DETECTION OF ONLINE FAKE NEWS USING OPTIMIZED ENSEMBLE DEEP LEARNING MODELS

SUHASINI V, Dr . N.VIMALA

Research Scholar, Department of Computer Science, L.R.G Government Arts College (W), Tirupur

E-Mail: suhasinikanishk2811@gmail.com

Assistant Professor, Department of Computer Science, L.R.G Government Arts College (W), Tirupur

E-Mail: : vimalananjappan@gmail.com

Abstract

Fake news is a real problem in today's world, and it has become more extensive and harder to identify. Major challengein fake news detection is to detect it in the early phase. Another provocation in fake news detection is the unavailability or the shortage of labeled data for training the detection models. These research works proposes novel fake news detection frameworks that can addressthese challenges. The proposed framework exploits the information from the social contexts to detect fake news. This methodology works as a fused Deep Learning Ensemble model developed with ML models, which is then used on a publicly available dataset to predict if a news is true or not. The proposed model appraised with the popular Deep Learning models, while performance metrics such as accuracy, recall, specificity, precision, and f1-score used to measure the performance of the proposed model. Results presented showed that the proposed model outperformed other popular Deep Learning models.

Keywords: Online Fake News, CNN, DNN, LSTM

DOI Number: 10.14704/NQ.2022.20.15.NQ88396 **NeuroQuantology2022;20(15): 3941-3957**

1. Introduction

Fake news detection is a subtask of text classification and is often defined as the task of classifying news as real orfake. The term 'fake news' refers to the false or misleading information that appears as real news. It aims to deceive or mislead people. Fake news comes in many forms, such as misleading headlines, malicious intention to mislead the public, false information regardless of the motive behind, hoax, parody,satire, rumor, deceptive news etc.,.Traditionally, people get news from trusted sources, media outlets and editors, usually following a strict code of practice. In the late twentieth century, the internet has provided a new way to consume, publish and share information with little or no editorial standards. Lately, social media has become a significant source of news for many people [1]. According to a report by Statistica, there are around 3.6 billion social media users (about halfthe population) in the world. There are obvious benefits of social media sites and networks in news dissemination, such as instantaneous access to information, free distribution, notime limit, and variety. However, these platforms are largely unregulated. Therefore, it is often difficult to tell whether some news is real or fake.

Recent researches show that the speed at which fake news travels is exceptional, and the outcome is its wide scale spread. A clear example of this is the spread of anti-vaccination misinformation and the rumor that incorrectly compared the number of registered voters in 2018 to the number of votes cast in US Elections 2020. The implications of such news are seen during the anti-vaccine movements that prevented the global fight against COVID-19 or in post-election unrest [2].

This section presents an overview of dataset, preprocessing techniques, and description of the deep learning model used for classification. Figure 1 represents the proposed methodology.For the data preprocessing, several steps like tokenization, lemmatization, and stop word removal were used. After preprocessing, NLP processing is applied for feature extraction process. Then, various DL models are implemented with extracted model.

The remainder of this paper is outlined as follows. The related literature is discussed in Section 2. Section 3 presents the methodology, result analysis is presented in Section 4 and the paper is succinctly in Section 5.

2. Literature Review

Numerous neural networks and models based on machine learning have been applied to detect fake news. Models were developed with features designed for specific datasets.Aphiwongsophon and Chongstitvatana et al.,(2018) based their models on identifying fake news using selected data sourced from Twitter. It is likely that these approaches will fall victim to dataset bias and possibly perform poorly on a different category of news. Gilda et al., explored some traditional machine learning approaches. Ahmed et al. investigated and compared six different classification techniques using n-gram analysis on a single dataset using feature extraction. The models were evaluated independently and the linear support vector machine classifier achieved the best score [3]. Yun and Ahnet al., (2018) detected fake news in Korea with machine learning and text mining using a two-step approach. Initially, the news contents are converted to values by applying text mining, and then classifiers are trained on these values [4]. However, several advanced learning models are not applied although they have excelled in text classification.

Research using deep learning to identify fake news works has accomplished encouraging results. Rashkin et al.(2017) used linguistic feature analysis and achieved the remarkable outcome of Long Short-Term Memory.Factchecking is indeed a challenging task but that various lexical features can contribute to our understanding of the differences between more reliable and less reliable digital news sources [5]. Singhaniaet al.(2017) applied a three-level attention network incorporating sentences, words, and headlines. 3HAN provides an understandable output through the attention weights given to different parts of an article, which can be visualized through a heatmap to enable further manual fact checking.[6]. Thota et al. (2018) presented a neural network to forecast the stance using the headline and the body of the article.The exponential increase in production and distribution of inaccurate news presents an immediate need for automatically tagging and detecting such twisted news articles [7].William Yang Wang constructed a hybrid model using a convolutional neural network that outclassed other traditional learning models.This paper presented a benchmark dataset named Liar and investigated using current models. The evaluation hints at how different types of models perform on data that is structured. Also, some models were prone to being overfit [8].

