Social Network Analysis and Mining (2025) 15:21
https://doi.org/10.1007/s13278-025-01412-3

ORIGINAL ARTICLE

ScrutNet: a deep ensemble network for detecting fake news in online text

Aryan Verma[1,8] · P. Priyanka[2] · Tayyab Khan[3] · Karan Singh[4] · Lawal .O. Yesufu[5] · Mazeyanti Mohd Ariffin[6] · Ali Ahmadian[7,8]

Received: 2 September 2024 / Revised: 13 January 2025 / Accepted: 17 February 2025

Abstract

The expeditious propagation of fake news through online social media platforms has cropped up as a captious challenge, undermining the credibility of information sources and affecting public trust. Accurate detection of fake news is imperative to maintain the integrity of online content but is constrained by availability of data. This research aims to detect fake news from online articles by proposing a novel deep learning ensemble network capable of effectively discerning between genuine and fabricated news articles using limited data. We introduce ScrutNet, which leverages the synergistic capabilities of a bidirectional long short-term memory network and a convolutional neural network, which have been meticulously designed and fine-tuned for the task by us. This comprehensive ensemble classifier captures both sequential dependencies and local patterns within the textual data without requiring very large datasets like transformer based models. Through rigorous experimentation, we optimise the individual model parameters and ensemble strategy. The experimental results showcase the remarkable efficacy of ScrutNet in the detection of fake news, with an outstanding precision of 99.56%, 99.43% specificity, and an F1 score of 99.49% achieved on the partition test of the data set. Comparative analysis against state-of-the-art baselines demonstrates the superior performance of ScrutNet, establishing its prominence as a generalised and dependable fake news detection mechanism.

Keywords Fake news detection · Bi-directional LSTM · CNN · Social media

Abbrevations

ARE	Average relative error
CNN	Convolutional neural network
Bi-LSTM	Bidirectional long short-term memory
TF-IDF	Term frequency-inverse document frequency
BOW	Bag of words
GloVe	Global vectors for word representation
NLP	Natural language processing

✉ Aryan Verma
s2512060@ed.ac.uk

✉ Ali Ahmadian
ahmadian.hosseini@gmail.com

P. Priyanka
dr.priyanka@nith.ac.in

Tayyab Khan
tayyabkhan.cse2012@gmail.com

Karan Singh
karancs12@gmail.com

Lawal .O. Yesufu
Lawal.Yesufu@Faculty.Hult.Edu

Mazeyanti Mohd Ariffin
mazeyanti@utp.edu.my

[1] School of Mathematics, University of Edinburgh, Edinburgh, Scotland EH8 9YL, UK

[2] Department of Computer Science and Engineering, National Institute of Technology Hamirpur, Hamirpur, Himachal Pradesh 177005, India

[3] School of Computer & Systems Sciences, Jawaharlal Nehru University, New Delhi 110067, India

[4] Department of Computer Science and Engineering, IIIT Sonepat, Sonipat 131001, India

[5] Hult International Business School, Dubai, UAE

[6] Positive Computing Research Cluster, Universiti Teknologi Petronas, 32610 Seri Iskandar, Perak, Malaysia

[7] Faculty of Engineering and Natural Sciences, Istanbul Okan University, Istanbul, Turkey

[8] Jadara Research Center, Jadara University, Irbid 21110, Jordan

Published online: 20 March 2025

TPR True positive rate
API Application programming interface

1 Introduction

burgeoning popularity of social media and online platforms as a pervasive medium for news consumption has witnessed a notable surge in recent years. This upward trend can be attributed numerousous factors of significance. Firstly, the ubiquitous presence of smartphones and internet connectivity has empowered individuals to remain interconnected and informed while on the move (Manzoor et al. 2019). With a mere tap on their devices, users can access a rich repository of news articles, tweets, and live updates from diverse sources. The unparalleled convenience and accessibility proffered by social media platforms render them an unparalleled choice for individuals seeking expeditious and real-time news updates (Ahmad et al. 2020; Park and Chai 2023).

