International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 8, No. 3, 2021, pp. 1-5.

ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

An Efficient Approach for Spammer Detection on Twitter and their Behaviour Analysis

Aman Saxena [a], Sushma Yadav [b]

[a] *Department of Computer Science and Engineering, CloudCore Technical University, Surat, India.*
[b] *Department of Computer Science and Engineering, CloudCore Technical University, Surat, India.*

Abstract

Because of social media, internet use, and online marketing, fake news is a serious problem. It seeks to draw in customers and present false information. Since there is currently no reliable method for differentiating between fake and legitimate news, deep learning (DL) is a potential academic subject for predicting bogus news. To solve this, an efficient detection strategy is needed, such as the construction of a Dual-Stage Deep Capsule Auto-Encoder model for Twitter data. Outperforming existing methods, this work suggests a novel use of deep neural networks for user geolocation and bogus news detection. Simulation planning assesses the efficacy of proposed methods. A major issue in public health, politics, the economy, and business is the consumption and dissemination of erroneous or biased content brought on by the pervasiveness of digital information. Internet users need to deal with this problem. The method tackles the problem of identifying fake news in contemporary culture by modeling user behavior and geolocation on social media platforms using psychological elements.

Keywords: Twitter, Fake News, Deep Learning.

DOI:10.71086/IAJSE/V8I3/IAJSE0816
Received: 05 June 2021/Revised: 07 July 2021/Accepted: 09 August 2021/Published: 30 September 2021

Introduction

These days, microblogging has become a common way to have conversations and find information. With more than 500 million active users, Twitter is one of the most well-known microblogging sites in recent years (Nadiammai & Hemalatha, 2014). A tweet, retweet, account users, account bio, geolocation, language, source, profile image uniform resource locator (URL), number of followers, number of friends, location, status URL, and entity string are just a few of the many pieces of information that may be found on Twitter. This data is linked to users, messages, and tweets. Every Tweet has a creator, a message, a unique ID, the time it was posted, and sometimes geo-specific metadata that the user disclosed. Every user has an account bio, a number of followers, an identification, and a Twitter name. The fundamental atomic building blocks of everything on Twitter are tweets. Another name for tweets is "status updates." The lengthy list of "root-stage" attributes in the tweet component contains crucial attributes like content, timestamp, and identity. In this case, identity is represented by a distinct identifier. Additionally, a timestamp for the tweet's universal time coordinated (UTC) time indicates when it was created, and the content contains the status update's actual unified resource locator-8 (UTF-8) text (Kachirski & Guha, 2003).

The user describes the message and includes metadata from their public Twitter account. Anybody or anything can be a user. They may be referred to, appear in bulk, tweet, retweet, add quotes to tweet, follow others, make lists, and have a home timeline. The values of these user metadata are remarkably consistent. Less than 140 characters (now 280 characters) make up the brief twitter content, which is mixed with inferred traces. Twitter bios are brief—less than 160 characters—and accurately describe the user's actual activities. This allows one to market the user's genuine identity. Users can enter personal information in Twitter's bio-box, which is limited to 160 characters (Jyothsna et al., 2011). Block Diagram of Online Social Networking Sites shown in Fig. 1.

