International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Designing A Model for Fake News Detection in Social Media Using Machine Learning Techniques

A Kumari Shalini [1], Dr.Sameer Saxena[2], Dr. B Suresh Kumar[3]

Submitted: 22/10/2022 **Revised:** 27/12/2022 **Accepted**: 28/01/2023

Abstract: Fake identity is a critical problem nowadays of social media. Fake news is rapidly spread by fake identities and bots that generate the trustworthiness issue on social media. Identifying the profiles and accounts using soft computing algorithms is necessary to improve the trustworthiness of social media. In this work, we proposed Recurrent Neural Network to identify fake identities on social media. Initially, we extract data from social media such as Twitter.com using Twitter API. Hybrid feature extraction has been done based on the characteristics of data. It generates training rules which are associated with a fake and legitimate profiles generated by a human. In pre-processing and filtration process, all bot entries are eliminated using a policy-based approach. To generate strict rules that improve the classification accuracy, the training of dataset primarily focuses on attributes such as friends count, the total number of followers, tweet counts, re-tweets count, etc., The Recurrent Neural Network (RNN) categorizes each profile based on training and testing modules. This work focuses on classifying bots or human entries according to their extracted features using machine learning. Once the training phase is completed, features are extracted from the dataset based on the term frequency on which the classification technique is applied. The proposed work is very effective in detecting malicious accounts from an imbalanced dataset in social media. The system provides maximum accuracy for the classification of fake and real identities on the social media dataset. It achieves good accuracy with Recurrent Neural Network (RNN) using the different activation functions. The system improves the classification accuracy with the increase in the number of folds in cross-validation. In experiment analysis, we have done testing on synthetic and real-time social media datasets; We achieve around 96% accuracy on the real-time Twitter dataset while 98% accuracy on synthetic social media datasets.

Keywords*: Social media, fake news detection, machine learning, recurrent neural network, , feature extraction and selection, classification*

1. Introduction

The social networking sites like Facebook, Twitter is the most significant ways of internet communication and collaboration. Fake information spreading by bots is another major problem of social media, around 30% of posted information is fake or bogus every day from malicious web applications or bots [4]. So, it is important to improve the trustworthiness of social media by detecting fake news and bots timely. Analysing user's profiles information and identifying its trustworthiness using various soft computing techniques is the way of eliminating fake news distribution. To commit numerous cybercrimes such as profile hacking, identity hacking, session hijacking, malicious linking, mail bombs, and so on helps criminals build false identities. Mostly bots or humans may create such kinds of fake identities. In general, the fake identities of bots target large numbers of individuals on social media. Fake information or accounts can spread forged information rapidly without any verification policy, which is the big drawback of social media.

This work also demonstrates a feature of Fake news detection that classifies the news article as trustworthy or fake. We examine fake news identification using various models and classifiers and predict unconventional models and classifiers' efficiency. We check which prototype will give more precision and incorporate the news into real or fake. We also produce computational sources and illustrations for the work of fake news apprehension. We recommend a Twitter dataset, which consolidates real and fake news and implements that dataset using various methods like machine learning, natural language processing, and deep learning. Using these datasets, we examined distinct exploratory analyses to distinguish semantic properties broadly present in unreliable content. We furthermore correlate our fake news identification standards precision with expected accuracy to put our consequences in aspect. We use Natural Language Processing, machine learning, and deep learning procedures to complete our models and distinguish which models will give higher efficiency. Fake accounts can be created by humans, computers, or even cyborgs. A cyborg

1Research Scholar
Amity University, Jaipur shalinijaiswal.c@gmail.com
2Associate Professor
Amity University, Jaipur saxena1@jpr.amity.edu
3Associate Professor
Sanjay Ghodawat University,Kolhapur
sureshkumarbillakurthi@gmail.com

account is half-human and half-bot. A person physically initiates the account, but a bot handles all subsequent activities. There are differences between bots versus human profiles. When profiles are fake and not hacked by authorized users, bots are referred to as Sybil profiles. On the other hand, fake human identities are referred to as "trolls" when used to marginalize another user's reputation. In this paper, we examine whether readily accessible and designed features that have been effective in detecting false documents created by bots or machines using machine learning algorithms can also be used to detect fake documents created by humans.

