International Journal of Informatics and Communication Technology (IJ-ICT)
Vol. 15, No. 1, March 2026, pp. 179~188
ISSN: 2252-8776, DOI: 10.11591/ijict.v15i1.pp179-188

NLP-based fraudulent biomedical news identification using LSTM-SGD deep learning algorithm

Siva Dhievaraj, Agusthiyar Ramu
Department of Computer Science and Applications, SRM Institute of Science and Technology, Ramapuram, Chennai, India

Article Info

Article history:

Received Oct 17, 2024
Revised Oct 15, 2025
Accepted Nov 5, 2025

Keywords:

Chi-squared and lasso
Deep learning
Feature extraction
Natural language processing
Stochastic gradient descent
TF-IDF

ABSTRACT

Concern over bio medical fake news is rising, particularly as false information about illnesses, medical procedures, and public health regulations becomes more prevalent. It is essential to recognize such false information, and deep learning (DL) algorithms can offer a potent remedy, especially when paired with sophisticated natural language processing (NLP) methods. This technique improves the model's capacity to ignore frequently used but uninformative terms and concentrate on important terminology. The model's capacity to concentrate on the most pertinent phrases for fake news identification is enhanced by the use of chi-squared, a statistical test that ascertains the dependency between various variables and aids in the removal of unnecessary data. By reducing less significant characteristics to zero, the Lasso approach, a kind of regression, is used for feature selection, guaranteeing that the model only utilizes the most predictive features for classification. A crucial step in getting the data ready for DL models is feature extraction, which turns unprocessed text into numerical data. After the structured data has been analyzed, algorithms like as stochastic gradient descent (SGD), long short-term memory (LSTM) may determine whether or not an article is accurate. The authenticity and dependability of medical information provided across platforms may be ensured by effectively identifying biomedical fake news by fusing DL with sophisticated NLP techniques.

Corresponding Author:

Siva Dhievaraj
Department of Computer Science and Applications, SRM Institute of Science and Technology
Ramapuram, Chennai, India
Email: sivad@srmist.edu.in

1. INTRODUCTION

In this study, we experimented with transformers, deep learning (DL) models, and machine learning classifiers in three different ways. To extract contextual characteristics from articles, we used word embedding in each trial. According to the findings of our experiments, DL models performed more accurately than machine learning classifiers and the bidirectional encoder representations from transformers (BERT). Furthermore, the findings indicated that the long short-term memory (LSTM) and gated recurrent unit (GRU) models had about equal accuracy. Using machine or DL models in conjunction with an enriched linguistic feature set, we demonstrated that we can detect bogus news with high accuracy [1]-[4].

The research revealed an evolutionary trend toward the present usage patterns of pre-trained natural language processing (NLP) in a variety of scenarios. This study focuses on neural networks, sequence-to-sequence models, multitasking NLP models, word embedding, and attention processes. Several techniques and algorithms were used in this process. These advancements appear to have benefited the healthcare industry greatly. There is a lot of promise in using NLP for tourism communication. NLP's creative and strong skills may contribute to the advancement of tourism-related services. It is impossible to overstate the benefit that technology provides to the travel and tourism sector as it develops and advances. At the conclusion, a number of this review's drawbacks are noted. There are ramifications for tourist messaging from this study [5]-[8]. The study will examine the underlying mathematics of these models and how they have been used in NLP. The research will also look at the most recent developments in DL and NLP while recognizing the major obstacles and new developments in the area. Lastly, the research will clarify the latest DL frameworks and libraries as well as the resources that are accessible. The study's findings and implications show that LLMs struggle with pragmatic language because they over-rely on statistical learning techniques and fail to fully understand social norms, context, implicature, and assumption. Last but not least, we have provided a comprehensive analysis of NLP domains where DL models like DBN, CNN, and RNN are effectively used and establish new performance boundaries, including speech recognition, machine translation, text classification, POS tagging, NER, and question answering [9]-[12].

This research proposes a framework that detects hate speech on social media by integrating machine learning with NLP. Subsequently, the gathered textual data is analyzed using a powerful ensemble-based DL optimization approach tailored for NLP tasks. This technique uses an efficient learning process to categorize the content into hateful, insulting, and neutral words in order to identify hate speech on social media platforms. Metrics like recall, f-score, precision, and overall accuracy are then used to assess the system's quality. The system achieved 98.71% accuracy with the lowest deviations of mean square error (0.019), cross entropy loss (0.015), and logarithmic loss L-0.0238, respectively [13]-[17]. This paper practice a computer-based technique that uses ML and NLP from electronic clinical notes to automatically identify gout flares. In two separate sets of 100 patients, rheumatologists reviewed 1,264 and 1,192 clinical notes, respectively, out of 16,519 individuals, which served as the training and assessment data sets. To capture various facets of gout flares, we used distinct NLP searches. Each note's final categorization decisions were determined by the machine learning system using the NLP search results as inputs. We found that 18,869 clinical notes for 16,519 patients with a gout diagnosis and a prescription were gout flare-positive (sensitivity 82.1%, specificity 91.5%) for a urate-lowering drug: 5,954 patients had one or two flares, 1,402 patients had three or more flares (sensitivity 93.5%, specificity 84.6%), and 9,163 patients had no flares (sensitivity 98.5%, specificity 96.4%) [18]-[25].

