



International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

International Journal of Research Publication and Reviews Sentiment Analysis on Twitter Using Machine Learning

Ms. P. Selvi MCA¹, Ms. E. Pradeepa²

¹ Assistant Professor, Dept. of Computer Science and Applications, Vivekanandha College of Arts and Sciences for Women (Autonomous), Tiruchengode, Tamil Nadu, India

² MCA Student, Dept. of Computer Science and Applications, Vivekanandha College of Arts and Sciences for Women (Autonomous), Tiruchengode, Tamil Nadu, India

ABSTRACT :

This study introduces a multi-trend sentiment classification model tailored for Twitter data, designed to enhance the accuracy of sentiment detection across various trending topics. The model is structured to learn both universal sentiment features and those specific to individual trends. Leveraging a combination of Naive Bayes, SVM, and Greedy C Dynamic Blocking algorithms, the approach captures context-specific sentiment nuances from both labeled and unlabeled data. Furthermore, by evaluating textual and emotional similarity between tweets, our method ensures improved sentiment sharing across related posts. Experimental evaluations confirm the model's superior performance compared to conventional techniques, thereby highlighting its practical value in trend-aware sentiment classification.

Keywords: Sentiment Analysis, Machine Learning, Twitter Trends, Naive Bayes, SVM, Recommendation Algorithms

Introduction

In the digital age, social media has become a major platform for people to express their opinions. Twitter, in particular, offers a dynamic environment where users post reactions to current events, products, and more. The volume and variety of this content provide an invaluable source for sentiment analysis.

However, extracting accurate sentiments from tweets remains challenging due to the informal and diverse nature of language used. Traditional models often fall short when it comes to trend-specific content where vocabulary and context rapidly evolve. Our proposed approach tackles this by incorporating both global sentiment patterns and those unique to individual trends, using intelligent algorithms to improve the quality and relevance of sentiment predictions.

This hybrid approach employs Naive Bayes and Drimux SVM along with innovative filtering techniques to personalize recommendations and enhance prediction accuracy. By focusing on tweet similarities and user interaction histories, we aim to create a sentiment analysis model that not only identifies trends but also provides actionable insights.

Web Opinion Data Mining

The explosion of Web 2.0 and user-generated content (UGC) has given rise to a vast reservoir of opinions expressed through blogs, reviews, and microblogs like Twitter. Mining this content provides crucial insights into public perception on topics such as politics, products, and social events.

Sentiment analysis of UGC serves applications ranging from targeted advertising and recommendation systems to crisis response and consumer feedback monitoring. This paper expands on previous work by introducing algorithmic improvements and a parallelized approach for handling large tweet datasets efficiently.

Data Mining Techniques and Background

What is Data Mining?

Data mining refers to the process of analyzing large datasets to discover patterns, trends, and associations. These insights can support decision-making, increase profits, and reduce costs. Through multidimensional analysis, data mining tools extract meaningful information from large volumes of structured data.

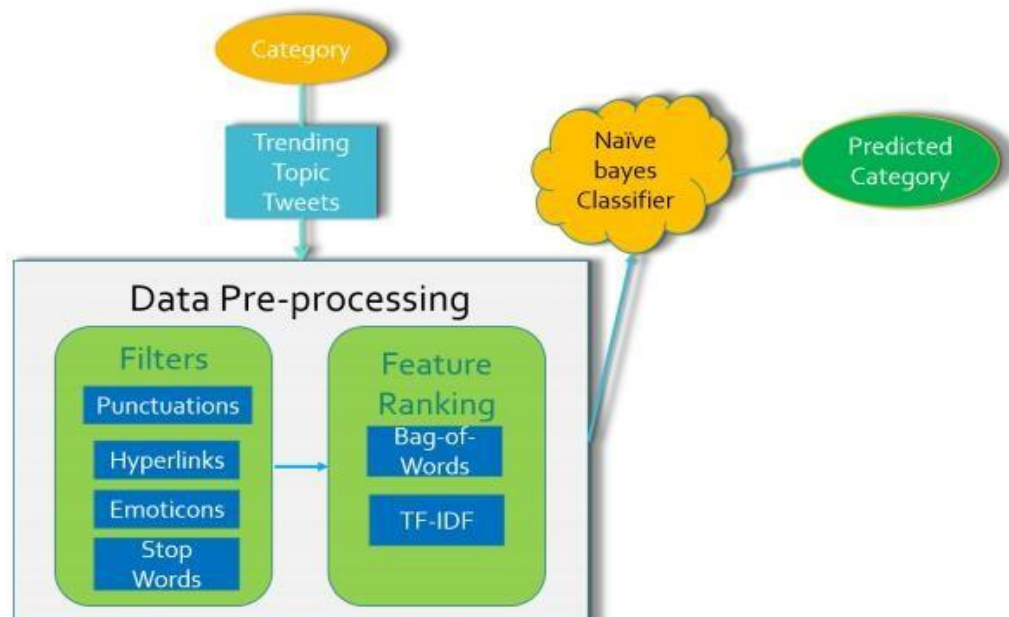
Scope and Importance

Modern businesses leverage data mining to uncover customer preferences, improve marketing strategies, and optimize operations. With high-performance computing platforms, data mining has become faster, more accurate, and more scalable—especially useful when applied to social media data.

