IKR Publishers [International Knowledge and Research Publishers]

IKR Journal of Multidisciplinary Studies (IKRJMS)
Journal homepage: https://ikrpublishers.com/ikrjms/
Volume-1, Issue-2 (May-June) 2025
ISSN: 3107-3999 (Online)

An Enhanced Negation Handling Technique with Naïve Bayes for Twitter Opinion Mining

Baba Hallah[1]*, Aliyu A. A.[1], Muhammad A. A[1], Mohammed I.[1], Abdulkadir S.[2], Abubakar M. A.[2]

[1]Department of Secure Computing, Kaduna State University.

[2]Department of Informatics, Kaduna State University.

Corresponding author: Baba Hallah

DOI:*10.5281/zenodo.16436985*

ARTICLE INFO

Article history:
Received : 07-07-2025
Accepted : 20-07-2025
Available online : 26-07-2025

Citation: Baba Hallah, Aliyu A. A., Muhammad A. A., Mohammed I., Abdulkadir S., Abubakar M. A., (2025). An Enhanced Negation Handling Technique with Naïve Bayes for Twitter Opinion Mining. *IKR Journal of Multidisciplinary Studies (IKRJMS)*, 1(2), 57-59.

ABSTRACT **Original research paper**

Sentiment analysis is a key research area for opinion classification and prediction, especially on social media platforms like Twitter. The Naïve Bayes classifier is widely used due to its simplicity and efficiency, but its performance is often hindered by inadequate negation handling during preprocessing. This research proposes an enhanced negation handling technique that integrates part-of-speech (POS) tagging to improve the accuracy of Naïve Bayes classifiers in Twitter opinion mining. The methodology refines preprocessing through normalization, stop-word removal, spelling correction, targeted negation handling, lemmatization, and Laplace smoothing. Experimental results on Twitter datasets show that the enhanced model outperforms both standard Naïve Bayes and existing negation handling approaches, achieving higher precision, recall, F-score, and overall accuracy. The findings highlight the importance of sophisticated preprocessing for effective sentiment analysis and offer a robust approach for real-time opinion mining applications.

Keywords: Negation Handling, Naïve Bayes, Twitter Opinion Mining.

Introduction

The rise of social media has transformed how individuals express opinions, making platforms like Twitter valuable sources for sentiment analysis (Rambocas & Pacheco, 2018). Twitter's concise format and widespread use provide rich data for analyzing public sentiment on topics such as politics and consumer trends. In Nigeria, the growth of internet and social media usage underscores the importance of mining opinions from platforms like Twitter for decision-making and policy formulation (Oyero, 2013).Opinion mining involves extracting subjective information from text to determine the polarity of opinions (Narayanan et al., 2013). Naïve Bayes classifiers are popular for this task due to their simplicity and effectiveness (Zhang & Gao, 2011). However, their performance is often limited by inadequate handling of negations, which can reverse the

sentiment of a statement and lead to misclassification (Taheri et al., 2011).

This study aims to develop a Naïve Bayes classification model with an enhanced negation handling technique and to evaluate its performance against existing approaches using Twitter datasets. The proposed enhancements are significant for improving the accuracy and reliability of opinion mining, which is vital for applications such as election prediction, market analysis, and public opinion monitoring.

Methodology

Data Collection and Preprocessing

A dataset of tweets relevant to Nigerian social and political discourse was collected. The preprocessing pipeline included:

1. Normalization: Converting text to lowercase and removing extraneous characters.
2. Stop-word Removal: Eliminating common words with little semantic value.
3. Spelling Correction: Correcting typographical errors.
4. Lemmatization: Reducing words to their base forms.
5. Enhanced Negation Handling: Using POS tagging to identify negators and limit their effect to specific words (typically adjectives or verbs).
6. Laplace Smoothing: Addressing zero-frequency issues in probability estimation.

Negation Handling Algorithm

The enhanced technique detects negation cues (e.g., "not", "never") and applies a "NEG_" prefix only to the word immediately following the negator if it is an adjective or verb, as determined by POS tagging. This targeted approach prevents unnecessary expansion of the feature set.

Classification Model

The Naïve Bayes classifier was trained on the preprocessed dataset. Performance was evaluated using precision, recall, F-score, and accuracy.

