# Comparative Analysis of Linguistic Features in Computer-Generated and Original Reviews

Himakar Paradi
Nl95695@umbc.edu
Department of Computer Science
University of Maryland, Baltimore County

Ravi Varma Gottumukkala
Ha27498@umbc.edu
Department of Computer Science
University of Maryland, Baltimore County

Rohith Kankipati
Rohithk1@umbc.edu
Department of Computer Science
University of Maryland, Baltimore County

Gadu Dinesh Kumar Naidu
Dineshg1@umbc.edu
Department of Computer Science
University of Maryland, Baltimore County

## Introduction

In today's digital age, internet reviews influence consumer selections in a variety of businesses, including retail, hospitality, and eating. However, the growing use of computer-generated evaluations presents a serious challenge. These evaluations, which are frequently generated using complex algorithms such as GPT-2, can be difficult to discern from genuine ones submitted by actual consumers. This can lead to consumers being deceived by fake comments, resulting in bad purchasing decisions and a loss of trust in online review sites.

Identifying the differences between computer-generated and original evaluations is critical to the integrity of online feedback systems. By distinguishing between these two sorts of evaluations, we can assist consumers make more informed judgments while still maintaining the credibility of review sites. This study seeks to identify distinguishing linguistic and literary patterns that can help people identify computer-generated evaluations without the need for advanced technical understanding. Our ultimate goal is to develop a user-friendly guide that highlights these crucial qualities, allowing the typical individual to identify inauthentic reviews.

This solution will benefit many participants, including consumers who depend on honest feedback to make purchasing decisions, businesses who rely on genuine evaluations to grow their reputation, and online platforms that seek to retain user confidence and transparency. Our study will help to create a more trustworthy digital marketplace by improving the ability to detect computer-generated evaluations, allowing customers to make educated decisions with confidence.

## Related Work

Detecting fake reviews has been a significant area of research, with numerous studies exploring various methods to identify deceptive content. One of the foundational works in this domain was conducted by Ott et al. (2011), who investigated the linguistic features of deceptive opinion spam. Their research demonstrated that deceptive reviews often have distinct linguistic patterns compared to genuine reviews, providing a basis for further exploration in this field.

Mukherjee et al. (2013) expanded on this by examining behavioural footprints of opinion spammers. Their study utilized metadata and user behaviour to spot spammers, combining these with textual analysis to improve detection accuracy. This work highlighted the importance of considering multiple dimensions, including user activity patterns, in identifying fake reviews. Further extending this line of inquiry, Mukherjee (2015) integrated linguistic, behavioural, and statistical modelling to enhance the detection process, showing that a multi-faceted approach could significantly improve detection performance.

Dematis et al. (2018) explored the exploitation of spam indicators and reviewer behaviour characteristics, focusing on specific indicators that signal spamming activities. Their research identified key behavioural traits of spammers, such as review frequency and consistency, and combined these with textual features to detect fake reviews. Similarly, Hájek and Sahut (2022) mined behavioural and sentiment-dependent linguistic patterns from restaurant reviews, finding that sentiment analysis combined with linguistic features could effectively distinguish between genuine and deceptive reviews.

In the context of recent advancements, F. Abri et al. (2020) provided a comprehensive analysis of linguistic features for detecting fake reviews. Their study utilized advanced machine learning techniques to analyse various linguistic attributes, showing that certain textual features, such as lexical diversity and syntactic patterns, are strong indicators of deception. Wang et al. (2020) proposed a method based on multiple feature fusion and rolling collaborative training, combining linguistic, behavioural, and user interaction features to detect fake reviews more accurately. Their work underscored the effectiveness of integrating multiple types of data to enhance detection algorithms.

While significant progress has been made in detecting fake reviews, many studies primarily focus on distinguishing between human-written fake reviews and genuine reviews. Our project, however, addresses the unique challenge of differentiating between computer-generated (CG) reviews and original (OR) reviews. This involves analysing specific linguistic and textual patterns that are characteristic of AI-generated content. Our approach not only builds on the foundational work by incorporating linguistic and behavioural features but also introduces a manual guide aimed at helping typical internet users identify CG reviews without needing advanced technical knowledge. This practical application fills a gap in the current research, providing a valuable tool for everyday users to navigate the growing landscape of online reviews.