Ruchansky et al. (2017) built a CSI (capture, score, and integrate) model that used text, article response, and characteristics of the users' behavior. Ajao et al. developed a framework for classifying and identifying fake news in Twitter posts using a hybrid of neural networks. The tactic intuitively identified pertinent features without considering prior knowledge [9]. Lu and Li developed (GCAN) (2020) Graph-aware Co-Attention Networks to determine if a tweet is fake by using the associated sequence of retweet users. Khan et al. analyzed the performance of dissimilar approaches on three datasets and showed that Naive Bayes can achieve a similar result as neural network models when working with a dataset containing under 100 thousand articles [10]. Vijayaraghavan et al.(2020) applied different models to detect fake news and state that neural networks generally perform consistently and serve as a powerful universal approximator. However, the loss and accuracy come after using too many epochs and thus the issue of overfitting comes into play. In addition, a simpler model using logistic regression also delivered good performance results. Consequently, it does not necessarily follow that the more complicated the model, the better the performance. Furthermore, deep learning is "time-consuming and resource-consuming" [11].

Roy et al., (2018) developed models built on a Bidirectional Long Short -Term Memory and Convolutional Neural Network. The output from both of these models was input into a Multilayer Perceptron Model to obtain the final result [12]. Al-Ash et al. (2019) used a random forest classifier a hybrid approach for fake data detection using an ensemble model [14]. Ahmad et al. (2020) which consists of a decision tree classifier as an ensemble classifier to detect Indonesian fake news. An ensemble approach is adopted to help resolve the identified gap. A blended machine learning ensemble model developedto predict if a news report is true or not.[13]. Reddy et al. (2019) presented explored different textual properties in an ensemble approach to detect fake news.Particular domain has to explore multiple aspects before giving a verdict on the truthfulness of an articles[15]. Gutierrez-Espinoza et al (2020) evaluated the performance of ensemble learning using different machine learning techniques for classification in order to identify bogus online information and this approached to a collection of fake restaurant reviews and to detect deceptive information better than conventional machine learning algorithms. [16].

Mahabub(2020) used a distinct method for detecting fake news in developing an ensemble voting classifier that incorporates many familiar machine learning algorithms and also intelligentdetection system is proposed to deal with news classification both real and fake tasks.[17].

Kaur et al. (2020) designed a voting model with multiple levels in automating the detection of fake news by experimenting with several models. Itis to find out which classification model identifies phony features accurately using three feature extraction techniquesTerm Frequency–Inverse Document Frequency (TF–IDF), Count-Vectorizer (CV) and Hashing-Vectorizer (HV)[18]. Li et al. (2019) applied a pipeline to identify fake news by taking into consideration the headline and article text in a stacked ensemble. In all these studies, the ensemble approach yielded better performance when compared to the individual model in the detection of deceptive information [27]. Therefore, an advanced ensemble approach is adopted to detect fake news. The strategy will integrate blending and machine learning with natural language processing to extend and improve the current approaches.

Researchers have studied numerous algorithms for text classification that give good performance. However, some algorithms perform better on some datasets but may give even an average performance on other datasets. Therefore, instead of using a single classifier, it is better to use a group of classifiers and take a collective or team decision, rather than basing a decision on an individual classifier. This approach called an ensemble approach overcomes the weakness of one classifier by the strength of other classifiers and gives better performance than an individual classifier. The diverse nature of the approach and keeping the variance under control contribute greatly to its success. Furthermore, ensemble learning can result in more robust schemes of classification.

3. Methodology

This section presents the datasets, proposed framework, explanation of the algorithms, and the metrics that are used for performance evaluation. The datasets have been selected for our experiments which include news from a range of different categories and a combination of fake and real. It is publicly available and easily accessible on the web. Categorization of news as “fake news” can be “a very challenging and time-consuming task”. Hence, existing benchmark datasets are used in this work which is obtained from Kaggle.It comprises 25117 short labelled statements from kaggle.com. There are two labels for rating the truthfulness and false information.It focused on classifying news as true or fake. For binary classification of the news, to transform these labels into two labels. This dataset largely focuses on politics that contain statements of republicans and democrats, in addition to a substantial quantity of posts from social media. Each data point consists of a title, text, subject, and date. The text is the actual news article, and the subject or category is any one of Middle East, government news, US news, world news, politics news, left-news.In the proposed framework, as shown in Figure 1, are extending the current literature by introducing fused learning techniques incorporating blending. News articles from several domains are classified as true or fake by working with different feature sets. Blending techniques with Term Frequency Inverted Document Frequency and n-grams are used in proposed deep learning approach.The following operations will be carried out during the preprocessing of the dataset:

