Is GenAI Male ot Femaile? Gender Prediction of Generated Tweets using GenAI

Jalal S Alowibdi

Department of Computer Science and Artificial Intelligence, College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia

Abstract
With the use of Generative AI (GenAI), Online Social Networks (OSNs) now generate a huge volume of content data. Yet, user-generated content in OSNs, aided by GenAI, presents challenges for analyzing and understanding its characteristics. In particular, tweets generated by GenAI on request by authentic human users present difficulties in determining the gendered variation of the content. The vast amount of data generated from tweets' content warrants investigation into the gender-specific language used in these tweets. This study explores the task of predicting the gender of text content in tweets generated by GenAI. Through our analysis and experimentation, we have achieved a remarkable 90% accuracy in attributing gender-specific language to these tweets.
Keywords:
Generative AI; Artificial Intelligence; Linguistic Patterns; Text Classification; GenAI-generated; Human authored; Gender-Specific

1. Introduction

With the majority of people using Online Social Networks (OSNs), these platforms are overwhelmed with a massive volume of text content teeming with diverse perspectives, opinions, and sentiments. Twitter, now rebranded as X, has over 350 millions and 100 millions active users monthly and daily respectively resulting in 100 millions tweets and billions of words daily (Alowibdi et. al. 2024). Therefore, understanding and analyzing the characteristics of this content, especially those generated with the help of Generative AI (GenAI), presents a significant challenge, specifically, tweets generated by GenAI on request by authentic human users present difficulties in determining the gendered variation of the content. Predicting gender from content is significantly different from face recognition for gender, even though both use similar classification techniques. While face recognition relies on visual cues and patterns that are often distinct and easily detectable, predicting gender from text content involves analyzing linguistic and stylistic characteristics that are much more intricate. The challenge in text-based gender prediction lies in the variability and complexity of language, where individual writing styles can vary widely regardless of gender (Alowibdi et. al. 2013; alowibdi et. al 2013). This makes it harder to achieve accurate predictions compared to face detection, where the visual features are more consistent and easier to classify. Our research explores the interesting task of predicting the gender of text content in tweets generated by GenAI.

In addition, GenAI applications have advanced significantly, replicating human language and cognitive patterns with increasing sophistication (OpenAI 2024). GenAI has revolutionized content creation, offering a tantalizing glimpse into algorithmically generated text that mimics human-like language patterns. While GenAI algorithms demonstrate remarkable proficiency, human-authored content emanates from the depths of individual thought processes, reflecting the complexities of human cognition and emotion (Ali et. al. 2024; Garcia et. al. 2023; Kumar et. al. 2024; Gu et. al. 2024; Yan et. al. 2024; Sun et. al. 2022). This progress presents a technological and societal challenge in accurately predicting gender-specific language in tweets generated by these algorithms. We investigate the small language details, contextual clues, and delicate hints that reveal the gender differences embedded in a tweet. By examining sentence structure, semantic coherence, and other linguistic features, we aim to predict gender-specific tweets generated by GenAI on request by authentic human users. Understanding the gender-specific language in GenAI-generated content holds profound implications for content creation on OSNs. Our fundamental research questions are: What is the impact of using GenAI to produce gender-specific tweets compared to tweets authored by gender-specific humans? What linguistic features, such as syntax, vocabulary, and grammar, are most indicative of gender-specific language in tweets generated by GenAI? How accurately can machine learning models predict the gender of GenAI-generated tweets compared to human-authored tweets? What role do sentiment and emotional expression play in distinguishing gender-specific language in GenAI-generated tweets from that in human-authored tweets? How does the use of hashtags differ between male and female language in GenAI-generated tweets, and what impact does this have on tweet engagement and authenticity? What are the patterns of user interaction with gender-specific language in GenAI-generated tweets compared to those authored by humans? These questions guide our exploration, driving us to uncover the underlying dynamics of tweet generation and consumption in OSNs.

