Detecting Fake News through Deep Learning and Natural Language Programming

B. SANTOSH KUMAR

Associate Professor, Department of MCA, Wesley PG College, Secunderabad, India.

Abstract - In the era of information abundance, the proliferation of fake news poses a significant challenge to reliable information dissemination. This paper presents an innovative approach to combatting the spread of misinformation by leveraging GloVe (Global Vectors for Word Representation) in tandem with deep learning techniques for fake news detection. GloVe is employed to encode semantic meaning and relationships between words, providing a robust foundation for understanding context within textual data. The study commences with the preprocessing of textual information, utilizing GloVe embeddings to transform words into numerical vectors while preserving their contextual significance. These embeddings are then integrated into a deep learning model, designed to capture intricate patterns and nuances indicative of fake news. The model is trained on a diverse dataset encompassing both legitimate and deceptive content, optimizing its ability to discern subtle linguistic cues associated with misinformation. Evaluation metrics such as precision, recall, accuracy, and F1 score are employed to assess the model's performance in detecting fake news.

Keywords: Fake news detection, GloVe (Global Vectors for Word Representation), Deep learning techniques, Semantic meaning

I. INTRODUCTION

The rapid spread of false information and deceptive news has become a significant issue in the modern digital environment, infiltrating various online platforms such as social media, digital news outlets, and other digital domains at an alarming speed [1]. The proliferation of misleading information, often crafted with the intention of influencing public sentiment, shaping political narratives, or disseminate lies, presents a substantial danger to the exchange of ideas within society, public confidence, and democratic mechanisms. The identification and mitigation of misinformation has emerged as a critical issue for prompt action and novel approaches [2]. The complexities and diverse nature of identifying and mitigating false news pose significant obstacles [3]. The rapid production and widespread circulation of false information on many internet platforms provide challenges in discerning between accurate reporting and misleading content [4]. The conventional approaches to human fact-checking and content verification have challenges in keeping up with the large quantity and ever-changing nature of false information, highlighting the need for more effective and adaptable strategies to promptly identify and address this issue.

The integration of Natural Language Processing (NLP) with deep learning models has emerged as a powerful combination for tackling the intricate challenges associated with the identification of false news [5]. The use of natural language processing (NLP) methods allows for the identification and extraction of cognitive details, linguistic structures, and contextual signals that are present within written text [6]. Deep learning models are well recognized for their ability to autonomously acquire complex representations from large datasets [7]. These models have the potential to identify subtle linguistic details and underlying patterns in many forms of textual information, such as news articles, social media postings, and online materials.

The potential for a transformative impact in false news identification is evident via the merging of natural language processing (NLP) and deep learning models. These models possess the capability to efficiently handle and evaluate large amounts of textual data, enabling them to potentially reveal concealed connections, linguistic anomalies, and misleading trends that are often associated with false information [8]. NLP-driven deep learning models attempt to discern and distinguish between reliable, verifiable material and deceptive, misleading information by examining textual characteristics, doing sentiment analysis, and using conceptual comprehension. The integration of advanced technology aims to improve the precision, scalability, and effectiveness of detecting false information, therefore protecting the credibility of information shared on digital platforms.

Using the synergies between NLP and deep learning models, this research explores the field of false news identification. The objective of using these advanced technologies is to identify misleading patterns and linguistic anomalies that are inherent in fabricated news. NLP-driven deep learning models aim to discern between trustworthy, genuine material and deceptive, fabricated information by using sophisticated text analysis techniques, semantic comprehension, and sentiment analysis. The primary objective is to improve the precision, capacity, and effectiveness of false news identification, hence promoting a more knowledgeable digital environment that is resistant to the impact of misleading information.

II. LITERATURE

Pritika Bahad et al [9] proposed a novel approach for detecting false news, using a Bi-directional LSTM-recurrent neural network as the underlying model. The performance of the model is evaluated using two publicly accessible datasets consisting of unstructured news items. The findings indicate that the Bi-directional LSTM model outperforms other approaches, such as CNN, vanilla RNN, and unidirectional LSTM, in terms of accuracy when it comes to detecting false news.

Ye-chan Ahn et al [10] proposed a study which delineates the challenge of retrieving relevant sentences from an input sentence within the Fact Data Corpus, presumed to be factual, and subsequently assessing the veracity of the extracted sentence concerning the input sentence. Across multiple Natural Language Processing (NLP) tasks, a specialized pre-training model tailored for Korean language is developed using the cutting-edge BERT (Bidirectional Encoder Representations from Transformers). Leveraging this model, fine-tuning processes are executed to align with the dataset aimed at identifying Korean fake news.