Secondly, the interactive nature of social media platforms has contributed substantially to their escalating prominence as a conduit for news dissemination (Mridha et al. 2021). Users can actively engage with news content through actions such as liking, commenting, and sharing, thereby fostering increased participation and discourse (Ushashree et al. 2023). This interactivity fosters a sense of community and involvement surrounding news stories, thereby cultivating a personalised news experience. Moreover, social media algorithms are adept at tailoring content based on individual preferences, thereby culminating in a curated news feed that aligns with users' proclivities and convictions (Mishra et al. 2022). This enhanced customisation further augments the allure of online content platforms such as social media as a fount of news, as individuals can effortlessly access content that resonates with their worldview (Jain et al. 2022; Jaiswal et al. 2023).

Collectively, the mounting popularity of social media and online platforms as a medium of news dissemination can be ascribed to their unparalleled convenience, accessibility, interactivity, and personalisation. As technology continues to progress and connectivity permeates further, it is plausible that social media will continue to exert a momentous influence on the manner in which news is consumed and propagated (Tsai 2023; Shaik 2023). However, it remains imperative for users to exercise critical discernment and verify the credibility of sources to ensure the ingestion of accurate and reliable information amidst the vast sea of content offered by these platforms (Nasir et al. 2021). The pervasive rise and widespread popularity of social media platforms have ushered in a new era of communication, but lurking within this digital realm lies an ominous threat-fake news and the spread of false information. As social media has revolutionised the way we consume and share information, it has simultaneously opened the floodgates for the rapid propagation of false or misleading content (Zhang and Ghorbani 2020). This issue gained popularity in 2016 during the US presidential elections and is dealt with seriously (Allcott and Gentzkow 2017). There are multiple false allegations and claims made by the online articles which might influence people in the decision-making of various domains (Rasool et al. 2019). This phenomenon is propelled by the inherent virality and ease of sharing on these platforms, where information can be disseminated globally within seconds, often without undergoing rigorous fact-checking processes (Mahir et al. 2019).

The main contributions of this research study are:

1. We proposed novel model ScrutNet, an ensemble of Bi-LSTM and CNN, optimized for capturing both sequential dependencies and local text patterns in fake news detection.
2. We performed ablation analysis with multiple features and ScrutNet achieved 99.56% accuracy, 99.43% specificity, and 99.49% F1-Score, outperforming existing baselines in comprehensive experiments.
3. We demonstrated scalability, reliability, and real-time applicability through rigorous evaluations and comparative analyses with existing state-of-the-art methods.

2 Problem statement

The ramifications of falsely fabricated news articles and wrong information are manifold and profound. The erosion of public trust in reliable sources of information is a troubling consequence, as individuals find themselves grappling with the daunting task of distinguishing between credible and unreliable sources. Beyond trust, the impacts of fake news extend to societal divisions and the amplification of existing biases. False narratives and manipulated information can exacerbate polarisation, fomenting animosity and deepening societal rifts (Jang and Kim 2018). This not only hampers constructive dialogue but also obstructs the path towards finding common ground and forging unity. In extreme cases, it can even fuel violence and social instability. Economically, fake news can tarnish reputations, manipulate financial markets, and undermine the democratic process by influencing public opinion and electoral outcomes (Pennycook and Rand 2021). In conclusion, the rise of social media platforms has brought forth an unsettling consequence-the rampant spread of fake news and misinformation. To safeguard the integrity of information and fortify the foundations of an informed society, it is imperative that accountability from online platforms is maintained.

Utilisation of machine learning (ML) and deep learning (DL) methodologies has emerged as a cutting-edge approach

