Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Harnessing the Power of Multi-Classifiers: A Novel Adaptive Feature Neural Network Framework for Twitter Spam Detection

M.Arunkrishna[1*], Dr.B.Senthilkumaran[2]

[1]Research Scholar, PG & Research Department of Computer Science, Christhu Raj College (Affiliated to Bharathidhasan University), Tiruchirappalli - 620 012, Tamilnadu, India. ORCID ID: https://orcid.org/0000-0001-9310-9299, arunkrishna.murugan@gmail.com

[2]Research Advisor, PG & Research Department of Computer Science, Christhu Raj College (Affiliated to Bharathidhasan University), Tiruchirappalli - 620 012 , Tamilnadu, India. ORCID D: https://orcid.org/0000-0002-7111-1950, *Corresponding author

ARTICLE INFO

Received: 18 Dec 2024

Revised: 10 Feb 2025

Accepted: 28 Feb 2025

ABSTRACT

Currently, there are several social media platforms such as Facebook, Twitter, and Instagram on the internet that bring people together. Twitter, known for its wealth of information, is a popular choice among users for connecting with others and sharing updates. The platform utilizes Google Safe-browsing to identify and block spam URLs. The advanced API of Twitter allows for easy data manipulation, attracting various spammers to the platform. Previous studies have implemented ML (Machine Learning) techniques to combat Twitter spam, but their algorithms lack comprehensive evaluation and accuracy when dealing with large datasets. To address these challenges, this study proposes an adaptive feature neural network analysis. The proposed model is developed by most efficient method which differentiates between spam and non-spam tweets. In proposed model, the classifier is applied to a large dataset of 600 million public tweets and evaluated based on various metrics such as accuracy, True Positive Rate (TPR), False Positive Rate (FPR), and F-measure. Results indicate that the proposed technique shows robust performance in spam detection on Twitter.

Keywords: Neural Network, Accuracy, True Positive Rate, False Positive Rate, Online Social Networks, Twitter Spam Detection, Machine Learning.

INTRODUCTION

In today's world, OSNs (Online Social Networks) have become a popular tool used by millions of people on the Internet to communicate and collaborate with each other [1]. Twitter is a major social networking platform that attracts users globally by offering free services like microblogging, allowing messages or tweets up to 140 characters, following other users and celebrities, etc [2]. It can be accessed on smartphones, personal computers, and tablets, with many users sharing their thoughts, emotions, news, and special moments through tweets. Despite its convenience, Twitter also attracts criminals such as spammers due to its easy accessibility [3]. Various attacks, such as scams, phishing, and spamming, can target Twitter in an unauthorized manner, as shown in Figure 1 [4].

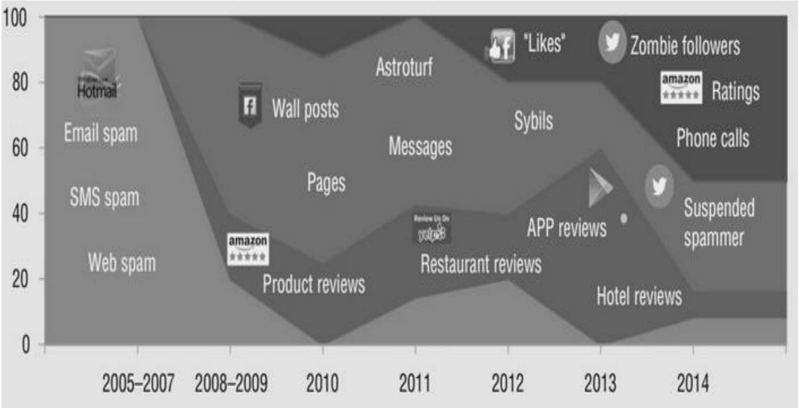

Figure 1. Different Types of Suspicious Attacks

Suspicious behaviour includes email, SMS, and web spam, along with suspicious reviews, messages, likes, followers, and pages. However, supervised learning approaches can be enhanced based on labelled examples availability [5, 6]. These approaches are beneficial for tasks like object detection, document analysis, and web page categorization, but labelling examples can be time-consuming and costly, requiring expertise.

Deep learning is an advanced form of supervised learning that resolves these issues effectively [7]. The decision tree classifier evaluates probabilities by distributing classes in the form of leaves on the branches of the tree, but this does not always improve the self-training algorithm's performance as desired [8]. Unlabeled data does not benefit from this algorithm, and predictions are made based on the class distribution at the tree's leaves, limiting the improvement in prediction ranking.