## Research Questions

1. What common characteristics distinguish computer generated reviews from genuine ones?

2. How can individuals use these features to determine whether online reviews are genuine?

## Approach

We used a structured approach to identify and distinguish between computer-generated (CG) and original evaluations. The first step was to identify the main features that distinguish CG evaluations from original OR reviews. We concentrated on linguistic and textual characteristics such as word count, sentence length, lexical diversity, readability, spelling and grammar errors, punctuation, and contextual importance.

We then collected a dataset that included both CG and OR reviews. Data collection was followed by pre-processing activities, such as cleaning the data to remove any unnecessary information and normalizing the text for consistency. This pre-processing was important in making sure the resulting analysis was correct and useful.

We obtained specific features based on the noticed attributes. These characteristics were then utilized to generate statistical measures like scores and ratios, which offered a quantitative foundation for identifying between CG and OR reviews. The next step was to organize these metrics into an easy-to-understand report. This paper outlined likely indicators for false reviews, allowing average internet users to identify CG reviews without significant technological understanding.

To further validate our findings, we conducted a survey to evaluate the effectiveness of our report. Participants were asked to classify a set of reviews as CG or OR before and after reading the report. The accuracy of their classifications was measured to determine the improvement in their ability to identify CG reviews. This step was essential in demonstrating the practical applicability of our research and its potential to aid everyday internet users in distinguishing between genuine and computer-generated reviews.


*Figure 1: Step-by-Step approach followed in the project*

## Methodology

*Dataset*

For our study, we utilized the OSF Reviews Dataset, which was presented in the paper by Salminen et al. (2022). This dataset includes a comprehensive collection of reviews, both original and computer-generated, gathered to examine and detect fake reviews of online products. The dataset consists of reviews for various products across different categories, providing a rich source of data for analysis. The authors generated this dataset by collecting genuine reviews from online platforms and creating fake reviews using advanced text generation models like GPT-2. This mix of genuine and artificial reviews allows for a robust analysis of distinguishing characteristics.

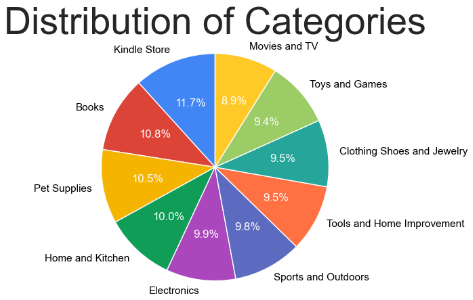


*Figure 2: Distribution of Dataset across various products*

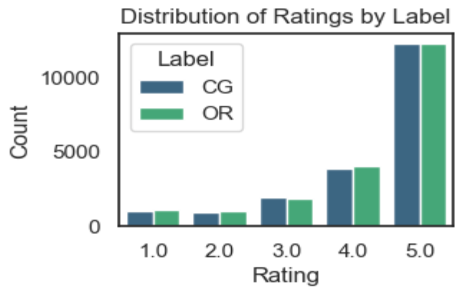


*Figure 3: Distribution of Ratings in dataset by label*

*Data Pre-processing*

The pre-processing consisted of several critical steps: cleaning the data to remove extraneous information such as HTML tags and special characters, normalizing the text by converting all characters to lowercase and standardizing punctuation, and tokenizing the text into individual words and sentences. We also used lemmatization to reduce terms to their basic forms, which improved our analysis of lexical diversity and readability. In addition, we removed common stop-words to provide room for more relevant vocabulary. These pre-processing methods ensured that the data was clean, consistent, and suitable for correct feature extraction and analysis, resulting in higher overall quality and reliability of our findings.

*Feature Extraction*

Feature extraction is a critical step in distinguishing between computer-generated (CG) and original reviews (OR). Each feature provides unique insights into the characteristics of the text, helping to identify patterns that can differentiate CG reviews from OR reviews.