in the detection and mitigation of fake news (Ibrishimova and Li 2020). Given the proliferation of falsified articles on online platforms, traditional manual fact-checking methods have proven to be inadequate. By harnessing the power of ML algorithms equipped with copious amounts of data, it is possible to automate the process of identifying and categorising fake news with improved efficiency and efficacy (Giachanou et al. 2019). The power of such methods are augmented with the methods which include sentiment analysis to improve the generalisation power of these algorithms (Ahmed et al. 2018; Ghanem et al. 2020). Some prevalent strategies involve the application of supervised ML algorithms on processed data with the help of natural language processing methods (Bhutani et al. 2019). These algorithms are trained on labelled datasets, differentiating between genuine and deceptive news articles. By extracting relevant features such as linguistic patterns, source credibility, and social media engagement metrics, models can be trained to accurately classify news articles as either authentic or fabricated. Through comprehensive data analysis, these models can discern common attributes associated with fake news, enabling automated detection (Ajao et al. 2019). Leveraging the prowess of extensive data processing, these models are capable of capturing intricate relationships among various features. For instance, recurrent neural networks (RNNs) can analyse sequential data, such as the online content and the posts and news articles that are streamed on social media, to unveil patterns indicative of misinformation. Additionally, CNN can scrutinise textual and visual content to uncover inconsistencies, biased language, or manipulated images frequently associated with fake news. Nevertheless, challenges persist in the realm of MLand DLfor fake news detection (Madani et al. 2022). The dynamic and evolving nature of misinformation necessitates a more accurate and generalised model to detect fake news.

3 Key contribution

1. Innovative approach: This research presents a pioneering approach, ScrutNet, which harnesses the capability of deep ensemble learning to effectively detect fake news. ScrutNet combines a Bi-LSTM network and a CNN, meticulously designed and parameter tweaked for the task. By leveraging the synergies between these models, ScrutNet constructs a comprehensive ensemble classifier that captures both the sequential dependencies and local patterns within textual data. This novel approach demonstrates the potential of deep ensemble learning in improving the accuracy of detecting the fake news.
2. Superior performance: Through extensive experimentation and optimization, ScrutNet achieves outstanding performance in detecting the fake news. The experimental results exhibit an exceptional accuracy of 99.56%, specificity of 99.43%, and a remarkable F1-Score of 99.49% on the test dataset. Comparative analyses against state-of-the-art baselines validate the superior performance of ScrutNet, underscoring its effectiveness in differing between real and fake synthesized news articles. This research contributes to advancing the field of detecting fake news by providing a highly accurate and reliable mechanism.
3. Holistic representation and informed decision-making: A key contribution of ScrutNet lies in its ability to comprehend textual content holistically, enabling informed decision-making. By considering both sequential dependencies and local patterns, ScrutNet enhances its understanding of the underlying information, leading to improved accuracy in identifying fake news. This comprehensive representation empowers users to make well-informed decisions based on trustworthy sources, thereby mitigating the negative impact of fake news dissemination.
4. Advancement in misinformation mitigation: ScrutNet's superior performance and accuracy represent a step forward in safeguarding the integrity of online information sources and fostering public trust. The determinations of this research add on to the development of robust mechanisms for detecting and mitigating fake news, ultimately supporting the reliability of online content and empowering individuals to navigate the digital landscape with confidence.

4 Literature survey

This section covers the recent advancement in the domain of detecting fake news using various computing methods. We did an exhaustive survey and found the techniques proposed by the research community for the task of fake news detection. There are various experiments with different datasets for detecting the deceptive patterns in the news articles and Twitter tweets. These experiments range from cleaning of data in different ways to utilising Natural Language Processing (NLP) techniques for the aim of augmenting the accuracy.

Umer et al. (2020) proposed a comparative study in which they trained the algorithm using features from raw data contrasted against the features extracted from the preprocessed data. The pre-processing applied by them increased the accuracy from 78% to wobbling 93.0%. Another work by Alsaeedi and Al-Sarem (2020) proposed some more preprocessing techniques such as lowercase removal, hashtag removal, Uniform Resource Locator (URL) removal, and mention character removal along with numeric chipping.

They also proposed supplementing the emoticons with descriptive words for them. After manual processing, researchers used NLP techniques such as stemming, lemmatisation, and stop word removal as an effort to gain accuracy. When we chop off the end of a word to achieve a base word, the process is called stemming. Normalisation of a word using morphological analysis and root form generation is known as lemmatisation (Padnekar et al. 2020). Rusli et al. (2020) in their study performed experiments for detecting fake news with and without the use of stemmers and stop word removal processes. Their method achieved a 0.8 F1 macro score without the use of stemmers and stop word removers and a score of 0.82, with their use. However, there are many other techniques that achieved good accuracy and F1 scores without the use of stemming and stop word removal processes (Granik and Mesyura 2017).