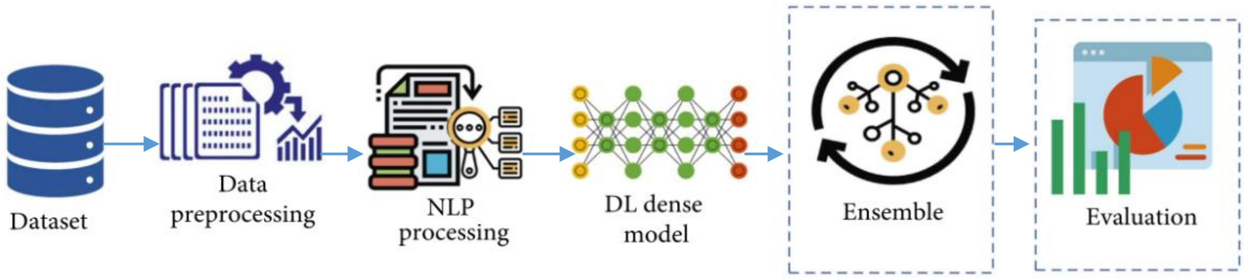

Figure.1. Proposed framework for Online Fake News detection

- Step 1: Preprocess: to remove noisy data such as stemming,tokenization,etc.,
- Step 2: Feature Extraction: the process of transforming raw data into numerical features and applying machine learning algorithms like SVM on extracted features.
- Step 3: Classifier: An algorithm that maps the input data to a specific category, which involves dividing a dataset into different segments (real or fake).
- Step 4: Ensemble model is developed by using Deep Learning and machine learning models like DNN, CNN and SVM to give the best outcomes.
- Step 5: CNN is preferred to extract the text characteristics and different layers of CNN includes network layer, pooling layer and normalization layer.
- Step 6: Construct the proposed model Ensemble Deep Learning with Machine Learning consists of four layers such as convolutional layer, pooling layer, fully connected layer and output layer.
- Step 7: Finally, best classification accuracy is obtained by newly developed CNN-LSTM with SVMensemble model to analyze for Fake or Real news classification.

4. Machine learning (ML) algorithms

Traditional machine learning-based methods have been used with NLP features because they need manual feature extraction. In this work, machine learning methods were built using the sci-kit-learn machine learning framework using Python programming language because this platform is flexible, robust, and easy to use. The following traditional machine learning algorithms were investigated in this methodology.

- **Support Vector Machine:** The SVM classifier is to find a hyperplane in an N-dimensional space (N — the number of features) that distinctly classifies the data points. SVM is one of the most popular Supervised

Learning algorithms, which is used for Classification as well as Regression problems.

- **Naive Bayes (NB):** The NB Classifier is the simplest form of Bayesian network but it can achieve high accuracy with a kernel density estimation implementation.
- **K-Nearest Neighbors (K-NN):** K-NN is a non-parametric method that can be used for both classification and regression tasks.
- **Logistic Regression (LR):** LR is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes.
- **Softmax:** It is a mathematical function that converts a vector of numbers into a vector of probabilities, where the probabilities of each value are proportional to the relative scale of each value in the vector.

5. Deep Learning (DL) Algorithms

Based on the nature of the features in the dataset, DL models were designed as discussed below. However, the DL with ML based classification modelachieve the best accuracy while compare with other models.

5.1 Deep Neural Networks (DNN)

A deep neural network is a neural network with more than two layers, some of whichare hidden layers. Deep neural networks use sophisticated mathematical modeling toprocess data in many different ways. A neural network is an adjustable model of outputs as functionsof inputs, which consists of several layers: an input layer, including input data; hidden layers, including processing nodes called neurons; and an output layer, including one or several neurons, whose outputs are the network outputs.

5.2 Recurrent Neural Networks (RNN)

Recurrent neural networks are a class of neural networks whose connections between neurons form a directed cycle, which creates feedback loops within the RNN. The main function of RNN is the processing of sequential information on the basis of the internal memory captured by the directed cycles. Unlike traditional neural networks, RNN can remember the previous computation of information and can reuse it by applying it to the next element in the sequence of inputs. A special type of RNN is long short-term memory (LSTM), which is capable of using long memory as the input of activation functions in the hidden layer.

The input data is preprocessed to reshape data for the embedding matrix. The next layer is the LSTM, which includes 200 cells. The final layer is a fully connected layer, which includes 128 cells for text classification. The last layer uses the sigmoid activation function to reduce the vector of height 128 to an output vector of one, given that there are two classes to be predicted (positive, negative).

5.3 Convolutional Neural Networks (CNN)

Convolutional Neural Networkis one of the descriptive network models in deep learning technology. CNN is grounded on artificial neural networks. Figure.1 shows the architecture of conventional CNN algorithm. For various classification tasks, the back-propagation algorithm is used to repeatedly upkeep or reduces the weight of corresponding features. The filters are used to auto extract features in the convolution layer to attain feature extraction. Among the various deep learning algorithms, CNN is preferred to extract the text characteristics of twitter data sequences because of its effectiveness in identification.

CNN has the ability to learn the features using its convolution property to the input data sequence with a high level of abstraction. The various layers of CNN include the network layers, the pooling layers and the normalization layers. A feature map is generated based on the raw input sequence in the convolution layer. With the adaption of a set of filter and weights, the pooling layer reduces the size of the generated feature map from the convolution layer making it more prominent for analysis. The normalization layer output comprises of the silhouette features as a single dataset to be feed to the classifier for the gait recognition purposes.