- **Word Count:** The total number of words in a review is a basic yet significant feature. It provides a general sense of the review's verbosity.
- **Average Word Length:** This feature measures the average length of words in a review, which can indicate the complexity of the language used. We computed this by dividing the total number of characters by the total number of words.
- **Sentence Count:** The number of sentences in a review helps in understanding the review's structure and coherence. This was calculated using sentence tokenization methods.
- **Sentence Length:** This feature examines the average number of words per sentence. It can reveal the complexity and readability of the text. We calculated this by dividing the total word count by the sentence count (Shah et al., 2023).
- **Rating Prevalence:** It involves analyzing the relationship between the textual content of a review and the rating score provided by the reviewer. This feature helps in understanding if the sentiment in the text aligns with the given rating. We used nlptown/bert-base-multilingual-uncased-sentiment model, a pre-trained sentiment analysis model available on Hugging Face (https://huggingface.co/nlptown/bert-base-multilingual-uncased-sentiment ).
- **Lexical Diversity:** Lexical diversity, measured as the ratio of unique words to the total number of words, indicates the richness of the vocabulary used in the review. We used NLP techniques to calculate this ratio (Chaabouni et al., 2020).
- **Readability Score:** Readability is important in identifying weather the review is easy or difficult to read. We calculated this score using readability methods such as Flesch Reading Ease and Flesch-Kincaid Grade Level. These formulas are widely used and validated in text analysis research (Kincaid et al., 1975).
- **Spelling & Grammar Mistakes:** The number of spelling and grammatical errors can be a strong indicator of review authenticity. We used a custom dictionary (python spellchecker library)and rule-based system (python language_tool_python library) to identify these mistakes (Mewada & Dewang, 2022).

- **Punctuation Usage:** The frequency and variety of punctuation marks can differentiate CG from OR reviews. We counted the occurrences of different punctuation marks. (Shah et al., 2023).
- **Specificity of Content:** Named entities and specific details, such as names, dates, and locations present in a review. We used SpaCy's Named Entity Recognition (NER) to identify and count these entities (Honnibal & Montani, 2017). This feature is important since it indicated the presence of specific and context-relevant information. (Chaabouni et al., 2020).
- **N-grams:** Frequent sequences of N words (N-grams) can reveal common phrases used in reviews. We analyzed N-grams using NLP techniques to identify these patterns (Shah et al., 2023).
- **Contextual Relevance:** This feature assesses how relevant the review content is to the product category. We used cosine similarity of BERT embeddings to measure this relevance (Devlin et al., 2019). Higher relevance scores indicate a closer match between review content and the product context (Bauman & Tuzhilin, 2014).
- **Pronoun and Modal Verb Usage:** The frequency of personal pronouns and modal verbs can indicate the personal and subjective nature of reviews. We computed this feature by counting occurrences of these words (McMahon et al., 2019).
- **Sentiment and Tone:** Sentiment polarity and subjectivity provide insights into the emotional tone of the reviews. We used the TextBlob library to perform sentiment analysis, which is based on established NLP methods (Loria, 2018).

*Report Generation*

After collecting and evaluating the features, we combined the results into an understandable report. This paper identified potential signs of computer-generated (CG) evaluations using linguistic and literary properties such as word count, sentence length, lexical variety, readability, and contextual significance.

*Evaluation by Survey*

To evaluate the effectiveness of our report, we conducted an evaluation survey involving students from the University of Maryland, Baltimore County (UMBC). The survey was created using Google Forms and included a set of 10 reviews, comprising 5 original (OR) and 5 computer-generated (CG) reviews. Participants were initially asked to classify each review as either CG or OR. Following this initial assessment, we presented the participants with our compiled report, which explained the key features and indicators of CG reviews. After reviewing the report, participants were asked to retake the survey to classify the same set of reviews again. This before-and-after comparison allowed us to measure the improvement in their ability to accurately identify CG reviews.

## Results

These are the findings of our analysis for each feature. The results are presented with corresponding graphs, charts, tables, and visual representations.

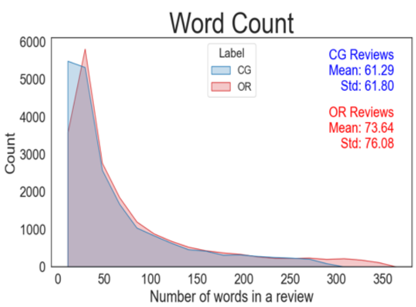


*Figure 4: Original reviews (OR) generally have a higher word count compared to computer-generated reviews (CG). This difference indicates that OR reviews tend to be more verbose and detailed.*

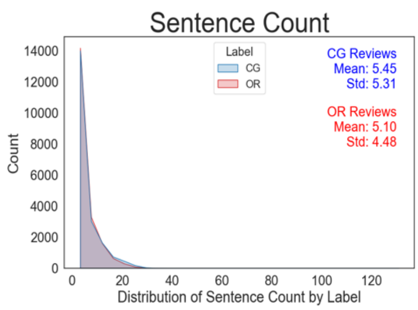


*Figure 5: CG reviews contain more sentences, but they are often shorter and less complex compared to OR reviews. This pattern reflects a more fragmented writing style in CG reviews.*