4.1 Feature extraction techniques

In the domain of fake news detection, the most go-to methods for feature extraction include term frequency-inverse document frequency (TF-IDF) and also the bag of words (BoW) that aim at improving accuracy. This approach is found to be suffering with information loss as the BoW model considers each presence of a news article as a potential document in the corpus. Also, in this technique, the contextual information of the words is lost (Thota et al. 2018). There has been an advancement from the BoW model by the arrival of 'word embedding' which is sort of a feature learning in which words are transformed to vectors of real-numbers. With the era of neural networks employing the task of fake news detection, there are several studies that initially found their power in fake news detection (Umer et al. 2020; Girgis et al. 2018). The Word2vec, which is a pre-trained model working on word embedding has an ability to get trained on large datasets (Kaliyar et al. 2020). There is also a model, Global Vectors for Word Representation (GloVe), which is currently in use by many research methods for obtaining better accuracy owing to the parallel implementation supported by GloVe. Further, there are studies suggesting repair of deep neural networks applicable to the current tasks.

4.2 Data splitting techniques

There were several attempts to discuss the ratio of the dataset to be used for training, validation, and testing. The simple and common ratios of splitting the data are 0.7:0.3 and 0.8:0.2. A study by Mandical et al. (2020) tested applying the data split ratio of 90:5:5 and 80:10:10, with the number of articles less than and greater than 10000, respectively. Another study by Jadhav and Thepade (2019) showed that 75:25 data split is optimal using the contrastive model performance. According to them, there is more variation in the model parameters exhibited by smaller sets of training data as compared to big training datasets. Also, performance metrics also demonstrate a larger variation when working with smaller testing datasets.

4.3 Feature extraction

There are attempts in the techniques of feature selection and extraction for effective fake news detection. Due to the nature of vast vocabulary coming from large corpuses of news articles, the dataset that is generated after feature extraction deals with the problem of dimensionality, for which feature selection algorithms are employed. The feature selection and extraction techniques are employed in text mining for various other tasks (Ozbay and Alatas 2020; Reddy et al. 2020). Usually in fake news detection, the language content and the visuals related to it are used as features [100]. The principal component analaysis (PCA) technique and Chi square techniques are basically employed for the feature reduction tasks in ML and DL methods. There are a number of studies which proved that using feature selection techniques renders more accurate models. A study by Umer et al. (2020) contrasted the PCA and Chi square feature selection applied on to different DL algorithms. They found that the feature-reduced data rendered better trained models by a wobbling margin of 20% in the F1 score and 4% in the accuracy, as compared to training the model on normal features. The neural networks have outperformed the traditional ML methods and also other techniques that combined these methods with NLP approaches. This is due to the powerful feature extraction ability of these models (Qawasmeh et al. 2019). The DL systems are able to explore the hidden features of the multimodal data facilities.

4.4 Use of CNN models

Kaliyar et al. (2020) proposed FNDNet, which is a customised deep CNN, to detect fake news using multiple hidden convolutional and pooling layers form the CNN. They claim the model is less prone for overfitting and achieved a test accurcay of 98.36%. Another work by Fernández-Reyes and Shinde (2018) developed a CNN, which they called StackedCNN. In this they inroduced two dimensional convolutional layers and used word embeddings, but still they got not so good performance in contrast to other state-of-the-art methods. Li et al. (2020) proposed a multilevel CNN and, along with it, used Sensitive word's weight calculating method (TFW). There method outperformed all the other methods. Alsaeedi and Al-Sarem (2020) added more layers to the CNN model which further lowered the performance by 0.014. They further recommended that the CNN model's best performance is noted

when dense units are 100 with window size 5. Further a dropout method of regularisation is used which lessens the problem of overfitting and is mostly used in the task of fake news detection (Ajao et al. 2018).