Ayat Abedalla et al [11] provided insights into the phenomenon of misinformation and the approaches used in detecting false information by using deep learning methodologies. The FNC-1 dataset has been used to develop many algorithms that aim to detect false news by analyzing the relationship between article titles and article bodies. The prevailing models mostly include Convolutional Neural Network (CNN), Long Short-Term Memory network (LSTM), and Bidirectional LSTM (Bi-LSTM).

Sherry Girgis et al [12] developed a classifier that can accurately assess the accuracy of news stories by analyzing their content only. The methodology used in this study is centered on the utilization of deep learning methods, with a particular focus on recurrent neural network (RNN) models such as vanilla RNN, gated recurrent unit (GRU), and long short-term memory (LSTM). These models are employed to effectively tackle the given challenge. The present work aims to elucidate the differences and evaluate the results achieved via the use of these models on the LAIR dataset, therefore demonstrating their efficacy in detecting misinformation.

Aswini Thota et al [13] presented a methodology that uses Deep Learning architectures to address the issue of identifying and detecting fabricated news articles. According to literature, the majority of individuals residing in developed countries would be exposed to a greater volume of misinformation compared to accurate information. The increasing prevalence of deceptive news calls for prompt action in the development of automated methods for classifying and detecting these altered news pieces. However, the automation of false news identification presents significant hurdles since models need to possess the ability to understand subtle details in natural language. In addition, a significant proportion of existing models designed for detecting fake news use a binary classification framework, which limits the model's capacity to assess the degree of connection between the disseminated news and factual news. In order to address these disparities, the present study presents a novel neural network framework designed to accurately forecast the attitude between a certain combination of title and article content.

Sneha Singhania et al [14] illustrated a deep learning-based automated detector which is employed using a three-level hierarchical attention network (3HAN) to swiftly and precisely identify fake news. The 3HAN model consists of three distinct levels dedicated to words, sentences, and the headline, thereby constructing a news vector that effectively represents an input news article. This hierarchical approach processes an article in a bottom-up hierarchical manner, systematically organizing information within the article. Notably, the headline serves as a discerning feature in fake news identification, and a limited set of words and sentences in an article carry more significance than others. The 3HAN model integrates three layers of attention, granting varied importance to different sections of an article based on their hierarchical structure.

Natali Ruchansky et al [15] introduced a novel model to enhance accuracy and automate prediction by amalgamating three critical aspects. This model takes into account the behavioral patterns of both users and articles, alongside the collective behavior of user groups disseminating fake news. Driven by these pivotal characteristics, the proposed model, named CSI (Capture, Score, and Integrate), comprises three distinct modules. The initial module, "Capture," employs a Recurrent Neural Network to capture temporal patterns in user engagement with a specific article, leveraging responses and text. The subsequent module focuses on learning source characteristics based on user behavior. Finally, the third module integrates these captured features to classify articles as authentic or fake. Rigorous experimental analysis conducted on real-world data showcases that the CSI model surpasses existing models in accuracy and effectively extracts meaningful latent representations of both users and articles.

III. PROPOSED MODEL

The identification of fake news utilizing a sequential neural network architecture that has LSTM layers, dense layers, and dropout, in conjunction with embeddings that are given by GloVe, is an effective method for determining whether or not textual material contains misleading information. The sequential model performs processing on the input text by using the temporal relationships that are recorded by LSTM layers. This allows the model to comprehend the sequential aspect of language. The addition of GloVe embeddings improves the model's capability of understanding the intricate semantic links that exist between different words. In order to extract higher-level characteristics even further, dense layers are used. This comprehensive architecture, which incorporates sequential processing, memory retention through LSTMs and semantic embeddings from GloVe, collectively contributes to a reliable and efficient system for detecting fake news. This system is capable of capturing intricate linguistic patterns that are indicative of misinformation.

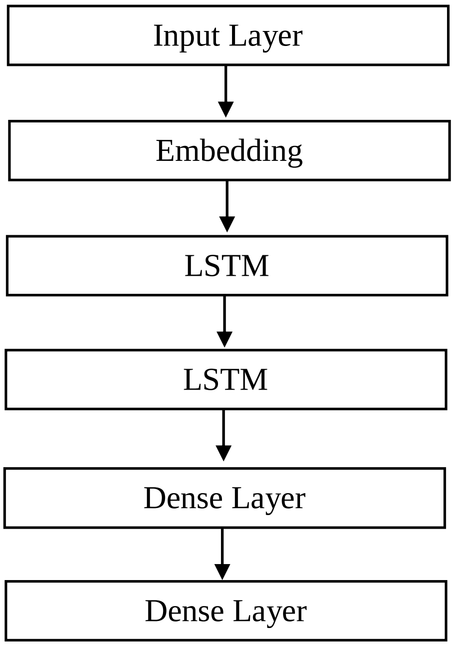

Figure 1: Proposed model Architecture

Neural network designs are essential for deriving meaningful patterns and characteristics from textual input in the context of natural language processing (NLP) and fake news identification. Each individual layer inside a neural network serves a distinct purpose, and the strategic combination of these layers may significantly augment the model's capacity to differentiate between authentic and fabricated news.

