Fourth International Conference on Electronics and Sustainable Communication Systems (ICESC - 2023)
DVD Part Number: CFP23V66-DVD; ISBN: 979-8-3503-0008-6

An Innovative Method for fake news classification using LSTM-RF Approach

1. Dr Prashant S Yadav, Lecturer, Computer Science and Engineering, MMIT Hathras, Uttar Pradesh, India. pyaduvanshimitusa@gmail.com

2. Dr K Subba Reddy, Computer Science and Engineering, Prakasam Engineering College, Kandukur, Andhra Pradesh, India. kurapatisr999@gmail.com

3. Dr S. Ashwini, Assistant Professor, Department of Computing Technologies, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology Kattankulathur Campus, Chennai, India. ashwinisekar.achu@gmail.com

4. Dr Santhosh .J, Assistant Professor, Department of Computer Applications, Sri Krishna Adithya College of Arts and Science, Coimbatore, Tamilnadu, India. santhoshj@skacas.ac.in

5. Dr Nirmala Devi M, Assistant Professor in English, St.Martin's Engineering college, Secunderabad, Telangana, India. nirmalmphil@gmail.com

6. Dr Kamlesh Singh, Professor, Department of Computer Science and Engineering, Graphic Era Hill University, Dehradun Campus, Dehradun, India. mailkamleshsingh@gmail.com

***Abstract* – The dissemination of news and coverage of events, both domestically and internationally, is now mostly the responsibility of print, web, and broadcast media. Because of the explosion of information available online, verifying the veracity of claims is more difficult than ever. It is extremely difficult to verify the veracity of fake news using conventional data processing techniques without seriously misinforming the public. Indeed, the difficulty for the government and the public is a matter of case by case argument. Several sites were built specifically for this purpose, each using its own unique logic and algorithm to determine which stories were more likely to be true or false. There needs to be a system in place for verifying the veracity of online claims, especially those that garner thousands of shares and clicks before being disproved by credible sources. In this proposed approach to an LSTM-RF classifier was used to determine whether or not a news story was fabricated. The three-step process that makes up the suggested method is as follows: preprocessing; feature extraction and feature selection; and LSTM-RF model training. In preprocessing, you'll clean the data, get rid of stop words, tokenize it, and stem it. The TF-IDF and Extra Tree Classifier are used for feature selection. The model is trained with LSTM-RF. The proposed method achieves superior results compared to LSTM and RF.**

***Keywords*—Random Forest (RF), Term Frequency – Inverse Document Frequency (TF-IDF), Extra Tree Classifier.**

I. INTRODUCTION

There are benefits and drawbacks to reporting news in the digital age. More and more channels of communication are becoming available. The internet's ubiquitous availability has led to the development of novel methods of disseminating information, some of which include blogs and social media. Convenience for consumers in accessing breaking news has increased. The current status of social media sites has both positive and negative impacts, such as the proliferation of fake and inaccurate news, because editorial boards do not always determine the integrity of the material uploaded. As it stands, social media platforms like these are excellent venues for discussing crucial issues like healthcare reform, government regulation, and school funding. Organizations typically use websites because of the monetary benefits they offer. The problems facing today's highly digitalized society cannot be reduced to a single factor. The spread of false information is one such instance. Disseminating false information is simple. The purpose of disseminating untruths is to destroy a person's good name. Disparaging statements can be made about anyone or anything. There are many places on the internet where the misinformation can be spread. The benefit of using online resources is that they give users quick and simple access to the most recent data. This opens the door for cybercriminals to use the same platforms to promote misinformation. Someone could be offended if they heard this. The general public has a bad habit of accepting whatever they read in the news at face value. It is incredibly difficult to identify hoaxes. The more quickly incorrect information spreads, the more people will come to believe it. People, communities, and even political parties can be swayed by false information. In order to trick readers, fake news reports will distort facts and data. Rhetorical strategies that mimic those of fake news but are motivated by an honest attempt to skew the truth. There have been instances of fake news erroneously using reliable authorities to support their assertions. The difficulties in identifying bogus news that have been discussed above have increased in recent years. Thanks to advancements in computing power and huge data processing, artificial intelligence (AI) techniques have recently seen a renaissance in popularity, and these approaches have shown promise in addressing the aforementioned difficulties with identifying fake news. There are numerous additional applications for artificial intelligence outside detecting hoaxes. Machine learning (ML) and deep learning (DL) techniques, both subfields of artificial intelligence, have seen widespread use in the battle against fake news. To this point, computer approaches for detecting fake news have relied primarily on satirical news providers like "The Onion" and fact-checking websites like "politiFact" and "Snopes." However, there are some potential issues and drawbacks with using these sources. If you're going to use satire as a source of fake information, you might as well think about how that might bring in some comedy or absurdity. Humor and irony have both been studied with the use of satirical news from "The Onion" as an example. However, it is difficult to obtain datasets that provide generalization across multiple domains due to the fact-checking websites' human-intensive nature and focus on a single domain of interest (such as politics). Unfortunately, there doesn't appear to be a centralized resource for information about the various approaches that have been presented to automate the detection of fake news on social media. As was discussed, systematic mapping is an improved alternative to the traditional literature review in displaying the state of the art in a research area. Studies like these use a predetermined method to seek out and thoroughly analyze relevant material. This literature mapping tries to comprehend this research topic in a fair, and auditable method by classifying, systematically locating, and analyzing realistic interoperability solutions. To this end, people have investigated some of the approaches taken in the past, some of the solutions in use now, some contexts in which these solutions have been put into practice, and some of the most significant challenges in the area of automatically identifying fake news. Preprocessing, Feature Extraction and Selection, and Model Training are the three main components of the proposed method.