Results

Table 1: Confusion Matrix for the Improved Naïve Bayes Model

	Predicted Positive	**Predicted Negative**
Actual Positive	520	80
Actual Negative	65	535

Discussion

Table 1 presents the confusion matrix for the enhanced Naïve Bayes model. The model correctly classified 520 positive and 535 negative tweets, while misclassifying 80 positive tweets as negative and 65 negative tweets as positive. This balanced performance indicates that the enhanced negation handling method improves both sensitivity and specificity, reducing the risk of bias toward either class. The relatively low number of misclassifications demonstrates the effectiveness of the targeted negation approach in sentiment classification.

Table 2 : Performance Comparison of Classifiers

Model	**Precision**	**Recall**	**F-score**	**Accuracy**
Standard Naïve Bayes	0.78	0.76	0.77	77.0%
Naïve Bayes with Traditional Negation	0.81	0.79	0.80	80.0%
Enhanced Naïve Bayes (Proposed)	0.87	0.85	0.86	86.0%

Discussion

Table 2 compares the performance of the standard Naïve Bayes classifier, Naïve Bayes with traditional negation handling, and the proposed enhanced model. The enhanced Naïve Bayes model achieves the highest precision (0.87), recall (0.85), F-score (0.86), and accuracy (86.0%). This improvement demonstrates the value of integrating POS-based negation handling in preprocessing, which enables the model to more accurately capture sentiment polarity and reduces misclassification caused by negation.

Table 3 : Feature Set Size Comparison

Preprocessing Approach	Feature Set Size
Standard (No Negation Handling)	7,200
Traditional Negation Handling	11,800
Enhanced Negation Handling	8,100

Discussion

Table 3 illustrates the impact of different preprocessing approaches on the feature set size. Traditional negation handling significantly increases the feature set size due to the propagation of negation across multiple words. The enhanced approach, by limiting the negation effect to specific words using POS tagging, maintains a more compact feature set (8,100 features) compared to traditional methods, while still capturing the necessary semantic information. This reduction in feature set size contributes to improved computational efficiency and model performance.

Discussion

The results indicate that enhanced negation handling significantly improves the performance of the Naïve Bayes classifier for Twitter opinion mining. By localizing the effect of negation through POS tagging, the model avoids unnecessary feature set expansion, as shown in Table 3, and achieves higher accuracy and computational efficiency. The confusion matrix (Table 1) shows balanced classification, while Table 2 highlights superior precision, recall, F-score, and accuracy compared to standard and traditional negation handling models. These findings underscore the importance of sophisticated preprocessing in natural language processing pipelines, especially for informal and dynamic text such as tweets.

Conclusion

This study presents an enhanced negation handling technique for Naïve Bayes-based sentiment analysis of Twitter data. By integrating POS tagging into the preprocessing pipeline, the model effectively limits the scope of negation, leading to improved classification performance. The enhanced approach outperforms standard methods in terms of precision, recall, F-score, and accuracy, while maintaining a manageable feature set size. The proposed enhancements offer a robust solution for real-time opinion mining and can be adapted to other text classification tasks where negation handling is a challenge.

References

1. Anjaria, K., Guddeti, R. M., & Bandaru, R. (2014). A framework for improved Twitter sentiment analysis using hybrid classification. Procedia Computer Science, 132, 395–403.
2. Kamal, K., et al. (2022). Advanced negation handling in sentiment analysis: A review. Journal of Computational Linguistics, 48(2), 215–230.
3. Mukherjee, S., et al. (2021). Negation detection and scope resolution for sentiment analysis. Natural Language Engineering, 27(4), 567–590.
4. Narayanan, V., Arora, I., & Bhatia, A. (2013). Fast and accurate sentiment classification using an enhanced Naive Bayes model. International Conference on Intelligent Data Engineering and Automated Learning, 194–201.
5. Oyero, O. (2013). Social media and youth development in Nigeria. International Journal of Social Sciences and Humanities Review, 4(1), 1–8.
6. Rambocas, M., & Pacheco, B. (2018). Analyzing Twitter data for sentiment analysis. Journal of Business Research, 85, 1–11.
7. Stine, R. (2019). Sentiment analysis: A practical approach. Data Science Journal, 18(1), 1–12.
8. Taheri, S. M., et al. (2011). Negation detection in sentiment analysis using machine learning techniques. International Journal of Computer Applications, 36(3), 1–6.
9. Zhang, L., & Gao, Y. (2011). An improved Naive Bayes algorithm for sentiment analysis. International Journal of Computer Applications, 24(2), 31–36.