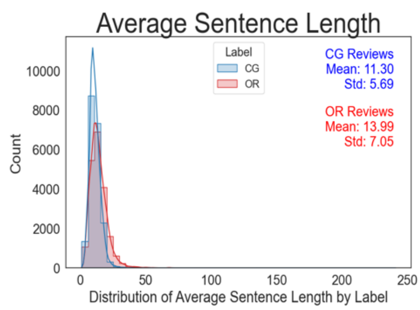


*Figure 6: The average sentence length is longer in OR reviews, indicating a more detailed and coherent structure.*

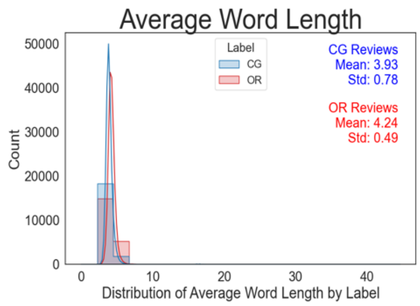


Figure 7: OR reviews use longer words on average than CG reviews. This suggests that OR reviews may use more complex vocabulary.

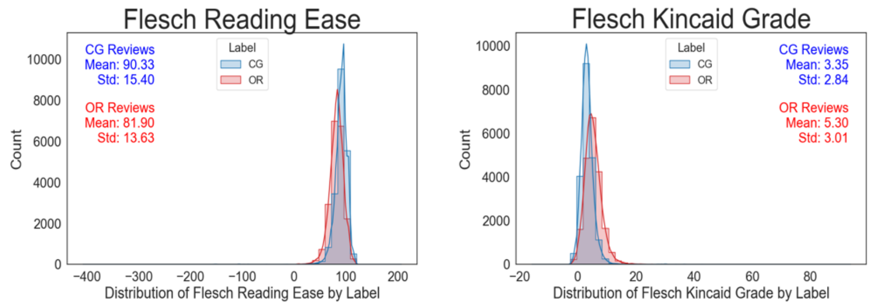


Figure 8: CG reviews score higher on readability tests, such as the Flesch Reading Ease and Flesch-Kincaid Grade Level, indicating they are easier to read but potentially less authentic.

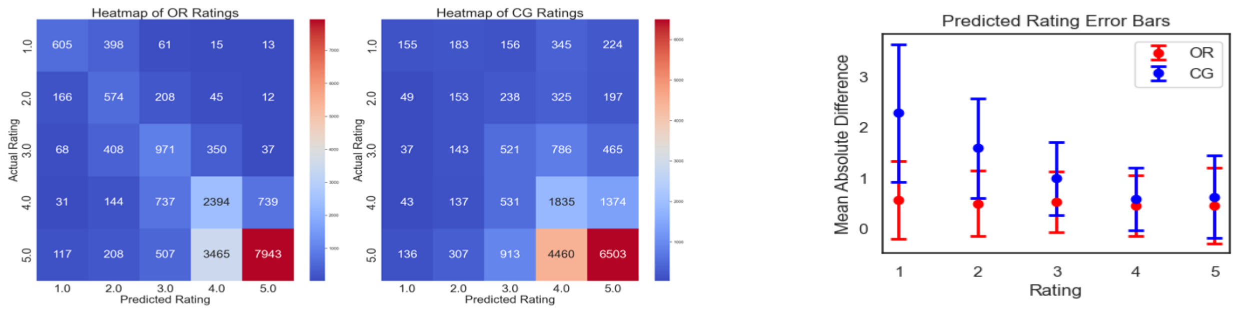


Figure 9: OR reviews show a stronger correlation between the review content and the given rating, while CG reviews tend to have more neutral or positive ratings regardless of the content.

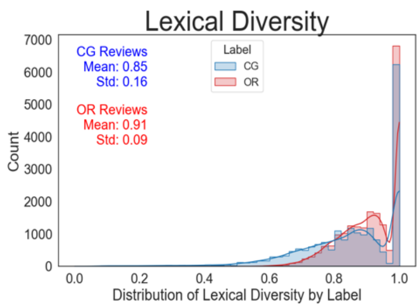


Figure 10: OR reviews exhibit higher lexical diversity, meaning they use a wider range of unique words compared to CG reviews.