To address this issue, a deep learning technique is employ to detect the spam on Twitter. The proposed research employs adaptive feature neural network analysis to efficiently detect the spam. The proposed adaptive feature neural network analysis is evaluated by 600 million public tweets. The performance of the proposed research is assess by the performance metrics of TPR, FPR, and F-measure.

1.1. *Objectives*

- To detect the spam tweets by proposed adaptive feature neural network analysis.
- To minimize the error rate during detection.
- To efficiently categorize the spam and non-spam tweets.

2. LITERATURE SURVEY

This represents the prevailing approaches which have been examined and investigated by the various studies are discussed in this sections.

The study [9] has implemented spam detection to identify spam on Twitter's social media platform, a semi-supervised spam detection framework known as S3D has been used. This framework consisted two key components such as a detection of spam module which operated in a model update and real-time module which has worked in batch mode. Four lightweight detectors have implemented in the module to identify spams, such as the blacklisted domain detector marking tweets with blacklisted links and the near duplicate detector flagging suspected duplicate tweets. The study [10] has used reliable non-spam detector labelled tweets from trustworthy users without spam content, while the remaining tweets have classified by multi classifiers. The system updated data based on previously labelled tweets in the detection window. The study has attained satisfactory performance. Similarly, the study [11] has demonstrated the impact of uneven distribution between spam and ham classifies on the spam detection rate. To address this issue, a new oversampling technique based on fuzzy logic has approached to generate synthetic data samples from limited samples by fuzzy-based information decomposition. Additionally, an ensemble learning technique has developed to classification accuracy by learning more accurate classifiers in three steps, even with imbalanced data. Different techniques such as FOS, random under sampling, and random oversampling have used to adjust the class distribution within the imbalanced data in the first step. The study [12] employed a technique has been generated for classification based on each redistributed dataset in the following step. Finally, a voting technique has been employed to combine predictions from each classification model. The study has used data from recent tweets on Twitter for analysis, and the results showed that the approached learning technique has majorly improved the rate of spam detection in datasets with imbalanced distribution. The study [13] has included deceitful information found in unwanted Twitter spam. Approximately 500 million tweets have been collected, with around 6% classified as spam, and analysis was conducted on the data. It was discovered through the analysis that many deceptive elements in spam serve different purposes to lure individuals to harmful websites. The response rate to Twitter spam outbreaks varied by region, with some areas experiencing higher rates. The effectiveness of different methods for detecting spam was improved by the findings discussed earlier.

Likewise, the study [14] has made a comparison between a wider range of traditional ML techniques to determine the most effective in terms of performance. The stability of these techniques has also assessed using a large amount of ground truth data. Additionally, the scalability of the techniques has been evaluated to improve real-time detection of Twitter spam. Various performance measures such as detection accuracy, TPR, FPR and F1 – Score have analyzed. The constancy of the algorithms has been tested using randomly selected training samples of different sizes. The impact of parallel computing on scalability has been investigated to reduce the training and testing time of the ML techniques have been compared. Literally, the study [15] has addressed the problem of Twitter Spam Drift, an in-depth analysis has conducted on the statistical characteristics of approximately one million spam and an equal number of non-spam tweets. Followed by, a novel approach called Lfun has been introduced in the study. The revised spam tweets have identified from unlabelled tweets using the proposed method. These identified tweets have then incorporated into the classifier training phase. The effectiveness of the proposed method has assessed through multiple experiments. The results of the experiments suggest that the Lfun technique introduced in the study has the potential to enhance the accuracy of spam detection in practical scenarios. Literally, the study [16] has implemented a semi-supervised approach in order to identify spam on Twitter. This approach has used a specialised ensemble-based technique with four different classifiers. The technique utilised various Probabilistic Data Structures, including OF (Quotient Filter), to query databases containing URLs, undesirable users, spam, and LSH (Locality Sensitive Hashing). LSH has used to search for similarities and classifiers at diverse stages, providing quick results with minimal computational effort. The effectiveness of the approached technique has assessed through a comparative analysis of PDS with similar data structures, measuring performance metrics such as precision, recall, and F score. Likewise, the study [17, 18] has explored the imbalance in class distribution when detecting spams on Twitter. Initially, the study compared several well-known techniques designed to tackle class imbalance in order to identify the most effective one. Then, a comparison has made between these techniques and traditional approaches used for spam detection on Twitter. The results of the experiment showed that using ensemble learning with a fuzzy approach has significantly improved classification performance when using real Twitter data as the ground truth.