This discourse aims to examine the functions and significance of the Input Layer, Embedding Layer, LSTM (Long Short-Term Memory) Layer, and Dense Layer within the domain of false news identification.

a) Input Layer:

The input layer is the first layer of a neural network. It receives the initial input data and passes it on to the next layer. The Input Layer serves as the primary layer of a neural network, responsible for receiving and processing the unprocessed input data. In the context of identifying false information, the input data would consist of textual content derived from news articles or social media postings.

The Input Layer is responsible for representing individual words or tokens in the text as distinct input nodes. The entrance point for the sequential flow of information into the neural network is provided by this component.

b) Embedding Layer:

The embedding layer is a fundamental component in deep learning models, particularly in natural language processing tasks. It is responsible for transforming input data. The primary function of the Embedding Layer is to transform the discrete word representations obtained from the input layer into dense vector representations of a predetermined and consistent dimension. The aforementioned layer acquires a distributed representation for words, effectively capturing and encoding semantic links that exist between them. Embeddings play a crucial role in facilitating the model's comprehension of the contextual semantics of words, hence enhancing its overall comprehension of the text. The incorporation of this aspect is of utmost importance in the identification of false information, as it enables the model to comprehend intricate linguistic characteristics that might potentially signify deceitful or misleading material.

c) Long Short-Term Memory (LSTM):

The LSTM layer is a kind of recurrent neural network (RNN) layer that is designed to address the vanishing gradient problem in traditional RNNs. Long Short-Term Memory (LSTM) networks are a specific variant of recurrent neural network (RNN) layers, renowned for their exceptional proficiency in processing sequential input, notably textual information. The purpose of its design is to effectively capture extended dependencies and temporal patterns within the given input sequence.

In the realm of false news identification, Long Short-Term Memory (LSTM) models may be used to enhance the model's comprehension of the contextual nuances and interrelationships among words inside a given phrase or paragraph. The capacity to take into account the sequential structure of language is very advantageous in identifying nuanced linguistic signs or patterns that might potentially signify the presence of disinformation or propaganda.

d) Dense Layer:

A dense layer, also known as a fully connected layer, is a fundamental component of neural networks. The Dense Layer, sometimes referred to as a completely linked layer, is frequently positioned at the terminal position inside the neural network design. The final classification or prediction is generated by taking the output from the preceding layers.

Within the context of fake news detection, the Dense Layer assumes the crucial role of amalgamating the acquired information from the embeddings and LSTM layers. This amalgamation facilitates the process of determining the likelihood of the input text being classified as either false or true news. The quantity of neurons inside this layer, as well as the selection of the activation function, have an impact on the model's capacity to effectively capture intricate connections within the given data.

To summarize, the Input Layer is responsible for receiving the unprocessed textual data, while the Embedding Layer is responsible for transforming individual words into vector representations that have semantic meaning. The LSTM Layer is designed to collect and understand the sequential relationships within the text, and the Dense Layer is responsible for amalgamating these extracted features to ultimately classify the input in the context of false news identification. The proposed architectural design capitalizes on the unique capabilities of each layer to augment the model's comprehension of the linguistic characteristics linked to deceptive or misleading information.

3.1 Global Vectors for Word Representation (GloVe)

GloVe refers to an unsupervised learning approach that is used to produce word embeddings. The GloVe model, which was created by scholars at Stanford University, utilizes vector representations to depict words inside a continuous vector space. This approach effectively captures the semantic associations between words by analyzing their patterns of co-occurrence in large collections of texts.

The following section outlines the fundamental elements and concepts that underpin the GloVe model.

