; entre 0,33 e 0,474, médio; e acima de 0,474, grande (KITCHENHAM; MADEYSKI; BRERETON, 2019).

4 RESULTADOS E DISCUSSÕES

Este capítulo apresenta os resultados do experimento e discute suas implicações à luz da hipótese H_0 e do referencial teórico. Descrevem-se as estatísticas descritivas e a análise inferencial conduzida conforme a Metodologia, com base na amostra de participação completa definida na Seção 3.5.

4.1 Organização dos Dados

As pontuações foram reorganizadas por condição experimental, conforme descrito na Seção 3.8. A diferença de escala entre os questionários (máximos de 20 e 21 acertos) deve ser considerada na interpretação da comparação agregada, uma vez que as pontuações foram analisadas sem normalização.

4.2 Estatísticas Descritivas

A Tabela 3 apresenta as estatísticas descritivas dos 20 participantes com participação completa. As médias das duas condições diferiram em apenas 0,15 ponto, com mediana idêntica em ambos os casos.

Tabela 3 – Estatísticas descritivas do desempenho nos questionários

Condição	n	Média	Mediana	DP	Mín.	Máx.
Com uso do ChatGPT	20	15,05	16,00	3,43	6,00	19,00
Sem uso do ChatGPT	20	14,90	16,00	3,97	7,00	20,00

Fonte: Elaboração própria, com base na planilha de coleta do experimento.

A Tabela 4 detalha, de forma exploratória, os resultados por conteúdo. Em Heap, a média dos estudantes que utilizaram o ChatGPT (13,00) foi ligeiramente inferior à dos que não utilizaram a ferramenta (13,18). Em Tabelas Hash, observou-se o padrão inverso (16,73 versus 17,00), porém com diferenças pequenas em ambos os casos. Esses valores descrevem o desempenho por tópico, mas não constituem a base da análise inferencial principal, conduzida com base na comparação agregada entre condições.

Tabela 4 – Desempenho médio por conteúdo e condição de uso do ChatGPT

Conteúdo	Condição	n	Média
Heap (máx. 20)	Com ChatGPT	9	13,00
Heap (máx. 20)	Sem ChatGPT	11	13,18
Tabelas Hash (máx. 21)	Com ChatGPT	11	16,73
Tabelas Hash (máx. 21)	Sem ChatGPT	9	17,00

Fonte: Elaboração própria, com base na planilha de coleta do experimento.

4.3 Análise Estatística da Hipótese H_0

Antes da comparação entre as condições, aplicou-se o teste de Shapiro–Wilk, conforme descrito na Metodologia. Para a condição com ChatGPT, obteve-se $W = 0,875$ e valor- $p = 0,014$; para a condição sem ChatGPT, $W = 0,917$ e valor- $p = 0,088$.

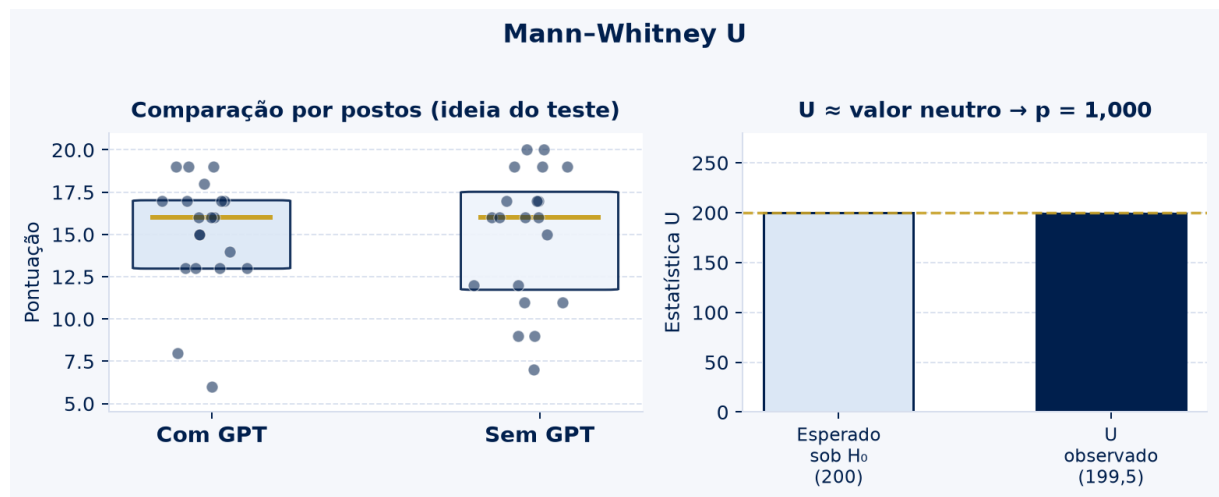
Para avaliar a hipótese H_0 , aplicou-se o teste de Mann–Whitney U à comparação agregada ($\alpha = 0,05$). O resultado é apresentado na Tabela 5. A Figura 4 ilustra a distribuição das pontuações por condição e a proximidade entre o valor de U observado e o esperado sob H_0 .

