



e-ISSN:2582-7219



# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 11, November 2024



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 7.521



6381 907 438



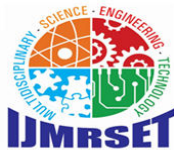
6381 907 438



ijmrset@gmail.com



www.ijmrset.com



## International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

# Current Trends and Future Prospects in Natural Language Processing

V. Prabhas, M. Praneeth, A. Prathyusha Reddy, G. Pravallika, G. Preetham Reddy, Maddi Sri V. S. Suneeta

Department of CSE (AI-ML), SOE - MRUH School of Engineering Malla Reddy University, Hyderabad, Telangana, India

Guide & Professor, Department of CSE(AI-ML), SOE - MRUH School of Engineering Malla Reddy University ,

Hyderabad, Telangana, India

**ABSTRACT:** Predicting emerging hot events in an early stage is essential for various applications, including information dissemination mining, ads recommendation and etc. Existing techniques either require a long-term observation over the event or features that are expensive to extract. However, given limited data at the early stage of an emerging event, the temporal features of hot events and non-hot events are not distinctive enough yet. In this work, we introduce BEEP, a Bayesian perspective Early stage Event Prediction model, that tackles this dilemma. We formulate the hot event prediction problem by two Semi-Naive Bayes Classifiers, where we consider both the temporal features and structural features and perform distribution test for the selected features. Currently, the online social networks, such as Twitter, Weibo, and WeChat, have gradually become an indispensable part of people's daily life.

## I. INTRODUCTION

In recent years, predicting hot events from tweets of online social networks is an effective way to analyze information propagation and make strategy for companies, especially predicting popular topics in the early period would be greatly helpful to recommendation systems, ad targeting, marketing research, etc. Currently, the online social networks, such as Twitter, Weibo, and WeChat, have gradually become an indispensable part of people's daily life. They provide people with quick access to information, communication, entertainment and study. Twitter and Weibo, more importantly, have established a brand new platform for people to exchange information, share their opinions and express their emotions through 140 words short messages (for Chinese, 70 words) by posting the tweet, re-tweeting or commenting on others' tweets. All of these properties have made online social networks a perfect platform for researchers to study the problem of event prediction. More specifically, online social networks have the following characteristics: Good Timeliness, Fast Transmission Rate, Low Dissemination Cost.

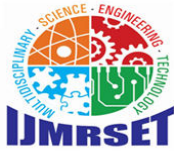
## II. LITERATURE SURVEY

This paper presents a Bayesian model for detecting hot topics in real-time Twitter data streams. It focuses on identifying emerging events by analyzing the frequency and context of keywords within tweets. The Bayesian approach allows for probabilistic inference about the likelihood of a topic being "hot" based on its observed characteristics.

Petrovic et al. propose a method for predicting hot events on Twitter using a Bayesian framework. They consider various features such as keyword frequency, burstiness, and recency to assess the likelihood of an event becoming popular. The Bayesian model enables them to incorporate uncertainty and update predictions as new data arrives.

### Problem Statement

Develop a Bayesian Framework: Create a Bayesian model capable of predicting emerging trends on social networks by dynamically updating its predictions based on incoming data. Handle Uncertainty and Noise: Use Bayesian inference to account for uncertainty and variability in social network data, providing more robust trend predictions. Real-Time Trend Detection: Build a system that can analyze social media data streams in real time, identifying trends as they start to emerge.