2. Literature Survey

According to Estee et al. [1] trained the classifier by applying used features for bot detection in order to identify fake accounts created by the human on Twitter. The training is based on supervised learning. They have tested for three different classifiers, i.e., Support Vector Machine (SVM) with linear kernel, Random Forest (RF) and Adaboost. For SVM classification, the SVM linear library in R software is used. Here the boundary based on feature vectors is created for classification. For the RF model, the RF library in R software is used. RF model creates variations of trees,

and mode of class outcome is used to predict identity deception. For boosting model, the Adaboost function in R is used. Adaboost is used along with decision trees where each feature is assigned a different weight to predict the outcome. These weights are iteratively adjusted, and output is evaluated for the effectiveness of identity deception prediction at each iteration. This process is repeated until the best result is obtained. Among these three classifiers, RF reached the best result.

The topic of knowing and manipulating social media user accounts for fake news detection According to [2]. First, the system analyses user sharing habits and community representation users that are more likely to spread false and real news to explain similarities between user-profiles and fake news; then the system conducts a comprehensive evaluation of formal and informal important to remind between such user groups, revealing their ability to help discern fake news against real news. In a false news classification mission, the framework shows these user account features utility to manipulate profile page features. The framework further legitimizes the use of these characteristics by evaluating the value of functionality. This work lays the groundwork for further analysis of online user profile functionality and improves the ability to identify fake news.

According to [3], a supervised neural multi-thread model can use either textual content and user contact lists to jointly classify the integrity of the theory and users positions. Tests on two datasets rumors datasets show that Rumor Detective outperforms present condition models and enjoys up to 14% performance benefit in the classification of rumor veracity and approximately a 6% increase in the percentage of user stance. To jointly know the speculation class and roles, a micro-learning strategy. Inter learning involves learning multiple tasks that exploit common mechanisms and communications together.

Information from user posts on two common social media platforms was collected in [4] template files: Twitter and Facebook. A user's depression level was identified based on his social media messages. Structured research or quasi-interview procedure (SDI) [1] is the traditional way of identifying a person's depression. Such approaches require a tremendous amount of knowledge from the individual. Mobile messaging sites such as Twitter and YouTube have become widespread platforms for sharing people's activities and opinions. Statistical analysis from tweets and posts shows the presence of the user's symptoms of major depression. Quantum computing is used in this study to process the discarded data obtained from SNS users. Natural Language Processing (NLP), categorized to diagnose depression more conveniently and accurately using the SVM and Naïve Bayes algorithm.

Perform a three-phase technique to detect press farming According to [5]. The method begins by clumping communities based on newly developed networks of collaboration. The framework then applies the method of Louvain group detection to community detection. Finally, the system performs a binary classification designed to detect. These findings from over a year-long study indicate that the frequency of click farming differs across GCSEs; most click farmers are low-rated; click-farming groups have relatively similar user relationships; more high-rated stores have a higher percentage of fake feedback. The creation of a web search to test to equal thickness of store pages and user pages. The success of GCSEs has gotten the interest of security professionals over the past couple of years. Review-spam identification can be regarded as an issue with classification problems or rating. There are many methods of detection supported by previous studies.

A model was suggested in [6] to explain social information pressure when individuals attempt to communicate their feelings through social networking tweets or emails. The Twitter datasets can be used for user activity and the Recognition of content by CNN for learning. For classification, CNN is being used. Since the identification of big data files with a standard method is very complicated, the practice file is represented using social media platforms. Class 0, Class1, and Class2 of the tweet application test. Class 0 indicates a positive level of

stress, Class 1 demonstrates unfavorable stress levels, and Class 2 refers to moderate stress.

According to [7], the interest in identifying false profiles and researching their activities has increased. Questions such as who the impersonators are? What are their distinctive features? And are they robots? It is going to emerge. To react, the framework begins this research by gathering data on Instagram from three significant groups, including "Politician," "News Agency," and "Sports Star." Four top checked accounts are selected within each group. The device detects 4K based on the users who replied to their published posts.