The importance of our work extends beyond academic research to include practical benefits for a variety

Manuscript received June 5, 2025
Manuscript revised June 20, 2025
https://doi.org/10.22937/IJCSNS.2025.25.6.8

of stakeholders. From OSNs platforms grappling with issues of content moderation to marketers tailoring their strategies, businesses targeting specific demographics, policymakers regulating AI usage, and users seeking authentic sources of information, the ability to predict gender-specific language in GenAI-generated tweets holds immense value. By providing insights into the distinctive characteristics of gendered language from each source, our research aims to empower individuals and organizations to make informed decisions in increasingly complex OSNs. Our contributions are manifold:

- We collected a dataset containing gender-specific GenAI-generated tweets, from users using ChatGPT, and human-authored tweets labeled gender (OpenAI 2024).
- We presented a novel approach and methodology for collecting dataset tagged with hashtags, utilizing a temporal approach to capture trending hashtags over different time periods. This ensures a balanced and representative sample of tweets.
- We employed a two-stage feature selection method to identify the most discriminative features for gender prediction. This involved analyzing term frequencies and applying the Chi-Square test to select features with high discriminative scores that significantly contribute to distinguishing gender-specific language in tweets.
- Through extensive experimentation with various Machine Learning (ML) classifiers, including Support Vector Machine (SVM), Naive Bayes (NB), Decision Tree (DT), Random Forest (RF), and Multi-Layer Perceptron (MLP), we validated the efficacy of our method. Our results demonstrate that we can accurately predict the gender of text content in tweets generated by GenAI.

Indeed, this article is outlined as follows: Section I provides the related works. Then, Section II explores the proposed work. Also, Section III introduces the experimental results of our work and lists the outcomes. Finally, we highlight the conclusion and the future work.

2. Related Works

Gender prediction on OSNs has been a significant area of research due to its applications in targeted marketing, personalized recommendations, and social studies. Early work in this domain focused on utilizing profile information and textual content to predict the gender of users. Peersman et al. explored methods for predicting age and gender in online social networks by analyzing user profiles and social media comments (Peersman et al. 2011). Their study highlighted the importance of linguistic and behavioral features in gender prediction tasks.

Also, Merler et al. extended this work by incorporating semantic analysis of social media images to predict gender (Merler et al. 2015). They found that combining visual and textual features significantly improved the accuracy of gender prediction models on social media platforms like Twitter. Another notable study by Çelik and Aslan utilized artificial intelligence to predict gender from social media comments, emphasizing the role of natural language processing techniques in enhancing prediction accuracy (Çelik and Aslan 2019).

In addition, Reddy et al. presented an N-gram approach for gender prediction, demonstrating how specific linguistic patterns can be used to distinguish between male and female language in social media content (Reddy et al. 2017). Similarly, Krüger and Hermann investigated the state-of-the-art in gender identification from texts, evaluating various online services and their effectiveness in gender prediction (Krüger and Hermann 2019). Bamman et al. examined gender identity and lexical variation in social media, highlighting how gender influences language use and communication styles (Bamman et al. 2014).

With the advancement of Gen-AI technologies like ChatGPT, the generation of text that depicts human language has become increasingly prevalent (OpenAI 2024). This raises questions about the ability of AI to replicate gender-specific language characteristics. The work by (OpenAI 2024) demonstrated the capability of ChatGPT to produce coherent and contextually relevant text, yet it also highlighted the challenges in ensuring the generated content accurately reflects gender-specific subtleties. Research by Gu discussed the ethical considerations and responsibilities involved in generative AI, particularly concerning the generation of biased or stereotypical content (Gu 2024). This work emphasizes the need for careful design and monitoring of AI systems to avoid perpetuating gender biases in generated content. García-Peñalvo and Vázquez-Ingelmo provided a comprehensive overview of the evolution and trends in generative AI, underscoring the significance of addressing biases and ensuring the ethical deployment of these technologies (García-Peñalvo and Vázquez-Ingelmo 2023).