Correspondingly, the study [19] implemented a hybrid technique to discovered the spam profiles by the combination of bio-inspired computing and social media analytics. To identify spams in the Twitter market, an improved algorithm has used which combined K-means with LFA (Levy flight Firefly Algorithm) and incorporated chaotic maps (known as FA or Firefly Algorithm). Data from 14, 235 Twitter profiles, comprising 18, 44701 tweets, have been collected and analyzed based on 13 different statistically important features resultant from social media analytics. The method of fuzzy C-Means clustering has used to detect overlying users in two separate clusters of spammers and non-spammers. Six different categories of FA integrated with chaotic maps and K-means have tested along with the results showed that the approached FA with chaos rapidly congregates to a functional result when using the Gauss map method. Similarly, the study [20] has demonstrated an approach for identifying spam on Twitter. According to the existing study, traditional spam detection techniques have not been effective on Twitter due to its unique characteristics. To address this issue, the study has employed specific Twitter features to detect spam. Using Twitter's API, data from 77,033 tweets by 50,490 users data has been used to evaluate the existing study. Spammers have been distinguished from legitimate users by Naïve Bayes for training the spam detection system. The study has achieved an accuracy and sensitivity in a better way. Besides, the study [21] has implemented K-L divergence to indicate the distribution of spams, and potential drifts have identified by the Multi-scale Drift Detection Test (MDDT). The performance of the study has been improved when the base classifier has retrained based on the detection results. The experimental outcome has showed that when a drift occurred, the K-L divergence technique consistently showed changes in patterns across characteristics. As a result, the final classification results have found to be more potential and improved in terms of performance metrics such as accuracy, recall, and F1-score.

Moreover, the study [22] has suggested a hybrid INB-DenStream which has merges the principles of DenStream and INB (Incremental Naïve Bayes). The effectiveness of INB-DenStream has demonstrated through evaluation of complexity in computation, general precision, purity, general recall, parameter sensitivity and F1-score. A comparative analysis has carried out with existing techniques such as CluStream, DenStream, and StreamKM ++, which has shown that the proposed approach outperforms them in terms of performance. Likewise, the study [23] has demonstrated a sophisticated Intelligent Twitter Spam Detection System to identify Twitter spam that could accurately identify spam profiles. The study has employed as a hybrid classifier that takes into account specific feature sets before evaluating collected tweets and verifies links using the Google Safe Browsing API to enhance security. As a result, the system improved the classification of collected tweets and offered intelligent detection of Twitter spam. Literally, the study [24] has introduced a framework for uncovering the latent potential of a specified issue, the artificial neural BSF filter has replaced the code design to allow the analysis of data using random characteristics and multilayer perceptron on customer object communication function. This empowered the study to perform effectively. It is essential to train hidden data and ANSF before Matrix factorization. A combination of CMfact MPeep has utilized to explore the adaptable factorization and interaction feature. To handle a huge number of fuzzy sets, PCA (Principal Component Analysis) based on a non-parametric statistical approach has been employed for reducing dimensionality with fuzzy image sets. The suggested ML method proves to be valuable in identifying similarities among fuzzy image sets.

3. PROPOSED WORK

3.1 Proposed Flow

Many tweets are collected from one of the most famous online social media, Twitter to categorise them into spam and ham tweets. The Twitter Spam Dataset is used for gathering data, which is then pre-processed.

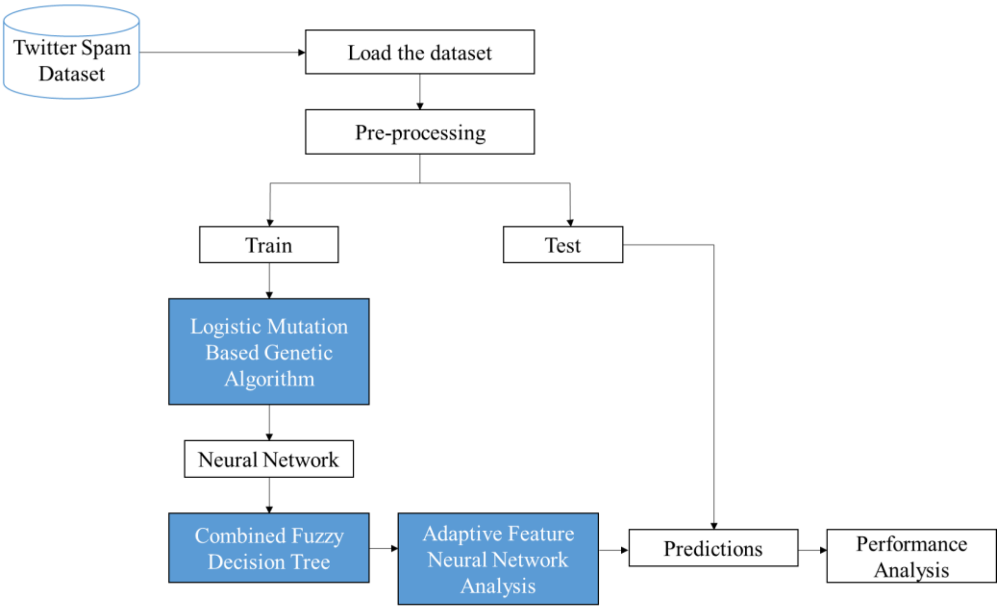
Twitter Spam Dataset
Load the dataset
Pre-processing
Train
Test
Logistic Mutation Based Genetic Algorithm
Neural Network
Combined Fuzzy Decision Tree
Adaptive Feature Neural Network Analysis
Predictions
Performance Analysis