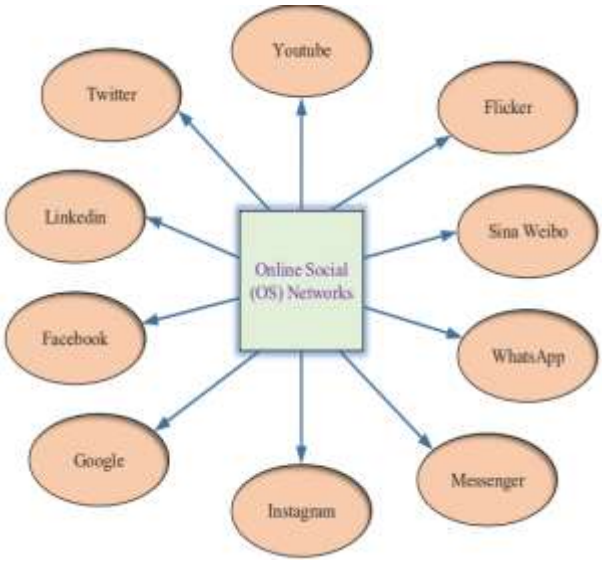

Fig. 1: Block Diagram of Online Social Networking Sites

When a user enters this box, it can be used for a variety of purposes, such as providing demographic data, such as age, marital status, place of residence, or occupation, as well as indicating a preference, interest, or other private information. Users employ particular phrases that represent their function or occupation in this bio description, and they also share some commonalities. Positive Twitter bios highlight their achievements and portray them as genuine and humanizing. It can provide a suggestion to follow them. Additionally, it displays details about the user's identity and interests, which is crucial to understand, much like tweets (Jabez & Muthukumar, 2015).

Related Work

By concentrating on methods for identifying fake news, Altheneyan & Alhadlaq, (2023) have brought scholars one step closer to an answer. The dataset FNC-1, which provides four criteria for categorizing false news, will be used in this study. The most advanced methods for spotting false news are evaluated and contrasted using machine learning and big data technology (Spark). A stacked ensemble model was constructed in this work using a decentralized Spark cluster. Following the feature extraction process with N-grams, Hashing TF-IDF, and count vectorizer, we used the proposed stacked ensemble classification model (Gogoi et al., 2010). In order to identify the Source of Fake News

(SOFNs) on Twitter, Hamdi et al., (2020) proposed a hybrid technique based on social graphs and user attributes. In order to categorize users with SOFNs and Source of Real News (SORNs) labels, the author employed the graph embedding technique. A threshold value of 5% was chosen, and users with a value below that threshold were categorized as SOFNs, while others were labeled as SORNs. The author relevance of each Node2vec feature that was taken out of the dataset was determined using the split and gain method with the use of the Light GBM model (Aljawarneh et al., 2018). When compared to existing state-of-the-art techniques that base their choices solely on user attributes, the authors' automatic false news identification system, which they developed by employing a graph embedding methodology, significantly improved fake news detection performance.

Gupta et al., (2015) looked into a number of spam detection algorithms designed for social networks. The difficulties of cross-domain spam detection techniques in the context of social networks were also examined by Mehrotra & Dhaka, among others. Similarly, a detailed examination of techniques for identifying spam profiles in social networks was part of Krithiga and Ilavarasan's study. A group of academics also looked at the design of SMS-based study on social network spam detection methods, pointing out the limitations of past studies in this field.

Methodology

Because they handle a lot of sensitive user data every day, many social media sites, including Twitter, have privacy and security flaws (Erbacher & Sobylak, 2002). Even though Twitter is a well-known worldwide internet network, these issues still exist. The community faces a serious security risk as a result of the platform's increasing popularity, which has made it a target for bad users who try to discredit legitimate Twitter users. Different groups of Twitter spammers have been identified via research. The first group consists of "fake users," or Twitter users whose real identities are unknown. These individuals utilize their false identities to impersonate actual Twitter users in order to spread spam content, usually directed at the impersonated user's followers or other Twitter users. Spam Detection Based on URL shown in Fig. 2.

