International Journal of Combined Research & Development (IJCRD) eISSN:2321-225X; pISSN:2321-2241 Volume: 14; Issue: 8; August- 2024

RUMOR DETECTION IN TWITTER AN ANALYSIS IN RETROSPECT

Sindhu K
PG, Student
Dept. of MCA
The Oxford College of Engineering,
Bommanahalli, Bengaluru- 560068
sindhumca2025@gmail.com

Dharamvir
Associate Professor
Dept. of MCA
The Oxford College of Engineering,
Bommanahalli, Bengaluru- 560068
hodmcatoce@theoxford.edu

ABSTRACT

Information may now be disseminated quickly and widely thanks to the meteoric rise of social media sites like Twitter. Nevertheless, these platforms are prone to the spread of unsubstantiated information, sometimes called rumors, due to the absence of content management and the dependence on crowdsourcing. The spread of false information, such as rumors, can have far-reaching effects on society and the economy. Using machine learning approaches, this research tackles the subject of rumor detection in Twitter data. We do a retrospective study using a dataset of tweets collected in 2009 to find characteristics and patterns that differentiate rumors from real information. While taking user attributes into account to improve identification accuracy, the suggested method mainly focuses on content-based features like word frequency and linguistic clues. We do this by evaluating the efficacy of several algorithms in identifying rumor and non-rumor tweets, such as k-Nearest Neighbor (k-NN) and Naive Bayes Classifier. Furthermore, a dedicated pre-processing technique is presented to clean and standardize the content of tweets while maintaining crucial information for feature extraction. The machine learning models' efficacy is greatly enhanced by this phase. To further guarantee the safety of training data, the project uses RSA cryptographic techniques; this allows for the safe exchange of models between entities like fact-checking and law enforcement. This study lays the groundwork for future real-time detection system development by shedding light on the significance of feature selection and pre-processing in automated rumor identification.

KEYWORDS: Social media, Twitter, Rumor detection, False information, Data Mining, Learning Algorithms, k- Nearest Neighbor (k-NN), Naive Bayes Classifier, Content-based features, User attributes, Real-time detection

INTRODUCTION

Online Social Media sites like Twitter are very commonly used for sharing information to other users in the social network. By design, Twitter is a micro-blogging service provider where users send information using short messages called tweets, which are typically characters long. Each user can subscribe to receive messages from other users by becoming

their followers. Similarly, other users will see this user as someone who follows them. Such follower-follower connections in a network allow for the quick transmission of data. Also, it can end up propagating rumors, which are unfounded reports. A rumor is a sense or a claim that is not based or deceptive to a person or a thing. Misinformation or disinformation, more widely known as false information, may be disseminated as a rumor. The inherently crowdsourced nature of twitter just exacerbates the situation. Identification of rumor (or the untrustworthy information in a dire state), through Twitter has become the topic of many major publications in the academic sphere. Our paper is devoted to solving the issue of rumor detection.

LITERATURE SURVEY

Author: Marcelo Mendoza, Carlos Castillo the information reliability of news stories spread through the popular microblogging site Twitter is examined. While the vast majority of tweets are factually accurate, studies have revealed that Twitter is sometimes used to disseminate falsehoods and disinformation, sometimes unknowingly. In this article, we zero in on automated techniques for determining the veracity of a certain collection of tweets. In particular, we use characteristics collected from microblog posts on "trending" subjects to determine whether or not they are genuine. The text of the postings, citations to othersources, and characteristics derived from individuals' posting or re-posting (re-tweeting) activity are considered.

Author- Jure Leskovec Communities arise when edges present a high density of nodes in particular networks of a practical problem, which may be the social, informational or technical kind. Lack of a reliable gold-standard ground-truth, assessment intricacies, intractability of algorithms and the proliferation of community definitions have served to make the problem of community identification in networks an elusive one. We examine here 230 large networks that include information, social relations and group memberships that are specified by nodes. It is these groupings that form our concept of ground truth communities. After this we offer a framework whereby, we can do far-reaching quantitative comparisons and assessments of different network community definitions. There are 13 popular definitions of network community that we apply to test their robustness, sensitivity, and quality. The thirteen definitions can be grouped into four categories as we will see.