II. Literature Survey

In recent years, academics have given more attention to the issue of identifying false content on social media. Using temporal-linguistic data derived from a user comment thread, new techniques have been created to identify bogus news. [1]find question phrases in user comments and use them to determine whether or not a piece of news is fake. Recurrent neural networks that store temporal-linguistic information from a stream of user comments are used by [2]to detect hoaxes. Building on the work of, [3] uses recurrent neural networks equipped with a soft-attention mechanism to extract specific temporal-linguistic properties [2]. The authors of the recent paper [4]offer a method for identifying fake news by building a framework at the Discourse level. In addition, they acknowledge helpful structure-related aspects that explain their discovery and deepen our understanding of disinformation. It is possible that using only linguistic indications to spot fake news is insufficient, hence these techniques are often paired with machine learning technologies. The cited works provide other examples of significant efforts to detect fake news utilizing linguistic basis approaches [5]. Due to their lack of applicability to the planned research, language-based approaches are only briefly discussed. They identified three types of deception in their study by[6]. They categorized the made-up stories as either having significant ramifications, having widespread implications, or being humorous. Disingenuous content is being utilized to impact public opinion at an alarming rate, because to the extensive reach of social media platforms like Facebook and Twitter. This has had a significant impact on voters and online retailers alike. Clickbaits were detected using a two-tiered text-based classifier by [7]. Aspects of morphology, grammar, style, vocabulary, and emotion were all employed[8] used satirical markers to differentiate between fake and authentic news. Their strategy, which also considered punctuation and grammar, achieved accuracy and recall rates of 90.21% and 86.1%, respectively, by focusing on the text's absurdity. [9]used Support Vector Machines with n-gram features. Using tf-idf for feature extraction and linear SVM for classification, they were able to get 92.5% accuracy on 50000 features. A novel n-gram model was created by [10]to automatically identify false comments and fake news. TFIDF was used for feature extraction, and LSVM, SVM, DT, KNN, LR, and SGD were used for classification. When compared to more conventional methods, the outcomes of their experiments are extremely encouraging. [11]gave a presentation that focused on several aspects of news articles, including citations and social media feeds. They present a novel set of features for automatic false news identification by contrasting the predicting abilities of NB, KNN, RF, XGBoost (XGB), and SVM. [12]explored the issue of rumor detection using a number of different context-aware deep learning models. To implement bidirectional long-term dependent memory, the proposed method uses a convolutional neural network. Based on their findings, they were able to label tweets as either rumor or fact. Due to the widespread availability of deliberately misleading or otherwise low-quality content, the detection of such stories has been a topic of intense interest among academics. Before moving further with this implementation, we researched a variety of reliable sources on Fake news detecting algorithms. In particular, studies of social network data [13] have aided our knowledge of how to identify fake news in written and visual formats. Overviews of the problem, its importance, its challenges, detection approaches, important feature identification, and datasets may be found in the literature [14]The most recent and noteworthy advancements in this field are discussed in the next section. In [15]performed research to determine who were the most influential people in a specific field of study. They used a whale-optimized algorithm to study data from social media platforms to find influential people's perspectives[15]. Various lexical features have been found to be helpful in prior studies in the field of fake news detection in identifying trustworthy and less trustworthy digital news sources. The foregoing information about the spread of fake news during the past decade and its use in various settings highlights the need of finding a solution to this issue. Disinformation tactics have a far greater impact because the majority of people use social media on a regular basis and are eager to share posts and news reports whose truth cannot be confirmed. According to the authors[16], it may take a lot of time and money to put the idea into effect. Using a bidirectional transformer strategy with a feed-forward classification layer, we demonstrated the benefits of transformer-based models over machine learning models. On the NewsFN dataset, the accuracy of fine-tuned BERT was found to be 97% in a study by[17], while that of XGBoost was only 89.4%. Another study used LSTM cells with the Attention model (LSTM-ATT) for content-based categorization and found it to be 83.3%