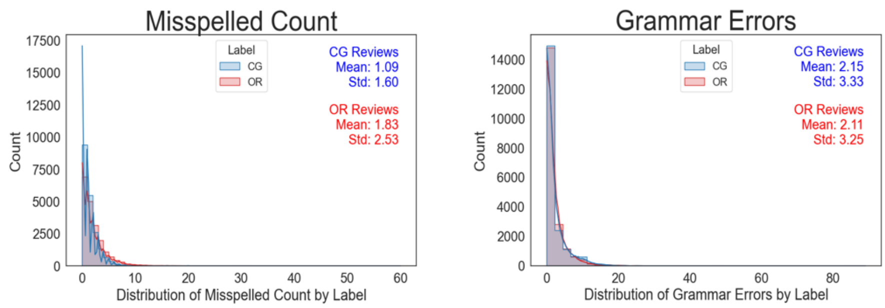


Figure 11: OR reviews tend to have more spelling and grammatical errors, which may be a sign of genuine human writing as opposed to algorithmically generated text.

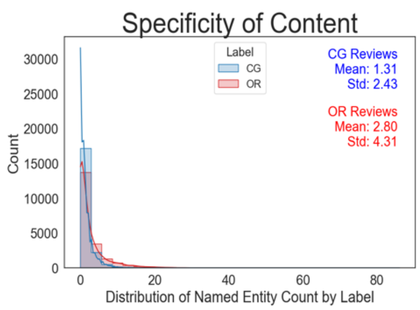


Figure 12: OR reviews include more specific details and named entities, such as names, dates, and locations, which contribute to their perceived authenticity.

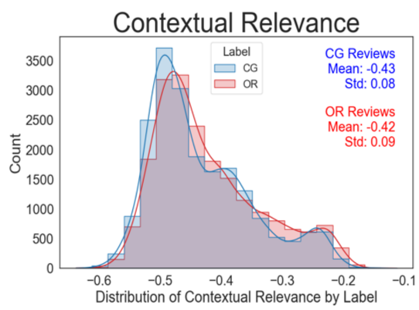


Figure 13: Both OR and CG reviews have similar contextual relevance scores, indicating that CG reviews are often contextually appropriate but may still lack other authentic markers.

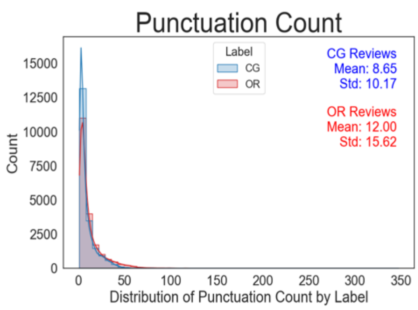


Figure 14: OR reviews use a wider variety of punctuation marks, suggesting a more nuanced and expressive writing style.

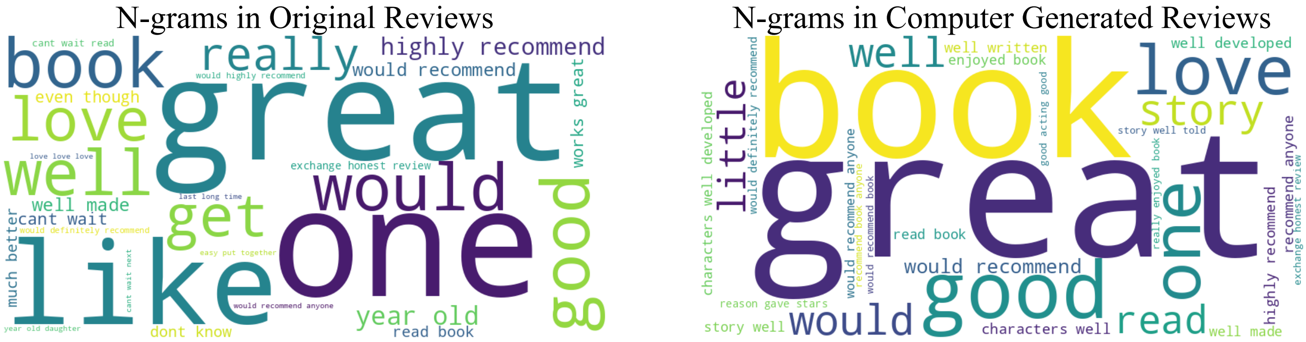


*Figure 15: Frequent N-grams (sequences of words) show different patterns in OR and CG reviews, with OR reviews using more varied and context-specific phrases.*