6. Proposed Ensemble Method

The proposed ensemble models includeDLand ML algorithms, which are described below,

Ensemble DL-ML

A drawback of RNN (LSTM) in general is the huge amount of hyper-parameter, which have to be set properly in order to achieve bestdetection performance. The neuron parameters whichare learned during training time, hyper-parameters are manually preselected that affect the architecture of the artificial network. In order to overcome this drawback CNN model is utilized in this work. The network is divided into various stages like input data is preprocessed and features are extracted which will ensure thesuccess of the subsequent steps, extract the features with CNN and finally binary class ML method is applied to classify the fake or real news. Basically, the CNN consists of four layers such as convolutional layer, pooling layer, fully connected layers and output layer.

Convolutional Layer: A convolution layer is a fundamental component of the CNN architecture that performs feature extraction, which typically consists of a combination of linear and nonlinear operations, i.e., convolution operation and activation function. Convolution is a specialized type of linear operation used for feature extraction, where a small array of numbers, called a kernel, is applied across the input, which is an array of numbers, called a tensor.

These layers store the output of the kernels from the previous layer which consists of weights and biases to be learned. The generated kernels that represent the data without an error is the point of the optimization function. In this layer, a sequence of mathematical processes is done to extract the feature map of the input data.This procedure is repeated applyingmultiple kernels to form an arbitrary number of feature maps,which represent different characteristics of the input tensors;

Different kernels canbe considered as different feature extractors. Two key hyper parameters that define theconvolution operation are size and number of kernels. Theformer is typically 3 × 3, but sometimes 5 × 5 or 7 × 7. Thelatter is arbitrary, and determines the depth of output featuremaps.

Nonlinear activation function: The outputs of a linear operation such as convolution are thenpassed through a nonlinear activation function. Althoughsmooth nonlinear functions, such as sigmoid or hyperbolictangent (tanh) function, were used previously because they aremathematical representations of a biological neuron behavior,the most common nonlinear activation function used presentlyis the ***Rectified Linear Unit (ReLU)***, which simply computes thefunction: $f(x) = \max(0, x)$.

Pooling Layer: This layer reduces over fitting and lowers the neuron size for the down sampling layer. This layer reduces the feature map size; reduce parameter numbers, training time, computation rate and controls over fitting. It extracts patches from the input feature maps, outputsthe maximum value in each patch, and discards all the other values. A max pooling with a filter of size 2 × 2 with astride of 2 is commonly used in practice. This downsamplesthe in-plane dimension of feature maps by a factor of 2.

Fully Connected Layer: The output feature maps of the final convolution/pooling layer is typically flattened, i.e., transformed intoa one-dimensional (1D) array of numbers (or vector), and connected to one or more fully connected layers, also known as dense layers, in which every input is connected to every output by a learnable weight. Once the features extracted by the convolution layers and down sampled by the pooling layers are created, they are mapped by a subset of fully connected layers to the final outputs of thenetwork, such as the probabilities for each class in classification tasks. The final fully connected layer typically has the same number of output nodes as the number ofclasses. Each fully connected layer is followed by a nonlinear function, such as ReLU. This layer is used to analyze the class probabilities and the output of this layer is the input of classifier. Here, ML binary classifier is applied to classify the various online social networks. The below figure shows the proposed CNN with ML based classification model,