accurate on the PolitiFact dataset and 79.4% accurate on the GossipCop dataset. The scientists compared their proposed model to others already on the market and found that it performed noticeably better than the rest [18]. With an essay and a set of classifications (such as false and real) to which a text can belong, an algorithm is tasked with delivering the correct class label, which is the fake news text classification problem. The most common use of false news classification of texts is in the field of social media data categorization. Interestingly, it has been shown [19]that approaches based on stylometry could compromise the ability to automatically detect fake news. This is due to the fact that robots can produce writings that are consistent regardless of subject matter, unlike human writers who are prone to bias and occasional errors across topics. As deep learning has gained traction, similar models to Convolutional and Recurrent Neural Networks have been developed for headline and text matching [20]. Using graph coattention networks on data about the news, but also about the writers and dissemination of the news, [21] provided a solution to a more realistic scenario for detecting bogus news on social media sites. Given the problem-dependency of individual document representations, we are motivated to investigate ensembles of representations that may entail different aspects of the represented text and so generalize better. Our approach involves an LSTM-RF Algorithm and a new dataset developed specifically for the task. The technique relies on TF-IDF and Extra Tree Classifier to select features from texts.

III. PROPOSED SYSTEM

The widespread influence of social media in modern society has piqued the interest of both academics and the general public in developing methods for automatically detecting fake news. Machine learning algorithms that take into account multiple aspects of news are now used by existing detection methods to identify false stories. A fundamental shortcoming of these approaches is their inability to spot fake news in their early stages, as the data required to do so is often either unavailable or inadequate. This makes early detection of fake news difficult. In order to remedy this problem, the authors of it present a new model for the timely identification of social media fake news by use of the identification of news propagation routes. Figure 1 represents the proposed models flow diagram.

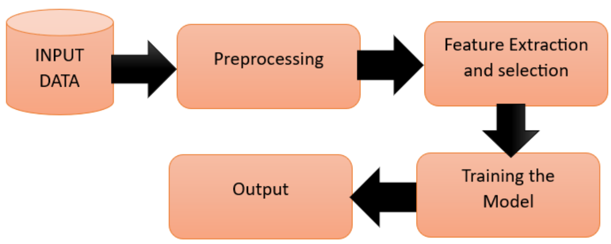

Fig. 1. Flow Diagram of the Proposed Model

A. *Preprocessing*

Because machines can't read human language, raw text must be converted to numbers before it can be used for training. The term "preprocessing" is used to describe the act of modifying or cleaning data before it is used to improve efficiency.

1) *Cleaning of Data*

It's the procedure of cleaning a dataset of unwanted information such as duplicates, errors in formatting, noise, and missing data. Productivity can be raised generally by maintaining tidy data.

2) *Removal of Stop-Words*

Common words that do not add value to the end product and can be ignored during training are called "stop words."

3) *Tokenization*

To analyze and eliminate all username-mentions, hashtags, and URLs from tweets, we developed a tokenizer.

