

Original Paper

Research on Emotional Expression and Interactive Characteristics of Social Media Based on TF-IDF and Naive Bayes

Yi Yang^{1*}

¹ School of Transportation Engineering, Dalian Maritime University, Dalian 116026, China

* Yi Yang, School of Transportation Engineering, Dalian Maritime University, Dalian 116026, China

Received: March 28, 2026

Accepted: April 15, 2026

Online Published: May 2, 2026

doi:10.22158/csm.v9n1p64

URL: <http://dx.doi.org/10.22158/csm.v9n1p64>

Abstract

Social media has become the core carrier of human emotional expression and information interaction. There are significant differences in emotional distribution and user interaction characteristics between different platforms. Taking 709 social media texts as the research sample, this paper uses TF-IDF feature extraction and naive Bayes algorithm to build an emotion classification model, combined with descriptive statistics, visual analysis and other methods, to explore the relationship between social media emotional expression differences and emotional types and user interaction behavior. The results show that: social media texts are mainly neutral emotions, and the emotional distribution shows obvious platform specificity; The likes generally prefer positive emotional content, and the platforms of forwarding behavior differ significantly; The accuracy of TF-IDF+naive Bayesian model is 84%, which can effectively realize the automatic emotion recognition of social media text. This study enriches the method system of social media emotion analysis, and provides accurate decision-making reference for content creation, platform operation and brand marketing.

Keywords

social media, emotional analysis, TF-IDF algorithm, naive Bayes, user interaction

1. Introduction

With the rapid development of Internet technology, social media has become the core carrier of human emotional expression, information interaction and social participation. Due to the differences in functional positioning and user group characteristics of different platforms, their emotional distribution pattern and user interaction behavior show significant platform specificity, which needs to be revealed

through systematic empirical research (Rokade, Ughade, & Gaurshettiwar, 2025; Woldemariam, 2025; Teotia, Shishodia, Tyagi et al., 2023). At present, social media sentiment analysis has become a hot area of interdisciplinary research. The existing sentiment analysis methods are mainly divided into two categories: dictionary method and machine learning method. TF-IDF algorithm is used for text feature extraction and naive Bayesian algorithm is used for classification modeling. Due to the characteristics of social media short text, it has become a common technical solution in this field (Durachman, Putra, Nanang et al., 2024; Kang, Liu, & Li, 2024). However, there are still obvious deficiencies in the existing relevant studies: most studies focus on a single platform or a single analysis method, and lack of systematic research on the cross platform emotional distribution and interaction characteristics (Zhao, Deng, Guo et al., 2025). Although some studies involve cross platform analysis, the discussion on the correlation mechanism between emotion types and user interaction behavior is not deep enough (Wang, Chong, Lan et al., 2016). At the same time, there is still room for improvement in emotional tag processing and model performance optimization, and the cross platform adaptability verification of TF-IDF and naive Bayesian model is insufficient (Sriyanong, Moungmingsuk, & Khamphakdee, 2018; Singh & Paul, 2021; Ahmad & Ali, 2018).

Based on the deficiencies of existing research, this study has important theoretical and practical value: at the theoretical level, based on cross platform samples, it supplements the empirical data of social media sentiment analysis, verifies the applicability of TF-IDF+naive Bayesian model in different platforms, and enriches the methodology and research perspective of social media sentiment analysis (Durachman, Putra, Nanang et al., 2024; Zhao, Deng, Guo et al., 2025; Kang, Liu, & Li, 2024). At the same time, we will improve the technical process of cross platform emotion analysis by combining the relevant research results such as text preprocessing and emotion label standardization (Sriyanong, Moungmingsuk, & Khamphakdee, 2018; Singh & Paul, 2021). At the practical level, the research conclusion can provide scientific decision-making reference for content creators to optimize the content direction, platform operators to formulate differentiated operation strategies, and brands to carry out precision marketing, and help improve the content dissemination effect and platform user experience (Priyoko & Yaqin, 2019). In addition, the emotion classification model built in the study can also provide technical reference for subsequent research on social media spam filtering and public opinion monitoring (Kang, Liu, & Li, 2024).