The key principles of anomalies in online communities were analyzed and implemented According to [8]. Also, sent the various types of occurrences and the category of occurrences is divided into four groups, depending on the presence of both the anomalies, based on the dynamic existence of the structure of the graph, based on the knowledge contained in the structure of the graph, based on facts, based on functional procedures in the structure of the graph and based on the pattern of relationships in the structure of the graph. The framework also illustrated the techniques for detecting irregularities. The research study presented an existing study that focused precisely on detecting irregularities in social media platforms via the Internet. The best method of identifying deviations in OSN is based on the form of irregularities by previous researchers.

According to [9], It creates a nationwide networking atmosphere where persons groups enjoy their company and activities or are involved in those other people's choices, interests and hobbies. While the messaging service has offered tremendous benefits to individuals, it is damaging individuals through various devious activities taking place through social media. This allows society to suffer major economic losses and also undermines global defense. Facebook, Twitter, Snapchat, etc., are all extremely vulnerable to ransom ware activities on social networks. Twitter is one of the largest media websites for online media, with over 2 billion tweets being posted to Facebook on aggregate every day by millions of users. Twitter is quickly intruded into disruptive operations with such flexibility and a large spread of using it. Malicious operations include the infiltration of malware, spam delivery, social attacks, etc. Spammers use the attack technique of social manipulation to deliver spam messages, spam URLs, etc. These have made twitter a perfect arena for unusual fake accounts to increase. The effect enables researchers to construct a model that analyzes, identifies, and recovers Twitter against untruthful behavior. The Twitter network is flooded with thousands or even millions of fake spam accounts that can jeopardize regular user's security and privacy. Improving the protection of legitimate customers and detecting spam profiles are the main components of the report.

A deep learning model structure that uses checkered past (RNN-LSTM) with template decoding for joint purpose identification and slot filling simulation, according to [10]. These results suggest significant progress over state-of-the-art conceptual frame investigative techniques in slot filling and intention analysis at the sentiment analysis. Linguistic frame identification has been commonly used in Chat-bots for language comprehension tasks, such as speech processing or, more commonly. Models are based on deep neural networks (RNN) or frameworks that impact the long storage used in more previous stuff (LSTM). The role of detection of intent is closely related to the issue of slot filling. At the level of the sentence, it is analyzed, which could be seen as an issue with classification. The source is the same in SF, except there is only feedback describing the mark of each paragraph's intention.

human rights-based approach to detecting automatic fraudsters by combining neighborhood functionality with the other features in [11] is documentation, information, and communication features. The newness of the suggested solution is to classify users based on their relationships with certain followers, provided that a user can avoid characteristics related to his behaviors. Still, it is difficult to avoid those based on adherents. Seventeen features describe, including basic behavioral defined characteristics and two new added features.

In the hope of progressing the effective identification of false identities invented by nature on Social Media Platforms (SMP), these same good occupational refer to a collection of fake human accounts, thus according [1]. Although very little has been done to recognize a real fake host of activities on SMPs, the framework looked at previous research solving similar issues. For obvious reasons, spam behavior discovered in emails and SMS shows similar bad intentions with fake comments propagating false rumors. Filtering is mostly reactive: the sender will be added to a blacklist only when a new threat is identified and verified. To blacklist recognized network Intrusion information and confinement are known hackers, similar techniques of dealing with spam have also been proposed on Twitter. In addition to filtering methodologies, rules for identifying fraudulent identities during identification have been developed. Examples of such regulations are predicated on words that are recognized to belong with junk mail and within texts.

Twitter [12] is used for reasons such as disinformation and the development of a plan. That is one of the fundamental challenges facing social media networks. The identification of suspicious accounts is, therefore, important. Methods focused on machine learning have

been used to identify fake accounts that could confuse individuals. For this reason, the produced schema was well before, and machine supervised learning determined the fake accounts. For the identification of fake accounts, random forests, regression models and back propagation machinery architectures are used. These classification outputs are measured, and the regression model has proven more effective than all the others.

