

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/376894231>

# A comparison between ChatGPT-3.5 and ChatGPT-4.0 as a tool for paraphrasing an English Paragraphs

Conference Paper · December 2023

CITATIONS

0

READS

7

6 authors, including:



Thaeer M Sahib

Universiti Sains Malaysia

10 PUBLICATIONS 13 CITATIONS

SEE PROFILE



Osamah Mohammed Alyasiri

Universiti Sains Malaysia

13 PUBLICATIONS 42 CITATIONS

SEE PROFILE



VII.

**INTERNATIONAL  
APPLIED SOCIAL  
SCIENCES  
CONGRESS  
(C-IASOS – 2023)**

**Full Paper  
Proceedings  
Book**



**L-Università  
ta' Malta**



SVEUČILIŠTE U RIJECI UNIVERSITY OF RIJEKA  
FAKULTET ZA MENADŽMENT U TURIZMU I UGOSTITELJSTVU  
FACULTY OF TOURISM AND HOSPITALITY MANAGEMENT  
OPATUA, HRVATSKA CROATIA



## A comparison between ChatGPT-3.5 and ChatGPT-4.0 as a tool for paraphrasing an English Paragraphs

Thaeer Mueen Sahib<sup>1,2,\*</sup>, Osamah Mohammed Alyasiri<sup>3,4</sup>, Hussain A.Younis<sup>4,5</sup>, Dua'a Akhtom<sup>4</sup>, Israa M. Hayder<sup>6</sup>, Sani Salisu<sup>4,7</sup>, Muthmainnah<sup>8</sup>

### Abstract

**Introduction:** In recent years, there has been a proliferation of online platforms and tools aimed at enhancing English language learning, including prominent ones such as TikTok, Instagram, and Duolingo. Among these, a noteworthy addition emerged in September 2022 in the form of an artificial intelligence-based platform, ChatGPT-3.5, which was subsequently updated to ChatGPT-4.0.

**Aim:** This paper investigates the performance of these platforms in rephrasing and proofreading a specific paragraph, with a particular focus on their ability to generate text devoid of pronouns.

**Method:** We selected a specific paragraph as the focal point of our study and conducted a comparative analysis between ChatGPT-3.5 and ChatGPT-4.0, utilizing different user accounts and initiating distinct chat sessions. Pronouns generated by each version were closely examined and highlighted.

**Findings:** While the proprietary software, ChatGPT-4.0, exhibited only marginal disparities in terms of paraphrasing and proofreading, it continued to encounter challenges in completely eliminating pronouns from the paragraph, despite our utilization of a single user account. Conversely, the open-source software, ChatGPT-3.5, demonstrated the ability to generate a paragraph free from pronouns when different user accounts were employed.

**Originality and significance:** This study offers valuable insights and practical recommendations for researchers looking to leverage ChatGPT in their English language endeavors, emphasizing the feasibility of selecting a pronoun-free paragraph, particularly when utilizing ChatGPT-3.5.

**Key Words:** ChatGPT-3.5, ChatGPT-4.0, Paraphrasing, Proofreading, English Language.

### 1. INTRODUCTION

Digital technologies have advanced rapidly in recent years, along with software and social media platforms that enable human communication. TikTok, Instagram, Facebook, and many others are examples of platforms that have been applied in various

---

<sup>1</sup>School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Penang 14300 Nibong Tebal, Malaysia

<sup>2</sup> Kufa Technical Institute, Al-Furat Al-Awsat Technical University, Kufa 540011, Iraq

<sup>3</sup> Karbala Technical Institute, Al-Furat Al-Awsat Technical University, Karbala 56001, Iraq

<sup>4</sup> School of Computer Sciences, Universiti Sains Malaysia, Penang 11800, Malaysia

<sup>5</sup> College of Education for Women, University of Basrah, Basrah 61004, Iraq

<sup>6</sup> Qurna Technique Institute, Southern Technical University, Basrah 61016, Iraq

<sup>7</sup> Department of Information Technology, Federal University Dutse, Dutse 720101, Nigeria

<sup>8</sup> Universitas Al Asyariah Mandar, Sulawesi Barat Indonesia

\*Corresponding author

fields and domains (Sahib & Ali, 2023;Thurnes, 2023). Language skill has become a crucial and competitive factor among these platforms (Lee, 2023).