The comparative analysis of human-authored and AI-generated text reveals several challenges in predicting gender-specific language. Alowibdi et al. explored the task of distinguishing between human-authored and GenAI-generated tweets, achieving high accuracy in identifying the source of tweets (Alowibdi et al. 2024). Their research underscores the complexity of modeling gender-specific language in generative AI content. Overall, the body of work in gender prediction for generative AI content highlights the progress made and the ongoing challenges in this field. As GenAI continues to evolve, it is crucial to develop robust methods for predicting and analyzing gender-specific language to ensure the ethical and accurate representation of gender in OSNs content.

3. Materials and Methods

A. *Motivation*

The advance development of GenAI applications has transformed content creation, enabling machines to generate text that closely resembles human language and is widely adopted by users. Noway, many users use GenAI for daily use of everythings. I noticed users use it to generate tweets for them. This technological leap raises important questions about the gender-specific traits of content shared on OSNs. Understanding the gender-specific in generated tweets by both GenAI-generated and human-authors has become a pressing concern, with significant implications for trust, transparency, and the overall integrity of OSN content. Consequently, our research seeks to address these challenges and develop robust methods for predicting the gender-specific in GenAI-generated tweets upon request of human-authors. This distinction is crucial because the increasing volume of GenAI-generated content on OSNs platforms like Twitter can obscure the differences between authentic human interactions and GenAI responses, potentially failing to identify gender-specific language. This lack of differentiation can diminish user trust, as gender-specific characteristics are often key to authenticity. If users start to doubt the genuineness of the content they encounter, their engagement and trust in these platforms may decline.

Also, it is crucial for enhancing personalization and user engagement on OSNs. By accurately identifying and generating gender-specific language, AI systems can tailor content more effectively to meet the preferences and expectations of diverse user groups. This level of personalization can lead to higher user satisfaction and increased engagement, as users feel more understood and valued. In addition, for businesses and marketers, gender-specific prediction in GenAI-generated tweets can significantly improve communication strategies. By tailoring messages to resonate with different gender groups, companies can enhance their marketing effectiveness and reach their target audiences more efficiently. Understanding the characteristics of gender-specific language allows for more compelling and persuasive communication, ultimately leading to better conversion rates and customer loyalty.

Gender-specific prediction in GenAI-generated content contributes to a more authentic user experience on OSNs. Users often expect content that aligns with their linguistic preferences and communication styles. By generating gender-specific content, GenAI systems can provide a more relatable and engaging experience for users, fostering a sense of community and belonging. Also, accurate gender-specific prediction in GenAI-generated tweets is also a matter of ethical AI deployment. Ensuring that GenAI systems respect and reflect gender differences responsibly is crucial for maintaining the trust and confidence of users. Ethical considerations, such as avoiding the reinforcement of harmful stereotypes and biases, are integral to the development and deployment of GenAI technologies. By focusing on gender-specific predictions, developers can contribute to the creation of fair and equitable GenAI systems. Yet, from a research perspective, exploring gender-specific prediction in GenAI-generated tweets provides valuable insights into the complexities of human language and communication. It allows researchers to better understand how gender influences language use and interaction patterns on OSNs. This knowledge can inform the development of more sophisticated and accurate GenAI models, contributing to advancements in the field of natural language processing (NLP) and AI.

Indeed, the ability to accurately predict and generate gender-specific language in GenAI-generated tweets is important for enhancing personalization, addressing bias, improving communication strategies, enhancing user experience, ensuring ethical AI deployment, and advancing research. As GenAI continues to evolve, prioritizing gender-specific prediction will play a crucial role in creating more inclusive, engaging, and ethical online environments.

B. *Dataset*

We started our journey to collect datasets from Twitter, by harvesting datasets from Twitter using two distinct approaches.