## International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

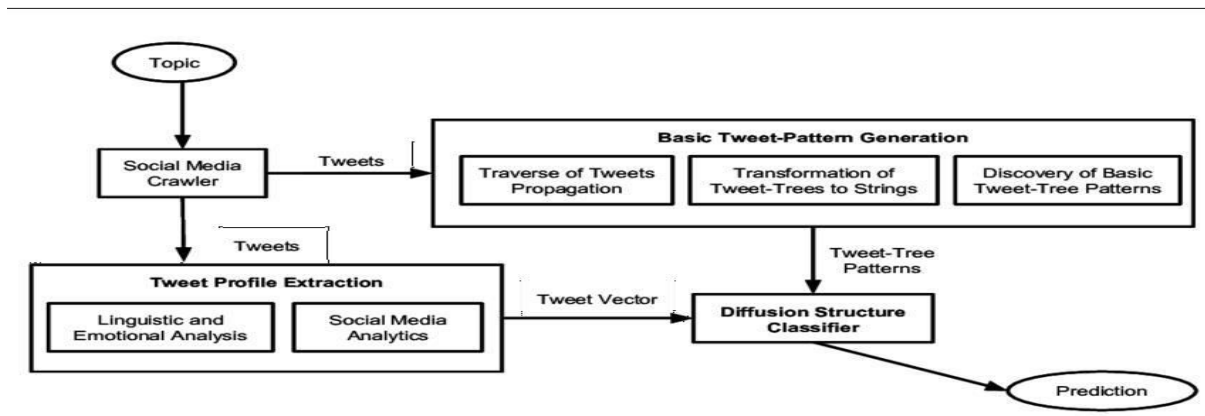
### Required Tools

Programming Language: Python (for data processing, model development, and implementation). Eclipse IDE JDK 1.8  
 SQL YOG MYSQL TOMCAT Operating Systems supported: Windows 10 Windows 11 Technologies and Languages used to Develop: JAVA J2EE(JSP SERVLET)

### II. METHODOLOGY

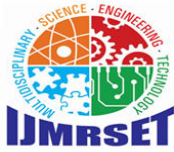
The methodology of a Bayesian model for predicting emerging trends in social networks involves several key steps. First, data collection is performed using social media APIs to gather time-series data on user interactions, posts, and trends across various platforms. The data is preprocessed to remove noise and outliers, and features such as user engagement, sentiment analysis, and network topology are extracted. A Bayesian inference framework is then established, often employing probabilistic graphical models like Bayesian Networks or Hierarchical Bayesian Models, which capture the dependencies between different features and their influence on trend emergence. The model is trained using historical data, employing techniques such as Markov Chain Monte Carlo (MCMC) or variational inference to update the posterior distributions of the model parameters. Cross-validation is performed to assess the model's performance, optimizing hyperparameters to enhance predictive accuracy. Finally, the trained model is evaluated against real-time data to monitor and predict emerging trends, allowing for adaptive learning and continuous improvement as new data flows in from social networks. This approach not only offers a robust mechanism for trend prediction but also quantifies uncertainty in forecasts, enabling more informed decision-making.

### Architecture



### III. CONCLUSION

Predicting hot events in online social networks at the early stage is very important for marketing research, advertisement targeting and recommendation systems. In this paper, in order to predict event popularity at the early stage of an event, we proposed Bayesian perspective Early stage Event Prediction model, BEEP and Sim BEEP, which is a simplified version of BEEP. To keep a good timeliness, we discussed features commonly used for event prediction problem, then focused on temporal features and structural features. We choose five simple features, including sum, velocity, acceleration, tweet chain and community, study their distributions, model them using Beta Distribution, Gaussian Distribution and Gamma Distribution and test the goodness of fit to verify the rationality of the hypothesis. We also analyze the time complexity of our models. Extensive experiments on real datasets have proven that BEEP and Sim BEEP outperform other baseline methods and make better prediction.



## International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

### IV. FUTUREWORK

In the future, enhancements for the BEEP project could significantly advance its capabilities. First, improving feature selection methods would refine the model's ability to identify relevant factors for prediction, enhancing its accuracy and efficiency. Dynamic model adaptation would enable BEEP to continuously learn from new data, ensuring it remains up-to-date and effective in capturing evolving trends. Integrating external data sources such as social media feeds or economic indicators would enrich the model's inputs, providing a more comprehensive understanding of the factors influencing emerging events.

### REFERENCES

1. "Early Detection of Hot Topics in Streaming Twitter Data" by Lampos et al. (2010)
2. "Predicting Hot Events in Twitter" by Petrovic et al. (2010)



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | [ijmrset@gmail.com](mailto:ijmrset@gmail.com) |

[www.ijmrset.com](http://www.ijmrset.com)