4) *Stemming*

Tokens can be "stemmed" back to their original forms by having their suffixes removed [22]. As an illustration, the token "playing" will be changed to "play" by dropping the "ing" at the end.

B. *Feature Selection and Extraction*

1) *TF-IDF*

One popular technique for determining a word's significance in a document is the Term Frequency-Inverse Document Frequency (TF-IDF) method, which makes advantage of the numerical representation of altered texts. This method of feature extraction is commonly employed in the field of Natural Language Processing (NLP). Since the computer can't read human language, the texts will be turned into numbers first [23][24]. A method for calculating the relative frequency of words in a text, TF-IDF. In addition to determining a word's frequency in a phrase, TF-IDF also ranks words according to their importance, as suggested by the name. The TF-IDF formulas are shown in Eqs. 1–4.

$$tf - idf = tf * idf \tag{1}$$

$$tf = \frac{count\ of\ s\ in\ a}{number\ of\ words\ in\ a} \tag{2}$$

$$df = occurence\ of\ s\ in\ M\ documents \tag{3}$$

$$idf = \log\left(\frac{M}{af} + 1\right) \tag{4}$$

A $tf - idf$ calculation, where tf = Term Frequency and idf= Inverse Document Frequency, is shown below. Term = s and Document = a

2) *Extra Tree Classifier*

It's a common tree classifier, like RF or DT, but we didn't use it to classify anything; instead, they used it as a feature selection approach to zero in on the most instructive qualities to input into our Classifiers.

C. Training the Model:

1) LSTM Algorithm

The long-short-term memory (LSTM) recurrent neural network (RNN) is a powerful tool for handling time-series problems. The problems of gradient growth and gradient disappearance in RNN can also be effectively addressed. Two gates in a long short-term memory cell control the information stored in state d_m. To regulate how much of the unit state d_{m-1} from the previous moment is carried over into the present moment d_m, the forgetting gate is one such mechanism. The other is the input gate, which determines what percentage of input k_m from the current network is appended to the current cell's state d_m value. The current LSTM output value c_m is dependent on the state of the output gate and the cell d_m. Encoders and decoders make up LSTM [25].

The reconstruction error is computed by the decoder, whereas isometric learning of the feature data input is accomplished by the encoder. Define the observation window size s , as a function of the input sample, $K_1, K_2, \cdots, K_S$, where $\{K_z\}_{z=1}^{S}$. Researchers will have $K_z = \{k_{z1}, k_{z2}, \ldots, k_{zs}\}$, ixtg if any of these are in their DNA. Formula (1) displays the encoder implied state vector for the z column of any sequence K_z at time $m\epsilon\{1,2,\ldots,s\}$

$$c_F^{mz} = b\left(Pk_{mz} + Qc_{(m-1)z}\right) \tag{5}$$

where at time instant $m-1$, the z coding unit's output state vector is $c_{(m-1)z}\epsilon Q^t$ Coefficient weight matrix Q of order $t \times a$ and $t \times t$; input matrix $k_{mz}\epsilon Q^t, P$. Typically, "$tane$" activation function $x(\cdot)$ is used. If they feed the encoder the vectors in each column of K_z, to get the following.

$$c_{mz} = b_s^{etc}\left(k_{mz}, c_{(m-1)z}\right) \tag{6}$$

where c_{mz} is the result of the z coding unit at time m and s is the encoding section's parameter set. The $tane$ activation function is the default for $x_s^{etc}(\cdot)$. Once the entire sequence has been fed into the encoder, the resulting sequence set will be $\{c_{mz}\}_{z-1}^{sz}$. The next steps then constitute the pooling process.

$$c_z = \frac{\sum_{x=1}^{s_z} c_{zx}}{s_z} \tag{7}$$

$$c_z = c_{m,sz}$$

$$c_z = \max_x \{c_{mz}\}_{z=1}^{sz} \tag{8}$$

the total number of rows in c_{mz} denoted by y. Pooled c_z then goes into the decoder. Formulas (6) and (7) can be derived by refactoring the inputs.

$$\tilde{c}_{mz} = b_\beta^{dec}\left(c_z, c_{(m-1)z}\right) \tag{9}$$

$$\tilde{k}_{mz} = \lambda\left(\tilde{c}_{s_k}\right) \tag{10}$$

where $\tilde{k}_{n_z}$ is the information that was reconstructed. The decoder's implicit state vector is represented by the notation $\tilde{c}_{s_k}$. The activation functions x_β^{dec} and are often set to "$tane$" Minimizing the function $\sum_{z=1}^{s} k_{m_z} - \tilde{k}_{n_z}^2$ yields the LSTM model.