1. Word Embeddings: The objective of GloVe is to encode words as vectors inside a continuous space, whereby the spatial proximity and orientation of vectors correspond to the semantic associations among words. Vector representations of words should exhibit similarity when the words have comparable meanings or are used in similar contexts.
2. Co-occurrence Matrix: The GloVe model is based on the concept that the semantic meaning of words may be deduced by examining their patterns of co-occurrence with other words. A co-occurrence matrix is generated using a substantial corpus, whereby each element in the matrix denotes the frequency at which a certain pair of words co-occur within a given context frame.
3. Objective Function: The primary aim of GloVe is to acquire word vectors that effectively represent the relative proportions of co-occurrence probability. The objective function is formulated to minimize the squared difference between the dot product of word vectors and the logarithm of their co-occurrence probability.
4. Matrix Factorization: Matrix factorization is a mathematical technique used to decompose a matrix into a product of two or more matrices. The optimization problem in GloVe may be seen as a kind of matrix factorization. The approach performs factorization on the co-occurrence matrix, resulting in the creation of two distinct sets of word vectors. One set represents words as row vectors, while the second set represents words as column vectors.
5. Scalar Weights: In the GloVe model, scalar weights are assigned to each word pair in the co-occurrence matrix. The assigned weights serve to accentuate the significance of infrequent co-occurrences while diminishing the influence of frequent pairings that provide less relevant insights.
6. Context Window: The GloVe model takes into account a context window around each target word in order to ascertain co-occurrences. The context window delineates the scope of words that are considered when computing co-occurrence statistics.
7. Enhanced Training Efficiency: GloVe has gained recognition for its superior training efficiency, particularly in comparison to other word embedding techniques. This approach successfully obtains semantically rich word representations while minimizing computing expenses.
8. Pre-trained Embeddings: Pre-trained embeddings refer to pre-computed vector representations of words or phrases that have been trained on a large corpus of text data. These embeddings capture semantic and syntactic information about the words, allowing them to be GloVe embeddings are often pre-trained on extensive corpora and then provided for many languages. These pre-trained embeddings may serve as an initial reference for researchers and practitioners in the field of natural language processing, enabling them to undertake various tasks like sentiment analysis, machine translation, and other related endeavors.

The use of GloVe embeddings has been prevalent in the domain of natural language processing owing to its efficacy in capturing semantic associations between words and their computational efficiency throughout the training process. Pre-trained GloVe embeddings are often used by researchers and developers as features in subsequent natural language processing (NLP) tasks, or alternatively fine-tuned on particular datasets to enhance performance in task-specific contexts.

IV. EXPERIMENTAL RESULTS

This section provides a comprehensive analysis of the results obtained from the simulations conducted using the proposed methodology. The dataset used in this study was acquired from Kaggle. The dataset underwent processing in accordance with the designated technique. The collection comprises over 40,000 pieces including both fabricated and authentic news. The objective of our study is to train our model with the intention of accurately predicting the veracity of a given news article, distinguishing between genuine and fabricated information. The dataset has two distinct sets of data, one containing fabricated news pieces and the other containing genuine news stories. Each dataset contains around 20,000 articles. Figure 2 shows the Sample of true data and Figure 3 shows the Sample false data from Dataset.

	title	text	subject	date
0	As U.S. budget fight looms, Republicans flip t...	WASHINGTON (Reuters) - The head of a conservat...	politicsNews	December 31, 2017
1	U.S. military to accept transgender recruits o...	WASHINGTON (Reuters) - Transgender people will...	politicsNews	December 29, 2017
2	Senior U.S. Republican senator: 'Let Mr. Muel...	WASHINGTON (Reuters) - The special counsel inv...	politicsNews	December 31, 2017
3	FBI Russia probe helped by Australian diplomat...	WASHINGTON (Reuters) - Trump campaign adviser ...	politicsNews	December 30, 2017
4	Trump wants Postal Service to charge 'much mor...	SEATTLE/WASHINGTON (Reuters) - President Donal...	politicsNews	December 29, 2017

Figure 2: Sample true data from Dataset

	title	text	subject	date
0	Donald Trump Sends Out Embarrassing New Year'...	Donald Trump just couldn t wish all Americans ...	News	December 31, 2017
1	Drunk Bragging Trump Staffer Started Russian ...	House Intelligence Committee Chairman Devin Nu...	News	December 31, 2017
2	Sheriff David Clarke Becomes An Internet Joke...	On Friday, it was revealed that former Milwauk...	News	December 30, 2017
3	Trump Is So Obsessed He Even Has Obama's Name...	On Christmas day, Donald Trump announced that ...	News	December 29, 2017
4	Pope Francis Just Called Out Donald Trump Dur...	Pope Francis used his annual Christmas Day mes...	News	December 25, 2017

Figure 3: Sample false data from Dataset

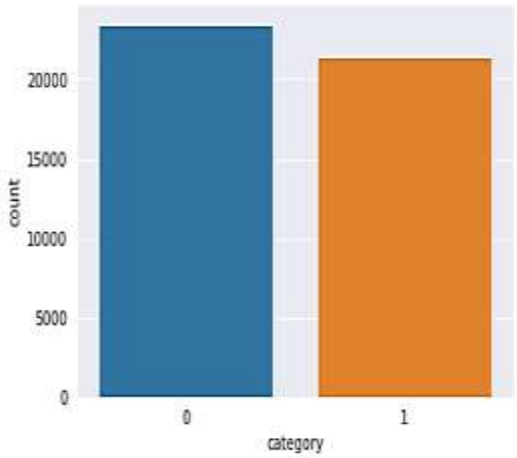

Figure 4: Count of the categories in dataset

Figure 4 shows Count of the categories in dataset"