The rest of this paper is organized into the following key sections: Section 2 reviews existing research on biomedical false news detection using DL techniques, emphasizing the various methodologies and strategies proposed by different scholars. Section 3 details the workflow of the proposed framework, encompassing stages such as data collection, preprocessing, feature extraction, and classification through advanced DL models. Section 4 offerings the findings of the biomedical false news identification process, as well as analysis and performance data. Section 5 concludes by outlining possible future research directions in the use of DL algorithms to improve the detection of biomedical false news, and it includes pertinent references.

2. RELATED WORK

Abdelminaam *et al.* [7] this research offers a better Using a deep neural network to identify false information. In terms of DL methods, the Modified-LSTM and Modified GRU have one to three layers. NB, LR, k closest neighbors, DT, RF, and SVM are the six machine learning approaches. Essential characteristics were extracted from the word Four benchmark datasets are used to test the suggested deep neural network methods of embedding feature extraction and the baseline machine learning model of TF-ID with N-gram. Findings using the suggested framework demonstrate great accuracy in identifying tweets that include COVID-19 information as well as those that do not. Compared to baseline machine learning models, our results show a significant boost in performance. Prachi *et al.* [8] this study uses a range of ML, DL, and NLP techniques, like, transformer-based bidirectional encoder representation, and extended short-term memory, to detect bogus news. To determine if the content is accurate, machine learning and DL algorithms are first trained on an open-source dataset for detecting false news. Tokenization, Regex, lemmatization, stop words and term frequency-inverse document frequency are some of the feature engineering techniques used to construct feature vectors. Sadeghi *et al.* [9] using natural language inference (NLI), we provide a novel method for identifying bogus news in this study. The proposed method, which focuses on assessing authenticity through a set of reliable news, leverages a human-like approach instead of depending solely on statistical aspects of the news's context or content. The NLI approach improves a number of deep and classical machine learning models, such as k-nearest neighbors, random forest, BiGRU, BiLSTM, Naïve Bayes, decision tree, random forest, and logistic regression. Tests on the FNID-FakeNewsNet and FNID-LIAR datasets demonstrate that the proposed method achieves 10.44% and 13.19% absolute improvements, respectively, with accuracies of 85.58% and 41.31%.

Magistris *et al.* [10] this paper suggests an autonomous FND system that provides a selection of articles from trustworthy sources and either confirms or refutes the questionable assertions. Numerous modules make up the system, which integrates a variety of dl, machine learning, and NLP approaches. To choose relevant articles and find those that support the tested claim and its views, several techniques are used. To confirm the accuracy of posts on the COVID-19 pandemic, vaccines, and therapies, the recommended method will be applied. Babar *et al.* [11] this file In this regard, DL provides a way to detect and counteract false information. The identification of bogus news is enhanced using a hybrid N-gram and LSTM model, which increases computation time, accuracy, and recall rate. This proposed method uses a classifier to categorize fake news. Because of its parallel and distributed platform foundation, it can use big data analytics to construct the DL model. This platform increases the accuracy of the suggested model and speeds up training and testing. The proposed approach separates content into. "fake news" and "real news" to create a system that can accurately and infrequently identify misleading information. Next, the outcomes are measured. The findings show that combining big data's deep neural network (DNN) with Spark architecture makes the suggested model extremely successful.

Alghamdi *et al.* [12] as to this study, identifying COVID-19 false news is one of the most important problems in the NLP industry. Several machine learning techniques and transformer-based models that have already been trained, includingCOVID-Twitter-BERT, are examined in this study for their efficacy in identifying COVID-19 bogus news. The actual world BiGRU on top of the CT-BERT model yields outstanding results, as demonstrated by the COVID-19 false news dataset, which has a state-of the-art F1 score of 98%. Our findings show that advanced machine learning algorithms can identify bogus news, which has significant ramifications for slowing the spread of misinformation on COVID-19.

Zhao *et al.* [13] an technique using DL algorithms was presented to detect false information and fake news inCOVID-19 news stories. The optimal approach is determined using DBNs, CNNs, and LSTM. Model performance indicators including accuracy, AUC score, and F1 score were analyzed in the study, indicates that LSTM and CNN are the most effective models for detecting COVID-19 false news, with up to 94% accuracy. This study concludes with an algorithm-based rating system for mainstream media credibility. According to the findings, big US media outlets are trustworthy sources of COVID-19 news and information.