2) Improved LSTM

The enhanced LSTM model includes an attention layer as a final component. Good results have been obtained when applying the LSTM and attention mechanism model to relationship categorization. Time series data prediction using the combined model, especially using public data, is still a work in progress. In this proposed approach to use a technique that incorporates LSTM with an attention mechanism to foretell the associated fake news.

3) Model based on RF

a) Fuzzy CMeans Clustering

The C-mean fuzzy clustering is calculated for a collection of data $K = \{k_1, k_2, \ldots k_s\}$. Researchers need to input the total number of categories (D), with $x = 1,2,\cdots,d$ being the center of each cluster t_x. The degree of membership of the b sample to class K_z is given by the membership function $\gamma_{zb} = \gamma_{K_z}(k_b)$, where k_b is the sample identifier. The membership-based clustering loss index function is given by Equation (11):

$$X = \sum_{x=1}^{b}\sum_{z=1}^{s}[\gamma_x(k_z)]^e k_z - t_x^2 \tag{11}$$

Where e is the smoothing factor (also known as the weighted index) and the degree to which samples are shared between fuzzy classes. The academic community is still divided about the best value for $\boldsymbol{e}$. when both calculation effort and precision are taken into account, the weighted index is typically 3.

If X's partial derivative with respect to t_x and $\gamma_x(k_z)$ is zero, then the following formulas (12) and (13) describe the conditions under which X must take on its minimal value:

$$t_x = \frac{\sum_{z=1}^{s}[\gamma_x(k_z)]^e k_z}{\sum_{z=1}^{s}[\gamma_x(k_z)]^e} \tag{12}$$

$$\gamma_x(k_z) = \frac{k_z - t_x^{\frac{2}{e}-1}}{\sum_{u=1}^{b} k_z - t_u^{-\frac{2}{h}-1}} \tag{13}$$

Therefore, the number of input categories d is necessary for generating c-mean fuzzy clustering. Set each cluster's center $t_x(x = 1,2,\cdots,d)$. Iteratively calculating $t_x(x = 1,2,\cdots,d)$ and $\gamma_{zb} = \gamma_{K_z}(k_x)$ from formulas (9) and (10), the corresponding membership function $\gamma_{zb} = \gamma_{K_z}(k_x)$ for sample k_b is obtained. Until the accuracy condition is met, the clustering center and membership function can be determined.

b) Optimal Cluster Number Determination

Cluster analysis using c-means fuzzy clustering requires the number of categories, c, to be determined in advance. Clustering is highly sensitive to the value of c. If there are too many clusters, similar samples will be split up into distinct categories. Data from distinct classes may be lumped together if the clustering number is too little. If the improper clustering number is used, the iteration may fail to converge and produce an incorrect clustering result. As a result, it's important to establish optimization criteria and compute the clustering number. The cluster count can be determined with relative ease by introducing outcome evaluation markers. The optimal number of clusters can be determined by increasing the number of clusters and evaluating the change in evaluation indices. Similarity within a class and dissimilarity between classes are typically used to evaluate clustering results for the evaluation index. This allows us to break down the evaluation indices into two distinct measures: intraclass similarity (X_f) and interclass similarity (X_q)

$$Y_{fz} = \frac{1}{s_z}\sum_{k\epsilon K_z} k - d_z \tag{14}$$

Where s_z the total number of records is in the group. The representative object is k. Class K_z revolves around the node d_x. The closer the samples are to the middle of the class and the smaller the value of Z_{fz}, the more similar they are:

$$z_{qzx} = p(d_z - d_x) z \neq x \tag{15}$$

where K_z's pivot point is located at d_z. Class K_x revolves around d_x. As Z_{qzx} increases, class similarity decreases and the center distance between adjacent classes increases. Therefore, full evaluation metrics can be collected.