An identification method was suggested According to [13] that can recognize fake and clone Retweet profiles. Based on several guidelines that can efficiently distinguish fake and authentic profiles, Facebook accounts are identified. For, methods are used for Profile Replication Detection. One uses tests of similarity, and the other uses the decision tree of C4.5. Main methods of correlations are considered in the Clustering Algorithm: Similarity of Features and Similarity of Social networks. C4.5 distinguishes clones by growing the tree by considering the benefit of knowledge. A contrast is made to verify how well these two approaches assist in recognizing clone profiles. Fake profiles and clones have become a very significant social threat. An attacker can easily hack such data and create false or clone profiles because information such as telephone number, email ID, school and college name, business name, location etc., are widely spotted on social media platforms.

Thematic analysis for [14] uses only accounts messages. Further classification is also provided by the system that distinguishes fake followers and spammy links from actual accounts. Max reliability in fraudulent vs. genuine accounts classification using TF-IDF functionality and GMM algorithm has been achieved in this study at 95.55% and 95.2% in all three forms of accounts using Word2Vec tools and XGBoost algorithm. Almost all of these studies are already achieving outstanding results, ranging from 85% to 99% accuracy. While mostly related to the common form of activity, these studies differ, be it fake followers or automated spam, how system collate and accomplish the malicious account data. Some of them, such as spam messages and automated spam, submit malicious links. A wider description is used, such as accounts that continue to retweet Italian applicants or work ads.

According to [15] contributed to identifying false and programmed accounts, leading to Instagram fake participation. As far as the machine knows, false and programmed accounts do not have a publicly accessible dataset. Two datasets for the monitoring of fraudulent and artificial accounts have been created for this purpose. Neural networks such as Naive Bayes, logistic regression, fully convolutional machines and deep learning are implemented to detect such accounts. In comparison, because of the sample's undesirable bias, a premium evolutionary algorithm is used to identify automated accounts. This method is used to deal with the inhomogeneity question in the fake database. For the problem of automatic and fake accounts detection, 86% and 96% are received, accordingly.

3. Proposed System Design

The proposed system architecture demonstrated Figure 1, with evaluation of fake news using machine learning techniques. This system uses datasets from numerous social networking websites to identify bogus accounts across the full dataset. The Twitter Application Programming Interface (API) is used during the first stage of data collection by the system. This takes data from tweet comments that have been recently accessed by users. The inability of social media programmes to identify automated accounts or fraudulent profiles is the primary issue. In this study, we combined the natural language processing (NLP) and machine learning (ML) techniques in an effort to resolve the aforementioned problems that are present in existing systems. To begin, we collect information from a variety of social media sites. After the data was successfully transferred, it was put away in the appropriate data sources and given data files. Since the data came from a variety of internet sources like Twitter, it is reasonable to expect that some of it will be fragmented. It is required to perform a preliminary processing step on this data using a certain sampling method and data filtration strategies. The method of sampling design is what is done here.

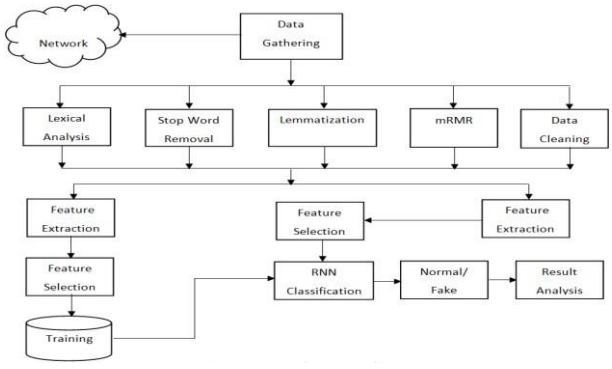

Fig. 1: proposed system architecture for fake news detection on social media

Pre-processing: During the pre-processing phase, the data were checked for validity using the established guidelines and regulations that were drafted by us. Each attribute has a minimum and maximum limit for the values it can take; whenever any of these limits are exceeded or violated, the system will immediately get rid of any instances that match those criteria. The complete pre-processing encompasses data collecting, data acquisition, data cleaning, data filtrations, normalizations, etc.

Data Cleaning: Data cleaning is a process to detect and change missing, wrong, inaccurate, or irrelevant data and then to duplicate, edit, or delete filthy or delicate information. False or incomplete assertions can be removed from records, tables, or systems using this method. The interactive use of scripting software or transaction processing are both viable options for the cleanup of data. The data have been filtered using the purposeful sampling approaches, and after the filtration process is complete, the data are balanced and free of the different classifiers that were present in the normalised dataset.