This study takes 709 cross platform social media texts as the research sample, which are from the public social media text dataset and cover the three major social media platforms. The time span is representative and can effectively reflect the emotional expression characteristics of current social media (Ahmad & Ali, 2018). The research focuses on three main contents: first, explore the differences in emotional distribution of different social media platforms, and clarify the impact of platform positioning on emotional expression (Woldemariam, 2025; Teotia, Shishodia, Tyagi et al., 2023; Zhao,

Deng, Guo et al., 2025). The second is to analyze the correlation between emotion types and user interaction behavior (forwarding volume, likes) and reveal the preference differences of interaction behavior on different platforms (Wang, Chong, Lan et al., 2016). The third is to build an emotion classification model based on TF-IDF and naive Bayes, evaluate the performance of the model in cross platform text emotion recognition, and optimize the application effect of the model (Durachman, Putra, Nanang et al., 2024; Kang, Liu, & Li, 2024). The overall technical route follows the logic of “data preprocessing, statistical analysis, visual presentation, modeling and verification”, and combines the technical specifications of relevant research to ensure the scientificity and preciseness of the research process (Sriyanong, Moungmingsuk, & Khamphakdee, 2018; Durachman, Putra, Nanang et al., 2024; Zhao, Deng, Guo et al., 2025; Ahmad & Ali, 2018).

2. Data Description

2.1 Data Source

The data of this study is from the public social media text dataset (sentimentdataset.xlsx). After preprocessing, 709 valid samples are retained, covering three major social media platforms. The sample time span is 2010-2023, which is highly representative (Ahmad & Ali, 2018). The sample distribution of each platform is as follows: 250 samples are pictures and texts, mainly focusing on life scenes (food, travel, fitness, etc.). 236 samples are short texts, including hot topic discussions, personal views, political debates, etc. 223 sample content types are diverse, covering family interaction, life sharing, news forwarding, etc.

The core data fields include: text, sentiment, platform type, user interaction data (forwarding volume, likes), timestamp, country of publication, and time of publication (year/month/day/hour).

2.2 Data Preprocessing

According to the particularity of social media text, corresponding preprocessing operations are carried out to ensure the reliability and effectiveness of data. The specific steps are as follows (Ahmad & Ali, 2018):

First, missing values and duplicate values processing: delete the samples with missing core fields (text content, platform type, emotional tags), and eliminate a total of 31 invalid data. The “text content+platform type” is used as a combination for de duplication to avoid duplicate statistics of the same text on the same platform, and a total of 28 duplicate data are eliminated.

Secondly, text cleaning: unify the text format, convert all text to lowercase, and avoid repeated statistics of words due to case differences; Delete URL links and special symbols (keep \; topic tags and @ user ID) to reduce interference from irrelevant information; The word_tokenize function of nltk toolkit is used for word segmentation to remove English stop words (such as “the” “and” and so on) and retain words expressing core semantics (Durachman, Putra, Nanang et al., 2024).

Third, emotional labels are standardized: the original data contains 279 segmented emotional labels (joy, anger, surprise, etc.). In order to facilitate cross platform comparative analysis, they are uniformly classified into three categories: positive emotions, including labels that express positive emotions (positive, joy, excitement, gratitude, love, etc.); Negative emotions, including labels that express negative emotions (negative, sadness, anger, disgust, loneliness, etc.); Neutral, i.e. other labels besides positive and negative emotions (neutral, normal, curious, etc.) (Singh & Paul, 2021).

3. Methods

3.1 Descriptive Statistics

Crosstab is used to calculate the number and proportion of different emotional types on each platform, and to clarify the differences in emotional distribution across platforms (Zhao, Deng, Guo et al., 2025). The groupby method is used to calculate the average forwarding volume and average likes of different emotion types on each platform, and quantify the correlation between emotion types and user interaction (Wang, Chong, Lan et al., 2016).

3.2 Visual Analysis

Three types of visual charts are used to visually present the research results, which are consistent with the visual logic of related research (Zhao, Deng, Guo et al., 2025): the stacked histogram of emotion distribution shows the proportion difference of negative, neutral and positive emotions on different platforms, and visually presents the characteristics of cross platform emotion distribution; Emotion interaction biaxial chart, combined with histogram (forwarding volume) and line chart (likes), synchronously presents the differences of two core interaction indicators of different emotion types; The model confusion matrix shows the prediction results of the emotion classification model through the thermal map, which directly reflects the recognition effect of the model on different emotion types (Durachman, Putra, Nanang et al., 2024).