The reviewed studies highlight the effectiveness of DL models like LSTM, GRU, and transformer-based architectures in detecting fake news, especially related to COVID-19. Techniques such as TF-IDF, n-grams, and word embeddings improve feature representation, while preprocessing enhances model accuracy. Some works utilize big data platforms and NLI for better scalability and reasoning. Overall, integrating advanced NLP, machine learning, and DL ensures robust and accurate misinformation detection.

3. PROPOSED WORK

It is vital to notice the rising incidence of bogus biomedical news since it poses a major risk to public health. The proposed system aims to correctly recognize and classify fake biomedical news using DL algorithms and methods for NLP. To understand and evaluate textual data, NLP, is essential shown in Figure 1. Several NLP approaches are used by the system to process news stories on biomedicine. The first step is to normalize the data and remove noise using text preparation methods such as stemming/lemmatization, tokenization, and stop word removal. Term Frequency-Inverse Document Frequency is used by the system. To translate the text into numerical characteristics after preprocessing. Key phrases in biomedical news may be effectively identified by using TF-IDF, which down weights common terms and gives greater weights to words that are distinct and pertinent to the document's context. Chi-Squared or Lasso (L1 regularization) are used for feature selection in order to improve model performance. The algorithm can detect statistically significant phrases by using the Chi-Squared test, which quantifies the relationship between words and the target label (genuine or false). However, by driving the coefficients of unnecessary characteristics to zero, Lasso reduces the dataset's dimensionality and keeps just the most crucial features for classification.

3.1. Text pre-processing

During the procedure of text analysis and preparing test data from raw text, text preprocessing is very important. Among the main steps in this process, one can mention stopword removal: it utilizes the NLTK stop words algorithm, which erases frequently used words such as the, and, and. While these words are known to appear frequently in language, they often have very few terminological connotations, and using them burdens the dataset. Deleting them helps to maintain greater meaningfulness in those remaining terms, which, in turn, improves efficiency in analysis. Another vital procedure entails lemmatization; the current experiment employs the NLTK Word Net Lemmatizer algorithms. Lemmatization removes inflection from a word and reboots it to the lowest stem, for example, running to run. This normalization step is helpful to remove any variation of words as part of the text data and, thus, the quality of feature extraction for further steps. In combination, the process of data preprocessing prepares this dataset for subsequent processing to be more efficient as well as less complicatedfeature extraction.