After the COVID-19 pandemic ended, the world faced a challenge to create a powerful platform based on artificial intelligence (AI) that could provide easy and sufficient knowledge for humanity. In response to this challenge, researchers developed ChatGPT-3.5 version and upgraded it to ChatGPT-4.0 (Hill-Yardin, 2023).

ChatGPT (ChatGPT Playground) Chat Generative Pre-Trained Transformer) is an OpenAI-developed AI tool that generates text depending on user input. It is programmed to comprehend natural language and produce intelligent and pertinent replies to user requests. It has been trained on vast quantities of data (Fitria, 2023; Halaweh, 2023).

ChatGPT has potential uses in many fields, such as software engineering, healthcare, marketing, and education. In education, ChatGPT can analyze writing for grammar, punctuation, spelling errors, and offer suggestions to improve the overall structure (Fraivan, 2023). Such as provide guidelines on crafting accurate queries helps reduce the likelihood of using incorrect language and including irrelevant references (Alyasiri et al., 2023).

Even though ChatGPT speeds up the academic and scientific writing process for writers, especially students and early-career researchers, one of its primary benefits is the rapid analysis of large amounts of data. Using approaches, researchers may examine many research articles in half the time required to read them (Zohery, 2023). From another point of view, a question arises about the quality of the texts that ChatGPT produces. Consequently, this paper has discussed some limitations.

A study conducted by (Shahriar et al., 2023) employed ChatGPT to compose a literature review on the function of digital twins in healthcare. Despite the promising outcomes, significant similarities were detected by a plagiarism checker when rephrasing sentences. In another study carried out by (Koos & Wachsmann, 2023) AI language models become more prevalent in academic writing and research, and text writing is likely to lose much of its relevance leading to a shift from text design to intelligent and creative steering of AI systems by users and sophisticated control of the generated text. Eventually, a study presented by (Halaweh, 2023) discussed three argument, first, ChatGPT provide high-quality outputs that have a high likelihood of passing plagiarism detection software, Second, AI can determine with high precision whether a sentence was created by a person or by OpenAI's classifier. Thirdly, ChatGPT is readily available to all users (Mohammed, O., Sahib, 2023)

Based on our knowledge, no study employed depth focusing on creative ways to help the researchers and users in enhancing their writing. This study makes a simple comparison between ChatGPT-3.5 and ChatGPT-4.0 to measure their abilities in terms of re-writing the paragraphs without pronouns. The paragraphs have been tested in both platforms based on two methods, the first included different user accounts, and the second examined in single user account but in a different new chat.

The remainder of this paper includes Section 2 research methodology, Section 3 results and discussion, and Section 4 conclusion.

## **2. RESEARCH METHODOLOGY**

To support the argument presented in this paper, various indicators and facts must be obvious and experimentally described through the steps in Figure 1.

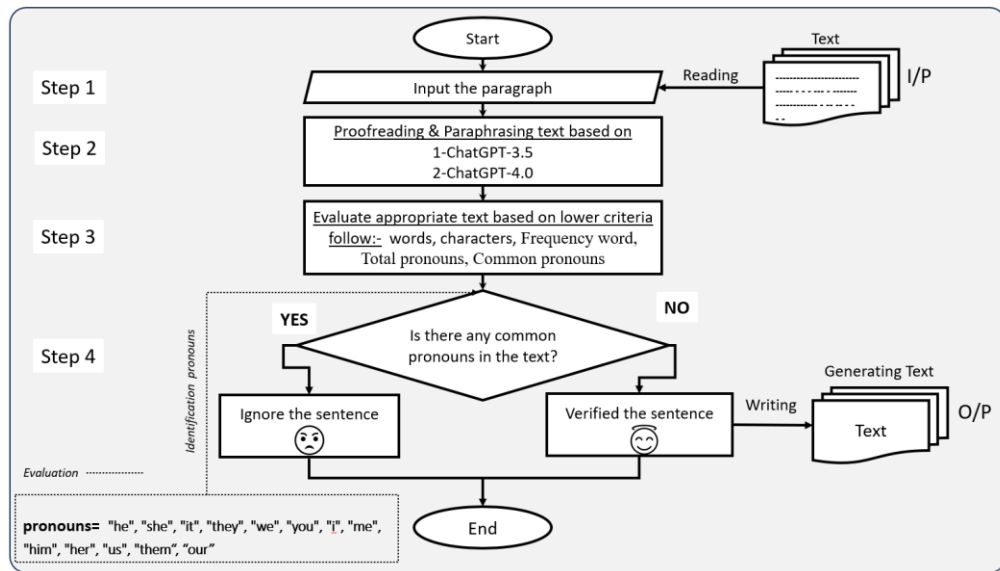


Figure 4. Flowchart of comparison between ChatGPT3.5 and ChatGPT-4.0 in terms of re-writing text without pronouns.