Tabela 5 – Resultado do teste de Mann–Whitney U na comparação agregada

Comparação	Estatística U	p -valor	Resultado
Agregada (com vs. sem ChatGPT)	199,50	1,000	Não rejeita H_0

Fonte: Elaboração própria, com base na planilha de coleta do experimento.

Figura 4 – Distribuição das pontuações e estatística U do teste de Mann–Whitney.



Fonte: Elaboração própria.

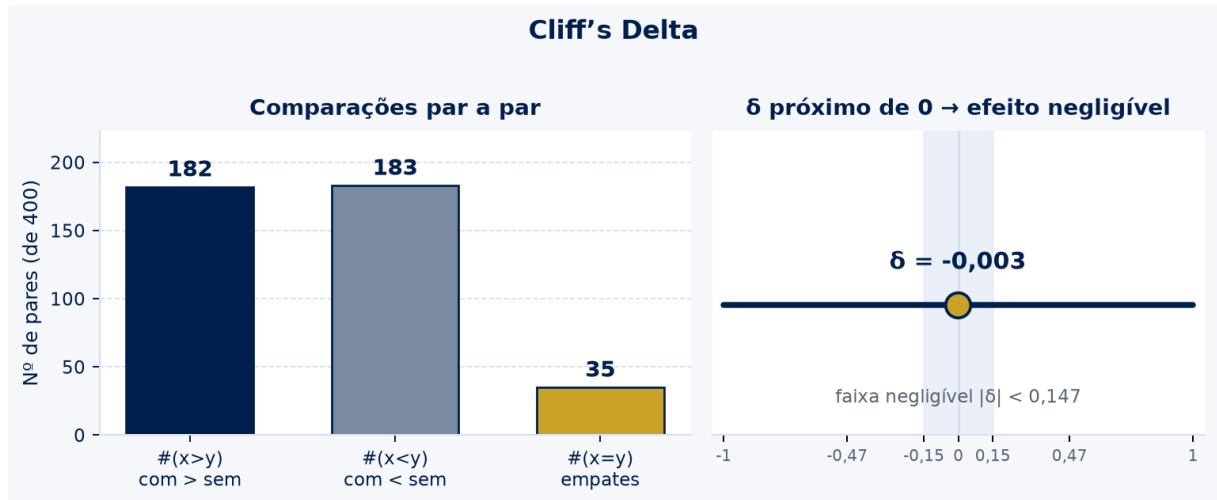
O teste de Mann–Whitney U não indicou diferença estatisticamente significativa ($p = 1,000$); portanto, não há evidências para rejeitar a hipótese H_0 . O Cliff's Delta na comparação agregada foi de $-0,003$, classificado como efeito negligível (Tabela 6). A Figura 5 sintetiza as comparações par a par e a magnitude do efeito.

Tabela 6 – Coeficiente Cliff's Delta e magnitude do efeito na comparação agregada

Comparação	Cliff's Delta (δ)	Magnitude
Agregada (com vs. sem ChatGPT)	$-0,003$	Negligível

Fonte: Elaboração própria, com base na planilha de coleta do experimento.

Figura 5 – Comparações par a par e magnitude do Cliff's Delta.



Fonte: Elaboração própria.