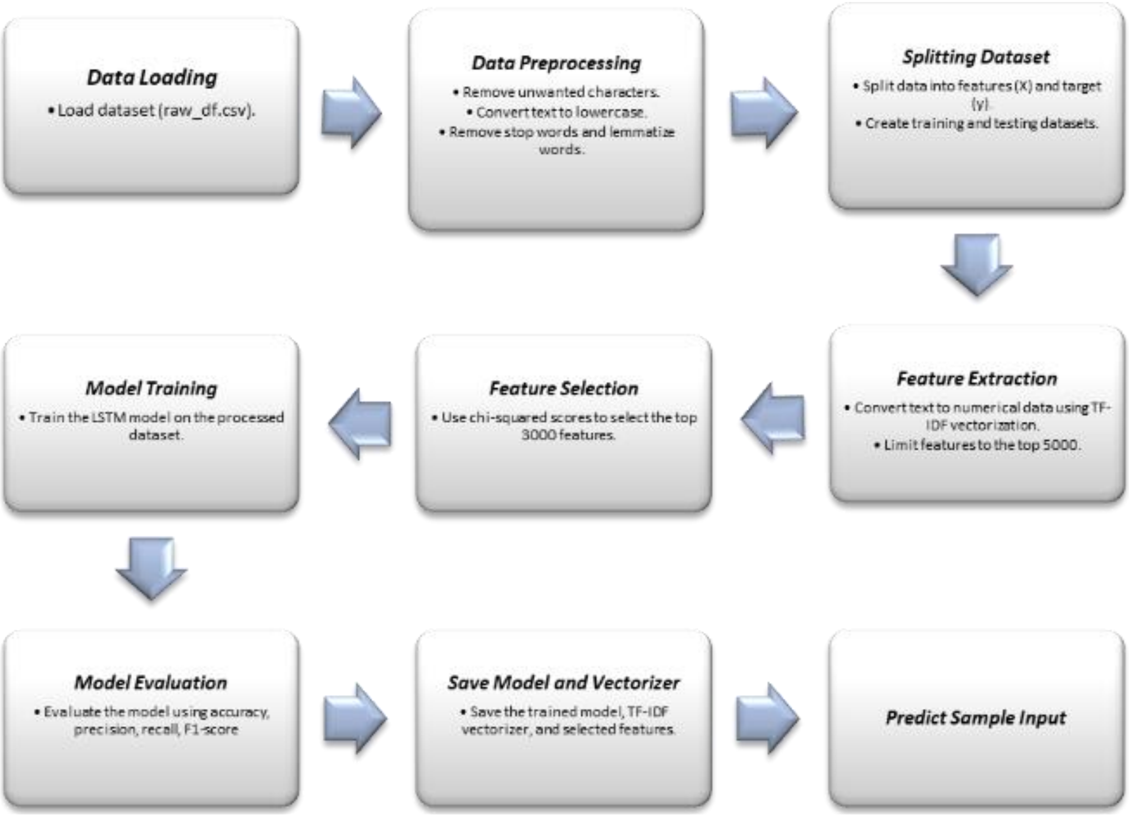

Figure 1. Proposed block diagram

3.2. Feature extraction with TF-IDF

Feature extraction is essential to the field of fake news identification because it converts unprocessed text input into a format that feature vectors can use in machine learning models. It's important to note the popular term frequency-inverse document frequency (TF-IDF) method in this context. In this instance, the TfidfVectorizer is used to implement the TF-IDF technique. This technique converts texts into numerical attributes by computing the relevance value of each term within a document in relation to a set of documents, underscoring its importance in our work. Term frequency (TF), a part of the TF-IDF technique, identifies the extent to which a term was used specifically in a document and how important it is in that document.

3.3. Feature selection

This paper covers feature selection as an essential step in feature subset selection for improving machine learning model performances. Classification algorithms include the Chi-Squared Test that divides this line into two parts and selects features most correlated with the target class, e.g., the division of real news from fake news. It determines an individual feature’s independence from the target via calculations of the observed frequency against the expected frequency. When it comes to choosing the features with high Chi-squared scores, the model complexity is decreased, computational speed is optimized, and any noise is rejected. In fake news detection, the method recognizes keywords and patterns, which make the model more accurate and efficient, especially for big text data sets.

To compute the Chi-Squared statistic, use the formula below:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{1}$$

Where:
O_i Observed frequency of the i-thfeature in the dataset.
E_i Calculated under the premise that the feature and the target class are independent, the expected frequency of the third feature.

3.4. long short-term memory (LSTM)

Particularly helpful are recurrent neural networks (RNNs), particularly LSTM networks, which may learn long-term dependencies in data due to the sequential structure of material, like text. The vanishing gradient problem is resolved by feed forward recurrent neural networks, which include LSTM networks when working with sequences in text. For the same reason, LSTMs are suitable for analyzing textual data in which the meaning of a word is relative to the preceding text in the document. Mathematically, the operations inside an LSTM can be described using the following formulas:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{2}$$

Which data from the prior time step will be eliminated is decided by the forget gate.

f_t The forget gate's output, W_f The weight matrix is, h_{t-1} is the concealed state that existed before, x_t is the input that is currently being used and b_f is the phrase used to describe prejudice. The sigmoid function σ squashes the values between 0 and 1.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

The input gate regulates the amount of fresh data that is added to the cell state. In this case, i_t stands for input gate output, and W_i, h_{t-1}, and x_t are as previously defined, with b_i being the bias.

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4}$$

The cell state C_t is updated by combining the previous state C_{t-1}, regulated by the new data and the forget gate $\tilde{C}_t$, controlled by the input gate.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

Which aspect of the cell state will be output at the current time step is decided by the output gate. $o_t$is the output, and W_o, h_{t-1}, and x_t are the current input, the weight matrix, and the prior hidden state, respectively, with b_o as the bias.

$$h_t = o_t \cdot \tanh(C_t) \tag{6}$$

The final output or hidden state h_t is a combination of the cell state C_t, processed by the output gate o_t and passed through the hyperbolic tangent activation function tanh, which squashes the values between -1 and 1.

3.5. Stochastic gradient descent (SGD)

In machine learning, Stochastic Gradient Descent, or SGD, is a very popular optimization approach, particularly for training models in TensorFlow and Keras. When the model is being trained to categorize features, the optimizer is a minimum in the loss function, which modifies the neural network values. The basis of SGD is reflecting the parameter (weight) of the model through iterations by computing the partial derivative of the given loss function to the parameter. This action using random values is the most efficient approach to identifying the weights that provide the lowest loss, on which the model for the prediction of outcomes relies. The formula to use in order to update weights in SGD is as follows:

$$w = w - \eta \cdot \nabla_w J(w) \tag{7}$$

Where:
w represents the model weights.
η is the learning rate, which establishes the size of the steps at every iteration.
∇_w is the gradient of the loss function $J(w)$with respect to the weights.

In fake news detection, the loss function typically used is binary cross-entropy, as the task involves classifying news as either "fake" or "real". The binary cross-entropy formula is:

$$J(w) = -\left(y \cdot \log(\hat{y}) + (1 - y) \cdot \log(1 - \hat{y})\right) \quad (8)$$

Where:
y is the true label (either 0 for real or 1 for fake news).
$\hat{y}$ is the predicted probability of the news being fake (i.e., the output from the model).