Firstly, we retrieve a dataset containing tweets associated with older hashtags, originating from real gender-specific labeled human-authored users. Simultaneously, we generate an equivalent dataset using a GenAI application such as ChatGPT, specifically instructing it to produce tweets based on the same hashtags without emotional bias, to observe how it interacts with the hashtags [5]. We created the tweets from two different sources of ChatGPT; A male person instructed ChatGPT to create the tweets for the assigned hashtags and a female person instructed ChatGPT to create the tweets for the assigned hashtags. While the tweets collected from human-authored sources encompass a mix of positive and negative sentiments depending on the content generated for specific assigned hashtags and are randomly collected, labeled and verified, those generated by GenAI are solely based on the provided hashtags using the two different sources mentioned above. This meticulous approach to dataset employs a dual-pronged strategy to ensure a robust foundation for our analysis, capturing a diverse array of tweets authored by genuine human users. Concurrently, we harness the capabilities of cutting-edge GenAI to generate a synthetic dataset mirroring the thematic scope of the collected tweets. By using identical hashtags for both datasets, we create a controlled environment for comparative analysis, enabling detailed insights into the dynamics of content generation.

Secondly, we extend beyond historical datasets to encompass contemporary trends. In parallel with our exploration of older hashtags, we pivot towards the latest trending hashtags on Twitter. This temporal approach enables us to capture real-time conversations and emergent themes, thereby enriching the breadth and depth of our dataset. By correlating the datasets spanning distinct temporal epochs, we aim to spot shifting patterns and trends in gender-specific human-authored and gender-specific GenAI-generated content creation. We have observed that GenAI now produces tweets with different contextual characteristics compared to those generated in the past. Therefore, similar to the initial phase, we acquire a dataset comprising tweets generated by real human users, alongside a corresponding dataset generated by GenAI for the same set of hashtags. Thus, two types of hashtags on Twitter are collected from two different sources and two different times to analyze the behavior of gender-specific for both GenAI-generated tweets and human-generated tweets. The hashtags selected are not related to any specific topic but are randomly picked from the trending list at the time or those that showed up during our exploration. Subsequently, all tweets related to the chosen hashtags are collected and stored in a database. Concurrently, all hashtags from the collected tweets are extracted and inputted into the male and female sources of the GenAI application to produce tweets. This results in two types of tweets: those generated by real human users with two classes of gender and those generated by GenAI with two classes of gender as well after being provided with the hashtags. The process unfolded through a systematic approach, encompassing various stages to ensure a comprehensive and balanced selection of data.

Therefore, we have collected gender-specific human-authored for 3000 tweets spanning more than 150 different hashtags, representing 8 spanning years. Subsequently, we tasked the GenAI application to generate an equivalent number of gender-specific GenAI-generated tweets for the same number of hashtags [21]. This approach yielded a balanced dataset comprising 6,000 tweets containing four different classes. These classes are: male human-authored tweets, female human-authored tweets, male GenAI-generated tweets and female GenAI-generated tweets.

C. *Proposed Approach*

The preprocessing step is crucial, before applying the features selection, to ensure the quality and relevance of the data used for analysis. We started by eliminating noise such as irrelevant characters, URLs, and special symbols from the tweets. Then, we split the tweets into individual words or tokens. We then remove common words that do not contribute to the meaning (e.g., "the", "is", "at") and reduce words to their base or root form to ensure consistency. After that, each tweet is converted into a bag-of-words model to represent the actual content. This model creates a set of words used in the tweet without considering the order. Finally, we analyze the sentiment of all the collected tweets to categorize them into gender-specific for both GenAI-generated and human-authored. After preprocessing, we conduct a comprehensive analysis of the collected datasets, focusing on extracting various textual features. Our analytical framework emphasizes the extraction and characterization of salient features inherent in textual data. Ultimately, we identified numerous characteristic features. Therefore, we applied a reduction technique using a feature selection method to the extracted features. Through rigorous feature engineering, we discovered that conventional feature selection methods could result in poor performance and, consequently, reduced accuracy.

To overcome these issues, we propose a novel approach of choosing Term Frequency (tf) and Document Frequency (df) with higher scoring by computing the most discriminative term or document. The goal is to select features that are most discriminative between tweets generated by GenAI and those authored by humans and then more specifically focusing on gender-specific characteristics tweets generated by one of GenAI-generated or human-authored. To achieve this, we propose an approach to leverages the Chi-Square statistic to evaluate the significance of each term's frequency in predicting gender classification, while also using probability equations to assess the differences between firstly, GenAI-generated tweets and human-authored, and secondly, male and female term usage. This will help us evaluate the independence of the terms across GenAI-generated and human-authored categories, then gender-specific categories.