**Step 1-** Configure four different user accounts from ChatGPT-3.5 platform (Open-source software) and a single user account from ChatGPT-4.0 platform (close source software), and an entire certain text four times into ChatGPT-3.5 version, the original text illustrated in Figure 2.

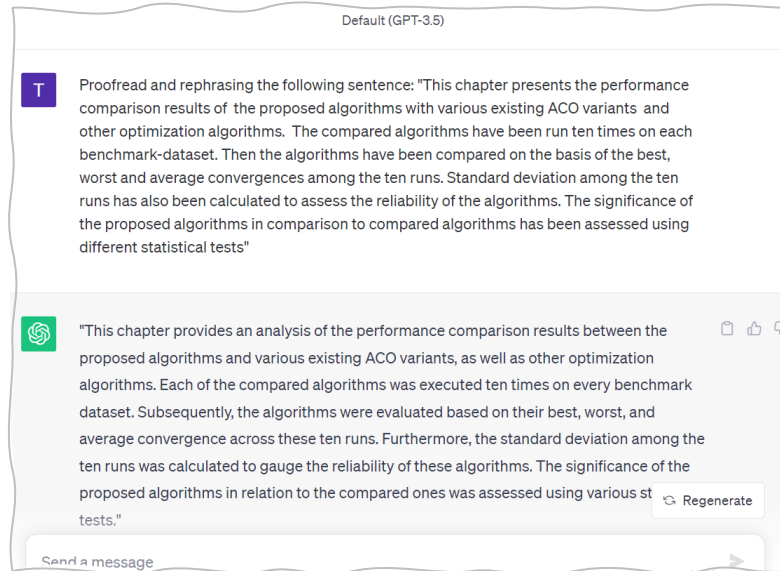


Figure 2. Sample of asking and response ChatGPT-3.5 about proofread and paraphrasing a paragraph.

**Step 2-** Paraphrasing and proofreading the text in terms of different user accounts. As well as inputting the same paragraph again, into ChatGPT-4.0 version for the same purpose. Eventually, entering the same paragraph into ChatGPT-3.5 version in terms of several new chats four times to examine what the text would look like. **Step 3-** Verified

the best paragraphs that match the criteria including evaluating appropriate text based on a lower criterion as follows: - words, characters, frequency word, total pronouns, and common pronouns.

**Step 4-** Comparing the new text generated by ChatGPT-3.5 and ChatGPT-4.0 versions in terms of texts free from pronouns. In this paper, the identify the list of most pronouns using Spyder (python 3.8) environment, detecting the pronouns with compare to the original text, then compute the total number of pronouns used in the entire paragraphs, configured as follows: common pronouns = ["he", "she", "it", "they", "we", "you", "i", "me", "him", "her", "us", "them"] as step 3 shown in Table 1.

*Table 1: Pesudocode identification the text without pronouns*

1	<code>text =Initialize the paragraph</code>
2	<code>text_lower = text.lower() # Convert text to lowercase for case-insensitive matching</code>
3	<code>common_pronouns = Initialize the pronounce</code>
4	<code>pronoun_count = 0 # Initialize a counter for pronoun occurrences</code>
5	<code>words = text_lower.split() # Split text into words and count pronouns</code>
6	<b>For</b> word := words <b>do</b>
7	<b>if</b> word := common_pronouns <b>then</b>
8	<code>pronoun_count += 1</code>
9	<b>End if</b>
10	<b>End for</b>
11	<code>Print (Total common pronouns in the text: {pronoun_count})</code>
12	<b>for</b> word := words <b>do</b>
13	<b>for</b> common_pronoun := common_pronouns <b>do</b>
14	<b>if</b> word==common_pronoun <b>then</b>
15	<code>Print ("The pronoun is",word)</code>
16	<b>End if</b>
17	<b>End for</b>
18	<b>End for</b>

The next section leads to describe the results and discussion.