To accelerate convergence, momentum is often incorporated into SGD, helping the algorithm escape local minima and smooth out updates. Momentum introduces a moving average of previous gradients to the weight update formula:

$$v_t = \beta v_{t-1} + (1 - \beta)\nabla_w J(w) \quad (9)$$

$$w = w - \eta \cdot v_t \quad (10)$$

Where:
v_t is the momentum term or velocity at time step t.
β is the momentum factor, which is usually set at 0.9.
This momentum-based approach enables SGD to "speed up" learning in the right direction and suppress fluctuations in other wrong directions.

4. RESULTS AND DISCUSSION

This section presents a comprehensive evaluation of the experimental results obtained through a series of data preprocessing, feature engineering, model training, and performance assessment steps. The implementation was carried out using Python programming, leveraging libraries such as NumPy, Pandas, scikit-learn, TensorFlow, Keras, Matplotlib, and Seaborn for data manipulation, model training, and result visualization. The primary objective of the experiment was to detect and classify fake news using DL techniques, specifically a LSTM based model, evaluated against traditional metrics and visual indicators. Figure 2 illustrate the impact of preprocessing, the effectiveness of the model, and the statistical robustness of the proposed framework.

Before preprocessing, the actual text data enter the model in a raw, unadulterated form that may include unwanted characters, words that may not have any relevance to the issue under study, and variance in the formatting that contributes to depending on the word count, which makes the range lower in the Before Preprocessing distribution.

The TF-IDF scores in the corpus both before and after preprocessing are shown in Figure 3. Unwanted characters, superfluous words, and formatting errors may be present in the raw, unadulterated text input that enters the model prior to preprocessing. These problems lead to a more scattered distribution in the "Before Preprocessing" representation by increasing the variability in the TF-IDF scores. Preprocessing results in a more refined and consistent distribution, referred to as the "After Preprocessing" distribution. Noise is greatly reduced at this level by actions like lemmatization, stopwords, single characters, and punctuation removal. The quality of feature extraction is enhanced by this cleaning procedure, which also makes the processed text more uniform and pertinent for analysis and model training.

Figure 4 is a graphic picture of the model classifying capabilities between classes. The ROC curve illustrates how the False Positive Rate, which is equal to one minus the specificity or the percentage of sound negative cases that are accurately identified as such at each potential continental index threshold, affects the True Positive Rate, or sensitivity, also known as the probability of detecting an actual case. With an efficient model, the obtained curve swings near the upper left-hand corner, this shows high accurate positive percentages and low false positive percentages. The area under the depicted ROC curve serves as the sole metric for assessing the performance of the model. An AUC value near 1 signifies a good level of prediction or discriminative ability, while a value approaching 0.5 shows random performance. As depicted in the figure, the model earned an AUC score of 0.85, which is suitable for correctly differentiating the target classes.

Figure 4 is a graphic picture of the model classifying capabilities between classes. The ROC curve illustrates how the False Positive Rate, which is equal to one minus the specificity or the percentage of sound negative cases that are accurately identified as such at each potential continental index threshold, affects the True Positive Rate, or sensitivity, also known as the probability of detecting an actual case. With an efficient model, the obtained curve swings near the upper left-hand corner, this shows high accurate positive percentages and low false positive percentages. The area under the depicted ROC curve serves as the sole metric for assessing the performance of the model. An AUC value near 1 signifies a good level of prediction or discriminative ability, while a value approaching 0.5 shows random performance. As depicted in the figure, the model earned an AUC score of 0.85, which is suitable for correctly differentiating the target classes.