4.4 Discussão dos Resultados

Os resultados da Seção 4.3 indicam que o uso do ChatGPT durante as atividades práticas não alterou, de forma mensurável, o desempenho avaliado pelos questionários. Esse achado converge com estudos que também não identificaram impacto significativo do ChatGPT no desempenho de estudantes em disciplinas de programação e estruturas de dados (KOSAR et al., 2024).

A elevada familiaridade prévia dos participantes com IA generativa (Tabela 1) pode ter atenuado as diferenças entre as condições, pois a restrição ao ChatGPT durante uma atividade não elimina a experiência acumulada com a ferramenta. Além disso, a comparação agregada combina pontuações de questionários com escalas distintas e sem normalização, o que reforça a leitura conjunta com a Tabela 4. As demais limitações do estudo são discutidas na Seção 4.5.

Do ponto de vista educacional, a ausência de diferença significativa não deve ser interpretada como prova de neutralidade do ChatGPT para a aprendizagem. Os questionários avaliaram desempenho por meio de questões objetivas de múltipla escolha, o que pode não capturar efeitos sobre autonomia, depuração ou dependência da ferramenta (KOSMYNA et al., 2025; DELIKOURA; FUNG; HUI, 2025).

Esses achados têm implicações distintas para os públicos envolvidos. Para os estudantes, o resultado sugere que o uso da ferramenta na prática, por si só, não garante melhor desempenho em avaliações objetivas posteriores sem apoio da IA, o que reforça a importância de estudar e revisar os conteúdos com autonomia. Para o docente, indica que permitir ou restringir o ChatGPT em uma atividade pontual pode não alterar de imediato as notas em questionários desse tipo, mas continua relevante orientar o uso crítico e combinar a ferramenta a instrumentos que avaliem implementação, justificação e depuração. Para pesquisadores que derem continuidade a esta linha, os dados motivam amostras maiores, análises complementares (por exemplo,

incluindo participantes com dados parciais em um único tópico) e desenhos que separem melhor o efeito da sessão experimental da experiência prévia com IA generativa.

Por fim, os achados reforçam a necessidade de investigações adicionais sobre o uso pedagógico do ChatGPT em Algoritmos e Estruturas de Dados, com amostras maiores, instrumentos mais abrangentes e observação de efeitos de longo prazo.

4.5 Ameaças à Validade

Algumas ameaças à validade devem ser consideradas na interpretação dos resultados. A primeira diz respeito ao tamanho da amostra, pois estudos em turmas específicas podem envolver um número limitado de participantes, o que reduz a generalização dos achados. Em particular, a análise principal restringiu-se aos 20 participantes com participação completa; não foi conduzida, neste TCC, uma análise complementar apenas com os dados de Heap dos 25 estudantes distribuídos, o que permanece como possibilidade para trabalhos futuros.

A segunda refere-se à diferença de dificuldade entre os conteúdos. Embora o desenho com grupos cruzados controle parte das assimetrias entre tópicos, ainda é possível que variações de dificuldade entre Heap e Tabelas Hash influenciem o desempenho no questionário, especialmente na comparação agregada entre condições. Além disso, diferenças prévias de conhecimento entre os estudantes podem afetar os resultados, mesmo com a inversão das condições.

Também existe uma ameaça relacionada ao controle do uso do ChatGPT. Embora o laboratório tenha sido monitorado durante as atividades, não é possível garantir a ausência total de uso indevido fora do ambiente controlado. Outra limitação envolve o instrumento de avaliação: questionários objetivos permitem a comparação direta entre condições, mas podem não captar toda a profundidade da aprendizagem, especialmente aspectos como raciocínio, justificativa de decisões e capacidade de depuração. Soma-se a isso a escolha do Mann–Whitney U agregado em um desenho que, no nível individual, é pareado, conforme discutido na Seção 3.8.

Por fim, os resultados devem ser interpretados considerando o contexto institucional, a linguagem de programação adotada, o perfil da turma e o período de aplicação do estudo.

5 CONCLUSÕES

Este Trabalho de Conclusão de Curso investigou o impacto do uso do ChatGPT no desempenho de estudantes da disciplina de Algoritmos e Estruturas de Dados, com foco em Heap e Tabelas Hash, por meio de um experimento com grupos cruzados e questionários objetivos aplicados ao final de cada atividade prática.