### 3. RESULTS AND DISCUSSION

As ChatGPT is a relatively new technology, it is highly likely that many faculty and students are uninformed of it or have only heard about it without using it. To ensure using best paraphrasing and proofreading properly. Entering the original text in ChatGPT-3.5 and ChatGPT-3.5, the best paragraph evaluated based on (lower criteria / fewer words, characters, frequency word, and total common pronouns), after proofreading and paraphrasing the response of ChatGPT-3.5 provides the different answers based on text analysis operations using natural language processing (NLP) as described in Table 2.

VII. International Applied Social Sciences Congress - (C-IASOS – 2023)  
Valletta –Malta, 13-15 November 2023

Table 2: Responses of Prephrasing and Proofreading by ChatGPT-3.5 in Terms of Different User Account

User Account (UA)	The response	Metrics of Text Analysis
1	This chapter <u>presents</u> the performance comparison results of the proposed algorithms <u>with</u> various existing ACO variants and other optimization <u>algorithms</u> . The compared algorithms <u>have been run</u> ten times on each benchmark-dataset. <u>Then</u> the algorithms <u>have been compared on the basis of the</u> best, worst and average convergences <u>among</u> the ten runs. Standard deviation among the ten runs has also been calculated <u>to assess the reliability of the algorithms</u> . The significance of the proposed algorithms in comparison to compared algorithms <u>has been</u> assessed using <u>different</u> statistical tests.	Words= <b>86</b> Characters = <b>586</b> Frequency word (FW)= <b>92</b> Total common pronouns= <b>0</b> Common pronouns= <b>None</b>
2	This chapter presents the performance comparison results of the proposed algorithms against various existing ACO variants and other optimization techniques. The compared algorithms were executed ten times on each benchmark dataset. Subsequently, the algorithms were assessed based on their best, worst, and average convergence across the ten runs. Additionally, we calculated the standard deviation among the ten runs to evaluate the algorithms' reliability. The significance of the proposed algorithms, in comparison to the other algorithms, was assessed using various statistical tests.	Words= <b>81</b> Characters = <b>588</b> Frequency word (FW)= <b>93</b> Total common pronouns= <b>0</b> Common pronouns= <b>None</b>
3	This chapter provides the results of a performance comparison between the proposed algorithms and various existing ACO variants, as well as other optimization algorithms. The compared algorithms were executed ten times on each benchmark dataset. Subsequently, the algorithms were evaluated based on the best, worst, and average convergence achieved in these ten runs. The standard deviation among these ten runs was also calculated to gauge the algorithms' reliability. To assess the significance of the proposed algorithms in comparison to the other algorithms, various statistical tests were employed.	Words= <b>87</b> Characters = <b>603</b> Frequency word (FW)= <b>87</b> Total common pronouns= <b>0</b> Common pronouns= <b>None</b>
4	This chapter provides a presentation of the performance comparison results between the proposed algorithms and various existing ACO variants, as well as other optimization algorithms. These compared algorithms underwent ten separate runs on each benchmark dataset. Subsequently, these algorithms were assessed based on their best, worst, and average convergences across the ten runs. Additionally, the standard deviation among these ten runs was calculated to evaluate the reliability of the algorithms. To gauge the significance of the proposed algorithms in comparison to the compared algorithms, various statistical tests were employed.	Words= <b>89</b> Characters = <b>639</b> Frequency word (FW)= <b>100</b> Total common pronouns= <b>0</b> Common pronouns= <b>None</b>

The Zigzag line is obvious in the text of user account 1 the most paraphrased words. All criteria have been obtained in the second column which represents the analysis of text metrics in terms of different user accounts. Where the lower metric value is the best text considered. Frequency distribution of words refers to the number of words repeated in the paragraph entered. The variations of matrices between four user accounts represent the quality of proofreading and paraphrasing as visualized in Figure 3.