In a validation study, Girgis et al. (2018) performed experimentation with different models, including CNN, LSTM, Vanilla, and Gated Recurrent Unit (GRU). There are multiple conlusions that they drew from this study.According to them, Vanilla suffered with vanishing gradient, which is solved by GRU, which further takes more time to train. Singhania et al. (2017) used a bi-directional GRU for word annotations. In an effort to do something different, Long et al. (2017) proposed an LSTM with additional features as the speaker profile. Their findings suggest that the speaker profiles can improve the fake news detection results. A study by Bahad et al. (2019) developed a method that used a RNN model, but this is found to be tackling the vanishing gradient problem. Now, for coming up with the solution of vanishing gradient problem, they proposed the utilisation of LSTM-RNN. This model is found to have higher precision as contrasted with the state of the art CNN. A further study in this direction done by Asghar et al. (2021) proposed a method which used a Bi-LSTM along with the power of CNN for the task of rumour identification. The Bi-LSTM is said to preserve the information in both directions. This approach is found to be computationally expensive. Sahoo and Gupta (2021) proposed merging the user profile and the features from content of news in order to detect the fake news. They used facebook Application Programming Interface (API) to crawl the content and some more new features described by them which required a more time-consuming process.

4.5 Use of BERT models

Recent advancements in fake news detection have explored innovative combinations of BERT with other techniques to enhance performance. Verma et al. (2023) fine-tuned BERT using CNN-based N-gram features on the Kaggle Fake News Dataset, achieving a +1.10% improvement over base BERT. Koru and Uluyol (2024) combined BERT with Word2Vec embeddings to detect fake news in the Turkish TR_FaRe_News dataset, reporting accuracies of 90-94%. Essa et al. (2023) integrated the BERT tokenizer with LightGBM and Word2Vec, obtaining 91.31% accuracy on the FNC-I and ISOT datasets. Additionally, Zhang et al. (2024) employed a Graph Neural Network with BERT and a co-attention mechanism, achieving 90.30% accuracy on the Twitter 15, Twitter 16, and PHEME datasets. These studies underscore the effectiveness of combining BERT (Jazi et al. 2024; Yousefpanah et al. 2024) with complementary methods for improved fake news detection.

A tabular analysis of these studies done for the purpose of fake news detection is represented in Table 1. Along with the performance measures, the datasets used to run these methods on are also displayed. Going through all the literature cited above, we found our method to be better in terms of performance, which is explained in later sections of this work.

Table 1 A summary of existing ML techniques for fake news detection with their reported accuracy. All these studies are aimed at classifying news articles from a corpus into fake and reliable

S. No	Method	Dataset	Accuracy	References
1	Passive aggressive method applied with TF-IDF generated vectors	Kaggle Fake News dataset	83.8%	Mandical et al. (2020)
2	Features of CNN and LSTM on GloVE vector representations	Kaggle Fake News dataset	94.71%	Agarwal et al. (2020)
3	TF-IDF vector representation with passive aggressive	LIAR Dataset	99.0%	Mandical et al. (2020)
4	GLoVE vector representation with BERT model	Kaggle Fake News dataset	98.90%	Kaliyar et al. (2021)
5	Bidirectional LSTM RNN applied with GLoVE vectors	Kaggle Fake News dataset	98.75%	Bahad et al. (2019)
6	GLoVE vectors with deep CNN network	Kaggle Fake News dataset	98.36%	Kaliyar et al. (2020)
7	CNN network with tensorflow embedding layer	Kaggle Fake News dataset	96.0%	Ajao et al. (2018)
8	TF-IDF with DSSM LSTM	Kaggle Fake News dataset	99.0%	Jadhav and Thepade (2019)
9	PCA technique and ensemble of CNN+LSTM	FNC-1 Dataset	97.8%	Umer et al. (2020)
10	Bidirectional LSTM with Word2vec	FNC-I Dataset	94.0%	Padnekar et al. (2020)
11	BERT trained on CNN based N-gram features	Kaggle Fake News Dataset	+1.10% than base BERT	Verma et al. (2023)
12	BERT with Word2Vec	Turkish TR_FaRe_News dataset	90-94%	Koru and Uluyol (2024)
13	BERT tokeniser with LightGBM with Word2vec	FNC-I and ISOT Dataset	91.31%	Essa et al. (2023)
14	Graph Neural Network with BERT and co-attention mechanism	Twitter 15, 16, and PHEME Dataset	90.30%	Zhang et al. (2024)

5 Data and methods

This section describes the main techniques involved in this research and presents details about the dataset utilized for this research work. Figure 1 illustrates a condensed representation of the experimental process conducted. Subsequent sections elaborate on the different stages depicted in Fig. 1, which outline the progression of the fake news detection approach development.