True = Category 1

False = Category 0

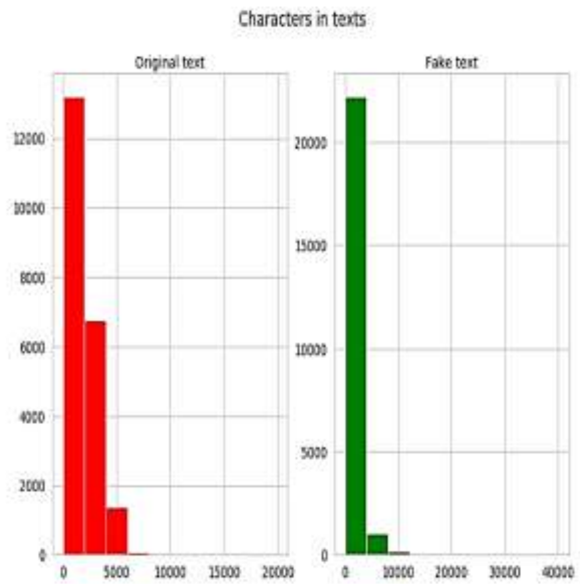

Figure 5: Characters in texts

There seems to be a discernible difference in the distribution of both variables. The frequency of texts of 2500 characters is highest in the original text category, while texts with around 5000 characters are most commonly found in the false text category. Figure 5 shows the characters in texts and figure 6 shows the words in texts.