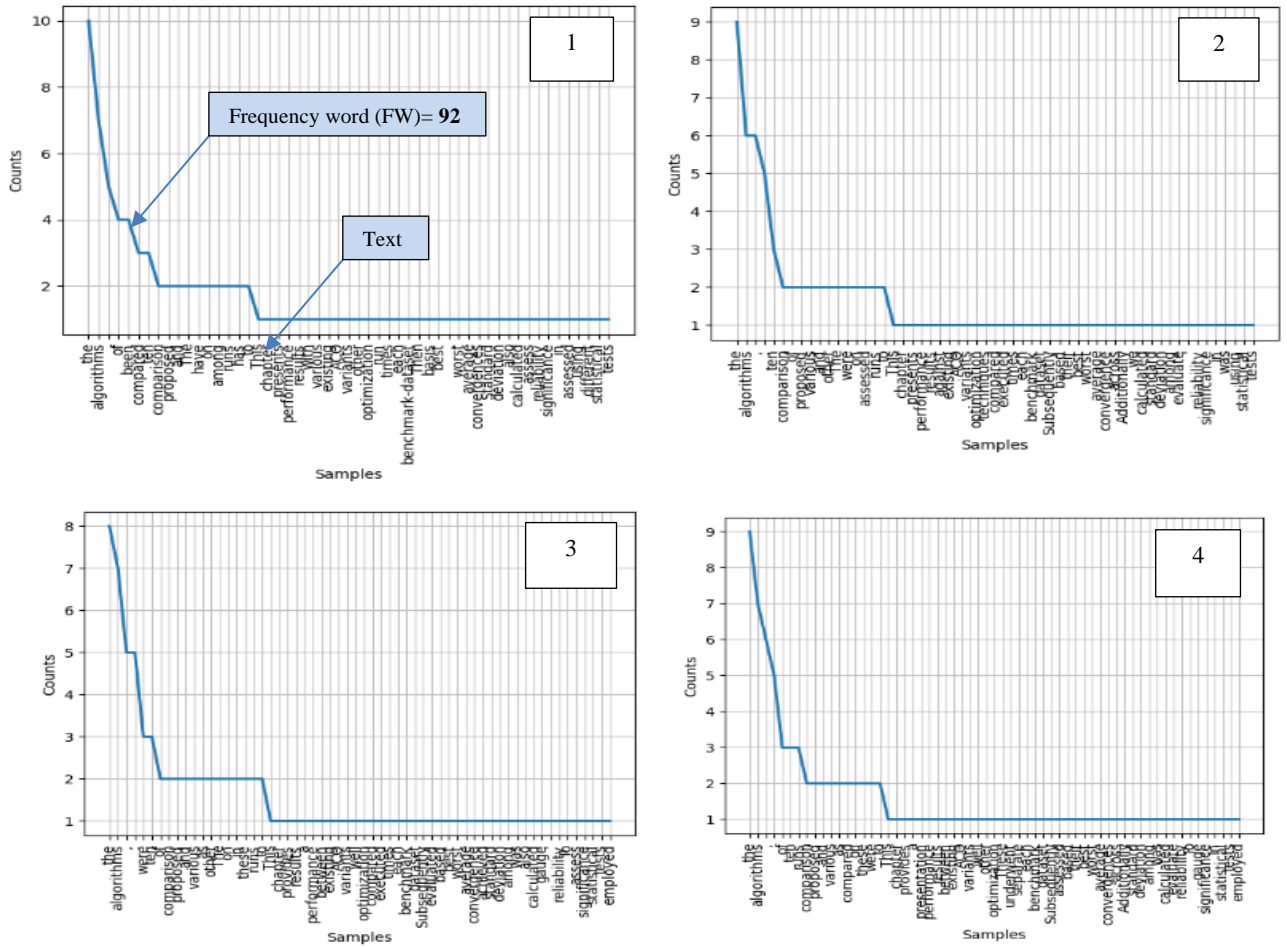


Figure 3. Frequency distribution of words in ChatGPT-3.5 as follows:  
 1) User account 1 in FW=92, 2) User account 2 in FW=93, 3) User account 3 in FW=78  
 4) User account 4 in FW=100.

According to the output paragraph in each user account, the best text is user account 2 because the significant criteria among the others, as well as its total common pronouns, is 0 which refers to this paragraph empty from pronoun elements as displayed in Figure 4.