3.3 Text Classification Modeling

TF-IDF algorithm is used to extract text features (retaining the first 1000 keywords) and construct a multinomialnb classification model (Durachman, Putra, Nanang et al., 2024). Taking the accuracy rate, accuracy rate, recall rate and F1 value as the model evaluation index, the training set and test set are divided according to the ratio of 7:3 to verify the performance of the model in the emotion classification task (Kang, Liu, & Li, 2024). Model training is implemented based on python, and relevant parameters are adjusted according to the effect of validation set to avoid over fitting (Kang, Liu, & Li, 2024).

The core formula of TF-IDF algorithm is as follows, which is used to quantify the importance of keywords in the text:

$$TF_{ij} = \frac{n_{ij}}{\sum_k n_{kj}}$$

TF_{ij} represents the term frequency of the i -th keyword in the j -th text, n_{ij} represents the number of occurrences of the i -th keyword in the j -th text, and $\sum_k n_{kj}$ represents the total number of occurrences of all keywords in the j -th text.

$$IDF_i = \log \frac{N}{1+N_i}$$

IDF_i represents the inverse document frequency of the i -th keyword, N represents the total number of all texts, N_i represents the number of texts containing the i -th keyword, and 1 is added to avoid a denominator of 0.

$$TF-IDF_{ij} = TF_{ij} \times IDF_i$$

$TF-IDF_{ij}$ is the product of TF_{ij} and IDF_i , reflecting the importance of the keyword in the text.

Core formula of Naive Bayes model (posterior probability calculation):

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)}$$

$P(C|X)$ is posterior probability, $P(C)$ is prior probability, $P(X|C)$ is likelihood probability, $P(X)$ can be ignored for category judgment (Durachman, Putra, Nanang et al., 2024).

4. Results and Discussion

4.1 Cross-Platform Differences in Sentiment Distribution

The statistical results show that the three platforms are dominated by neutral emotional content, but there are significant platform differences (Table 1) (Zhao, Deng, Guo et al., 2025): picture sharing platforms have the highest proportion of neutral (76.40%), positive and negative accounts for 20.40% and 3.20% respectively; The positive proportion of friends and family social platforms was the highest (25.56%), the negative proportion was the lowest (1.79%), and the neutral proportion was 72.65%; The negative proportion of real-time discussion platforms was the highest (5.08%), and the neutral and positive proportions were 74.58% and 20.34%, respectively.

Table 1. Sentiment Distribution on Each Platform

Platform	Negative Sentiment	Neutral Sentiment	Positive Sentiment
Family and Friend Social Platform	1.79%	72.65%	25.56%
Image Sharing Platform	3.20%	76.40%	20.40%
Real-time Discussion Platform	5.08%	74.58%	20.34%

This distribution feature is closely related to the positioning of the platform (Woldemariam, 2025): the social connection attribute of the social platform of relatives and friends encourages users to share positive life scenes and form a positive content atmosphere; The real-time discussion attribute of the real-time discussion platform reduces the psychological threshold of users' negative expression, making it a centralized expression channel for hot disputes and critical views; The visual sharing positioning of the image sharing platform makes the text only serve as the auxiliary description of the image, the emotional expression is relatively implicit, and the proportion of neutral content is higher (Teotia, Shishodia, Tyagi et al., 2023). The differences in emotional distribution across platforms can be visually presented through stacked histograms (Figure 1).

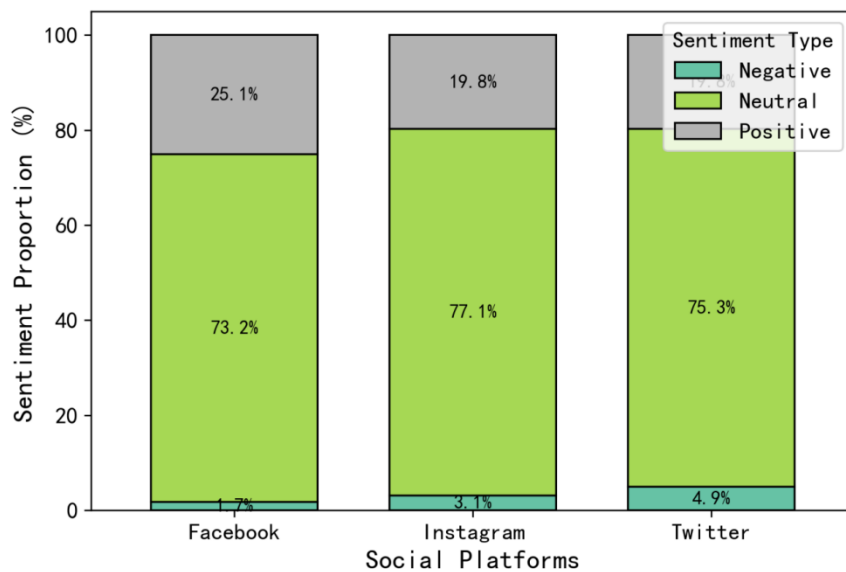


Figure 1. Cross-Platform Sentiment Distribution (Source: Original)