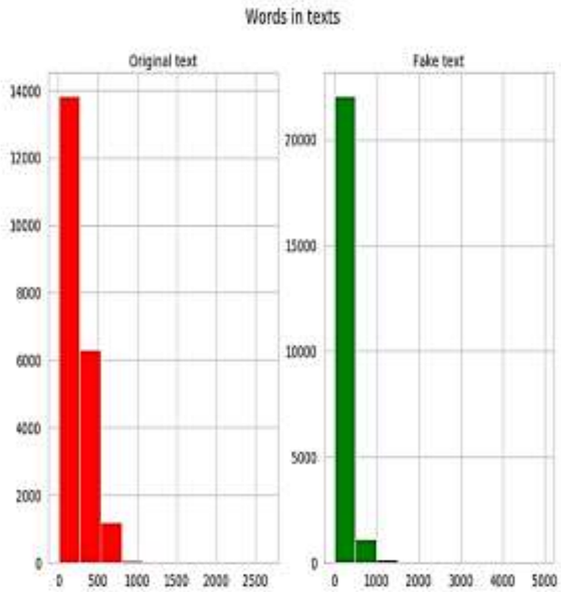

Figure 6: Words in texts

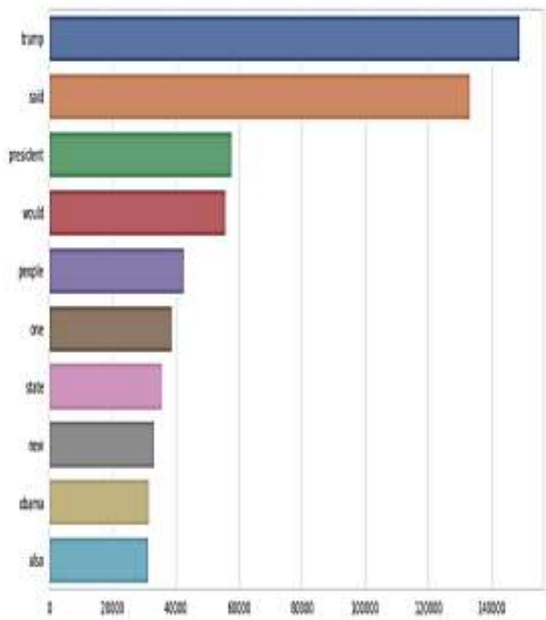

Figure 7: Unigram Analysis

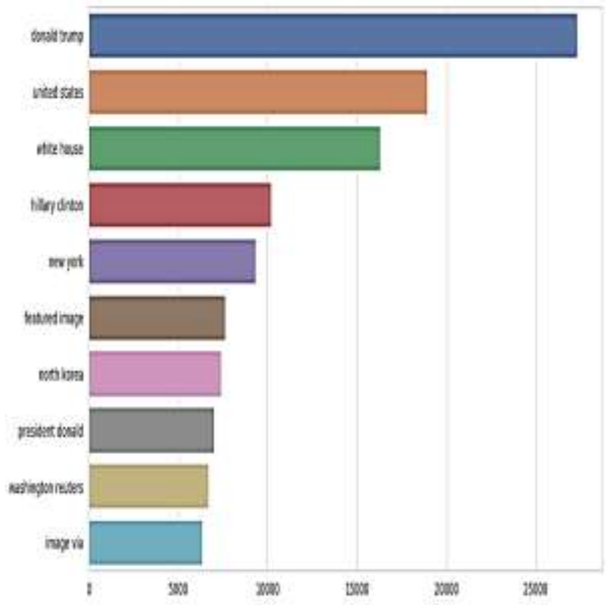

Figure 8: Bigram Analysis

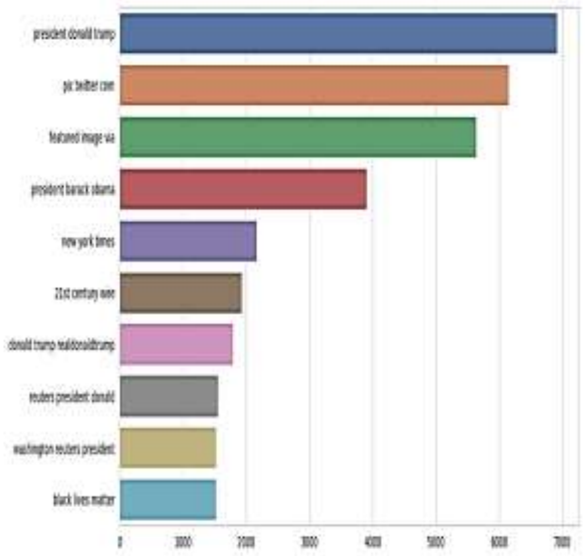

Figure 9: Trigram Analysis

Figure 7,8 and 9 shows the unigram, bigram and trigram analysis respectively. Unigram analysis, bigram analysis, and trigram analysis are often used approaches in the field of natural language processing to comprehend and identify significant patterns from textual material. The process of unigram analysis is the examination of individual words within a specified corpus, hence offering valuable insights about the frequency and distribution of these words. This method is essential for tasks such as text summarizing and sentiment analysis. The idea of bigram analysis expands upon this notion by taking into account pairs of successive words, so encompassing a greater amount of contextual information and aiding in the identification of phrases or collocations present within the text. Trigram analysis delves further into the study of language structures and connections by examining sequences of three words. The aforementioned analyses play a major role in tasks like as language modeling, whereby the accurate prediction of the subsequent word in a given sequence has significant importance. The combined use of unigram, bigram, and trigram studies allows for a comprehensive examination of linguistic characteristics, enabling diverse applications in the fields of natural language processing and text mining.

After the above analysis, using a tokenizer to represent words with numerical values while maintaining the original word-to-number mapping in the word index property, the text is tokenized. Notably, by specifically setting lowercase conversion to False, the text highlights the importance of maintaining word casing throughout tokenization. To further standardize input sizes for modeling, all news items are limited to 300 words. Longer articles are truncated, while those under 300 words are padded by appending zeros to the end of the sequence. Then glove.twitter.27B.100d.txt file taken as embedding file.