EXISTING WORK

Twitter and other online social media sites function on an unmoderated, crowdsourced model where anybody may publish or distribute anything without human intervention. Such platforms are extremely susceptible to the fast dissemination of false information and rumors due to the absence of credibility checks and moderation. The nature of Twitter, which promotes real-time publishing and sharing (retweeting), in

credible news and rumors by amplifying the dissemination of unconfirmed material. Despite a plethora of studies aimed at identifying and categorizing false or misleading information, present systems suffer from a number of processing accuracy and productivity issues. In addition, it is very uncommon for attribute or selection of features to not improve model performance in these systems. The goal of feature selection is to enhance efficiency of processing and model correctness by removing redundant or unnecessary data. People frequently mistake dimensionality reduction with the present approaches, which may not be successful enough. Dimensionality reduction and feature selection both seek to simplify data, however feature selection keeps the original features and dimensionality reduction makes new compounded features, which could make the data less interpretable. Preliminary Investigation: The primary goal of any project's development should be to provide a simple and easy-to-use framework for a small business to send and receive emails, complete with a search engine, contact book, and fun games. We will start the first action, which is preliminary investigation, after it is accepted by both the business and our project guide. The task consists of three sections. Ask for further information, do a feasibility study, and get the go-ahead. Request Clarification: The project request has to be reviewed to ascertain the exact requirements of the system once it has been approved by the organization and the project guide, and an investigation has been contemplated. Within this context, our solution is primarily targeted at enterprise customers whose systems are able to be networked over their local area network (LAN). Given modern man's hectic schedule, it is only fair that all necessities be supplied in a prefabricated form. So, the corresponding growth of the portal was brought into being, taking into mind the broad usage of the net throughout day-to-day life.

PROPOSED SYSTEM

To address the current framework's limitations, the suggested solution introduces a safe and effective architecture for rumor identification. This is accomplished by utilizing a modular architecture, with Module-1 being crucial in safeguarding the privacy of training data. This section uses the Rivest, Shamir, and Adleman (RSA) cryptographic technique to deal with the problem of illegal data access. A popular asymmetric encryption method for safe data transfer is RSA. Protected training data is encrypted using RSA so that only authorized users may decipher it. Meanwhile, verified groups like law enforcement or academic organizations can still access the data, allowing for coordinated rumor detection efforts without sacrificing privacy. In addition to protecting sensitive information, this method makes trained machine learning models more portable, facilitating their safe transfer and reuse.

Handling the model is expected to yield the better results

METHODOLOGY

The unit testing emphasis is on verifying the software's smallest building block, the module. By focusing on individual control pathways inside a module, unit testing may guarantee full coverage and maximize error detection. Each module is tested separately to make sure it works correctly when combined. Therefore, it is called Unit Testing. This testing, we check the interfaces between modules to sure to match the design specs and test each module separately. We have evaluated every critical processing step to ensure they provide the intended outcomes. All routes for handling errors have also been tested.

Integration testing is the best course of action if you are experiencing problems with software construction or verification. Following the program's incorporation, we do a series of high-order checks. The key purpose of this testing step is to build a program structure utilizing unit tested components in accordance with the design standards.

Using this method, constructing a program's framework is a gradual procedure. Beginning with the primary program module and working one's way down the control hierarchy is the proper approach to integrate modules. Both breadth-first and depth-first approaches are used to include the modules that are subordinate to the primary program module into the framework. The program is tested from the main module using this manner, and as the test goes lower, individual stubs are substituted.

EXPERIMENTAL RESULTS

Data Collection & Storage Historical Twitter datasets labeled as rumor or non-rumor were used. User metadata such as follower count, retweet count, and verification status were included. Data was stored in a MySQL database via XAMPP. Preprocessing Module Cleaned tweets by removing URLs, hashtags, mentions, and other noise. Extracted features like word count, hashtags, punctuation, uppercase usage, etc. Normalized and inputs for the machine learning models. Machine Learning Models We was applied naive bayes for content-based classification. K-Nearest Neighbour (KNN) was used for user-based feature classification

Security Implementation RSA encryption was integrated to secure sensitive training data and user credentials, enabling safe sharing with verified organizations. Rumor Detection Engine Combined content-based and user-based features for final classification. Calculated overall probability scores to determine whether a tweet was a rumor or non-rumor. User Interface Admin dashboard with modules for login, dataset upload, preprocessing, viewing datasets, feature extraction, probability calculation, and performance metrics. Displayed user reputation scores, classification results, and analysis reports. Testing & Validation Conducted unit, integration, and user acceptance testing. Evaluated precision, recall, F1-score, and accuracy.

the application functionalities such as data upload, preprocessing, viewing, and analyzing tweets. Enter the valid admin name and password, then click on the 'Login' button. The user will be successfully logged into the project

Fig. 1. Admin Login page

Select a valid twitter datasets file from the local storage by clicking on the button 'Choose file' and then click on the Upload Dataset