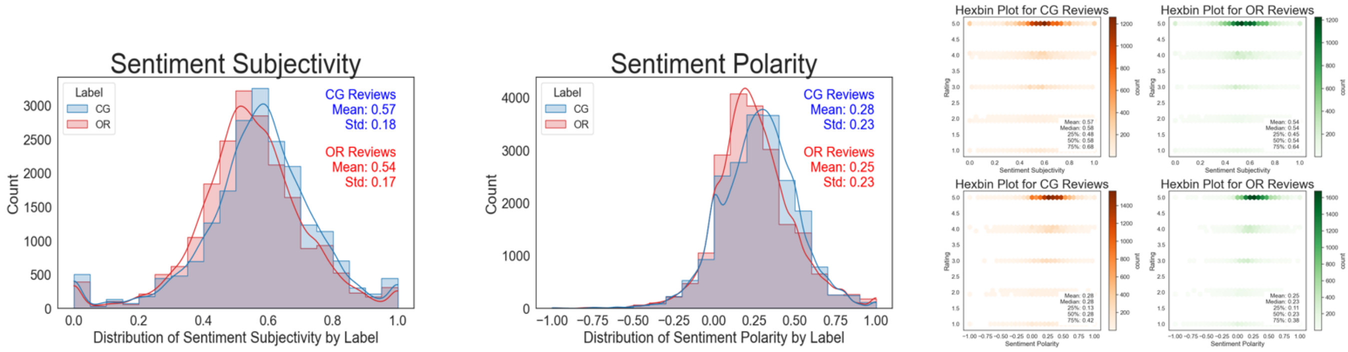


*Figure 16: Sentiment analysis shows similar polarity and subjectivity in both OR and CG reviews. Figures from left to right: Distribution of Sentiment Subjectivity across both labels, Distribution of Sentiment Polarity across both labels, Hexbin Plot comparing Sentiment Polarity and Subjectivity with Rating.*

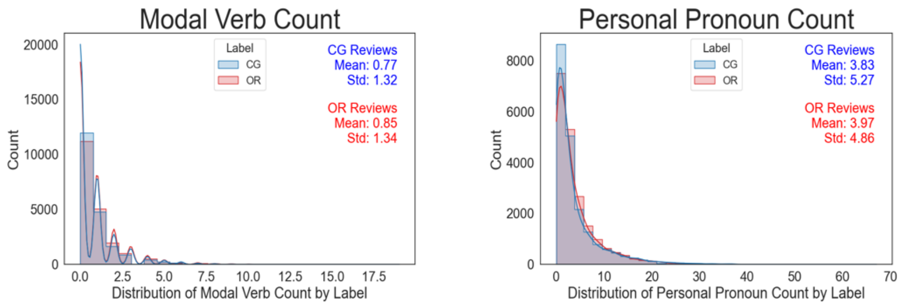


*Figure 17: OR reviews use more personal pronouns and modal verbs, reflecting a more personal and subjective writing style.*

*Summary Report Table*

The final report compiled our findings into an accessible format for typical internet users, highlighting key indicators of fake reviews and providing practical guidance on identifying CG reviews.

| Feature | CG Reviews | Original Reviews |
|---|---|---|
| **Word Count** | Lower | Higher |
| **Avg. Word Length** | Shorter | Longer |
| **Sentence Count** | More | Less |
| **Sentence Length** | Shorter | Longer |
| **Rating Prevalence** | Tends to be neutral | Almost none |
| **Lexical Diversity** | Similar | Higher |
| **Readability Score** | Lower | Higher |
| **Spelling & Grammar Mistakes** | Fewer | More |
| **Punctuation Usage** | Less | More |
| **Specificity of Content** | Less | More |
| **Contextual Relevance** | Equal | Equal |
| **Pronoun and Modal Verb Usage** | Less | Slightly greater |
| **Sentiment and Tone** | Equal | Equal |

| | | |
|---|---|---|
| **N-grams** | Great, love, well, enjoyed book, would recommend, highly recommend and other | One, great, good, would, read book, much better, can't wait and other. |

*Table 1: Final Report for distinguishing Computer Generated Reviews with Original Reviews*