5.1 Dataset used

We used a publicly available and benchmark dataset of fake news detection from Kaggle. The "Fake News" dataset (Veronica et al. 2018) from Kaggle is one of the most used datasets by the research community to come up with the research methodologies in the domain of fake news detection. This dataset has three main data files, train, test and submit. The train file is used by all the works as it is labelled and available in.csv (comma-separated values) format. There are a total 20800 entries with fields id, author, title, text, and label. A snapshot of the dataset is shown in Fig. 2. Taking a train-test split of the train.csv file, we use 16640 data rows to train our algorithm and 4160 data rows to evaluate it. We have concatenated title, author, and text to make one column on which further operations are applied. More information on the dataset and its acquiring strategies is available in the data availability statement of this research work.

5.2 Text cleaning

In this work we performed text cleaning, and all the operations are explained in this section. This is a crucial preprocessing endeavour within the realm of NLP, where the raw and unrefined textual data undergoes a metamorphosis into a refined and structured form. This intricate process encompasses an array of important tasks. Firstly, the replacement of contractions transpires, wherein contracted words, such as "don't," are expanded to their expanded counterparts, such as "do not," fostering consistency and optimising subsequent analysis (Verma et al. 2022). Secondly, the intricate chore of punctuation comes into play, encompassing the removal or substitution of punctuation marks like commas, periods, and question marks that often lack substantial contribution to the text's overall meaning. Thirdly, the intricate art of word splitting emerges, deftly separating compound words like "applepie" into its constituent parts, "apple" and "pie," facilitating more accurate analysis and interpretation. Next, the discerning endeavour of stopword removal transpires, extinguishing common and insignificant words like "the," "is," and "and" that possess minimal semantic weight and can potentially add noise to the analysis (Verma et al. 2023). Finally, the task of sign removal manifests, delicately discarding special characters, symbols, and emojis that wield little to no influence over the textual understanding and comprehension.

5.3 Lemmatization and tokenization

Within the context of our research paper, we incorporated two fundamental techniques, stemming and lemmatisation, as part of our text preprocessing pipeline (Singh et al. 2023).

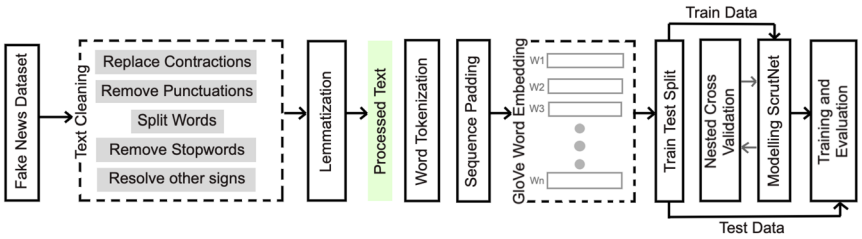

Fig. 1 Image showing an overview of the proposed approach for fake news detection

Fig. 2 Image showing a snapshot of the dataset used in this research work

	id	title	author	text	label
0	0	House Dem Aide: We Didn't Even See Comey's Let...	Darrell Lucus	House Dem Aide: We Didn't Even See Comey's Let...	1
1	1	FLYNN: Hillary Clinton, Big Woman on Campus - ...	Daniel J. Flynn	Ever get the feeling your life circles the rou...	0
2	2	Why the Truth Might Get You Fired	Consortiumnews.com	Why the Truth Might Get You Fired October 29, ...	1
3	3	15 Civilians Killed In Single US Airstrike Hav...	Jessica Purkiss	Videos 15 Civilians Killed In Single US Airstr...	1
4	4	Iranian woman jailed for fictional unpublished...	Howard Portnoy	Print \nAn Iranian woman has been sentenced to...	1

Stemming involves converting the words to come up with their root form by eliminating the prefixes and suffixes, thereby unifying different word variations into a single representation. This process aids in dimensionality reduction and streamlines subsequent analysis. For instance, words like "programming," "programmer," and "programs" would all be stemmed to "program." In contrast, lemmatisation takes a more sophisticated approach, considering the context and part of speech in the sentence. By mapping words to their dictionary form or lemma, lemmatisation enables more precise transformations. For example, verbs like "caring," "cares," and "cared" would be lemmatised to "car" while nouns like "dogs" would remain unchanged. By leveraging the power of lemmatisation, we successfully normalised our text data, reduced redundancy, and heightened the accuracy of our subsequent analyses.