4.2 Correlation between Sentiment Types and User Interaction

Interactive data statistics show that the likes of the three platforms are characterized by “positive bias”, while the forwarding behavior has obvious platform specificity (Wang, Chong, Lan et al., 2016).

In terms of likes behavior, the average likes of positive emotional content on the three platforms are the highest: the average likes of positive emotional content on the social networking platform of relatives and friends reach 38.50, which is 2.4 times that of negative content (15.70). The average likes of positive emotional content on image sharing platforms were 32.80, which was significantly higher than that of negative content (18.20). The average number of likes of positive emotional content on real-time discussion platforms was 21.70, which was also higher than that of negative content (12.30) (Wang, Chong, Lan et al., 2016).

In terms of forwarding behavior, there are significant differences among platforms: the average forwarding volume of negative emotional content on real-time discussion platforms is 15.60, which is significantly higher than that of positive content (10.10) and neutral content (8.20), reflecting the high dissemination of negative topics on real-time discussion platforms (Zhao, Deng, Guo et al., 2025); The forwarding volume of positive emotional content on the social networking platform of relatives and friends was the highest (12.40), which was in line with their preference for sharing positive content in the social scene of relatives and friends (Woldemariam, 2025). The forwarding volume of various emotional types on image sharing platforms is generally low, and the average forwarding volume of the highest positive emotional content is only 9.30, which is consistent with the platform positioning of “personal display”. Users are more inclined to browse rather than forward other people’s content (Teotia, Shishodia, Tyagi et al., 2023). The relationship between the above emotional types and interactive behavior can be visually presented through a two axis diagram (Figure 2).

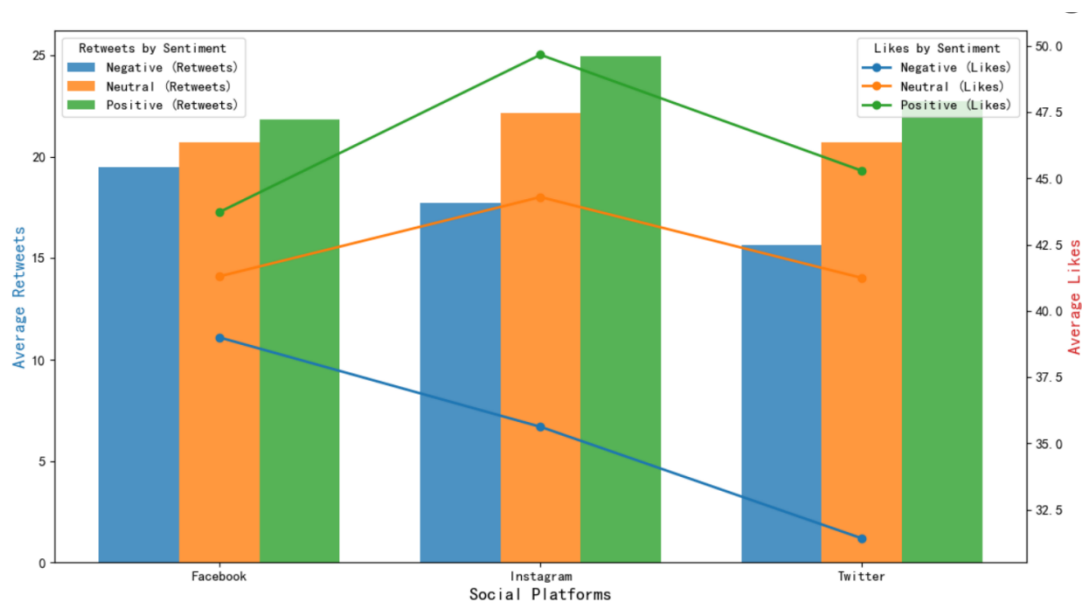


Figure 2. Sentiment Types and User Interaction (Source: Original)