Figure 2. The flow of the proposed work

Following to the pre-processing stage, the tweets are trained and tested. The training data is completed first, and then the trained tweets are input into neural networks. The proposed research employs logistic mutation based genetic algorithm to enhance the accuracy of the model and it leads to robust processing time when processing massive number of twitter dataset. These networks generate n number of NNs (Neural Networks), which are then associated with a fuzzy DT (Decision Tree) classifier. The classifier is then fused with adaptive feature neural network analysis. The tested tweets are evaluated using the proposed adaptive feature neural network analysis classifier to determine whether they are spam or non-spam. The predictions are made based on this classification. The flow of proposed research process is illustrated in Figure 2.

3.2 Proposed Adaptive Feature Neural Network Analysis

In this collaborative model, an adaptive feature neural network is initially employed with various aspects to identify spam tweets. The network contains an input layer, hidden layer and output layer. Every tweet is categorized into words. It utilizes various activation functions and combinations. The same training dataset is utilized in the Fuzzy Decision Tree, which is then normalized into real numbers within the range of 0 to 1. This approach incorporates potential fuzzy sets to achieve more adaptable membership functions. Each attribute in the FDT technique is linked with a fuzzy membership function. This combined technique enhances the efficiency and accuracy in detection of spam tweets. Adaptive feature NN mimics the processing of information by biological neurons, with learning through training and problem-solving through inference. Inference involves mapping input patterns to their respective output patterns. The application of adaptive feature neural network can vary based on its characteristics. In the kth iteration of the boosting process, the training set TR_k should be utilised. Let x_f represent the fth attribute, and $x_{f,i}$ as the value of x_f in the i-th sample of TR. All values of $x_{f,i}, i = 1 \dots N$, are assumed to be sorted in ascending order. If being a general interval on the universe U_f of x_f, and S_f being the set of examples in TR whose x_f values fall within I_f. A fuzzy partition over I_f is represented as $C_{I_f} = \{A_{f,1} \dots\dots A_{f,|C_{I_f}|}\}$, where C_{I_f} is the pth fuzzy set defined over a subinterval of I_f, and $\left|C_{I_f}\right|$ is the number of fuzzy sets in the partition. Each fuzzy set $A_{f,p}$ can be linked with the corresponding support set $S_{f,c}$: this is the subset of examples in TR that are part of $A_{f,p}$ or, more specifically, those which have a membership degree to $A_{f,c}$ greater than zero for the fth feature values.

$$S_{f,c} = \{X \mid X \in TR \; and \; \mu_{A_{f,c}}\left(x_f\right) > 0\} \qquad (1)$$

The support value, denoted as $|S_{f,c}|$, refers to the number of elements in the support set, and the fuzzy cardinality of $A_{f,c}$, denoted as:

$$\left|A_{f,c}\right| = \sum_{j=1}^{|S_{f,c}|} \mu_{A_{f,c}}(x_{f,j}) \qquad (2)$$

Where, $|A_{f,c}|$, is calculated as the sum of the membership values of the elements in $S_{f,c}$ based on their feature values. Each element is represented by $x_{f,j}$, where f is the feature number and j is the element number in $S_{f,c}$. The fuzzy entropy for a fuzzy set $A_{f,c}$ is determined as:

$$\mathrm{FEnt}(A_{f,c}) = -\sum_{m=1}^{M} \frac{|A_{f,c}, C_m|}{|A_{f,c}|} \; log_2 \left(\frac{|A_{f,c}, C_m|}{|A_{f,c}|}\right) \qquad (3)$$

Where $|A_{f,c}, C_m|$ represents the fuzzy cardinality of $A_{f,c}$ for elements with class label C_m. This equation calculates the relative weighted fuzzy entropy for all fuzzy sets in a fuzzy partition C_{I_f} over an interval I_f. The pseudo code 1 represents the training process of proposed adaptive feature neural network analysis.