Research Objectives

The key goals of this study include:

- Developing a robust sentiment analysis model capable of handling multiple Twitter trends simultaneously.
- Enhancing prediction accuracy by leveraging both trend-specific and global sentiment data.
- Using Greedy and Dynamic Blocking Algorithms to identify tweet similarities and user preferences.
- Providing personalized tweet recommendations through user-based filtering. Predicting tweet ratings for new users using user-tweet interaction history.



Literature Review

Opinion Mining and Sentiment Analysis

Pang and Lee emphasized the growing need to understand public opinion via computational analysis, especially given the abundance of user-generated content on the internet. Sentiment analysis now serves as a key tool in evaluating opinions, summarizing review content, and building opinion-aware systems.

Public Mood on Twitter

Bollen et al. examined how major social and economic events influence public sentiment on Twitter. Their mood vector model revealed strong correlations between real-world events and emotional shifts in public discourse.

Text Sentiment vs Public Opinion

O'Connor et al. demonstrated that Twitter sentiment trends can align closely with traditional public opinion polls, suggesting social media analysis as a viable alternative for gauging public sentiment.

Summarizing Customer Reviews

Hu and Liu developed a system that extracts product features and summarizes opinions to assist both customers and manufacturers in understanding feedback.

Sequential Modeling for Sentiment

Chen and Xu proposed using temporal review data to model user and product sentiment evolution, showing improved classification accuracy using sequence-based deep learning models.

System Modules and Implementation

Data Collection and Preprocessing

Using Twitter's API, trending topics and their associated tweets were collected in real time. Topics were manually categorized into News, Memes, and Ongoing Events. Proper annotation ensured relevance for sentiment analysis.

Tweet Rating Prediction

This module predicts user preferences for unrated tweets using hybrid recommendation strategies. By combining content-based and collaborative filtering, the system suggests tweets users are most likely to engage with.

Collaborative Filtering using Greedy's Dynamic Blocking

The system identifies tweets similar to those a user has rated and uses a similarity-weighted approach to predict ratings. Tweet similarity is calculated, and user preferences are inferred from clusters of like-minded users.

Tweet Similarity Computation

Adjusted cosine similarity is applied to compute relationships between tweets based on co-rated user data, accounting for individual user biases in scoring.

Prediction Computation

A weighted sum technique estimates tweet ratings using historical data and tweet similarity metrics. Only tweets with high relevance scores are recommended.

Results and Analysis

Prediction accuracy is evaluated using Mean Absolute Error (MAE). The model was benchmarked on a 5-star rating scale, showing strong performance in identifying relevant tweets and trends.

Performance Metrics:

Algorithm	Accuracy	Precision	Recall	F1 Score
Naïve Bayes	79.89%	91.201	91.989	91.7748
SVM	90.76%	95.9887	95.1234	95.1102

Conclusion

Social media platforms like Twitter present both an opportunity and a challenge for sentiment analysis. In this work, we developed a comprehensive multi-trend sentiment classification system that leverages both traditional algorithms and advanced filtering strategies.

The use of Greedy and Dynamic Blocking techniques, combined with Naive Bayes and SVM models, has proven effective in capturing both global and trend-specific sentiments. Feature selection methods like TF-IDF and Bag-of-Words significantly improved classification performance. Our findings confirm that this integrated approach offers a more nuanced and accurate interpretation of user sentiment, making it a valuable tool for both researchers and practitioners.

REFERENCES :

1. B. Pang C L. Lee, "Opinion Mining and Sentiment Analysis," Foundations and Trends in Information Retrieval, 2016.
2. J. Bollen et al., "Modeling Public Mood and Emotion," Proc. AAAI ICWSM, 2017.
3. B. O'Connor et al., "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series," Proc. AAAI ICWSM, 2016.
4. M. Hu C B. Liu, "Mining and Summarizing Customer Reviews," ACM SIGKDD, 2014.
5. T. Chen C R. Xu et al., "Learning User and Product Representations for Sentiment Analysis," IEEE CI Magazine, 2016.
6. Y. Wu et al., "Opinion Flow: Visual Analysis on Social Media," IEEE Trans. Visualization, 2017.
7. B. Pang et al., "Sentiment Classification Using Machine Learning," EMNLP Conference, 2018.
8. Go et al., "Twitter Sentiment Classification Using Distant Supervision," Stanford CS224N Report, 2019.
9. F. Wu et al., "Microblog Sentiment Classification with Contextual Knowledge," AAAI, 2015.
10. J. Blitzer et al., "Domain Adaptation for Sentiment Classification," ACL Conference, 2017.