*Survey Results*

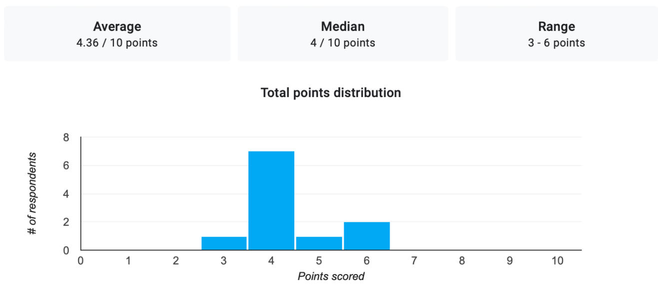


*Figure 18: Results after Initial Survey*

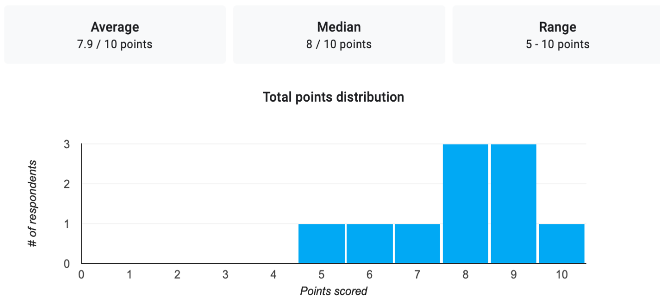


*Figure 19: Results after Final Survey*

By conducting an initial survey from 10 students of UMBC, we got an average score of 43.60% average with a median score of 4/10, and we have seen a significant improvement in accuracy to 79% with median score of 8/10 after providing the candidates with our report.

## Takeaway Message

*What We Learned*

We discovered that detailed analysis of linguistic and textual features can effectively distinguish computer-generated reviews from original ones. Key features like word count, sentence length, lexical diversity, and content specificity were particularly useful. Our surveys also showed that clear guidelines significantly improve users' ability to spot fake reviews, emphasizing the value of practical, user-friendly tools in enhancing online review authenticity.

*Mistakes We Made*

Initially, we overlooked the importance of thorough data pre-processing, leading to inconsistencies that required additional time to fix. We also struggled with optimizing computational resources for handling large datasets..

*What Could Have Been Done Better*

Our analysis could have been improved by identifying and incorporating more features, such as user behavior and metadata. Including these additional dimensions would likely enhance the robustness and accuracy of our detection methods.

*Next Steps*

The next step involves implementing machine learning algorithms to determine the weightage of each feature in classifying reviews. This approach will help quantify feature importance and create a more automated, scalable solution. Additionally, conducting more extensive user testing will further validate and refine our guide.8. Design Implications

## Design Implications

Based on our findings, several design implications can be proposed to improve the detection of computer-generated reviews and enhance the trustworthiness of online review platforms. One potential solution is the development of a browser extension or an integrated feature within review platforms that uses our identified linguistic and textual features to flag suspicious reviews. This tool would analyse new reviews in real-time, highlighting those that exhibit patterns consistent with computer-generated content, such as lower lexical diversity, higher readability scores, or fewer specific details. Additionally, the tool could provide users with a confidence score indicating the likelihood that a review is computer-generated.

To counteract future advancements in computer-generated text that may exploit our findings, continuous updating of the detection algorithm will be necessary. This includes incorporating more sophisticated linguistic features, such as advanced syntactic structures and sentiment shifts. By maintaining a dynamic and adaptable system, we can stay ahead of evolving text generation models and ensure the robustness of our detection mechanisms.

## Main Contributions of Your Project

Our project makes several significant contributions to the field of fake review detection. First, we identified and validated a set of linguistic and textual features that can effectively distinguish between computer-generated and original reviews. These features include word count, sentence length, lexical diversity, and the specificity of content. Second, we developed a practical and user-friendly guide based on these features to help typical internet users identify inauthentic reviews without needing advanced technical knowledge. Third, our evaluation through user surveys demonstrated the effectiveness of this guide in improving users' ability to distinguish between genuine and computer-generated reviews. Lastly, our research addresses a gap in the current literature by focusing specifically on the detection of AI-generated reviews, which is an emerging challenge in the digital age.

## Ethical Considerations

Our research brings up some important ethical questions. First, we need to make sure that our system doesn't wrongly target real reviews by actual users who might use certain words that seem suspicious. To prevent this, our system should be clear and allow users to challenge any flagged reviews. We also need to think about privacy when analyzing user content. Our method should follow data protection rules, making sure user data is kept anonymous and stored safely. Lastly, because our tool could affect how consumers see businesses, it's essential to make sure our system is accurate and fair to avoid unfair outcomes for both users and businesses.

## Limitations

While our project is good at telling apart computer-generated reviews from real ones, it has some weaknesses. One problem is that we rely on certain language features that might change as text generation technology gets better. This means we need to keep updating our detection methods. Another issue is that sometimes real reviews might be wrongly flagged as fake, which can make people lose trust in our system and the review platforms. Also, our study mainly looks at English reviews, so we don't know how well our method works in other languages. Finally, we tested our approach on a small group of people, so we need to do more testing with a wider range of participants to be sure our findings apply to everyone.