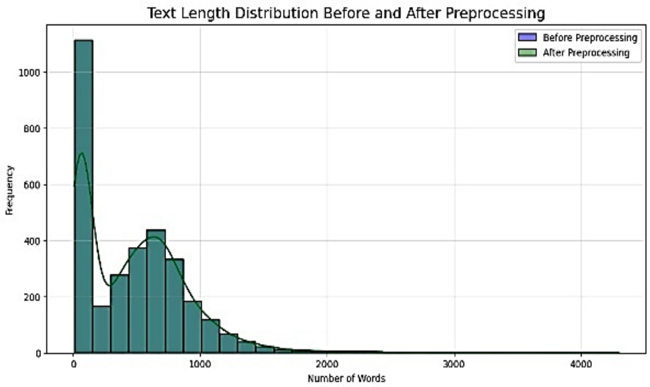

Figure 2. Before and after pre-processing

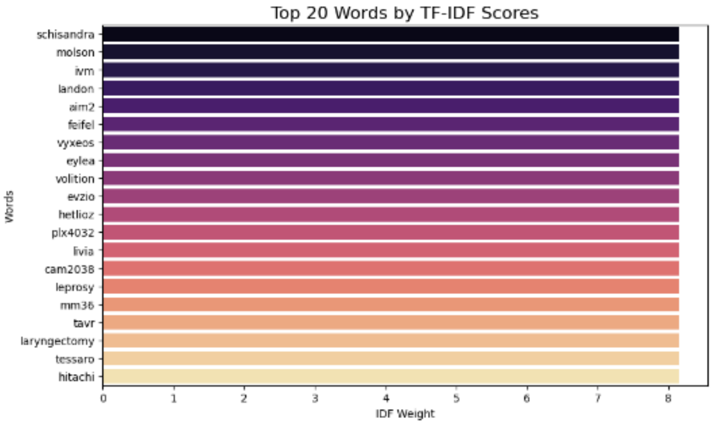

Figure 3. Displays the TF-IDE scores

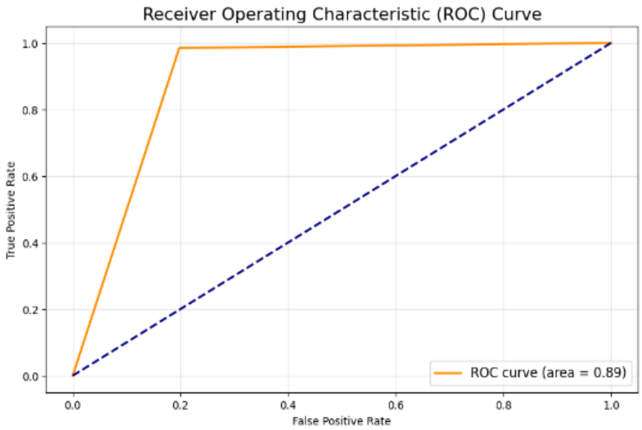

Figure 4. ROC curve of fake news model

The Figure 5 shows the changes in the training and validation accuracy for LSTM during the number of epochs during training. Firstly, both accuracies rise consistently with each epoch since the operation can indicate some level of learning and the capacity to overcome over-fitting. But, as can be seen in later epochs, in training from epoch 30-70, training accuracy increases, although validation accuracy either starts to level off or may even vary somewhat. This suggests overfitting, where the model discriminates very well in the training data but ultimately fails in the validation data set. This is perfectly illustrated in the visualization to call attention to tracking these metrics to manage learning and minimize overfitting in DL.

The loss (training and validation) over epochs is described in the Figure 6 showing the model's training. First, the training and validation loss is high because the neural network is not familiar with the data. The figure also points to epoch 10, and as the epochs advance, the training loss is getting smaller, and this means that the model is learning from the training data. The conclusion is the same regarding the validation loss – it is also less but could be slightly oscillating to indicate how the model performs on unseen examples. A confusion matrix as in Figure 7 highlights the distribution of actual versus predicted labels, offering insights into the model's behavior for each class.