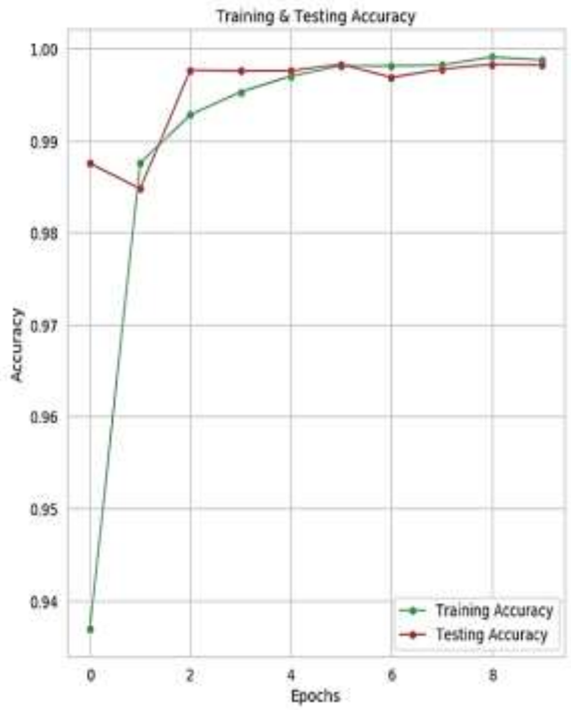

Figure 10: Training and Testing Accuracy

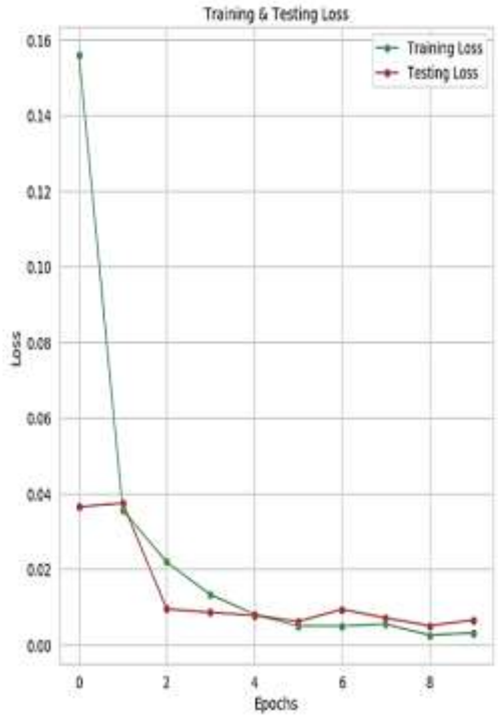

Figure 10: Training and Testing Loss

Figure 10 shows the training and testing accuracy and figure 11 shows the training and testing loss. Metrics like training accuracy and training loss are used to evaluate how well a machine learning model performs throughout the training process. The proportion of properly identified instances in the training dataset is known as training accuracy, and it shows how effectively the model has learnt from the available data. On the other hand, training loss measures how well the model aligns with the real labels compared to its predictions; lower values indicate greater alignment. Testing loss and accuracy assess how well the model generalizes to new data. In order to assess the model's accuracy in making predictions in real-world circumstances, testing accuracy evaluates how well the model performs on a different dataset that was not utilized during training. The prediction mistakes of the model on the testing data are measured by testing loss. A well-performing model should ideally show both high testing accuracy and low testing loss, which indicate strong generalization to new data, and high training accuracy and low training loss, which indicate effective learning.

Table 1: classification Report

	Precision	Recall	F1-score
Fake	1.00	1.00	1.00
Not Fake	1.00	1.00	1.00

A classification report, a popular machine learning assessment measure, is shown in Table 1 for a binary classification job that classifies news items as "Fake" or "Not Fake." For every class, the F1-score, accuracy, and recall metrics are given. The F1-score is the harmonic mean of precision and recall. Precision rates the accuracy of positive predictions, memory the capacity to properly recognize positive cases. The accuracy, recall, and F1-score of the two classes in this report—"Fake" and "Not Fake"—show perfect scores of 1.00, demonstrating perfect performance in the model's ability to categorize news stories as false or not fake. These high results point to a very good and trustworthy classification model for the job at hand.

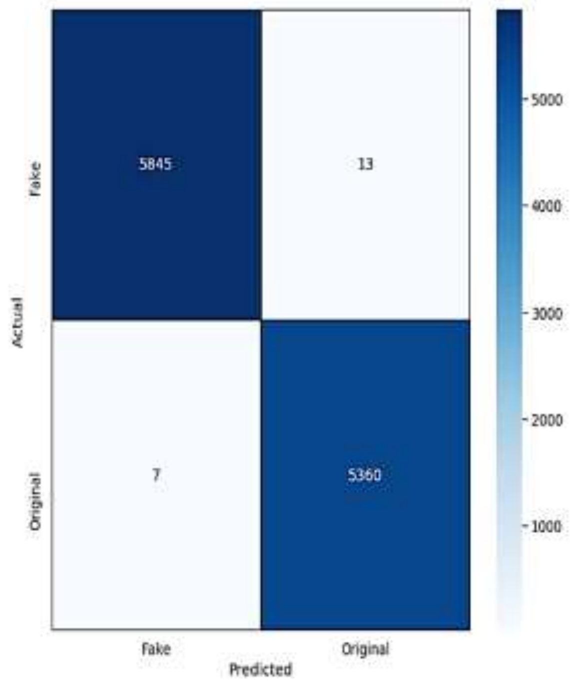

Figure 12: Confusion Matrix

The confusion matrix and the outcomes of a binary classification model for differentiating between "Original" and "Fake" cases are shown in Figure 12. The diagonal components show accurate predictions, showing that 5360 instances of original news and 5845 cases of fraudulent news are correctly categorized. Misclassifications are shown by the off-diagonal parts, which show 13 cases of original news that were mistakenly forecasted as fake and 7 cases of fraudulent news that were mistakenly predicted as original. Concisely displaying the distribution of true positive, true negative, false positive, and false negative predictions, the confusion matrix provides an efficient visual representation of a model's performance and offers important insights into the accuracy of the classifier as well as possible areas for improvement. The proposed model obtained "**99.82**" accuracy.