$$X_z = \sum_{z=1}^{d} \frac{s}{s_z} z_{fz} + \sum_{z=1}^{d}\sum_{x=z+1}^{d} \frac{1}{z_{qzx}} \tag{16}$$

Therefore, the following steps should be taken to determine the best possible cluster size: First, you'll need to establish a scale for the range of ID numbers. The standard value for the classification number d is $d\epsilon[2, \sqrt{s}]$, ffiffipt, where s is the total number of data points in the sample. For each d value run a clustering procedure. Third, evaluate the data from each cluster using some sort of index. Finding the inflection point and the distinct minimum points in the evaluation index provide the ideal clustering number by revealing the relationship between the data.

IV. RESULT AND DISCUSSION

The rapid rise in the production and spread of fake news has created a critical need for the automatic classification and detection of distorted news items. However, it is challenging to automate the detection of fake news because the model needs to be able to grasp nuanced written language. The model's ability to comprehend the level to which the reported news is related to the genuine news is also constrained because most existing fake news detection algorithms consider the issue as a binary classification task. It addresses these shortcomings by outlining an LSTM-RF architecture capable of reliably predicting future instances of fake news.

TABLE I. COMPARISON OF MODELS(%)

Models	ACCURACY	PRESICION	RECALL
LSTM	90.53	88.68	86.75
RF	94.72	92.68	89.53
LSTM-RF	96.42	95.29	92.63

Three trials were undertaken in this propose d approach to utilizing three distinct methods (LSTM, RF, and LSTM-RF) to identify the intruders. The results of the proposed model are shown in Table 1; An examination of the data reveals unprecedented precision; however, this does not equate to flawlessness. One should take precision "with a grain of salt," preferably in conjunction with other measurements. It is shown that standard algorithms like LSTM-RF can compete with even the most fundamental neural networks.

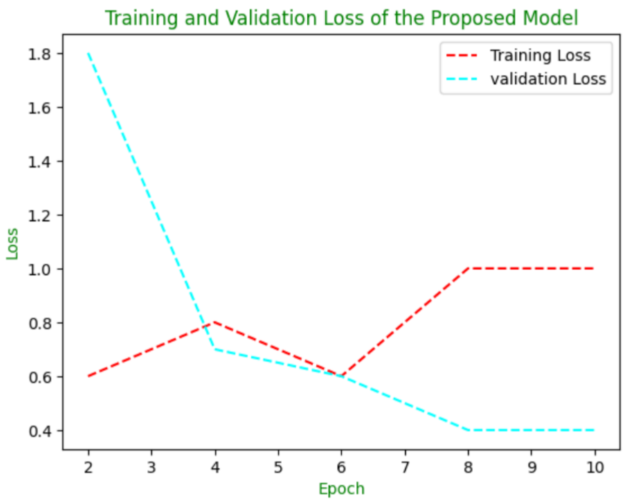

Fig. 2. Loss of the Proposed Approach

Model training took place over the course of 12 iterations, with a batch size of 130, as shown in Fig. 2. There was a loss of 0.4 in training and a loss of 0.94 in validation. Figure 2 depicts the loss and Figure 3 depicts the accuracy acquired during training and validation.

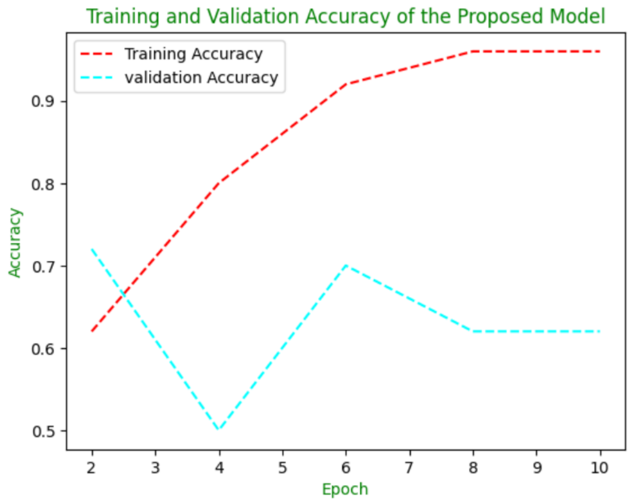

Fig. 3. Accuracy of the Proposed Model

Figure 3 depicts the training process for the model that took 10 iterations and 130 training samples to complete. The training accuracy acquired after 12 iterations was 96.34 percent, whereas the validation accuracy was 62.73 percent.