Pseudo Code 1: Training process of proposed adaptive feature neural network analysis

```
$X \rightarrow Input\ Training\ set$
$y \rightarrow Testing\ set\ Labels$
$f \rightarrow weights\ of\ layers$
$l \rightarrow no\ of\ layers\ in\ neural\ network\ 1..L$
$e_{i,j}^{(l)} \rightarrow error\ for\ all\ l, i, j$
$t_{i,j}^{(l)} \rightarrow 0\ .For\ all\ l, i, j$
$For\ i = 1\ to\ m$
  $a^l \rightarrow feedforward\ (x^{(i)}, f)$
  $d^l \rightarrow a(L) - y(i)$
  $t_{i,j}^{(l)} \rightarrow t_{i,j}^{(l)} + a_j^{(l)} . t_i^{l+1}$
$if\ j \neq 0\ then$
   $e_{i,j}^{(l)} \rightarrow \frac{1}{m} t_{i,j}^{(l)} + \lambda f_{i,j}^{(l)}$
$else$
```

4. PERFORMANCE ANALYSIS

This section explained the different performance metrics like accuracy, TPR, FPR, and F-measure. In the experimental findings, the proposed research employs a vast dataset. A total of 600 million tweets with URLs were gathered. Through the WSR service, the proposed research able to determine if a URL was malicious. Within this extensive dataset, 6.5 million malicious tweets were identified, making up around 1 % of all tweets.

4.1 Performance Measures

Performance evaluation involves determining the values of different performance metrics such as sensitivity, accuracy, precision, recall, F1-score, misclassification rate, and Jaccard-coefficient.

4.1.1. Accuracy

Accuracy is a measure of correct values and is referred to as the reciprocal of precision. It is evaluated by evaluated by equation (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (4)$$

4.1.2. TPR

TPR refers to the total number of correct positive results achieved for every positive sample available during analysis.

4.1.3 FDR

Similarly, FDR signifies the total number of incorrect positive results achieved for every negative sample available during analysis.

4.1.4. F1-measure

F-measure, also known as F1 score which measures by associates accuracy and recall using a specific formula.

$$F1\ Score = \frac{(2 * Precision * Recall)}{(Precision + Recall)} \qquad (5)$$

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

4.2 EDA (Exploratory Data Analysis)

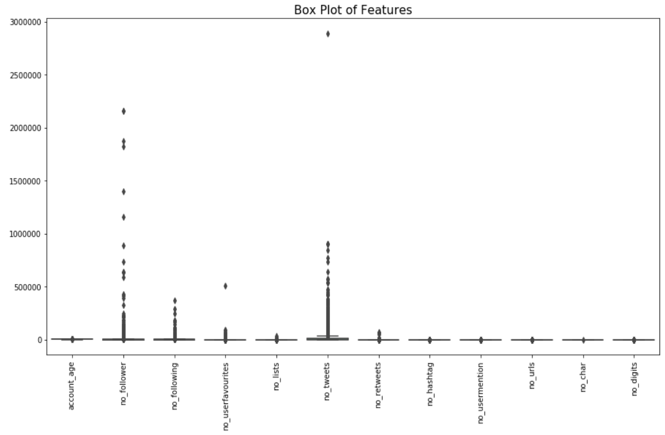

Figure 3. Box Plot of Features

Figure 3 depicts the box plot of features. The plots shows different features such as account_age, no_follower, no_following, no_userfavourites, no_lists, no_tweets, no_retweets, no_hashtag, no_usemention, no_urls, no_char and no_digits. Along with, the plot shows less outliers. Some of the major outliers in the plot are no_followers with maximum count of followers with more than 2,000,000. In no_tweets, represents tweet counts of 3,000,000. The least count is no_retweets.

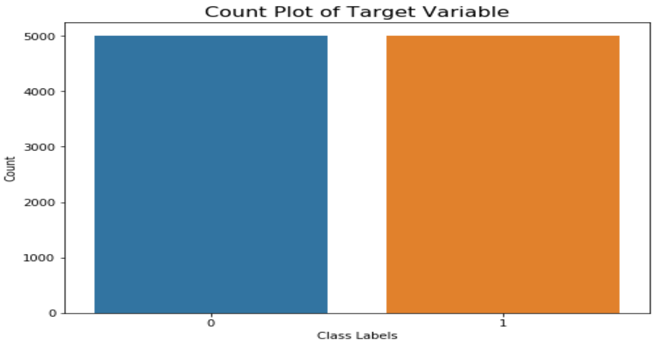

Figure 4. Count Plot of Target Variable

Figure 4 depicts the count plot of target variable. The x-axis shows the class labels of 0 and 1and the y-axis shows the count from 0 to 5000. Both class 0 and class 1 shows similar count of 5000, which represents that the proposed model class is balance. Thus, balanced dataset minimizes the risk of the bias during prediction particularly when using accuracy as performance metrics.