Em relação ao objetivo geral, os dados coletados não permitiram rejeitar a hipótese H_0 : o uso do ChatGPT durante as atividades práticas não produziu diferença estatisticamente significativa no desempenho medido pela pontuação nos questionários, e o tamanho de efeito foi negligível (Capítulo 4). Assim, no contexto analisado, não há evidências empíricas de que a ferramenta tenha alterado de forma mensurável o desempenho avaliado.

Esse resultado converge com estudos que também não identificaram impacto significativo do ChatGPT no desempenho de estudantes em disciplinas de programação e estruturas de dados (KOSAR et al., 2024), mas deve ser interpretado à luz das limitações discutidas na Metodologia, entre elas o tamanho amostral, o perfil de uso prévio de IA generativa e o escopo restrito dos questionários.

Do ponto de vista educacional, a ausência de diferença significativa não comprova neutralidade do ChatGPT para a aprendizagem. A ferramenta pode influenciar autonomia, dependência e engajamento cognitivo de formas não capturadas pelo instrumento adotado (KOSMYNA et al., 2025; DELIKOURA; FUNG; HUI, 2025).

Como desdobramentos futuros, sugere-se ampliar o tamanho amostral, empregar instrumentos que contemplem diferentes dimensões da aprendizagem, investigar efeitos de longo prazo e comparar estratégias de mediação docente no uso da ferramenta. Também seria relevante replicar o estudo em outros tópicos de Algoritmos e Estruturas de Dados.

REFERÊNCIAS

- BAIDOO-ANU, D.; ANSAH, L. O. Education in the era of generative artificial intelligence (ai): Understanding the potential benefits of chatgpt in promoting teaching and learning. **Journal of AI**, v. 7, n. 1, p. 52–62, 2023. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4337484>. Citado 2 vezes nas páginas 14 e 19.
- CORMEN, T. H.; LEISERSON, C. E.; RIVEST, R. L.; STEIN, C. **Introduction to algorithms**. 4. ed. MIT Press, 2022. Disponível em: <<https://mitpress.mit.edu/9780262046305/introduction-to-algorithms/>>. Citado 5 vezes nas páginas 14, 16, 17, 18 e 24.
- DELIKOURA, I.; FUNG, Y. R.; HUI, P. **From superficial outputs to superficial learning: Risks of large language models in education**. 2025. ArXiv preprint arXiv:2509.21972. Disponível em: <<https://arxiv.org/abs/2509.21972>>. Citado 4 vezes nas páginas 14, 20, 31 e 33.
- DWIVEDI, Y. K.; KSHETRI, N.; HUGHES, L.; SLADE, E. L.; JEYARAJ, A.; KAR, A. K.; BAABDULLAH, A. M.; KOOHANG, A.; RAGHAVAN, V.; AHUJA, M. et al. So what if chatgpt wrote it? multidisciplinary perspectives on opportunities, challenges and implications of generative conversational ai for research, practice and policy. **International Journal of Information Management**, v. 71, p. 102642, 2023. Disponível em: <<https://doi.org/10.1016/j.ijinfomgt.2023.102642>>. Citado 2 vezes nas páginas 14 e 20.
- GARCIA, M. B. Teaching and learning computer programming using chatgpt: A rapid review of literature amid the rise of generative ai technologies. **Education and Information Technologies**, p. 1–25, 2025. Disponível em: <<https://doi.org/10.1007/s10639-025-13452-5>>. Citado 2 vezes nas páginas 14 e 20.
- JAMIE, P.; HAJIHASHEMI, R.; ALIPOUR, S. **Utilizing ChatGPT in a Data Structures and Algorithms Course: A Teaching Assistant's Perspective**. 2024. ArXiv preprint arXiv:2410.08899. Disponível em: <<https://arxiv.org/abs/2410.08899>>. Citado 2 vezes nas páginas 14 e 20.
- KITCHENHAM, B.; MADEYSKI, L.; BRERETON, P. Problems with statistical practice in human-centric software engineering experiments. In: **Proceedings of the 23rd International Conference on Evaluation and Assessment in Software Engineering**. [s.n.], 2019. p. 134–143. Disponível em: <<https://dl.acm.org/doi/10.1145/3319008.3319009>>. Citado 3 vezes nas páginas 21, 24 e 28.
- KOSAR, T.; OSTOJIĆ, D.; LIU, Y. D.; MERNIK, M. Computer science education in chatgpt era: Experiences from an experiment in a programming course for novice programmers. **Mathematics**, v. 12, n. 5, p. 629, 2024. Disponível em: <<https://www.mdpi.com/2227-7390/12/5/629>>. Citado 4 vezes nas páginas 14, 20, 31 e 33.
- KOSMYNA, N.; HAUPTMANN, E.; YUAN, Y. T.; SITU, J.; LIAO, X.-H.; BERESNITZKY, A. V.; BRAUNSTEIN, I.; MAES, P. **Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task**. 2025. ArXiv preprint arXiv:2506.08872. Disponível em: <<https://arxiv.org/abs/2506.08872>>. Citado 3 vezes nas páginas 20, 31 e 33.