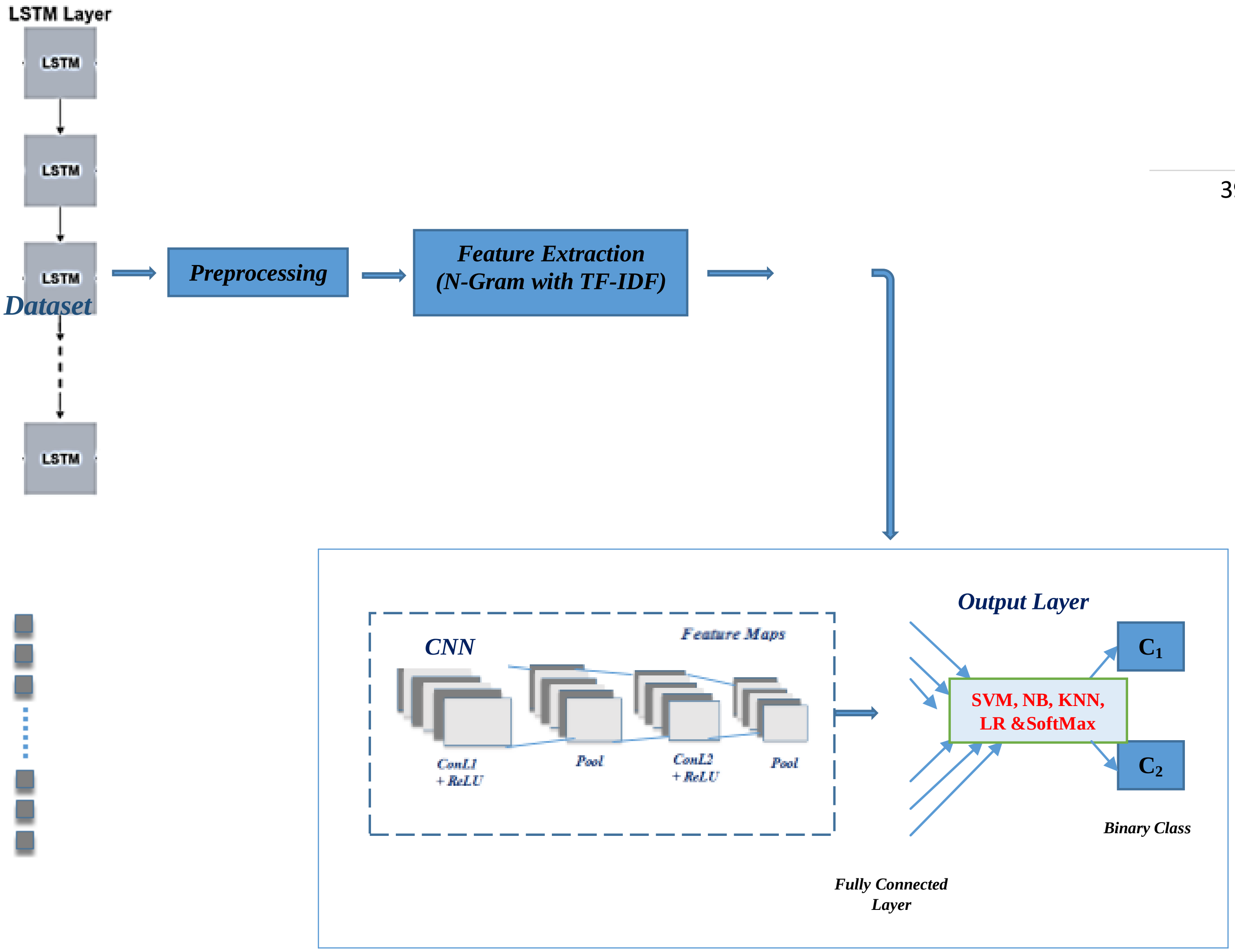

Figure.2.Proposed DL-ML based classification model

This work implements Twitter data using CNN-LSTM, a deep learning model for fake or real analysis that combines feature extraction methods such as TF-IDF, N-Grams and classify text using DL architecture. CNN cannot be utilized to generate long-distance dependencies from input text data due to the locality of the pooling and convolutional layers. However, a recurrent neural layer can effectively overcome this problem. Therefore, thelong-distance dependencies are captured using LSTM in our suggested model. Finally, thefully connected layer is passed on to machine learning classifiers such as SVM, NB, KNN, LR and SoftMax to classify the twitter into different categories (fake or real).To the best of our knowledge, this is the first time a deep CNN-LSTM hybrid model has been proposed for fake or real news classification. The suggested model's priority overhead is that CNN is utilized for feature optimization and LSTM layers are employed to capture long-term text data dependencies. In addition, the proposed model includes an additional LSTM layer, which improves the performance. Other current models employed the softmax function for classification, but we used typical machine learning classifiers. In a nutshell, our model includes the most recent word embedding, machine learning, deep learning, CNN, and recurrent neural network approaches. Figure 2 depictsthe working flow of the proposed method. The 2-dimensional output via ML binary class represents the 2 different categories (real or fake) of the detection problem and the results are discussed in following Section.

6.1 Proposed Algorithm Steps

Algorithm: Ensemble CNN-LSTM- SVM hybrid model algorithm.

Input: training text corpus.

Output: Ensemble model result

Step1: Parameter values initialization. Number of words a, word vector dimension w, sliding window size m, number of projection layer nodes j, iteration number Enum, sample error e, threshold i, logarithmic loss function Loss.

Step2: Feature Extraction (word2vec) uses N-Gram with TF-IDF model. Training to characterize words as vectors. Where a_j is the corresponding word, and c_i is i-th feature dimension corresponding to ($1\leq i\leq w$) word a_j.

$$a_j= [c_1, c_2, c_3 \ldots\ldots c_w]$$

Step3:Apply the CNN sliding window to get the window vector.

$$x_k= [a_k, a_{k+1}, a_{k+2}, a_{k+3} \ldots\ldots a_{k+m-1}]$$

Step4: Maximize the pooling after convolution togenerate word representation feature λ_k.The window vector feature matrix is expressed as $V_{k-m+1}, V_{k-m+2}, V_{k-m+3}, \ldots\ldots V_k$. Maximize each row of the featurematrix to obtain the maximum eigenvalue, then themaximized generated word feature is represented as:

$$\lambda_k = \max (V_{k-m+1}, V_{k-m+2}, V_{k-m+3}, \ldots\ldots V_k)$$

Step5: Each a corresponds to λt time t, input it into LSTM, and obtain the probability vector o_t.

$$o_t= \sigma (W_o h_{t-1} + U_0\lambda_t+b_0)$$

Step6: Apply the ML classifiers to classify the output of the LSTM. The probability of classifying
x as b is:

$$p(y^i = b|x^i, \theta) = \frac{e^{\theta_b^T x^i}}{\sum_{l=1}^{k} e^{\theta_b^T x^i}}$$

Step7: Calculate the loss function Loss.

$$\text{Loss } (y,P (y|x)) = -\log P(y|x)$$

7. Result and Discussion

Experimentation details

The whole experimentation is carried out in the 64-bit macOS system, Intel 2.6GHz 8-core i7 CPU, 16GB 2400MHz DDR4 memory, Radeon Pro 560X 4GB GPU. All the programs are implemented in the anaconda environment with python 3.8 programming.