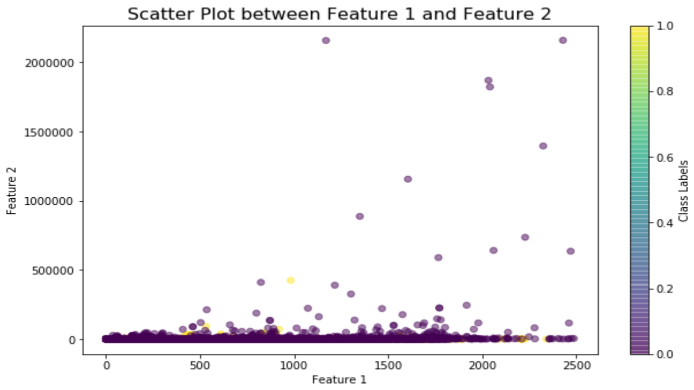

Figure 5. Scatter Plot between Feature 1 and Feature 2

Figure 5 depicts the scatter plot between feature 1 and feature 2. The x-axis shows the feature 1 and the y-axis shows the feature 2. The range of the plot is starts from 0.0 to 1.0. When compare to feature 2, feature 1 distributed evenly from the range of 0 to 2500. Feature 2 represents low values and its plots are low range and near to 0. While, the feature 2 shows extensive outliers. In feature 2, the outliers might be leveraging the distribution and this plot shows the relationship between feature 1 and feature 2 that might be related to predictive modelling.

4.3 Performance Analysis

Results have been achieved for several performance metrics such as precision, TPR, FPR, and F-measure, and these results are compared against multiple established classifiers to demonstrate the effectiveness of the proposed research.

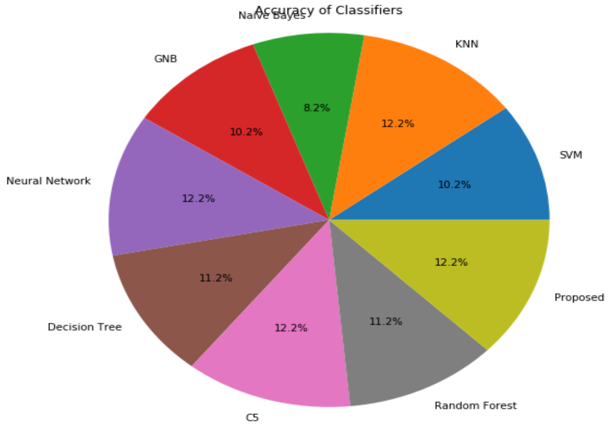

Figure. 6. Performance of Dataset 1

The accompanying Figure 6 illustrates the accuracy values, particularly for dataset 1. The precision of the proposed research is evaluated and contrasted with different established methods. It is evident from the comparison that the proposed technique yields the best accuracy score of 0.95.

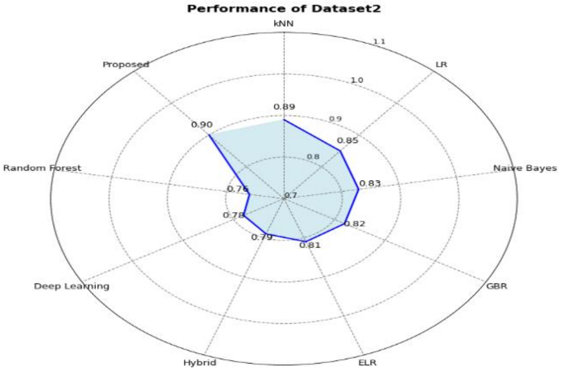

Figure 7. Performance of Dataset 2

Figure 7 depicts the accuracy of dataset 2. It shows the levels of precision, particularly for dataset 2. The accuracy of the suggested approach has been evaluated and contrasted with different current methods. Following the comparison, it is evident that the method described in the study achieves the highest accuracy rating of 0.94.