Therefore, for each term t, in the dataset, we calculate its frequency in GenAI-generated tweets TFt,G and in human-authored tweets TFt,H , and then, we calculate its frequency in male tweets TFt,M and in female tweets TFt,F in male tweets. After that, we calculate the total length of terms for both in GenAI-generated tweets TLG and in human-authored tweets TLH , and then, in male tweets TLM and in female tweets TLF . Then, we calculate the probability of each term t occurring in GenAI-generated and human-authored tweets as well as gender-specific of male and female tweets as follows:

$$P(t|G) = \frac{TF(t,G)}{TL(G)} \tag{1}$$

$$P(t|H) = \frac{TF(t,H)}{TL(H)} \tag{2}$$

$$P(t|M) = \frac{TF(t,M)}{TL(M)} \tag{3}$$

$$P(t|F) = \frac{TF(t,F)}{TL(F)} \tag{4}$$

Where Equation 1 is for GenAI-generated tweets, Equation 2 is for Human-authored tweets, Equation 3 is for Gender-specific-male tweets and Equation 4 is for Gender-specific-female tweets. Also, we calculate the difference in probabilities between GenAI-generated and human-authored tweets as well as gender-specific of male and female tweets for each term t as follow:

$$\triangle P(t)_{GH} = P(t|G) - P(t|H) \quad (5)$$
$$\triangle P(t)_{MF} = P(t|M) - P(t|F) \quad (6)$$

Where Equation 5 and 6 for GenAI-generated tweets and human-authored tweets and gender-specific of male and female tweets respectively. In addition, we calculate the expected frequency for term t in GenAI-generated and human-authored tweets as well as gender-specific of male and female tweets as follows:

$$E_G = \frac{\left(TF(t,G)+TF(t,H)\right)\times TL(G)}{TL(G)+TL(H)} \quad (7)$$
$$E_H = \frac{\left(TF(t,G)+TF(t,H)\right)\times TL(H)}{TL(G)+TL(H)} \quad (8)$$
$$E_M = \frac{\left(TF(t,M)+TF(t,F)\right)\times TL(M)}{TL(G)+TL(H)} \quad (9)$$
$$E_F = \frac{\left(TF(t,M)+TF(t,F)\right)\times TL(F)}{TL(G)+TL(H)} \quad (10)$$

Where Equation 7, 8, 9 and 10 for GenAI-generated tweets, human-authored tweets, gender-specific male tweets and gender-specific female tweets respectively. Furthermore, we calculate the Chi-Square statistic for each term t as follow:

$$\chi^2_{GH} = \frac{(TF(t,G)-E_G)^2}{E_G} + \frac{(TF(t,H)-E_H)^2}{E_H} \quad (11)$$
$$\chi^2_{MF} = \frac{(TF(t,M)-E_M)^2}{E_M} + \frac{(TF(t,F)-E_F)^2}{E_F} \quad (12)$$

Where Equation 11 and 12 for GenAI-generated and human-authored tweets and gender-specific of male and female tweets respectively. Moreover, we combine the Chi-Square statistic and the probability difference to form a composite feature selection criterion as follows:

$$Discriminative\ Score(t) = |\triangle P(t)| + \chi^2_{GH} \quad (13)$$
$$Discriminative\ Score(t) = |\triangle P(t)| + \chi^2_{MF} \quad (14)$$

Where Equation 13 and 14 is for GenAI-generated and human-authored tweets and gender-specific of male and female tweets respectively. Finally, we select the top terms based on their discriminative scores for GenAI-generated and human-authored classification as well as gender-specific of male and female classification using the following pseudo code on the left.