V. CONCLUSION

Users rely heavily on these networks as their primary source for news consumption and dissemination. In an effort to affect user opinion, social media often praises the most momentous actions in the globe, including political ones. Although social media's extensive use is crucial for spreading awareness, the reliability of its content is still up for debate. As a result, the study's findings emphasize the significance of proposing a strategy for recognizing and labeling disinformation campaigns. A misleading news story's content is the most important aspect in getting people to believe it. By creating a linguistic model, we may learn what kinds of content are likely to produce language-driven features. This language model is used to extract syntactic, grammatical, emotional, and readability features of individual news stories. A language-driven model requires a means of coping with the bother of painstakingly built features in order to overcome the curse of dimensionality. Therefore, the LSTM-RF model is used to achieve optimal performance in identifying misinformation. Natural language processing is used to pre-process data because dealing with languages complicates the effort of detecting fake news. It use TF-IDF and Extra Tree Classifier for feature selection and extraction. Finally, LSTM-RF was used to train the model. The suggested method achieves an accuracy of roughly 96.43%, which is significantly higher than that of the two baseline methods (LSTM and RF).

REFERENCES

[1] Z. Zhao, P. Resnick, and Q. Mei, "Enquiring Minds : Early Detection of Rumors in Social Media from Enquiry Posts Categories and Subject Descriptors Detection Problems in Social Media," *Proc. 24th Int. Conf. world wide web*, pp. 1395–1405, 2015.

[2] B. J. Jansen, "Detecting rumors from microblogs with recurrent neural networks," *Proc. 25th Int. Jt. Conf. Artif. Intell. (IJCAI 2016)*, vol. 3818–3824, no. Ijcai, 2016, [Online]. Available: https://ink.library.smu.edu.sg/sis_research/4630%0AThis

[3] T. Chen, X. Li, H. Yin, and J. Zhang, "Call attention to rumors: Deep attention based recurrent neural networks for early rumor detection," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11154 LNAI, pp. 40–52, 2018, doi: 10.1007/978-3-030-04503-6_4.

[4] H. Karimi and J. Tang, "Learning hierarchical discourse-level structure for fake news detection," *NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf.*, vol. 1, pp. 3432–3442, 2019, doi: 10.18653/v1/n19-1347.

[5] G. Gravanis, A. Vakali, K. Diamantaras, and P. Karadais, "Behind the cues: A benchmarking study for fake news detection," *Expert Syst. Appl.*, vol. 128, pp. 201–213, 2019, doi: 10.1016/j.eswa.2019.03.036.

[6] V. L. Rubin, Y. Chen, and N. J. Conroy, "Deception detection for news: Three types of fakes," *Proc. Assoc. Inf. Sci. Technol.*, vol. 52, no. 1, pp. 1–4, 2015, doi: 10.1002/pra2.2015.145052010083.

[7] O. Papadopoulou, M. Zampoglou, S. Papadopoulos, and I. Kompatsiaris, "A Two-Level Classification Approach for Detecting Clickbait Posts using Text-Based Features," 2017, [Online]. Available: http://arxiv.org/abs/1710.08528

[8] V. Rubin, N. Conroy, Y. Chen, and S. Cornwell, "Fake News or Truth? Using Satirical Cues to Detect Potentially Misleading News," *Proc. Second Work. Comput. approaches to Decept. Detect.*, pp. 7–17, 2016, doi: 10.18653/v1/w16-0802.

[9] S. S. Ahmed, Hadeer, Issa Traore, "Detection of online fake news using n-gram analysis and machine learning techniques," *First Int. Conf. Intelligent, Secur. Dependable Syst. Distrib. Cloud Environ.*, pp. 127–138, 2017, doi: 10.1007/978-3-319-69155-8.

[10] H. Ahmed, I. Traore, and S. Saad, "Detecting opinion spams and fake news using text classification," *Secur. Priv.*, vol. 1, no. 1, p. e9, 2018, doi: 10.1002/spy2.9.