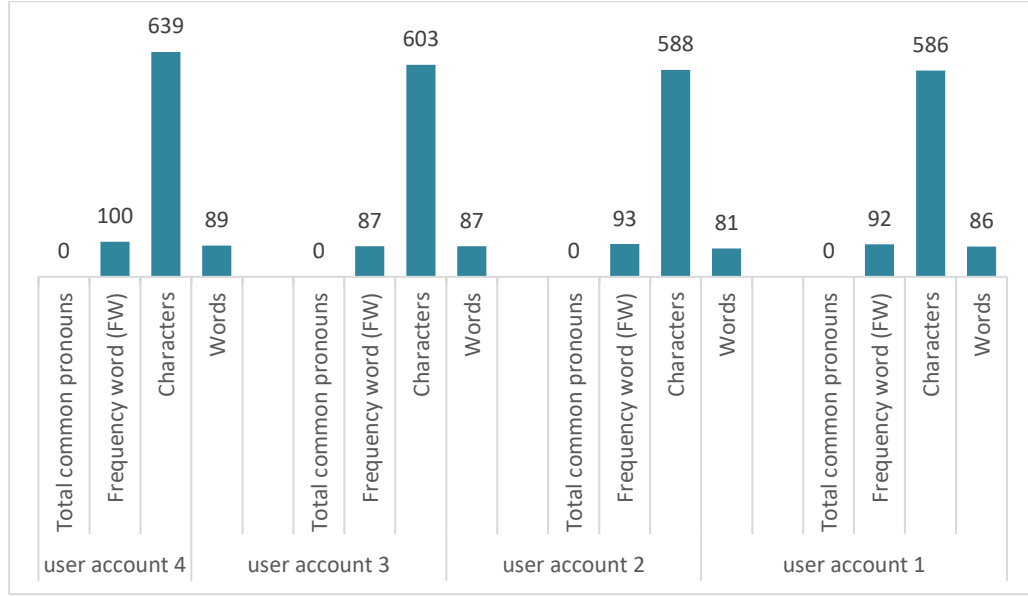


Figure 4. Histogram illustrate the best paragraph between different user account

As regards the second experiment followed the same procedure as the previous test but entering the text in the ChatGPT-4.0, this test suffered from the existing two pronouns in addition to the number of words equal to 92. Also, the number of characters being very high 620 with compared to the paraphrasing and proofreading by ChatGPT-3.5. as clear in Table 3.

Table 2: Responses of Paraphrasing and Proofreading by ChatGPT-4.0 in Terms of Different User Account

User Account	The response	Text Analysis Metrics
1	This chapter <u>provides</u> a comparative analysis of the performance of our proposed algorithms against various existing variants of Ant Colony Optimization (ACO) as well as other optimization algorithms. Each algorithm was executed ten times on each benchmark dataset. Performance metrics such as best, worst, and average convergence rates were evaluated across these ten runs. Additionally, the standard deviation was calculated for each set of ten runs to gauge the reliability of the algorithms. To assess the significance of <u>our</u> proposed algorithms relative to the compared ones, various statistical tests were employed.	Words= <b>91</b> Characters = <b>620</b> Frequency word = <b>102</b> Total common pronouns= <b>2</b> Common pronouns= <b>our, our</b>

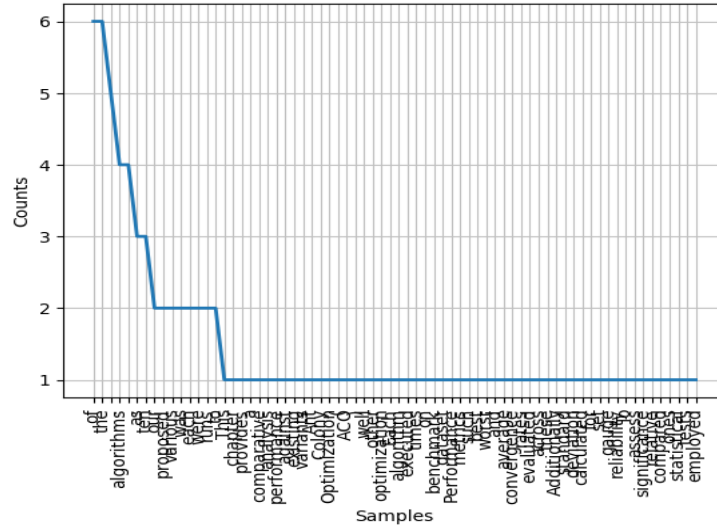


Figure 5. Frequency distribution of words in ChatGPT-4.0 in terms of single user account

Eventually, we are involved in the same paragraph in a single user account with a different new chat in ChatGPT-3.5, as results in Table 4.