After calculating the discriminative scores for each term, we proceed to select the most discriminative features that will be used in our GenAI-generated and human-authored prediction model and then gender-specific prediction model.

```
//Calculate Term Frequencies
calculate_term_frequencies(tweets, class):
  for tweet in tweets:
     words = preprocess(tweet)
     total_length += len(words)
     for word in words:
        if word not in term_freq:
           term_freq[word] = 0
        term_freq[word] += 1
  return term_freq, total_length

//calculate Proba.&Chi-Square&Disc. Score
calculate_discriminative_score(term_freq_c1,
 term_freq_c2,total_length_c1, total_length_c2):
  scores = {}
For term in set(term_freq_m).union(term_freq_f):
     tf_m = term_freq_m.get(term, 0)
     tf_f = term_freq_f.get(term, 0)
     p_m = tf_m / total_length_m
     p_f = tf_f / total_length_f
     delta_p = abs(p_m - p_f)
     e_m = (tf_m + tf_f) * total_length_m /
(total_length_m + total_length_f)
     e_f = (tf_m + tf_f) * total_length_f / (total_length_m
+ total_length_f)
     chi_square = ((tf_m - e_m)**2 / e_m) + ((tf_f -
e_f)**2 / e_f)
     scores[term] = delta_p + chi_square
  return scores
//Feature Selection
def select_top_features(scores, top_n):
  sorted_terms = sorted(scores.items(), key=lambda
item: item[1], reverse=True)
  return sorted_terms[:top_n]
```

We categorize these features into three different types to evaluate their effectiveness:

1. **Top 500 Features:** We first select the top 500 most discriminative features based on their scores. These features are expected to have the highest impact on distinguishing between male and female language in GenAI-generated tweets and human-authored tweets.
2. **Top 1000 Features:** In the second category, we extend our selection to the top 1000 most discriminative features. By including a larger set of features, we aim to capture more characteristics variations in gender-specific language. This broader selection helps ensure that detailed but potentially important linguistic patterns are not

overlooked.

3. **All Selected Features:** Finally, we compile a comprehensive set of all the features that were identified as discriminative, regardless of their rank. This complete set includes every term that demonstrates a statistically significant difference in usage between male and female categories. Using this extensive set allows us to fully explore the complexity of gender-specific language in GenAI-generated tweets and human-authored tweets and provides a robust basis for our predictive models.

By categorizing our features in this manner, we can systematically evaluate the impact of feature selection on the performance of our gender prediction models. This approach ensures that our analysis is both thorough and detailed, allowing us to identify the most effective features for distinguishing gender-specific language in GenAI-generated tweets and human-authored tweets.

4. Experimental Results

A. ***Evaluation***

We trained five different ML classifiers: SVM, NB, DT, RF and MLP. We applied 10-fold cross-validation on the selected features to ensure robust evaluation. The performance of each model was assessed using accuracy, precision, recall, and F1-score metrics. The results, as shown in Figure 1, Figure 2 and Figure 3, illustrate that as the number of features increases, the performance of all classifiers improves, with MLP consistently achieving the highest scores. This underscores the importance of feature selection in enhancing the model's ability to predict gender-specific language in GenAI-generated tweets.

These results highlight the effectiveness of the MLP model in capturing the details of gender-specific language in GenAI-generated tweets, making it the best-performing classifier among those tested. To evaluate both performance and accuracy, we tested the classifiers on three different feature sets: Feature 500, Feature 1000, and All Features. The results for each experiment are as follows:

- **Feature 500:** when trained on a feature set consisting of 500 features, The MLP classifier demonstrates the highest performance across all metrics, achieving an accuracy of 83%, SVM follows with slightly lower scores of accuracy 81%, RF obtained 78%, DT scored 80%, and the NB classifier shows the lowest performance in this feature set of an accuracy of 76%.
- **Feature 1000:** when trained on a feature set consisting of 1000 features, the MLP continues to outperform the other classifiers, achieving an accuracy of 86% accuracy, SVM shows strong performance as well, with slightly lower scores of an accuracy 84%, RF obtained 80%, DT scored 81%, and NB show improved performance compared to the 500-feature set, indicating that increasing the number of features enhances model performance of an accuracy of 77%.
- **All Features:** when trained on a feature set consisting of all features, the MLP achieves the highest scores across all metrics, with an accuracy of an accuracy at 90%, followed by SVM at 87%, RF at 84%, DT at 85%, and NB though improved, remains the lowest performer of an accuracy 80%.