LEBLANC, R. J.; SOBEL, A.; DIAZ-HERRERA, J. L.; HILBURN, T. B. **Software engineering 2014: curriculum guidelines for undergraduate degree programs in software engineering**. IEEE Computer Society, 2015. Disponível em: <<https://www.acm.org/binaries/content/assets/education/se2014.pdf>>. Citado na página 14.

MACNEIL, S.; TRAN, A.; MOGIL, D.; BERNSTEIN, S.; ROSS, E.; HUANG, Z. Generating diverse code explanations using the gpt-3 large language model. In: **Proceedings of the 2022 ACM Conference on International Computing Education Research**. [s.n.], 2022. p. 37–39. Disponível em: <<https://doi.org/10.1145/3501709.3544280>>. Citado na página 20.

MUÑOZ, S. A. S.; GAYOSO, G. G.; HUAMBO, A. C.; TAPIA, R. D. C.; INCALUQUE, J. L.; AGUILA, O. E. P.; CAJAMARCA, J. C. R.; ACEVEDO, J. E. R.; RIVERA, H. V. H.; ARIAS-GONZÁLES, J. L. Examining the impacts of chatgpt on student motivation and engagement. **Social Space**, v. 23, n. 1, p. 1–27, 2023. Disponível em: <https://www.researchgate.net/publication/370894242_Examining_the_Impacts_of_ChatGPT_on_Student_Motivation_and_Engagement>. Citado na página 20.

NETO, W. Toward evaluating the impact of chatgpt on learning data structures and algorithms: an experiment with undergraduate students. In: **Anais do Simpósio Brasileiro de Engenharia de Software (SBES), Congresso Brasileiro de Engenharia de Software (CBSOFT)**. São Paulo: [s.n.], 2026. Aceito para publicação. DOI/URL definitivo a ser inserido quando disponível. Citado na página 14.

QIN, C.; ZHANG, A.; ZHANG, Z.; CHEN, J.; YASUNAGA, M.; YANG, D. **Is ChatGPT a general-purpose natural language processing task solver?** 2023. ArXiv preprint arXiv:2302.06476. Disponível em: <<https://arxiv.org/abs/2302.06476>>. Citado 2 vezes nas páginas 14 e 20.

RAHMAN, M. M.; WATANOBÉ, Y. Chatgpt for education and research: Opportunities, threats, and strategies. **Applied Sciences**, v. 13, n. 9, p. 5783, 2023. Disponível em: <<https://www.mdpi.com/2076-3417/13/9/5783>>. Citado 2 vezes nas páginas 14 e 19.

SHESKIN, D. J. **Handbook of parametric and nonparametric statistical procedures**. Chapman and Hall/CRC, 2003. Disponível em: <<https://www.taylorfrancis.com/books/mono/10.4324/9780203489536/handbook-parametric-nonparametric-statistical-procedures-david-sheskin>>. Citado 3 vezes nas páginas 21, 24 e 28.

WOHLIN, C.; RUNESON, P.; HÖST, M.; OHLSSON, M. C.; REGNELL, B.; WESSLÉN, A. **Experimentation in software engineering**. Springer Science and Business Media, 2012. Disponível em: <https://books.google.com/books/about/Experimentation_in_Software_Engineering.html?id=QPVsM1_U8nkC>. Citado 2 vezes nas páginas 21 e 23.