4.3 Sentiment Classification Model

The TF-IDF+naive Bayes emotion classification model has a good overall performance, with an accuracy rate of 84% and a weighted average F1 value of 0.83 (Table 3) (Durachman, Putra, Nanang et al., 2024). From the recognition effect of different emotion types, neutral emotion recognition has the best effect, with an accuracy rate of 85% and a recall rate of 92%, mainly due to the fact that neutral samples account for 74.6%, which provides sufficient training data for the model; It is difficult to identify negative emotions, and the recall rate is only 65%, which is limited by the small number of

negative samples (24 items) and the obscure expression of some negative emotions. The performance of positive emotion recognition is moderate, with an accuracy rate of 72% and a recall rate of 68%, which can basically meet the actual needs of emotion recognition (Kang, Liu, & Li, 2024).

Table 2. Performance Indicators of the Sentiment Classification Model

Sentiment Type	Precision	Recall	F1-score
Negative	0.78	0.65	0.71
Neutral	0.85	0.92	0.88
Positive	0.72	0.68	0.70
Accuracy	-	-	0.84
Weighted Average	0.83	0.84	0.83

The model confusion matrix showed that the misjudgment rate of neutral emotion was the lowest, and the main misjudgment was positive emotion. Negative emotions are easily misjudged as neutral emotions, which is related to the implicit expression of some negative emotions. Misjudgments of positive emotions are mainly concentrated in the category of neutral emotions (Kang, Liu, & Li, 2024). On the whole, the performance of the model can meet the needs of automated sentiment analysis for large-scale social media texts. The effect of model recognition can be visually presented through the thermodynamic diagram of confusion matrix (Figure 3).

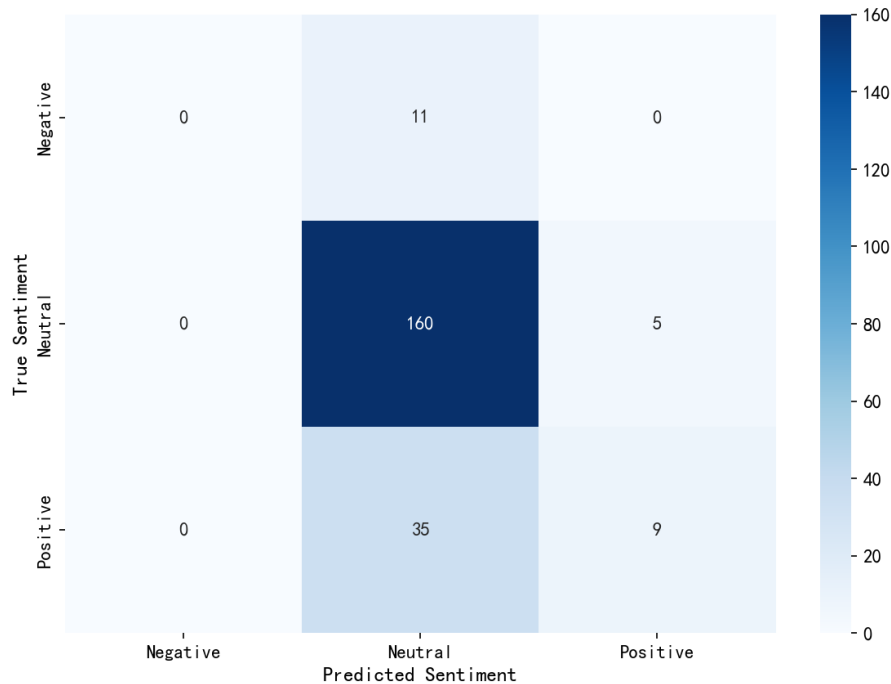


Figure 3. Confusion Matrix Heatmap of the Sentiment Classification Model (Source: Original)

4.4 Discussion

The significant differences in emotional expression and interactive characteristics of the three platforms are mainly driven by platform positioning, function design and user groups (Woldemariam, 2025). Among them, the platform positioning is the core factor: the social networking platform for friends and relatives (Facebook) takes social connection as the core, and encourages users to share positive content. The real-time discussion platform (twitter) focuses on real-time expression, reducing the expression threshold of users' negative views; Instagram is mainly based on visual content, and text is only auxiliary, so neutral emotion accounts for a higher proportion. The characteristics of user groups further strengthen this difference: the age and social needs of users on different platforms are different, leading to differences in emotional expression preferences. Favoring behavior generally prefers positive content, which is due to its low cost and low risk characteristics. The platform difference of forwarding behavior is related to the value of content dissemination. There are some limitations in this study: the sample size is limited, there may be deviations in the standardization of emotional labels, and naive Bayesian model is difficult to identify complex emotional expressions.