5.4 GloVE embeddings

We performed a comparison between various NLP methods such as TF-IDF, Bag-of-words, Word2vec, and GloVe for determining the best method in terms of performance for our application. A summary of the analysis between these methods is shown in Table 2. Finally, after comparing the above methods, we select GloVe word embeddings. GloVe is a prominent method that generates the embeddings for the words, which are representations in the form of dense vectors capturing both semantic and syntactic relationships among words in NLP tasks (Verma et al. 2023). Unlike traditional approaches that rely solely on local context, GloVe leverages global statistics by a matrix constructed for co-occurrence of words from large text corpora.

The co-occurrence matrix represents the likelihood of two words appearing together in a given context. By factorising this matrix, GloVe aims to learn word representations that encode the meaning and relationships between words. The training process involves minimising an objective function that is dependent on weighted least squares of the difference between the word vectors' dot product and the logarithm of probability of co-occurence. Here in the formula 1, Z is the co-occurrence matrix, and it $Z_{i,j}$ reports the count of p-th and q-th word co-occurrence. Now the cost function would be:

$$J = \sum_{p,q=1}^{N} f(Z_{p,q})(w_p^T \tilde{w}_q + b_p + \tilde{b}_q - logZ_{p,q})^2 \quad (1)$$

where N is vocabulary size, w is for word vectors, b for biases, $W + \tilde{w}$ is the output of this algorithm. The model architecture of GloVe is shown in Fig. 3. A one-hot word acts as an input to this algorithm, and inner products of word vectors in the shape of a vector are the output. The updation of embedding matrices is carried out with gradient of loss function in formula 1.

GloVe embeddings possess several desirable properties. They capture both local and global word relationships by considering co-occurrences across different distances, effectively incorporating semantic information. Moreover, they offer computational efficiency during training, making them scalable for large-scale corporations. Additionally, GloVe embeddings exhibit linear substructures, enabling algebraic operations that approximate semantic relations, such as vector analogies. Though Table 2 summarises the advantages

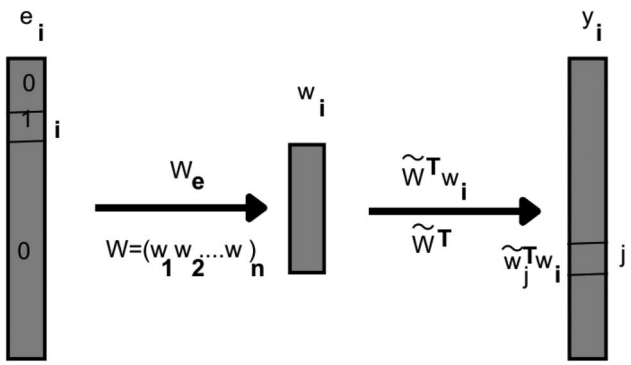

Fig. 3 Figure showing model architecture of GloVe embedding

Table 2 Summary of comparison drawn between various NLP methods for choosing best method to convert the tokens to vectors

S. No	NLP Method	Advantages	Disadvantages
1	TF-IDF	This method takes both less important and more important words into consideration. For large corpus, this method is slow. Also, position, co-occurrences and semantics are not considered	This method ignores the context and semantics of words, failing to capture the relationships and meanings between terms in a document
2	Bag-of-Words	This method is very easy to implement and understand	Again, the semantic relations and position are ignored
3	Word2Vec	It maintain the semantic meaning of the words contained within text. The information of context is preserved and vector size is small	Unavailability of common representations at sub-word level and it can't deal with unfamiliar words
4	GloVe	Better than above three, as it doesn't rely on local statistics	Word vectors are generated with help of global word co-occurrence

of GloVe, the following are given a comparison of recent approaches with GloVe.