B. ***Observation***

Our experiments demonstrate that it is possible to accurately distinguish between male and female language in GenAI-generated tweets. We have observed several notable differences and characteristics between these two gender-specific language patterns in GenAI-generated content. GenAI-generated tweets, when instructed to mimic male language, often exhibit certain linguistic patterns that are distinct from those instructed to mimic female language. Male GenAI-generated tweets tend to use more assertive and direct language as in “We need to tackle this problem head-on and come up with a solution quickly.”, while female GenAI-generated tweets often include more collaborative and empathetic expressions as in “Let's work together to find a solution that benefits everyone involved.”. This differentiation in language style reflects societal norms and stereotypes related to gender communication. Additionally, we observed that male GenAI-generated tweets frequently employ more technical jargon and formal language, particularly in professional contexts as in “Integrating the latest AI technologies will significantly enhance our operational efficiency.”, whereas female GenAI-generated tweets often incorporate more personal and relational language as in “I'm excited about the possibilities that new AI technologies bring to our work.”. This pattern is consistent with the tendency for male language to focus on information and task-oriented communication, while female language emphasizes relationship building and emotional expression.

In terms of emotional tone, male GenAI-generated tweets generally exhibit a more neutral or objective tone as in “Implementing the new software update should optimize our system's performance and increase efficiency.”, whereas female GenAI-generated tweets often convey a wider range of emotions, including empathy, warmth, and support as in “I really appreciate the team's support and dedication; it makes all the difference.”. This difference in emotional expression can impact the perceived authenticity and relatability of the tweets, with female GenAI-generated content potentially resonating more with audiences seeking emotional connection. Moreover, we found that male GenAI-generated tweets are more likely to include language related to authority, competition, and independence as in "Achieving these milestones ahead of schedule demonstrates our efficiency and dedication.", while female GenAI-generated tweets tend to use language associated with cooperation, nurturing, and community as in "Our collective efforts and shared vision have led to this wonderful accomplishment.". This distinction is evident in the choice of words, phrases, and overall tone used in the tweets.

Figure 1. Performance Metrics for Feature 500.

Figure 2. Performance Metrics for Feature 1000.

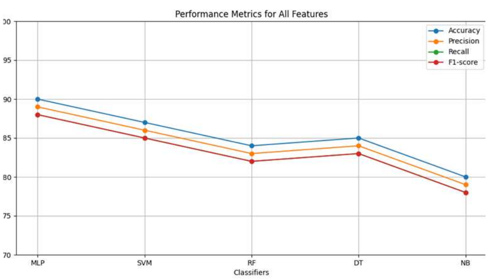

Figure 3. Performance Metrics for all Features.

In addition, we have observed many differences and characteristics between Male-GenAI-generated tweets and Female-GenAI-generated tweets. as shown in Figure 4 and Figure 5, we have noticed that Female-GenAI-generated tweets used less common words compared to Male-GenAI-generated tweets.

Figure 4. Most used words by GenAI-Male (left) and GenAI-Female (right).

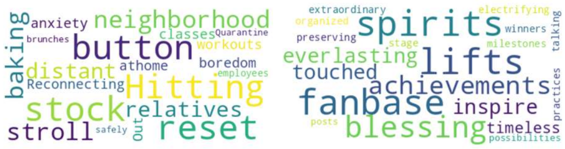

Figure 5. Least used words by GenAI-Male (left) and GenAI-Female (right).

5. Conclusion

Our research on predicting the gender-specific language in GenAI-generated tweets has highlighted significant differences between male and female language patterns, both in content and style. Our comprehensive analysis and the use of diverse machine learning models have validated the efficacy of our approach, with the MLP model consistently outperforming others in capturing the details of gender-specific language in GenAI-generated content.