Feature Extraction: The following provides an explanation of each of the three distinct methods of extracting the features that have been implemented. The features extracted from input data such as lemma-based features, Rule based features and Hybrid collective features.

Feature Selection: Once feature extraction has been done, using few quality thresholds, we optimized a feature set called feature selection. The weighted term frequency technique has been used to optimize features and forwarded to the training module.

Classification: Finally, the system detects each transaction, either fake or real using a supervised classification technique. Moreover, the system also demonstrates a fake news classification of social media datasets. In this work, we carried out RNN and LSTM as a supervised classification algorithm. Then supervised machine learning is applied to train the classifier. Here class labelled data is present at the beginning. RNN algorithm for Deep learning is applied to determine fake news identification on social networks where multiple decision trees are created using randomly selected features from the feature set, and the majority output class of all the decision trees is taken as output of the RNN.

ALGORITHM DESIGN

Classification using Random Forest Algorithm

Input: Extracted features of testing instances set Data [i..... n], Train data policies PSet[1] T[n]

Output: Fake or Real profile.

Steps:

1. For each (Data [i] into Data) choose n attributes randomly from Data [i] using below formula,

$$Treeset(k) = \sum_{k=1}^{n} attribute\ [D[i]k \ldots . . D[n]n]$$

2. For each (PSet [i] from PSet),

$$Train[m] = \sum_{m=1}^{n} attribute\ [T[i]k \ldots . . T[n]n]$$

3. Evaluate train and test instances using below formula,

$$Treeset[k].weight = similarity(Treeset[k] \sum_{m=1}^{n} train\ [m]$$

4. If (Treeset[k]: weight > Th),
   Treeset[k].class □ Train[m]:class
   Break;
5. return Treeset[k].class

Recurrent Neural Network

Training

At the end of the day, the task-specific meanings that were produced by all of the indicated multipurpose components are fed into a number of fully connected layers which are also undertaking.

$$\hat{y}(m) = softmax\_(W(m)\ h(m) + b(m))$$

Here $\hat{y}(m)$ implications of extrapolating for the work process m, W(m) is the frequency that one must know in order to proceed, and b(m) is a bias time. Our all-encompassing cost function can be conceptualised as an unique sequence of cost usefulness for all crossings..

$$ø = \sum_{m=1}^{M} . \lambda_m)$$

Now, λ_m is the respective weight for specific (m) task.

It should be noted that labeled data will come from entirely different repositories for the training of each task. The teaching is then done in a deterministic fashion by looping over the assignments:

1. Choose a job at random.
2. From this assignment, pick a random training example.
3. Update the specifications for this task with regards to this illustration by making a gradient phase.
4. Oh, go to 1.

Therefore, we should use a calibrating approach to further improve the output with each task after the mutual learning stage.

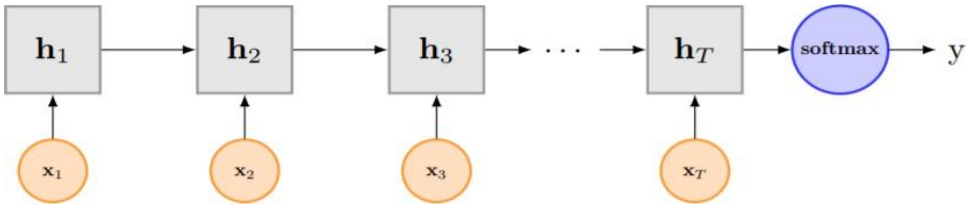

Fig. 2: Recurrent Neural Network for Classification

Algorithm for system testing

Input: Normalized training dataset *Train_Data[]*, Normalized testing dataset *Test_Data[]*, defined threshold *qTh*

Output: Result set as output with *{Predicted_class, weight}*

Step 1: Read all test data from *Test_Data[]* using below function for validating to training rules, the data is normalized and transformed according to algorithms requirements

$$test\_Feature(data) = \sum_{m=1}^{n} (Attribute\_Set[A[m] \dots\dots A[n] \quad Test\_Data)$$