Table 4: Responses of Paraphrasing and Proofreading by ChatGPT-3.5 in Terms of Single

New Chat	The Responses	Metrics of Text Analysis
1	"This chapter <u>unveils</u> the performance comparison findings between the proposed algorithms and several existing ACO variants, as well as other optimization algorithms. The compared algorithms underwent ten separate runs on each benchmark dataset. Subsequently, these algorithms were assessed based on the best, worst, and average convergence across the ten runs. To gauge algorithm reliability, <u>we</u> also calculated the standard deviation among the ten runs. Furthermore, <u>we</u> evaluated the significance of the proposed algorithms in comparison to the <u>aforementioned ones</u> through various statistical tests."	Words= <b>83</b> Characters = <b>600</b> Frequency word (FW) = <b>94</b> Total common pronouns= <b>2</b> Common pronouns= <b>we, we</b>
2	"This chapter <u>showcases</u> the results of performance comparisons between the proposed algorithms and various existing ACO variants, as well as other optimization algorithms. These comparative analyses involved running the algorithms ten times on each benchmark dataset. Subsequently, the algorithms were evaluated based on their best, worst, and average convergence across the ten runs. Additionally, <u>we</u> calculated the standard deviation among these ten runs to gauge the algorithms' reliability. To gauge the significance of the proposed algorithms relative to the compared algorithms, different statistical tests were employed."	Words= <b>86</b> Characters = <b>626</b> Frequency word (FW) = <b>98</b> Total common pronouns= <b>1</b> Common pronouns= <b>we</b>
3	"This chapter showcases the results of a performance comparison between the proposed algorithms and various existing ACO variants as well as other optimization algorithms. Each of these algorithms has been executed ten times on individual benchmark datasets. Subsequently, <u>we</u> assess these algorithms by examining their best, worst, and average convergence over the ten runs. Additionally, <u>we</u> calculate the standard deviation among these ten runs to gauge the reliability of the algorithms. Finally, <u>we</u> evaluate the significance of the proposed algorithms relative to the compared algorithms using various statistical tests."	Words= <b>89</b> Characters = <b>622</b> Frequency word (FW) = <b>99</b> Total common pronouns= <b>3</b> Common pronouns= <b>we, we, we</b>
4	"This chapter <u>provides</u> a presentation of performance comparison results for the proposed algorithms <u>when pitted against</u> a range of existing ACO variants and other optimization algorithms. Each of these algorithms <u>underwent</u> ten <u>separate</u> runs on various benchmark datasets. Subsequently, these runs were compared based on their best, worst, and average convergence <u>across</u> the ten <u>iterations</u> . Furthermore, a calculation of the standard deviation among these ten runs was performed to evaluate the algorithms' reliability. To gauge the significance of the proposed algorithms in relation to the compared algorithms, various statistical tests were conducted."	Words= <b>92</b> Characters = <b>652</b> Frequency word (FW) = <b>103</b> Total common pronouns= <b>0</b> Common pronouns= <b>None</b>

Table 4 exhibits the minimum number of words is 83 in the new chat 1, in opposite directions, suffering from two pronouns (we, we), the new chat 2 and new chat 3 both suffer from pronouns. We consider the new chat 4 to be the best paragraph because not have any pronouns based on its Metrics total common pronouns=0 and common pronouns=None. Figure 5 The variation of the Frequency word shows the increasing Frequency word, the words, and characters to avoid using pronouns anymore as shown in Figure 6.