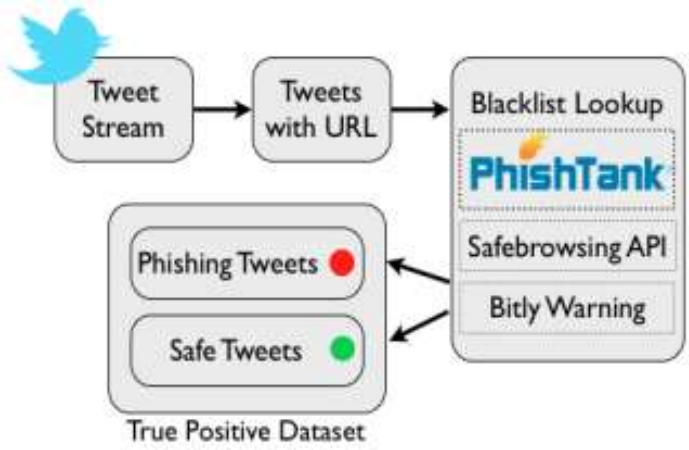

Fig. 2: Spam Detection Based on URL

This includes 11411 tweets from the Fake News Detection on Twitter EDA dataset, 14050 tweets from the ISOT dataset, and 7613 tweets from the FakeNewsNet dataset. Eighty percent of each dataset is utilized for training, and the remaining twenty percent is used for testing. Each dataset contains both real and artificial data. Both the presented noises and the quality of the data are improved throughout this preprocessing step. The embedded word analysis method is used to represent words in text analysis.

Experimental Results

By examining the performance metrics, the proposed false news/review prediction algorithm's efficacy is evaluated. This section describes the performance evaluation of the proposed prediction model using three distinct datasets. Furthermore, a comparative analysis is conducted to illustrate the efficacy of the recommended tactics. During the training and testing stages, the suggested model's accuracy and loss are examined for every dataset. Training phase accuracy outperforms testing phase accuracy in the ISOT dataset at epoch 100. Then, between the ages of 150 and 300, it goes on similarly. Similar to this, the training phase loss is consistent for both training and testing throughout epochs after reaching its minimum at epoch 100.

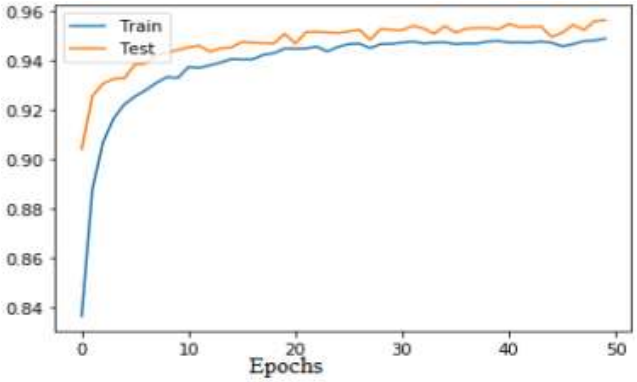

Fig. 3: Accuracy of Proposed Model

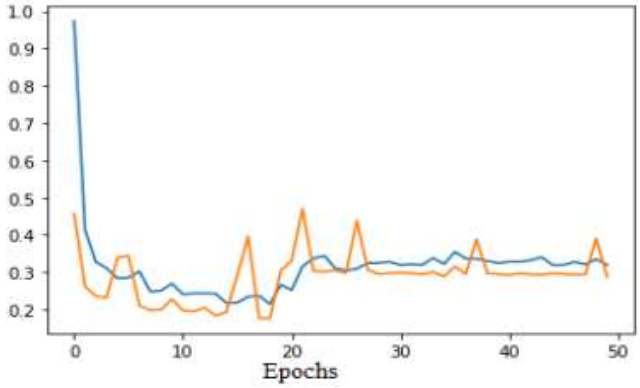

Fig. 4: Loss of Proposed Model

The accuracy and loss values of the Fake News Detection on Twitter EDA dataset are optimal for testing and training. The loss during the training phase is significantly decreased by the 50th epoch. The FakeNewsNet dataset's accuracy rises during the training phase starting at the 75th epoch and stays constant between the 100th and 300th epochs. For every dataset, the accuracy of the testing phase is marginally lower than that of the training phase. Fig. 3 and 4 display the performance analysis of recommended and implemented strategies.

The ROC analysis is essential for evaluating the potential of a prediction model. This analysis shows how accurately the prediction model classifies the classes of the given input data. The RIC analysis is measured using three datasets. The comparison study shows that the recommended method outperforms the existing methods in terms of ROC. ROC Curve Analysis shown in Fig. 5.