Performance Metrics and Evaluation:

The proposed architecture makes use of deep learning (DL) methods to more precisely categorize fake news from online social media data. The table below displays the partitioned datasets used to train and test the method.

Table.1. Number of datasets used in total for training and testing.

S.No	Total Number of tweets	Training Data	Testing Data
01	25117	20094	5023

The proposed architecture is tested using a dataset in which the suggested method correctly classifies the appropriate categories. The details of the adjusted parameter are presented in the section below. To evaluate accuracy, recall, specificity, precision and f1-score are generated. The performance evaluations are calculated by mathematical equations and it provides the metrics needed to evaluate them.TP is True Positive Values, TN is True Negative Values, FP is False Positive and FN is False Negative Values.

Results and Findings: In this Section, the superiority of the suggested architecture and the effectiveness of other current models are displayed. The performance of several DL architectures, including accuracy, recall, specificity, precision and f1-score is shown in the tables below when determining whether news is fake or real.

Performance Evaluation Metrics

In this project, Deep learning Models performance are evaluated based on different performance evaluation metrics such as Accuracy (acc), recall (*R*), precision (*P*), specificity (*S*) and F1-score (*F*1).

- Accuracy: Accuracy of an algorithm is represented as the ratio of correctly classified by the model (TP+TN) to the total numbers (TP+TN+FP+FN).

$$ACC = (TN+TP) / (TN+FP+FN+TP)$$

- *Recall*: It is the ratio of the true positives classified correctly by the model. TP and FN values are used to calculate Recall.

$$R = TP / (TP+FN)$$

- *Specificity: It* is the ratio of the true negatives (those not from pathology) classified correctly by the model. TN and FP values are used to calculate specificity.

$$S = TN / (TN+FP)$$

- *Precision*: This parameter measures the proportion of anticipated positives that are true positives. Therefore, it is dependent on true positive (TP) and false positive (FP) values.

$$P= TP / (TP+FP)$$

- *F1-Score: It* is an overall measure of the model's accuracy that combines precision and recall. It is the double of the ratio of the multiplication of $F1$-score precision and recall metrics to their totals.

$$F1 = 2 \times (P \times R) / (P+R)$$

TP is True Positive Values, TN is True Negative Values, FP is False Positive and FN is False negative values.

Table.2. Performance Analysis with DL Algorithms

Algorithm Details	Performance Metrics				
	Accuracy	Recall	Specificity	Precision	F1-Score
DNN	87	85	88	86	86
LSTM	88	86	89	87	87
CNN	89	88	89	89	88
DNN -LSTM	90	91	90	91	91
CNN-LSTM	91	92	91	92	92

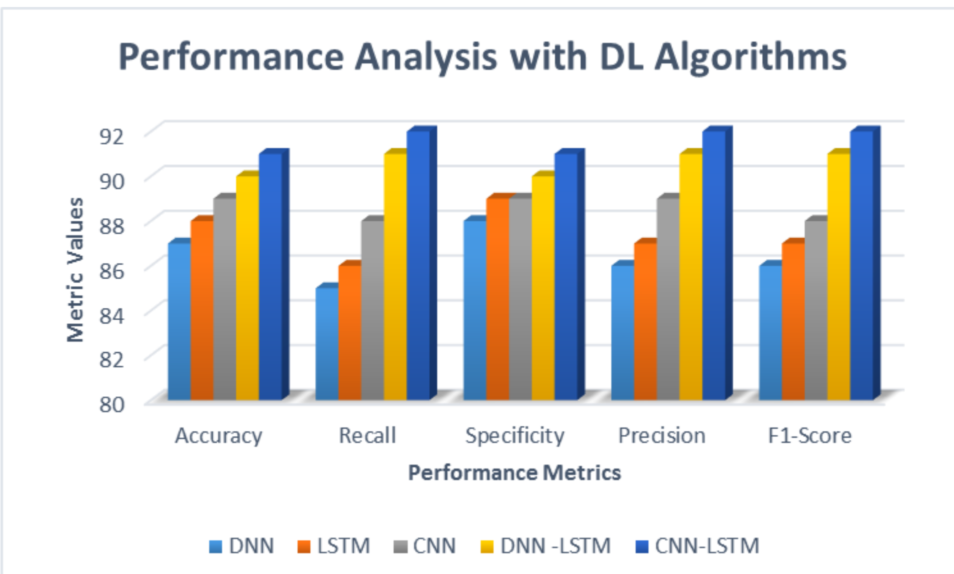

Figure.3.Comparative analysis of DL Algorithms

Table.3. Performance Analysis with ML-DL Algorithms

Algorithm Details	Performance Metrics				
	Accuracy	Recall	Specificity	Precision	F1-Score
CNN-LSTM-NB	93	94	93	92	92
CNN-LSTM-KNN	92	93	92	91	91
CNN-LSTM-LR	93	94	93	92	92
CNN-LSTM-SoftMax	94	95	94	93	93
CNN-LSTM-SVM	96	97	96	95	95