Step 2 : select the features from extracted attributes set of $test\_Feature(data)$ and generate feature map using below function.

$$\text{Test\_FeatureMap [t}\dots\dots\text{ n]} = \sum_{x=1}^{n} (t) \square\ test\_Feature(\text{x})$$

$Test\_FeatureMap\ [x]$ are the selected features in pooling layer. The convolutional layer extracts the features from input and passes to pooling layer and those selected features are stored in *Test_FeatureMap*

Step 3: Now read entire taring dataset to build the hidden layer for classification of entire test data in sense layer,

$$train\_Feature(data) = \sum_{m=1}^{n} (Attribute\_Set[A[m] \dots\dots A[n] \quad Train\_Data)$$

Step 4 : Generate the training map using below function from input dataset

$$\text{Train\_FeatureMap [t}\dots\dots\text{n]} = \sum_{x=1}^{n} (t) \square\ train\_Feature(\text{x})$$

$Train\_FeatureMap[t]$ is the hidden layer map that generates feature vector for build the hidden layer. That evaluate the entire test instances with train data.

Step 5 : After generating the feature map we calculate similarity weight for all instances in dense layer between selected features in pooling layer

$$Gen\_weight = CalcWeight\ (Test\_FeatureMap \,||\, \sum_{i=1}^{n} Train\_FeatureMap[i])$$

Step 6 : Evaluate the current weight with desired threshold

$$if(Gen\_weight >= qTh)$$

Step 7 : $Out\_List.add\ (trainF.class, weight)$

Step 8 : Go to step 1 and continue when Test_Data == $null$

Step 9 : Return Out_List

RESULTS AND DISCUSSION

In another investigation, the probability of fake news detection using supervised machine learning classification using machine learning techniques. In the first section, we downloaded data from the Twitter account using Twitter API; around 2000 samples evaluated the proposed model using supervised learning algorithms. The three data splitting mechanism has use as 5, 10 and 15-fold cross-validation.

Table 1: Dataset description downloaded using twitter API

Total Size	**2000**
Training Samples	**1400**
Testing Samples	**600**

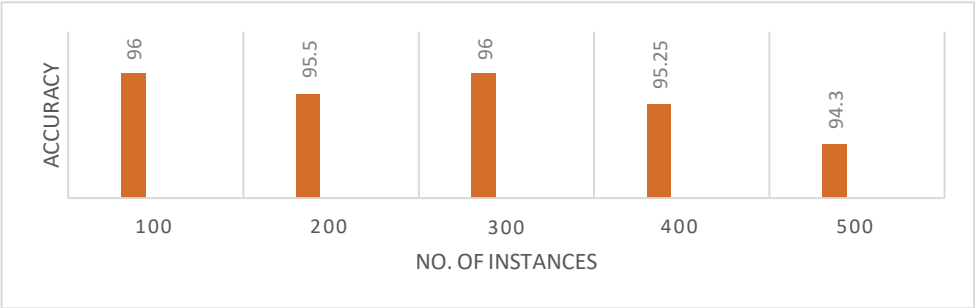

Fig. 3: Accuracy of System with No. of Events

Figure 3 illustrates the system's prediction accuracy with various numbers of samples; the detection accuracy has been evaluated using a modified Naïve Bayes classification algorithm. The most current expected sample uses a planning set and a test set for organization or grouping. Input function modules and their associated class labels are composed of the training package. An classification model is generated using this arrangement or learning set to organize the input courses into corresponding template files or labels. Then a test set is used by gleaning the class labels of orthonormal courses to validate the model. A variety of neural networks are used to identify reviews, such as Simple Bayes (NB), NLP, and Support Courses Machines (SVM) [6]. Word Absence or Presence, Term Replication, inadequacy, n-grams, and Sections of Expression can be used for semantic classification. The semantic identification of words, phrases, sentences, and reports may be used to discover these components. The polarization is linguistic-orientated info, and it can be positive or negative. For particular problems with extremely reliant characteristics, the Naive Bayes works admirably. Some other models in which efficient techniques are used for feature determination [1], weight estimation, and classification should be proposed by some current approaches. The novel design relies on Probabilistic calculation. Here, by making use of special features and representation features, the classification weights are well established. The representation feature is the knowledge that infers a class, and the data that helps in specific classes is the same characteristic. They figured out the likelihood of any identification for the Probabilistic method and used these weights.