V. CONCLUSION

In conclusion, this study has presented a pioneering approach to combat the formidable challenge of fake news proliferation by integrating GloVe (Global Vectors for Word Representation) with deep learning techniques for detection.

The model showcased exceptional performance, achieving an impressive accuracy of 99.82% in discerning between legitimate and deceptive content. This high level of accuracy underscores the efficacy of the proposed approach in effectively identifying subtle linguistic cues indicative of misinformation. The utilization of GloVe embeddings, encoding semantic meaning and relationships between words, proved instrumental in enhancing the model's understanding of contextual nuances within textual data. The deep learning model, trained on a diverse dataset, demonstrated its adaptability and robustness in capturing intricate patterns associated with fake news.

VI. REFERENCES

[1] Saikh, Tanik, Amit Anand, Asif Ekbal, and Pushpak Bhattacharyya. "A novel approach towards fake news detection: deep learning augmented with textual entailment features." In *Natural Language Processing and Information Systems: 24th International Conference on Applications of Natural Language to Information Systems, NLDB 2019, Salford, UK, June 26–28, 2019, Proceedings 24*, pp. 345-358. Springer International Publishing, 2019.

[2] Verma, Abhishek, Vanshika Mittal, and Suma Dawn. "FIND: Fake information and news detections using deep learning." In *2019 twelfth international conference on contemporary computing (IC3)*, pp. 1-7. IEEE, 2019.

[3] Yang, Yang, Lei Zheng, Jiawei Zhang, Qingcai Cui, Zhoujun Li, and Philip S. Yu. "TI-CNN: Convolutional neural networks for fake news detection." *arXiv preprint arXiv:1806.00749* (2018).

[4] Roy, Arjun, Kingshuk Basak, Asif Ekbal, and Pushpak Bhattacharyya. "A deep ensemble framework for fake news detection and classification." *arXiv preprint arXiv:1811.04670* (2018).

[5] Fernández-Reyes, Francis C., and Suraj Shinde. "Evaluating deep neural networks for automatic fake news detection in political domain." In *Advances in Artificial Intelligence-IBERAMIA 2018: 16th Ibero-American Conference on AI, Trujillo, Peru, November 13-16, 2018, Proceedings 16*, pp. 206-216. Springer International Publishing, 2018.

[6] Jadhav, Shrutika S., and Sudeep D. Thepade. "Fake news identification and classification using DSSM and improved recurrent neural network classifier." *Applied Artificial Intelligence* 33, no. 12 (2019): 1058-1068.

[7] Monti, Federico, Fabrizio Frasca, Davide Eynard, Damon Mannion, and Michael M. Bronstein. "Fake news detection on social media using geometric deep learning." *arXiv preprint arXiv:1902.06673* (2019).

[8] Amine, Belhakimi Mohamed, Ahlem Drif, and Silvia Giordano. "Merging deep learning model for fake news detection." In *2019 International Conference on Advanced Electrical Engineering (ICAEE)*, pp. 1-4. IEEE, 2019.

[9] Bahad, Pritika, Preeti Saxena, and Raj Kamal. "Fake news detection using bi-directional LSTM-recurrent neural network." *Procedia Computer Science* 165 (2019): 74-82.

[10] Ahn, Ye-Chan, and Chang-Sung Jeong. "Natural language contents evaluation system for detecting fake news using deep learning." In *2019 16th international joint conference on computer science and software engineering (JCSSE)*, pp. 289-292. IEEE, 2019.

[11] Abedalla, Ayat, Aisha Al-Sadi, and Malak Abdullah. "A closer look at fake news detection: A deep learning perspective." In *Proceedings of the 3rd International Conference on Advances in Artificial Intelligence*, pp. 24-28. 2019.

[12] Girgis, Sherry, Eslam Amer, and Mahmoud Gadallah. "Deep learning algorithms for detecting fake news in online text." In *2018 13th international conference on computer engineering and systems (ICCES)*, pp. 93-97. IEEE, 2018.

[13] Thota, Aswini, Priyanka Tilak, Simrat Ahluwalia, and Nibrat Lohia. "Fake news detection: a deep learning approach." *SMU Data Science Review* 1, no. 3 (2018): 10.

[14] Singhania, Sneha, Nigel Fernandez, and Shrisha Rao. "3han: A deep neural network for fake news detection." In *Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14-18, 2017, Proceedings, Part II 24*, pp. 572-581. Springer International Publishing, 2017.

[15] Ruchansky, Natali, Sungyong Seo, and Yan Liu. "Csi: A hybrid deep model for fake news detection." In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pp. 797-806. 2017.