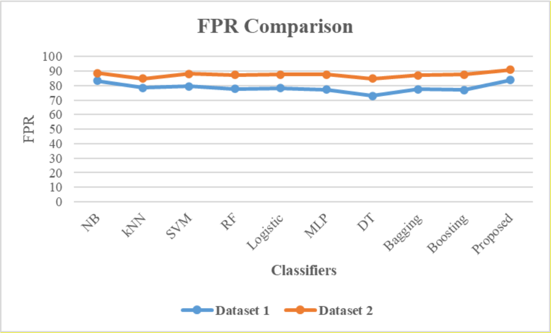

Figure 8. FPR Comparison with Dataset 1 and Dataset 2

Figure 8 depicts the comparison of FPR. It shows the FPR values for both datasets 1 and 2. The proposed work's False Positive Rate (FPR) is calculated and then compared to existing methods. The comparison reveals that the proposed technique has a lower FPR of 6 and 5 for dataset 1 and dataset 2, respectively, when compared to other methods.

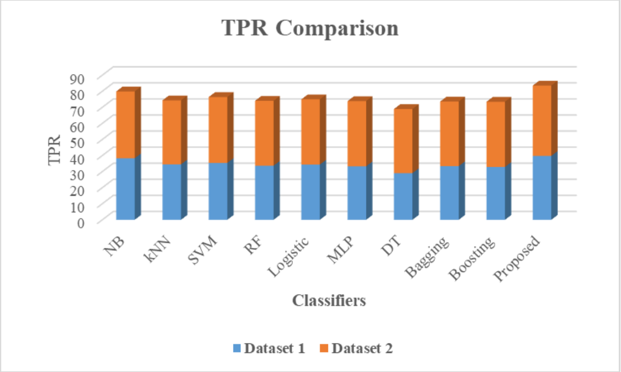

Figure 9. TPR Comparison with Dataset 1 and Dataset 2

Figure 9 depicts the comparison of TPR. It shows the TPR values which represent the assessment standards for the datasets provided (both 1 and 2). The TPR of the new method is calculated and then contrasted with different current approaches. Upon comparison, it is evident that the new technique achieves best TPR values of 92 and 93 for dataset 1 and 2 respectively.

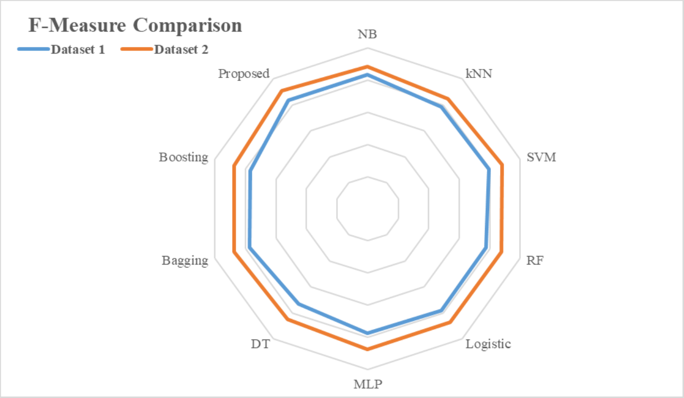

Figure 10. F-Measure Comparison with Dataset 1 and Dataset 2

Figure 10 depicts the comparison of F-Measure. The F-measure values show that for both datasets 1 and 2, the proposed work achieved lower F-measure values of 93.89 and 88.56, respectively, compared to other existing techniques.

6. CONCLUSION

Several social media platforms are currently accessible on the network, including Instagram, Twitter and Facebook to connect with others. Among these, Twitter has emerged as a significant platform. Users on Twitter share a wide range of content, including tweets, thoughts and articles. Due to the availability of advanced APIs which enable Twitter to access and update data, the platform appeals spammers of different kinds. The tweets gathered in this proposed research has classified using the Innovative ANN-FDT algorithm. This approach has helped in distinguishing between spam and ham tweets. The effectiveness of the new algorithm has evaluated using metrics such as precision, TPR, FPR and F-measure. The findings suggest that the algorithm's performance improved with the new technique.

REFERENCES

[1] S. B. Abkenar, M. H. Kashani, M. Akbari, and E. J. Mahdipour, "Learning textual features for Twitter spam detection: A systematic literature review," *Expert Systems with Applications*, vol. 228, p. 120366, 2023.

[2] M. Arunkrishna and B. J. Mukunthan, "Evolutionary Traits In Digital Spam: History, Types, Techniques and Anti Spam Solutions," *International Journal of Advanced Science Technology* vol. 29, no. 3s, pp. 825-834, 2020.

[3] B. J. Mukunthan, "Detection of Malicious Data in Twitter Using Machine Learning Approaches," *Turkish Journal of Computer Mathematics Education,* vol. 12, no. 3, pp. 4951-4958, 2021.

[4] N. Sun, G. Lin, J. Qiu, and P. J. Rimba, "Near real-time twitter spam detection with machine learning techniques," *International Journal of Computers Applications,* vol. 44, no. 4, pp. 338-348, 2022.