References

[1] Alowibdi, J. S. (2024). A human-authored or GenAI-generated: Who is creating the content. Eng. Technol. Appl. Sci. Res., 15(4), August 2024.

[2] Merler, M., Cao, L., & Smith, J. R. (2015). You are what you tweet… pic! Gender prediction based on semantic analysis of social media images. 2015 IEEE International Conference on Multimedia and Expo (ICME). IEEE.

[3] Çelik, Ö., & Aslan, A. F. (2019). Gender prediction from social media comments with artificial intelligence. Sakarya University Journal of Science, 23(6), 1256-1264.

[4] Alowibdi, J. S., Buy, U. A., Yu, P. S., Ghani, S., & Mokbel, M. (2015). Deception Detection in Twitter. Social Network Analysis and Mining, 5, 1-16.

[5] Peersman, C., Daelemans, W., & Van Vaerenbergh, L. (2011). Predicting age and gender in online social networks. Proceedings of the 3rd International Workshop on Search and Mining User-Generated Contents, 37-44.

[6] Reddy, T. R., Vardhan, B. V., & Reddy, P. V. (2017). N-gram approach for gender prediction. In 2017 IEEE 7th International Advance Computing Conference (IACC) (pp. 860-865). IEEE.

[7] Krüger, S., & Hermann, B. (2019). Can an online service predict gender? On the state-of-the-art in gender identification from texts. 2019 IEEE/ACM 2nd International Workshop on Gender Equality in Software Engineering (GE). IEEE.

[8] Bamman, D., Eisenstein, J., & Schnoebelen, T. (2014). Gender identity and lexical variation in social media. Journal of Sociolinguistics, 18(2), 135-160.
[9] OpenAI. (2024). ChatGPT (March 15 version) [Large language model]. https://chat.openai.com.
[10] Gu, J. (2024). Responsible Generative AI: What to Generate and What Not. arXiv preprint arXiv:2404.05783.
[11] García-Peñalvo, F., & Vázquez-Ingelmo, A. (2023). What do we mean by GenAI? A systematic mapping of the evolution, trends, and techniques involved in Generative AI.
[12] Alowibdi, J. S., Buy, U. A., & Yu, P. S. (2013, August). Language Independent Gender Classification on Twitter. In Proceedings of the 2013 12th International Conference on Machine Learning and Applications (Vol. 1, pp. 365-369). IEEE.
[13] Alowibdi, J. S., Buy, U. A., & Yu, P. S. (2013, December). Empirical Evaluation of Profile Characteristics for Gender Classification on Twitter. In 2013 12th International Conference on Machine Learning andApplications (Vol. 1, pp. 365-369). IEEE.
[14] Kumar, R., & Mindzak, M. (2024). Who Wrote This? Detecting Artificial Intelligence–Generated Text from Human-Written Text. Canadian Perspectives on Academic Integrity, 7(1).
[15] OpenAI. (2024). ChatGPT (March 15 Version) [Large language model]. https://chat.openai.com.
[16] Said, Y., Barr, M., & Ahmed, H. E. (2020). Design of a Face Recognition System based on Convolutional Neural Network (CNN). Eng. Technol. Appl. Sci. Res., 10(3), 5608–5612.
[17] Sun, J., Liao, Q. V., Muller, M., Agarwal, M., Houde, S., Talamadupula, K., & Weisz, J. D. (2022). Investigating Explainability of Generative AI for Code through Scenario-Based Design. In 27th International Conference on Intelligent User Interfaces, 212-228.
[18] Virmani, D., Girdhar, P., Jain, P., & Bamdev, P. (2019). FDREnet: Face Detection and Recognition Pipeline. Eng. Technol. Appl. Sci. Res., 9(2), 3933–3938.
[19] Yan, L., Martinez-Maldonado, R., & Gasevic, D. (2024). Generative Artificial Intelligence in Learning Analytics: Contextualising Opportunities and Challenges through the Learning Analytics Cycle. In Proceedings of the 14th Learning Analytics and Knowledge Conference, 101-111.