[11] J. C. S. Reis, A. Correia, F. Murai, A. Veloso, F. Benevenuto, and E. Cambria, "Supervised Learning for Fake News Detection," *IEEE Intell. Syst.*, vol. 34, no. 2, pp. 76–81, 2019, doi: 10.1109/MIS.2019.2899143.

[12] M. Z. Asghar, A. Habib, A. Habib, A. Khan, R. Ali, and A. Khattak, "Exploring deep neural networks for rumor detection," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 4, pp. 4315–4333, 2021, doi: 10.1007/s12652-019-01527-4.

[13] S. A. Alkhodair, S. H. H. Ding, B. C. M. Fung, and J. Liu, "Detecting breaking news rumors of emerging topics in social media," *Inf. Process. Manag.*, vol. 57, no. 2, pp. 1–23, 2020, doi: 10.1016/j.ipm.2019.02.016.

[14] A. Bondielli and F. Marcelloni, "A survey on fake news and rumour detection techniques," *Inf. Sci. (Ny).*, vol. 497, pp. 38–55, 2019, doi: 10.1016/j.ins.2019.05.035.

[15] L. Jain, R. Katarya, and S. Sachdeva, "Opinion leader detection using whale optimization algorithm in online social network," *Expert Syst. Appl.*, vol. 142, p. 113016, 2020, doi: 10.1016/j.eswa.2019.113016.

[16] Q. Li and W. Zhou, "Connecting the Dots Between Fact Verification and Fake News Detection," *COLING 2020 - 28th Int. Conf. Comput. Linguist. Proc. Conf.*, pp. 1820–1825, 2020, doi: 10.18653/v1/2020.coling-main.165.

[17] A. Aggarwal, A. Chauhan, D. Kumar, M. Mittal, and S. Verma, "Classification of Fake News by Fine-tuning Deep Bidirectional Transformers based Language Model," *EAI Endorsed Trans. Scalable Inf. Syst.*, vol. 7, no. 27, pp. 1–12, 2020, doi: 10.4108/eai.13-7-2018.163973.

[18] J. Lin, G. Tremblay-Taylor, G. Mou, D. You, and K. Lee, "Detecting Fake News Articles," *Proc. - 2019 IEEE Int. Conf. Big Data, Big Data 2019*, pp. 3021–3025, 2019, doi: 10.1109/BigData47090.2019.9005980.

[19] T. Schuster, R. Schuster, D. J. Shah, and R. Barzilay, "The limitations of stylometry for detecting machine-generated fake news," *Comput. Linguist.*, vol. 46, no. 2, pp. 499–510, 2020, doi: 10.1162/COLI_a_00380.

[20] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi, and B. W. On, "Fake news stance detection using deep learning architecture (CNN-LSTM)," *IEEE Access*, vol. 8, pp. 156695–156706, 2020, doi: 10.1109/ACCESS.2020.3019735.

[21] Y. J. Lu and C. Te Li, "GCAN: Graph-aware co-attention networks for explainable fake news detection on social media,"

Proc. Annu. Meet. Assoc. Comput. Linguist., pp. 505–514, 2020, doi: 10.18653/v1/2020.acl-main.48.

[22] A. Gupta, R. Sukumaran, K. John, and S. Teki, "Hostility Detection and Covid-19 Fake News Detection in Social Media," *arXiv Prepr. arXiv2101.05953*, 2021, [Online]. Available: http://arxiv.org/abs/2101.05953

[23] S. B. S. Mugdha *et al.*, "A Gaussian Naive Bayesian Classifier for Fake News Detection in Bengali," *Emerg. Technol. Data Min. Inf. Secur. Proc. IEMIS 2020*, vol. 2, no. May, pp. 283–291, 2021, doi: 10.1007/978-981-33-4367-2_28.

[24] S. Kumar and T. D. Singh, "Fake news detection on Hindi news dataset," *Glob. Transitions Proc.*, vol. 3, no. 1, pp. 289–297, 2022, doi: 10.1016/j.gltp.2022.03.014.

[25] X. Liu, "Music Trend Prediction Based on Improved LSTM and Random Forest Algorithm," *J. Sensors*, vol. 2022, 2022, doi: 10.1155/2022/6450469.