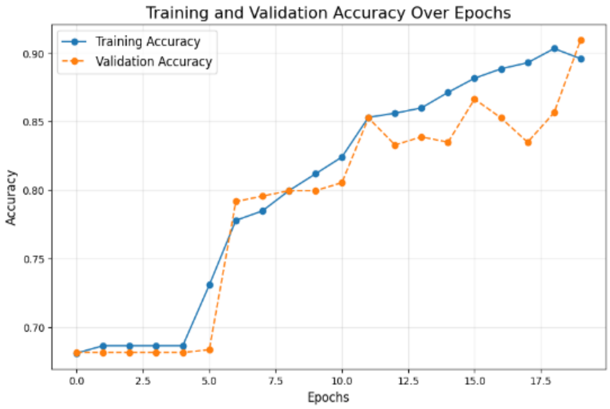

Figure 5. Training and validation accuracy over epochs

Figure 6. Training and validation loss over epochs

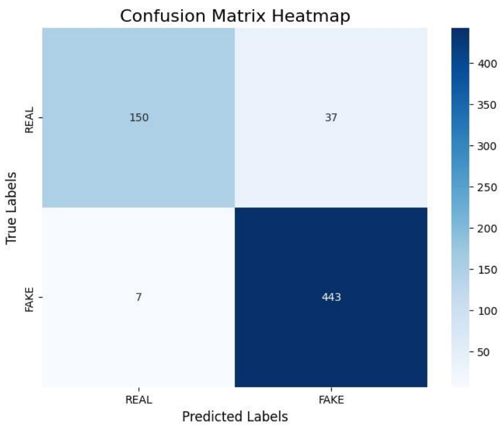

Figure 7. Plot confusion matrix as a heatmap

5. CONCLUSION

Therefore, the increase in biomedical fake news indicates there is a need for better techniques to detect fake news, especially in sensitive areas such as diseases, treatments and health regulations. The proposed solution is a combination of DL algorithms coordinated with modern NLP. The identification of primary features through pre-processing, the TF-IDF approach used in determining relevant terms, and the Chi-square methods used in or in the selection of proper features improve the effectiveness of the detection of fake news. Moreover, the approach guarantees the use of only the most essential features in achieving enhanced focus and, consecutively, better performance. Works such as feature extraction, which converts text into numerical data that is understandable to DL models, including CNNs and RNNs, can help in the classification of real and fake articles. By combining these advanced methodologies, the approach will enhance the reliability and authenticity of information disseminated in biomedical informatics to meet the essential need of containing the prevalence of fake news in healthcare.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Siva Dhievaraj	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Agusthiyar Ramu	✓		✓	✓			✓			✓	✓		✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] J. Y. Khan, M. T. I. Khondaker, S. Afroz, G. Uddin, and A. Iqbal, "A benchmark study of machine learning models for online fake news detection," *Machine Learning with Applications*, vol. 4, p. 100032, Jun. 2021, doi: 10.1016/j.mlwa.2021.100032.

[2] B. Al-Ahmad, A. M. Al-Zoubi, R. Abu Khurma, and I. Aljarah, "An evolutionary fake news detection method for COVID-19 pandemic information," *Symmetry*, vol. 13, no. 6, p. 1091, Jun. 2021, doi: 10.3390/sym13061091.

[3] G. Jain, D. Mittal, D. Thakur, and M. K. Mittal, "A deep learning approach to detect COVID-19 coronavirus with X-Ray images," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 4, pp. 1391–1405, Oct. 2020, doi: 10.1016/j.bbe.2020.08.008.

[4] F. Sun, J. Sun, and Q. Zhao, "A deep learning method for predicting metabolite–disease associations via graph neural network," *Briefings in Bioinformatics*, vol. 23, no. 4, Jul. 2022, doi: 10.1093/bib/bbac266.

[5] S. Savaş, "Detecting the stages of Alzheimer's disease with pre-trained deep learning architectures," *Arabian Journal for Science and Engineering*, vol. 47, no. 2, pp. 2201–2218, Sep. 2021, doi: 10.1007/s13369-021-06131-3.

[6] J. S. Chorba *et al.*, "Deep learning algorithm for automated cardiac murmur detection via a digital stethoscope platform," *Journal of the American Heart Association*, vol. 10, no. 9, May 2021, doi: 10.1161/jaha.120.019905.

[7] D. S. Abdelminaam, F. H. Ismail, M. Taha, A. Taha, E. H. Houssein, and A. Nabil, "CoAID-DEEP: an optimized intelligent framework for automated detecting COVID-19 misleading information on Twitter," *IEEE Access*, vol. 9, pp. 27840–27867, 2021, doi: 10.1109/access.2021.3058066.

[8] N. N. Prachi, M. Habibullah, M. E. H. Rafi, E. Alam, and R. Khan, "Detection of fake news using machine learning and natural language processing algorithms," *Journal of Advances in Information Technology*, vol. 13, no. 6, 2022, doi: 10.12720/jait.13.6.652-661.

[9] F. Sadeghi, A. J. Bidgoly, and H. Amirkhani, "Fake news detection on social media using a natural language inference approach," *Multimedia Tools and Applications*, vol. 81, no. 23, pp. 33801–33821, Apr. 2022, doi: 10.1007/s11042-022-12428-8.

[10] G. De Magistris, S. Russo, P. Roma, J. T. Starczewski, and C. Napoli, "An explainable fake news detector based on named entity recognition and stance classification applied to covid-19," *Information*, vol. 13, no. 3, p. 137, Mar. 2022, doi: 10.3390/info13030137.