The proposed RNN algorithm describes the new classification approach using deep learning algorithms. The system collects real-time social media datasets as well as synthetic datasets simultaneously. The proposed data placement mechanism can also provide the fast detection of fake accounts from application GUI.

Figure 4 shows the comparative analysis of the system. The proposed algorithm is compared with the existing algorithm such as Navie Bayes, Random Forest & Support Vector Machine. Here we can say that the performance of the system is better as compared to the existing system. The proposed system is also compared with ReLu, TanH & Sigmoid activation functions. Out of these three activations functions, the Sigmoid gives better accuracy compared to ReLu & Tanh.

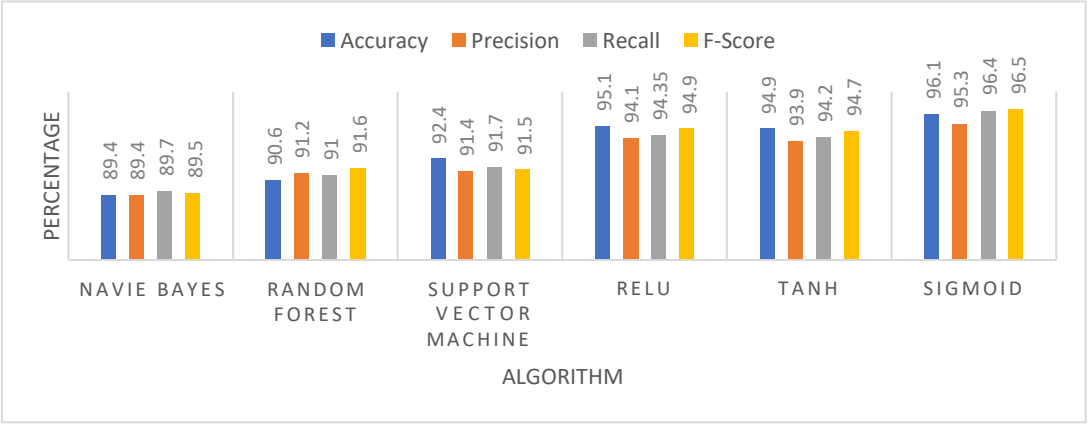

Fig. 4: Comparative Analysis of the System

The below figure 4 demonstrate the comparative analysis for fake account detection with some state-of-art method. The recurrent neural network [1] gives low accuracy of 87.11% for fake news detection while our proposed RNN reduce 95.5% detection accuracy on real-time Twitter data. In this experiment, we evaluate various machine learning as well as deep learning algorithms for bot detection and fake account detection.

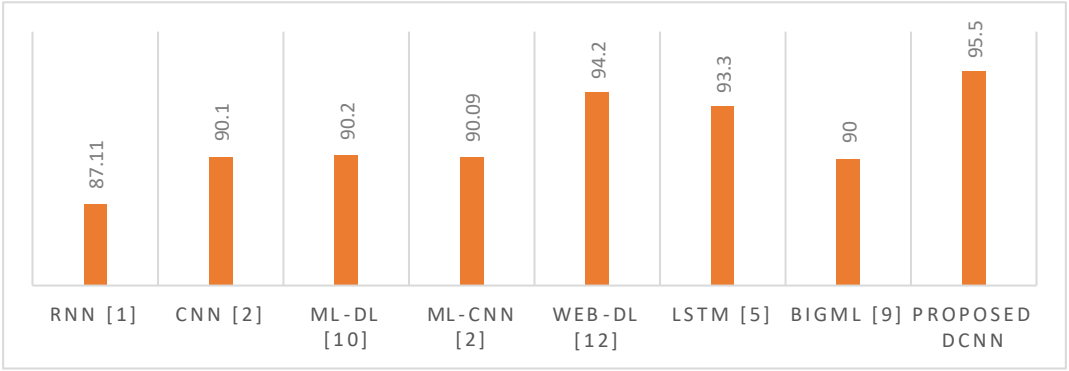

Fig. 5: Comparative Analysis of proposed algorithm with various state-of-art systems

Finally base on the overall experiment and analysis we conclude the proposed system with a modified reconnect neural network gives higher accuracy with three different SoftMax functions such as sigmoid, Tanh and RELU etc.