[5] J. Tanha, M. Van Someren, and H. J. I. J. o. M. L. Afsarmanesh, "Semi-supervised self-training for decision tree classifiers," *International Journal of Machine Learning Cybernetics,* vol. 8, pp. 355-370, 2017.

[6] M. Arunkrishna and B. J. Mukunthan, "Review on Classification of Anti-Spam Solutions: Approaches, Algorithms Demystified," *Studies in Indian Place Names,* vol. 40, no. 60, pp. 4449-4458, 2020.

[7] A. Ghosh and A. J. Senthilrajan, "Comparison of machine learning techniques for spam detection," *Multimedia Tools Applications,* vol. 82, no. 19, pp. 29227-29254, 2023.

[8] M. Arunkrishna and B. J. Mukunthan, "Applicability of Machine Learning in Spam Detection Systems," *Annals of the Romanian Society for Cell Biology,* pp. 12800-12808, 2021.

[9] N. Imam, B. Issac, and S. M. J. Jacob, "A semi-supervised learning approach for tackling Twitter spam drift," *International journal of computational intelligence applications,* vol. 18, no. 02, p. 1950010, 2019.

[10] B. S. M. Arunkrishna, "Efficient Spam Detection on X (formerly Twitter) : A Hybrid Artificial Neural Network and Fuzzy Decision Tree Approach," *Journal of Propulsion Technology,* vol. 45, 2024.

[11] C. Zhao, Y. Xin, X. Li, Y. Yang, and Y. J. Chen, "A heterogeneous ensemble learning framework for spam detection in social networks with imbalanced data," *Applied Sciences,* vol. 10, no. 3, p. 936, 2020.

[12] A. T. Kabakus and R. J. Kara, "“TwitterSpamDetector”: a spam detection framework for Twitter," *International Journal of Knowledge Systems Science,* vol. 10, no. 3, pp. 1-14, 2019.

[13] L. Das, L. Ahuja, and A. Pandey, "Analysis of twitter spam detection using machine learning approach," in *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)*, 2022, pp. 764-769: IEEE.

[14] M. J. I. J. o. I. T. Diqi, "TwitterGAN: robust spam detection in twitter using novel generative adversarial networks," vol. 15, no. 6, pp. 3103-3111, 2023.

[15] H. Tajalizadeh and R. J. Boostani, "A novel stream clustering framework for spam detection in Twitter," *IEEE Transactions on Computational Social Systems,* vol. 6, no. 3, pp. 525-534, 2019.

[16] K. K. Devi and G. J. Kumar, "Stochastic Gradient Boosting Model for Twitter Spam Detection," *Computer Systems Science Engineering* vol. 41, no. 2, 2022.

[17] B. Mukunthan and M. Arunkrishna, "Spam detection and spammer behaviour analysis in Twitter using content based filtering approach," in *Journal of Physics: Conference Series*, 2021, vol. 1817, no. 1, p. 012014: IOP Publishing.

[18] N. El-Mawass, P. Honeine, and L. J. Vercouter, "SimilCatch: Enhanced social spammers detection on twitter using Markov random fields," *Information processing management,* vol. 57, no. 6, p. 102317, 2020.

[19] K. S. Adewole, T. Han, W. Wu, H. Song, and A. K. J. Sangaiah, "Twitter spam account detection based on clustering and classification methods," *The Journal of Supercomputing,* vol. 76, pp. 4802-4837, 2020.

[20] A. T. Kabakus and R. J. Kara, "“TwitterSpamDetector”: a spam detection framework for Twitter," *International Journal of Knowledge Systems Science,* vol. 10, no. 3, pp. 1-14, 2019.

[21] X. Wang, Q. Kang, J. An, and M. J. Zhou, "Drifted Twitter spam classification using multiscale detection test on KL divergence," *IEEE Access,* vol. 7, pp. 108384-108394, 2019.

[22] H. Tajalizadeh and R. J. Boostani, "A novel stream clustering framework for spam detection in Twitter," *IEEE Transactions on Computational Social Systems,* vol. 6, no. 3, pp. 525-534, 2019.

[23] S. S. Vellela, A. Chaganti, S. Gadde, P. Bachina, and R. J. Karre, "A Novel Approach for Detecting Automated Spammers in Twitter," *Mukt Shabd,* vol. 11, pp. 49-53, 2022.

[24] P. Sivakumar, M. Balasubramani, R. Sowndharya, B. D. Priya, W. D. Priya, and M. J. Syamala, "Twitter spam drift detection by semi supervised learning approach using YATSI algorithm," *International Journal of System Assurance Engineering Management,* pp. 1-9, 2024.