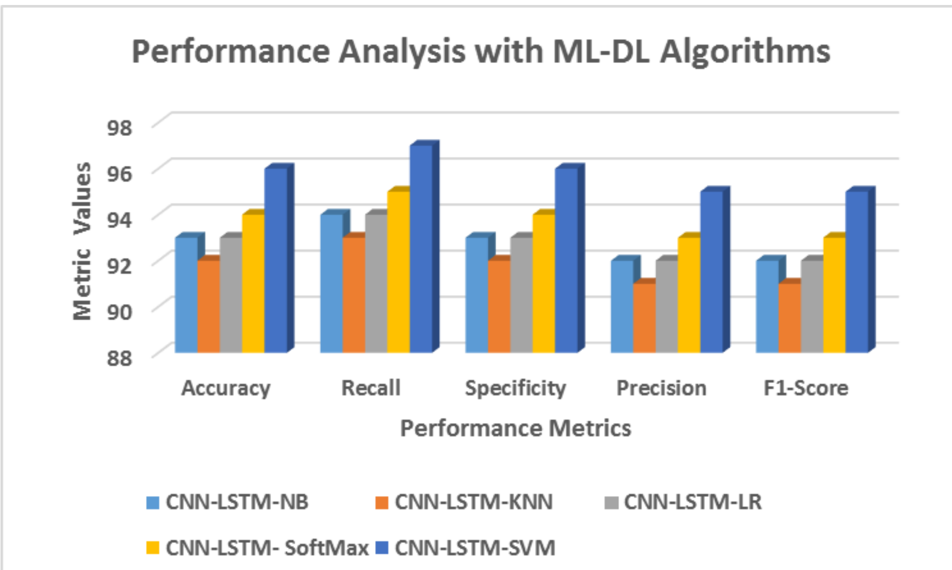

Figure.4.Comparative analysis of ML-DL (Ensemble) Algorithms

The various experiments are conducted on twitter datasets to confirm the results of the CNN–LSTM based ML model against existing models. The benchmark dataset includes 25117 user data gathered from social media source (Twitter) which is downloaded from the Kaggle website. The repository dataset contains 20094 and 5023 testing and training data. The proposed classification accuracy was compared to existing approaches. The research used various machine learning algorithms such as SVM, NB, KNN, LR & SoftMax with n-gram and TF-IDF features. Table 3 compares theachieved results of the proposed social media analysis with the existing techniques, in terms of accuracy, recall, specificity, precision and f1-score. The proposed CNN–LSTM based SVM method attained the highest performance with 96, 97, 96, 95 and 95for the accuracy, recall, specificity, precision and f1-score respectively, for the Twitter dataset. The two classes, which display the findings of classification trials, include the fake or real. The proposed technique has the highest accuracy of 96% in classifying data.

8. Conclusion

This paper has presented the application of deep learning models according to TF-IDF vectors as features (n-gram level TF-IDF) for the goal of discovering fake news. Thus, could arrive at the conclusion that the blending ensemble is the model that performs the best on the online social media datasets after developing the classifiers and running the experiments.This has been validated by employing a variety of metrics to measure performance. The results of the ensemble are compared with other methods. To enhance the current methods, this work proposed an ensemble model for detecting fake news.Due to the effectiveness of CNN and maintaining the long-term dependencies of LSTM with ML models, the suggested model outclassed well-known techniques on a twitter dataset, improving accuracyby up to 5%.Future plans include experimenting with other and larger datasets and varying the type, combination, and number of base models for the ensemble and also consider examining current trends on social media connected to fake news to incorporate them in our model for detection.

Reference:

1. Jamal Abdul Nasir, Osama Subhani Khan, IraklisVarlamis, Fake news detection: A hybrid CNN-RNN based deep learning approach,International Journal of Information Management Data Insights, Elsevier, Volume 1, Issue 1, 2021, 100007, ISSN 2667-0968.

2. K. Patel, Swati and I. U. Hassan, "Identifying Fake News with Various Machine Learning Model," 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2021, pp. 1-5.

3. S. Aphiwongsophon and P. Chongstitvatana, “Detecting fake news with machine learning method,” in Proceedings of the 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, pp. 18–21, Chiang Rai, Thailand, July 2018.

4. T. U. Yun and H. Ahn, “Fake news detection for Korean news using text mining and machine learning techniques,” Journal of Information Technology Applications and Management, vol. 25, pp. 19–32, 2018.

5. H. Rashkin, E. Choi, J. Y. Jang, S. Volkova, and Y. Choi, “Truth of varying shades: analyzing language in fake news and political fact-checking,” in Proceedings of the 2017

Conference on Empirical Methods in Natural Language Processing, pp. 7–11, Copenhagen, Denmark, September 2017.

6. Sneha Singhania, Nigel Fernandez, and Shrisha Rao. 3han: A deep neural network for fake news detection. In International Conference on Neural Information Processing, pages 572–581. Springer, 2017.