4. Conclusion

We proposed a fake account detection system from social media using machine and deep learning algorithms. The major problem of detecting fake identity profiles on social media has been solved using various machine learning techniques such as SVM, RF, NB, and RNN. Among these techniques, RNN with Sigmoid reaches the best performance with an accuracy of 96.10%. Also, we notice that the performance of the system varies with the classification technique and dataset used. Furthermore, the experimental result shows that the proposed approach achieved satisfactory results evaluated with the current state of art methods used in the literature. This research collects data from Twitter social media web domains for the detection of fake profiles. But still, research has ample space for improvement as most of the parameters such as metadata shared by the user, the trustworthiness of shared data, etc., parameters are not considered helpful to know the account's status is fake or real.

References

[1] Van Der Walt, Estée, and Jan Eloff. "Using machine learning to detect fake identities: bots vs humans." IEEE Access 6 (2018): 6540-6549.

[2] Shu, Kai, et al. "The role of user profiles for fake news detection." Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2019.

[3] Islam, Mohammad Raihanul, Sathappan Muthiah, and Naren Ramakrishnan. "RumorSleuth: joint detection of rumor veracity and user stance." 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE, 2019.

[4] Al Asad, Nafiz, et al. "Depression Detection by Analyzing Social Media Posts of User." 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON). IEEE, 2019.

[5] Li, Neng, et al. "Fake reviews tell no tales? dissecting click farming in content-generated social networks." China Communications 15.4 (2018): 98-109.

[6] Meshram, Shweta, Rajesh Babu, and Jayanth Adhikari. "Detecting Psychological Stress using Machine Learning over Social Media Interaction." 2020 5th International Conference on Communication and Electronics Systems (ICCES). IEEE, 2020.

[7] Zarei, Koosha, Reza Farahbakhsh, and Noël Crespi. "Typification of impersonated accounts on instagram." 2019 IEEE 38th International Performance Computing and Communications Conference (IPCCC). IEEE, 2019.

[8] Elghanuni, Ramzi H., Musab AM Ali, and Marwa B. Swidan. "An Overview of Anomaly Detection for Online Social Network." 2019 IEEE 10th Control and System Graduate Research Colloquium (ICSGRC). IEEE, 2019.

[9] Gheewala, Shivangi, and Rakesh Patel. "Machine learning based Twitter Spam account detection: a review." 2018 Second International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2018.

[10] Daha, Fatima Zohra, and Saniika Hewavitharana. "Deep neural architecture with character embedding for semantic frame detection." 2019 IEEE 13th International Conference on Semantic Computing (ICSC). IEEE, 2019.

[11] Fazil, Mohd, and Muhammad Abulaish. "A hybrid approach for detecting automated spammers in twitter." IEEE Transactions on Information Forensics and Security 13.11 (2018): 2707-2719.

[12] Aydin, Ilhan, S. E. V. İ. Mehmet, and Mehmet Umut Salur. "Detection of Fake Twitter Accounts with Machine Learning Algorithms." 2018 International Conference on Artificial Intelligence and Data Processing (IDAP). IEEE, 2018.

[13] Sowmya, P., and Madhumita Chatterjee. "Detection of Fake and Clone accounts in Twitter using Classification and Distance Measure Algorithms." 2020 International Conference on Communication and Signal Processing (ICCSP). IEEE, 2020.

[14] Pakaya, Farhan Nurdiatama, Muhammad Okky Ibrohim, and Indra Budi. "Malicious Account Detection on Twitter Based on Tweet Account Features using Machine Learning." 2019 Fourth International Conference on Informatics and Computing (ICIC). IEEE, 2019.

[15] Akyon, Fatih Cagatay, and M. Esat Kalfaoglu. "Instagram Fake and Automated Account Detection." 2019 Innovations in Intelligent Systems and Applications Conference (ASYU). IEEE, 2019.