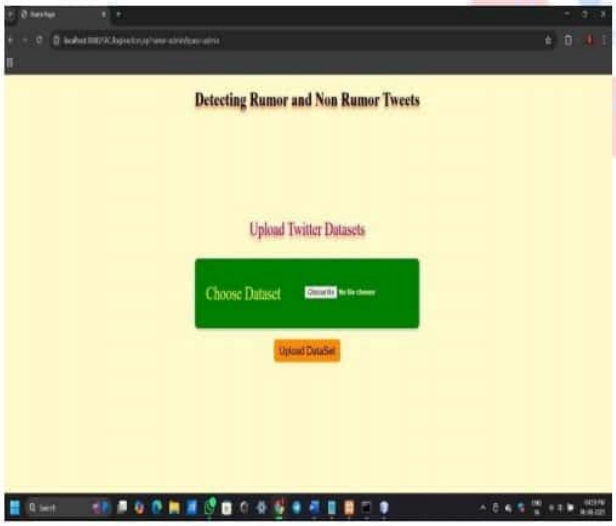

Fig. 2. Upload data base page

The Data Collection page displays a structured view of tweets collected during a specific event. It includes key metadata such as tweet ID, user handle, and tweet

content. This dataset is foundational for building a rumor detection system, where each tweet will eventually be labeled (manually or automatically) and for rumor detection in social media.

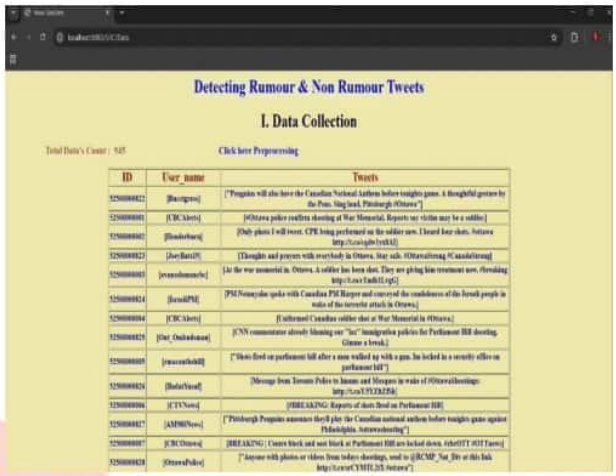

Fig. 3. Probability calculation

This image presents the Performance Metrics section Detecting Rumor and Non-Rumor Tweets.

Fig. 4. Output of agent page

CONCLUSION

Information production, dissemination, and consumption have all been impacted by the proliferation of social media, especially Twitter. There are many positive aspects to the democratization of information, but one negative is that it has facilitated the fast dissemination of unconfirmed material, or rumors. The purpose of this research was to create a rumor detection system that uses machine learning and feature-based analysis to tackle this important problem. The project used a methodical strategy that prioritizes accuracy, security, and accessibility across every phase of development, from collecting information and preprocessing through model training and deployment. Word frequency, punctuation, hashtags in order URLs, and all capitalization were some of the content-based criteria used to evaluate tweets. This approach proved that content characteristics alone may give substantial predictive power for rumor categorization, in contrast to conventional methods that rely mostly on user behavior.

Finally, the project was a success in creating a reliable, safe, and effective rumor detection system. In particular, it shows how crucial preprocessing, selecting features, and mixture modeling methods are for spotting false information. In addition to detecting rumors, this technology lays the groundwork for real-time applications in digital forensics, public information safety, and social media monitoring.

REFERENCES

[1] B. Poblete, M. Mendoza, and C. Castillo. Trustworthiness of Twitter Posts. Pages 675684 in the 2011 edition of the Proceedings of the 20th International Conference on the World Wide Web.

[2] Mendoza, B, and Castillo, C, [2]. We can't trust what we RT on Twitter; the platform is in crisis. Held during the 1st Annual Social Media Analytics Workshop (SOMA 10) on July 25, 2010.

[3] W. Friedrich, B. Doerr, and M. Fouz. Revelries in online communities propagate with a rate of sub logarithmic delay. Published in 2011 on page 2130 in the 43rd Annual ACM Symposium on Theory of Computing (STOC).

[4] Tobias Friedrich, Benjamin Doerr, and Mahmoud Fouz. The Reasons Why Social Media Rumors Go Viral So Quickly. Volume 55, Issue 6, June 2012, pages 70–75 of the

[5] ACM Communications Magazine (CACM) Homepage archive In their work, V. Qazvinian, E. Rosengren, Dragomir R. Radev, and Q. Mei contributed substantially. It has been circulating: How to Spot False Information on Microblogs. Included at the 2011 Conference on Empirical Methods in Natural Language Processing (2011).

[6]. Yang, F., Liu, Y., Yu, X., and Yang, M. The Sina Weibo rumor detector works automatically. Within the 2012 ACM SIGKDD Workshop on Mining Data Semantics, Proceedings.

[7]. Y. Wang, S. Kwon, M. Cha, K. Jung, W. Chen, and W. Chen [7]. Features of a Microblog Network for the Dissemination of Rumors. Pages 299–308 of the Social Information Index, LNCS 8238.