[11] M. Babar, A. Ahmad, M. U. Tariq, and S. Kaleem, “Real-time fake news detection using big data analytics and deep neural network,” *IEEE Transactions on Computational Social Systems*, vol. 11, no. 4, pp. 5189–5198, Aug. 2024, doi: 10.1109/tcss.2023.3309704.
[12] J. Alghamdi, Y. Lin, and S. Luo, “Towards COVID-19 fake news detection using transformer-based models,” *Knowledge-Based Systems*, vol. 274, p. 110642, Aug. 2023, doi: 10.1016/j.knosys.2023.110642.
[13] P. Zhao, X. Chen, and X. Wang, “Fake news and misinformation detection on headlines of COVID-19 using deep learning algorithms,” *International Journal of Data Science*, vol. 5, no. 4, p. 316, 2020, doi: 10.1504/ijds.2020.10038925.
[14] W. C. Serena Low *et al.*, “An overview of deep learning techniques on chest X-Ray and CT scan identification of COVID-19,” *Computational and Mathematical Methods in Medicine*, vol. 2021, pp. 1–17, Jun. 2021, doi: 10.1155/2021/5528144.
[15] A. S. Jaradat *et al.*, “Automated monkeypox skin lesion detection using deep learning and transfer learning techniques,” *International Journal of Environmental Research and Public Health*, vol. 20, no. 5, p. 4422, Mar. 2023, doi: 10.3390/ijerph20054422.
[16] L. Sawalha and T. C. Akinci, “Shallow learning versus deep learning in natural language processing applications,” in *Shallow Learning vs. Deep Learning*, Springer Nature Switzerland, 2024, pp. 179–206.
[17] G. Huang, Y. Li, S. Jameel, Y. Long, and G. Papanastasiou, “From explainable to interpretable deep learning for natural language processing in healthcare: How far from reality?,” *Computational and Structural Biotechnology Journal*, vol. 24, pp. 362–373, Dec. 2024, doi: 10.1016/j.csbj.2024.05.004.
[18] M. Q. Alnabhan and P. Branco, “Fake news detection using deep learning: a systematic literature review,” *IEEE Access*, vol. 12, pp. 114435–114459, 2024, doi: 10.1109/access.2024.3435497.
[19] Q. Chen *et al.*, “Artificial intelligence in action: addressing the COVID-19 pandemic with natural language processing,” *Annual Review of Biomedical Data Science*, vol. 4, no. 1, pp. 313–339, Jul. 2021, doi: 10.1146/annurev-biodatasci-021821-061045.
[20] P. K. Roy and S. Chahar, “Fake profile detection on social networking websites: a comprehensive review,” *IEEE Transactions on Artificial Intelligence*, vol. 1, no. 3, pp. 271–285, Dec. 2020, doi: 10.1109/tai.2021.3064901.
[21] K. Kamalakannan, G. Anandharaj, and M. A. Gunavathie, “Performance analysis of attributes selection and discretization of Parkinson’s disease dataset using machine learning techniques: a comprehensive approach,” *International Journal of System Assurance Engineering and Management*, vol. 14, no. 4, pp. 1523–1529, Jun. 2023, doi: 10.1007/s13198-023-01960-x.
[22] M. Viviani, C. Crocamo, M. Mazzola, F. Bartoli, G. Carrà, and G. Pasi, “Assessing vulnerability to psychological distress during the COVID-19 pandemic through the analysis of microblogging content,” *Future Generation Computer Systems*, vol. 125, pp. 446–459, Dec. 2021, doi: 10.1016/j.future.2021.06.044.
[23] L. Mathew and V. R. Bindu, “Efficient classification techniques in sentiment analysis using transformers,” in *International Conference on Innovative Computing and Communications*, Springer Singapore, 2021, pp. 849–862.
[24] V. Uma and V. Ganesh, “A novel classification framework using transformer-based encoding and low-rank tensor fusion for enhanced classification and efficiency,” *Advance Sustainable Science Engineering and Technology*, vol. 7, no. 2, p. 2502013, Apr. 2025, doi: 10.26877/jgz0xe27.
[25] A. B. F. Khan, K. Kamalakannan, and N. S. S. Ahmed, “Integrating machine learning and stochastic pattern analysis for the forecasting of time-series data,” *SN Computer Science*, vol. 4, no. 5, Jun. 2023, doi: 10.1007/s42979-023-01981-0.

BIOGRAPHIES OF AUTHORS

Mr. Siva Dhievaraj is a research scholar in the Department of Computer Science at SRM Institute of Science and Technology, Ramapuram Campus. He holds degrees in Physics, Computer Applications, and Computer Science Engineering (B.Sc., MCA, M.Tech., Ph.D.) from Bharathidasan University and SRM IST. His teaching and research interests span Cloud Computing, Machine Learning, and Programming. He has published several research papers in reputed journals, focusing on topics such as fake news detection, spammer identification, cloud data security, and predictive modeling using machine learning. He can be contacted at email: sivad@srmist.edu.in.

Dr. Agusthiyar Ramu is a Professor and the Head of the Department of Computer Applications at SRM Institute of Science and Technology, Ramapuram Campus. He specializes in Data Mining and Artificial Intelligence, areas where he has made significant contributions through his research and teaching. Dr. Agusthiyar earned his Ph.D. in Data Mining from Anna University in 2017, and his M.Phil. in Computer Science from Pride, Periyar University, Salem, in 2008. He also holds an MCA in Computer Applications from Madurai Kamaraj University, awarded in 2001, and a B.Sc. in Mathematics from Dr. Zakir Hussain College, Ilayangudi, affiliated to Madurai Kamaraj University, awarded in 1998. In his academic career, Dr. Agusthiyar has taught various courses including Data Mining and Warehousing and Software Engineering. His research interests lie in data mining, artificial intelligence, and machine learning, where he continues to explore and innovate. As the Head of the Department, he plays a crucial role in shaping the curriculum and guiding the department towards academic excellence. He can be contacted at email: hod.dca.rmp@srmist.edu.in.