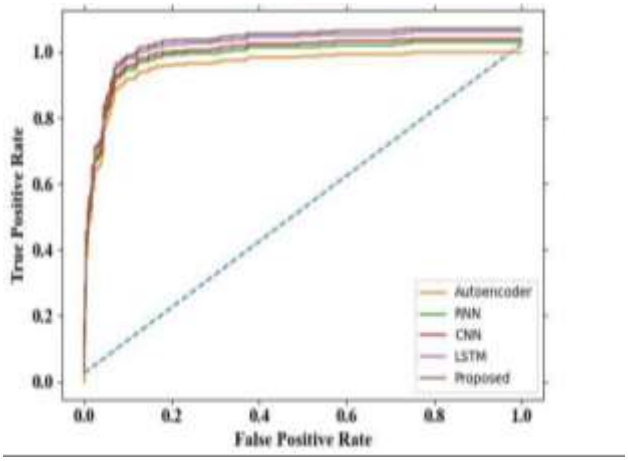

Fig. 5: ROC Curve Analysis

In terms of ROC, the existing auto-encoder performs worse than others. Because of the flawed decoding, the autoencoder performs worse in categorization. Better prediction results are produced by the system due to the effectiveness of the proposed model. The findings of the study and experiments showed that the proposed prediction model outperforms other deep learning techniques.

Conclusions

Social media's growth offers regular people worldwide a number of advantages. People can contact with one another and exchange evaluations, opinions, and other information via the Internet. Because more people are using the internet, social media has a vast amount of information. On the other hand, social media information is not very reliable. Because there are several false news reports that have an impact on society and generate confusion for many people. In order to prevent spreading false information, it is important to expose fake news in the current environment. The hybrid DL approach was employed in the proposed effort to identify the fake news. The suggested approach includes four different stages for detecting false news: data collection, preprocessing, embedded word analysis, and classification. First, the ISOT dataset, the Fake News Net dataset, and the Fake News Detection on Twitter EDA dataset are analyzed, and false information is distinguished based on genuine and fraudulent platform reviews.

References

Nadiammai, G. V., & Hemalatha, M. J. E. I. J. (2014). Effective approach toward Intrusion Detection System using data mining techniques. *Egyptian Informatics Journal*, *15*(1), 37-50. https://doi.org/10.1016/j.eij.2013.10.003

Kachirski, O., & Guha, R. (2003). Effective intrusion detection using multiple sensors in wireless ad hoc networks. In *36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the* (pp. 8-pp). IEEE. https://doi.org/10.1109/HICSS.2003.1173873

Jyothsna, V. V. R. P. V., Prasad, R., & Prasad, K. M. (2011). A review of anomaly based intrusion detection systems. *International Journal of Computer Applications*, *28*(7), 26-35.

Jabez, J., & Muthukumar, B. J. P. C. S. (2015). Intrusion Detection System (IDS): Anomaly detection using outlier detection approach. *Procedia Computer Science*, *48*, 338-346. https://doi.org/10.1016/j.procs.2015.04.191

Gogoi, P., Borah, B., & Bhattacharyya, D. K. (2010). Anomaly detection analysis of intrusion data using supervised & unsupervised approach. *Journal of Convergence Information Technology, 5*(1), 95-110.

Aljawarneh, S., Aldwairi, M., & Yassein, M. B. (2018). Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model. *Journal of Computational Science*, *25*, 152-160. https://doi.org/10.1016/j.jocs.2017.03.006

Gupta, S. (2015, October). An effective model for anomaly IDS to improve the efficiency. In *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)* (pp. 190-194). IEEE. https://doi.org/10.1109/ICGCIoT.2015.7380455

Erbacher, R. F., & Sobylak, K. (2002, September). Improving intrusion analysis effectiveness. In *2002 Workshop on Computer Forensics*.