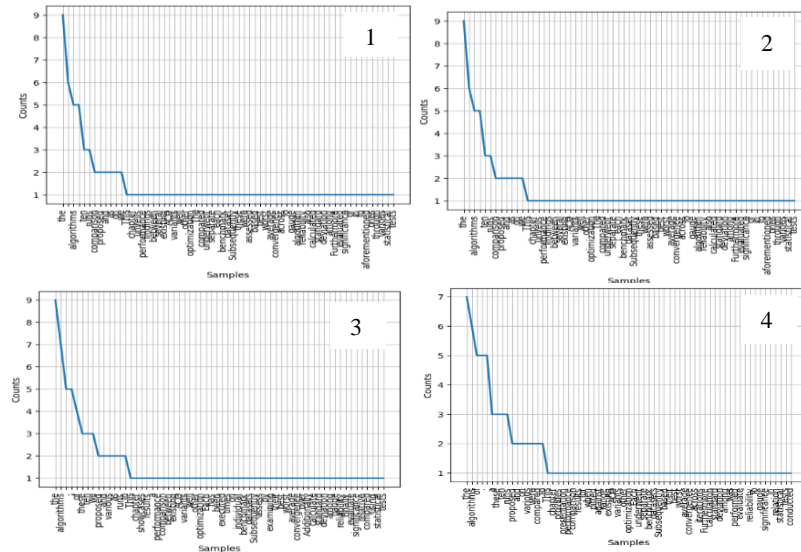


Figure 6. Frequency distribution of words in ChatGPT-3.5 in terms of single user account different new chats as follows: 1) New chat 1 in FW=94 2) New chat 2 in FW=98 3) New chat 3 in FW=99, 4) New chat 4 in FW=103.

#### 4. CONCLUSION

Embedding and improving AI language models, such as ChatGPT, in many disciplines presents issues and concerns for academic institutions responsible for maintaining research quality and ensuring fair evaluations. It is vital to strike a balance between the potential benefits of AI technologies and the maintenance of academic standards to enable students to learn essential skills for a rapidly evolving professional context. We conclude that the close source software (ChatGPT-4.0) doesn't have a big difference in terms of paraphrasing and proofreading, although we used just a single user account, still suffering from the existence of three pronouns in the paragraph. While the free source software (ChatGPT-3.5) but using different user accounts we can create a paragraph without pronouns, this technique is very beneficial for scientific academics when talking about paraphrasing and proofreading a particular paragraph, for example paraphrasing a certain abstract in a specific manuscript, in this paper recommend that any user can adding the word (paraphrasing without pronouns) to get a perfect paragraph.

## REFERENCES

- Alyasiri, O. M., Salisu, S., & Younis, H. A. (2023). ChatGPT Revisited: Using ChatGPT-4 for Finding References and Editing Language in Medical Scientific Articles. Available at SSRN: <https://Ssrn.Com/Abstract=4621581> Or, November. <https://doi.org/http://dx.doi.org/10.2139/ssrn.4621581>
- Fitria, T. N. (2023). Artificial Intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay. *Journal of English Language Teaching*, 6(1), 44–28. <https://journal.unnes.ac.id/sju/index.php/elt/article/view/64069>
- Fraivan, M. et al. (2023). A Review of ChatGPT Applications in Education, Marketing, Software Engineering, and Healthcare: Benefits, Drawbacks, and Research Directions. *ArXiv Preprint ArXiv*, 74(1934), 535–546.
- Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, 15(2). <https://doi.org/10.30935/cedtech/13036>
- Hill-Yardin. (2023). A Chat(GPT) about the future of scientific publishing. *Brain Behav Immun*.
- Koos, S., & Wachsmann, S. (2023). Navigating the Impact of ChatGPT/GPT4 on Legal Academic Examinations: Challenges, Opportunities and Recommendations. 6(2), 255–270. <https://doi.org/10.20473/mi.v6i2.45270>
- Lee, Y. J. (2023). Language learning affordances of Instagram and TikTok. *Innovation in Language Learning and Teaching*, 17(2), 408–423. <https://doi.org/10.1080/17501229.2022.2051517>
- Mohammed, O., Sahib, T. M. (2023). ChatGPT Evaluation: Can It Replace Grammarly and Quillbot Tools? *British Journal of Applied Linguistics*, 3((2)), 34–46. <https://doi.org/doi.org/10.32996/bjal.2023.3.2.4>
- Sahib, T. M., & Ali, A. H. (2023). *ChatGPT in Waste Management : Is it a Profitable*. February, 2021–2023.
- Shahriar et al. (2023). Let’s Have a Chat! A Conversation with ChatGPT: Technology, Applications, and Limitations. *Artificial Intelligence and Applications*, April, 1–16. <https://doi.org/10.47852/bonviewaia3202939>
- Thurnes, C. M. (2023). Tiktok and Instagram – using social media for teaching innovation. *ISPIM Conference Proceedings*, 1–5.
- Zohery, M. (2023). ChatGPT in Academic Writing and Publishing: A Comprehensive Guide. *Artificial Intelligence in Academia, Research and Science: ChatGPT as a Case Study*. (Frist Edition), Achatago Publishing., April. <https://doi.org/10.5281/zenodo.7803703>