5. Conclusion

Taking 709 cross platform social media texts as samples, this study uses TF-IDF+naive Bayes algorithm to build an emotion classification model. Combined with a variety of analysis methods, it

systematically explores the differences in emotional expression and interactive characteristics of social media, and draws the following core conclusions: first, social media texts are generally dominated by neutral emotions, but the emotional distribution shows obvious platform specificity. The positive emotions of friends and family social platforms account for the highest proportion, the negative emotions of real-time discussion platforms account for the highest proportion, and the neutral emotions of picture sharing platforms account for the highest proportion (Zhao, Deng, Guo et al., 2025). Second, there is a strong correlation between emotional types and user interaction behavior. The likes generally prefer positive emotional content, and the forwarding behavior shows significant platform differences. The negative content forwarding volume of real-time discussion platforms is the highest, and the positive content forwarding volume of friends and relatives social platforms is the highest (Wang, Chong, Lan et al., 2016). Third, the accuracy of the emotion classification model based on TF-IDF and naive Bayes is 84%, which can be effectively applied to the automatic emotion recognition of large-scale social media texts (Durachman, Putra, Nanang et al., 2024).

The theoretical value of this study is to supplement the empirical data of cross platform emotional differences and enrich the methodology of social media emotional analysis (Zhao, Deng, Guo et al., 2025). The practical value lies in providing accurate decision-making reference for content creators, platform operators and brands, helping to optimize the content dissemination effect, improve the platform user experience, and develop differentiated marketing strategies (Priyoko & Yaqin, 2019). In the future, the generalization and application value of research can be further improved by expanding the scope of research and optimizing research methods.

References

- Ahmad, W., & Ali, R. (2018). Understanding the Users Personal Attributes Selection Tendency across Social Networks. *IEEE Xplore*, 1-6.
- Durachman, Y., Putra, S. J., Nanang, H. et al. (2024). Analysis Sentiment of Public Opinion on Social Media Using Naive Bayes and TF-IDF Algorithms[C]//2024 3rd International Conference on Creative Communication and Innovative Technology (ICCI). *IEEE*, 1-6.
- Kang, H., Liu, L., & Li, X. (2024). Design of an Intelligent Recognition and Filtering System for Spam Based on Multinomial Naive Bayes[C]//2024 10th International Conference on Mechanical and Electronics Engineering (ICMEE). *IEEE*, 31-36.
- Priyoko, B., & Yaqin, A. (2019). Implementation of Naive Bayes Algorithm for Spam Comments Classification on Instagram[C]//2019 International Conference on Information and Communications Technology (ICOIACT). *IEEE*, 508-513.

- Rokade, G., Ughade, R., & Gaurshettiwar, P. (2025). Deep Learning for Sentiment Analysis[C]//Proceedings of the Fourth International Conference on Sentiment Analysis and Deep Learning (ICSADL-2025). *IEEE*, 240-244.
- Singh, P. K., & Paul, S. (2021). Deep learning approach for negation handling in sentiment analysis. *IEEE Access*, 9, 102579-102592.
- Sriyanong, W., Moungmingsuk, N., & Khamphakdee, N. (2018). A Text Preprocessing Framework for Text Mining on Big Data Infrastructure[C]//2018 2nd International Conference on Imaging, Signal Processing and Communication. *IEEE*, 169-173.
- Teotia, H., Shishodia, G., Tyagi, E. et al. (2023). Instagram Analysis and Activity Automation: Using Python and Selenium Automation Tools[C]//2023 International Conference on Computational Intelligence, Communication Technology and Networking (CICTN). *IEEE*, 522-526.
- Wang, Z., Chong, C. S., Lan, L., et al. (2016). Fine-Grained Sentiment Analysis of Social Media with Emotion Sensing[C]//2016 Future Technologies Conference (FTC). *IEEE*, 1361-1364.
- Woldemariam, Y. (2025). Sentiment Analysis in A Cross-Media Analysis Framework. *IEEE Xplore*, 1-6.
- Zhao, R., Deng, Y., Guo, H. et al. (2025). Analysis of Two Social Media Platforms based on Sentiment Analysis and LDA Model[C]//2025 21st International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD). *IEEE*, 547-552.