### Conclusion

In conclusion, our study tackles a big problem in the digital world: telling apart fake computer-generated reviews from real ones. By spotting key language and text features, we created a guide that helps regular internet users identify fake reviews. Our research not only adds to the academic knowledge of detecting fake reviews but also provides real benefits for consumers, businesses, and online platforms. Although there are some limitations, our project sets a solid base for future research and better detection tools. It's crucial to keep updating our methods to keep up with advances in text generation technology and to make sure our findings are used ethically and effectively in the real world.

By addressing fake reviews and proposing adaptable solutions, our research helps build a more trustworthy and reliable digital marketplace.

**References**


1. F. Abri, L. F. Gutiérrez, A. S. Namin, K. S. Jones, and D. R. W. Sears, "Linguistic Features for Detecting Fake Reviews," in 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA), Miami, FL, USA, 2020, pp. 352-359. doi: 10.1109/ICMLA51294.2020.00063.
2. J. Wang, H. Kan, F. Meng, Q. Mu, G. Shi, and X. Xiao, "Fake Review Detection Based on Multiple Feature Fusion and Rolling Collaborative Training," IEEE Access, vol. 8, pp. 182625-182639, 2020. doi: 10.1109/ACCESS.2020.3028588.
3. Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011). Finding deceptive opinion spam by any stretch of the imagination. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (pp. 309-319).
4. Mukherjee, A., Kumar, A., Liu, B., Wang, J., Hsu, M., Castellanos, M., & Ghosh, R. (2013). Spotting opinion spammers using behavioral footprints. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 632-640).
5. Mukherjee, Arjun. (2015). Detecting Deceptive Opinion Spam using Linguistics, Behavioral and Statistical Modeling. In Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 21-22). doi: 10.3115/v1/P15-5007.
6. Dematis, I., Karapistoli, E., & Vakali, A. (2018). Fake Review Detection via Exploitation of Spam Indicators and Reviewer Behavior Characteristics. In Tjoa, A., Bellatreche, L., Biffl, S., van Leeuwen, J., & Wiedermann, J. (Eds.), SOFSEM 2018: Theory and Practice of Computer Science, Lecture Notes in Computer Science, vol 10706. Springer, Cham. doi: 10.1007/978-3-319-73117-9_41.
7. Hájek, P., & Sahut, J.-M. (2022). Mining behavioural and sentiment-dependent linguistic patterns from restaurant reviews for fake review detection. Technological Forecasting and Social Change, 177, 121532. doi: 10.1016/j.techfore.2022.121532.
8. Mewada, A., & Dewang, R. (2021). Research on False Review Detection Methods: A state-of-the-art review. Journal of King Saud University - Computer and Information Sciences, 34. doi: 10.1016/j.jksuci.2021.07.021.
9. Salminen, J., Kandpal, C., Kamel, A. M., Jung, S., & Jansen, B. J. (2022). Creating and detecting fake reviews of online products. Journal of Retailing and Consumer Services, 64, 102771. https://doi.org/10.1016/j.jretconser.2021.102771.
10. Shah, A., Ranka, P., Dedhia, U., & Prasad, S. (2023). Detecting and Unmasking AI-Generated Texts through Explainable Artificial Intelligence using Stylistic Features. Journal of Advanced Research in Computer Science.
11. Chaabouni, R., Kharitonov, E., Bouchacourt, D., Dupoux, E., & Baroni, M. (2020). Compositionality and Generalization In Emergent Languages. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 4427–4442). Association for Computational Linguistics.
12. Mewada, A., & Dewang, R. K. (2022). Research on False Review Detection Methods: A state-of-the-art review. Journal of King Saud University-Computer and Information Sciences, 34(9), 7530-7546.
13. Bauman, K., & Tuzhilin, A. (2014, October). Discovering Contextual Information from User Reviews for Recommendation Purposes. In CBREcsys@ recsys (pp. 2-9).
14. McMahon, N., Dekan, G., & Amthauer, H. (2019). Analyzing Pronoun Usage in Yelp Reviews.
15. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.
16. Honnibal, M., & Montani, I. (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. To appear.
17. Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel. Naval Technical Training Command Millington TN Research Branch.
18. Loria, S. (2018). TextBlob: Simplified Text Processing. TextBlob Documentation.