7. A. Thota, P. Tilak, S. Ahluwalia, and N. Lohia, "Fake news detection: a deep learning approach," SMU Data Science Review, vol. 1, p. 10, 2018.

8. William Yang Wang. " liar, liar pants on fire": A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648, 2017.

9. N. Ruchansky, S. Seo, and Y. C. Liu, "A hybrid deep model for fake news detection," in Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pp. 6–10, Singapore, November 2017.

10. Y. J. Lu and C. T. Li, "GCAN: graph-aware co-attention networks for explainable fake news detection on social media," 2020, https://arxiv.org/abs/2004.11648.

11. S. Vijayaraghavan, Y. Wang, Z. Guo et al., "Fake news detection with different models," 2020, https://arxiv.org/abs/2003.04978.

12. A.Roy, K. Basak, A. Ekbal, and P. Bhattacharyya, "A deep ensemble framework for fake news detection and classification," 2018, https://arxiv.org/abs/1811.04670.

13. H. S. Al-Ash, M. F. Putri, P. Mursanto, and A. Bustamam, "Ensemble learning approach on Indonesian fake news classification," in Proceedings of the 3rd International Conference on Informatics and Computational Sciences, pp. 29-30, Semarang, Indonesia, October 2019.

14. P. B. P. Reddy, P. K. Reddy, G. V. M. Reddy, and K. Mehata, "Fake data analysis and detection using ensembled hybrid algorithm," in Proceedings of the 2019 3rd International Conference on Computing Methodologies and Communication, pp. 27–29, Erode, India, March 2019

15. A. Ahmad, M. Yousaf, S. Yousaf, and M. O. Ahmad, "Fake news detection using machine learning ensemble methods," Complexity, vol. 2020, Article ID 8885861, 11 pages, 2020.

16. L. Gutierrez-Espinoza, F. Abri, A. S. Namin, K. S. Jones, and D. R. Sears, "Fake reviews detection through ensemble learning," 2020, https://arxiv.org/abs/2006.07912.

17. A. Mahabub, "A robust technique of fake news detection using ensemble voting classifier and comparison with

other classifiers," SN Applied Sciences, vol. 2, pp. 1–9, 2020.

18. S. Kaur, P. Kumar, and P. Kumaraguru, "Automating fake news detection system using multi-level voting model," Soft Computing, vol. 24, no. 12, pp. 9049–9069, 2020.

19. P. P. Shelke and A. N. Korde, "Support Vector Machine based Word Embedding and Feature Reduction for Sentiment Analysis-A Study," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020.

20. M. Wongkar and A. Angdresey, "Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler: Twitter," 2019 Fourth International Conference on Informatics and Computing (ICIC), IEEE, pp. 1-5, doi: 10.1109/ICIC47613.2019.8985884.

21. M. I. Ramadhon, A. Arini, F. Mintarsih and I. M. Malik Matin, "N-Gram and K-Nearest Neighbor Algorithm for Sentiment Analysis on Capital Relocation," 2021 9th International Conference on Cyber and IT Service Management (CITSM), IEEE, pp. 1-6, doi: 10.1109/CITSM52892.2021.9587919.

22. N. Alvi and K. H. Talukder, "Sentiment Analysis of Bengali Text using CountVectorizer with Logistic Regression," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021, pp. 01-05, doi: 10.1109/ICCCNT51525.2021.9580017.

23. S. Cao and P. Gao, "LSTM-Gate CNN network for Aspect Sentiment Analysis," 2020 5th International Conference on Information Science, Computer Technology and Transportation (ISCTT), pp. 443-447, IEEE, doi: 10.1109/ISCTT51595.2020.00084.

24. Kussul N, Lavreniuk M, Skakun S, Shelestov A (2017) Deep learning classification of land cover and crop types using remote sensing data. IEEE Geosci Remote SensLett 14(5):778–782.

25. Sherstinsky A (2020) Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. Phys D Nonlinear Phenom 404:132306.

26. Borah, G., Nimje, D., JananiSri, G., Bharath, K.P., Kumar, M.R. (2021). Sentiment Analysis of Text Classification Using RNN Algorithm. In: Kumar, S., Purohit, S.D., Hiranwal, S., Prasad, M. (eds) Proceedings of International Conference on Communication and Computational Technologies. Algorithms for Intelligent Systems. Springer, Singapore.

27. S. Li, K. Ma, X. Niu et al., "Stacking-based ensemble learning on low dimensional features for fake news

detection," in Proceedings of the 2019 IEEE 21st International Conference on High Performance Computing and Communications, pp. 10–12, Zhangjiajie, China, August 2019.

28. S. Tammina and S. Annareddy, "Sentiment Analysis on Customer Reviews using Convolutional Neural Network," 2020 International Conference on Computer Communication and Informatics (ICCCI), 2020, pp. 1-6, doi: 10.1109/ICCCI48352.2020.9104086.

29. H. Ahmad, M. U. Asghar, M. Z. Asghar, A. Khan and A. H. Mosavi, "A Hybrid Deep Learning Technique for Personality Trait Classification From Text," in IEEE Access, vol. 9, pp. 146214-146232, 2021, doi: 10.1109/ACCESS.2021.3121791.