GloVe has advantages over TF-IDF in NLP:

1. Semantic representation: GloVe captures semantic relationships, while TF-IDF lacks semantic meaning.
2. Contextual information: GloVe considers global context and provides a broader understanding across documents, unlike TF-IDF, which focuses on individual documents.
3. Dimensionality reduction: GloVe has fixed dimensionality, reducing complexity compared to TF-IDF's high-dimensional and sparse vectors.
4. Transfer learning: Pre-trained GloVe embeddings generalise well, benefiting from semantic knowledge, whereas TF-IDF is dataset-specific.
5. Algebraic operations: GloVe allows algebraic operations to approximate semantic relationships, which is not possible with TF-IDF.

GloVe has advantages over Word2Vec in NLP:

1. Global information: GloVe captures global word co-occurrence patterns, while Word2Vec focuses on local contexts.
2. Intuitive vector arithmetic: GloVe allows for meaningful algebraic operations, while Word2Vec may not consistently preserve vector relationships.
3. Improved rare word performance: GloVe performs better on infrequent words due to larger training corpora and global context.
4. Scalability and efficiency: GloVe is computationally efficient and scalable, while Word2Vec requires more resources and training time.
5. Pretrained Embeddings: GloVe offers readily available pretrained embeddings with broader coverage compared to Word2Vec.

5.5 RNN

We introduced ScrutNet with the help of some base networks, which come from the family of RNN and CNN. RNN are a class of neural network architectures specifically designed for modelling the data, which is sequential in nature. Unlike feedforward neural networks, these networks incorporate feedback connections that allow information to be propagated not only forward but also back to previous steps in the sequence (Guleria et al. 2023). The RNN digests an input at every step, combines it with the previous step function output of hidden state, and an output is produced as a resultant. This recurrent nature enables the network to maintain an internal memory, allowing it to capture and utilise information from earlier steps. This memory mechanism allows RNNs to effectively process and learn from temporal dependencies in sequential data (Vardhan et al. 2023). However, standard RNNs are prone to vanishing or exploding gradients during training, which can impede their ability to learn long-term dependencies. To overcome the problem, advanced RNN architectures, such as LSTM and GRU, are currently in use.

5.5.1 LSTM

LSTM networks, which are a transformed version of RNN, are designed to overcome the limitations of traditional RNNs in capturing the dependencies that are found in sequential data for long terms. LSTM networks employ memory cells with specialised gating mechanisms, including input (i), forget (f), cell candidate (g),and output gates (o), to regulate the information and data flow throughout the neural units, as demonstrated in Fig. 4. The model input weights (Z_i, Z_f, Z_g, Z_o), recurrent weights (U_i, U_f, U_h, U_o), and biases (y_i, y_f, y_g, y_o) are determined by the parameter matrices Z, U, and y.

The input gate of an LSTM network determines the relevance of incoming information. It combines the input to the current unit and the output of the hidden state from the previous step, passes it through an activated sigmoid unit, and produces an activation vector that controls the input flow into the memory cell as shown in formula 2. The decision of which information is to be passed and what to be retained is done by the forget gate. The input is the current state and hidden state from the previous step and throws it to the sigmoid function, and outputs a forget vector. The forget vector takes each element and performs multiplication with the cell state of previous memory, to selectively forget irrelevant information as shown in formula 3. The output gate manages the memory cell output. This gate applies sigmoid to the input given to it and hidden state from the previous step and generates an output vector that influences the output of the LSTM as shown in formula 5.

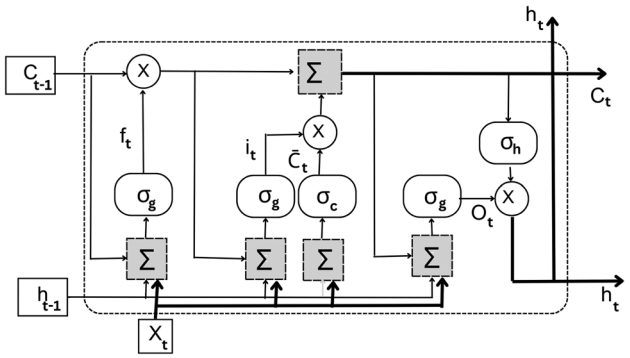

Fig. 4 Figure representing the unfolded model architecture of the LSTM cell

The state of the memory cell updates on the reception of outputs from the input gate, forget gate, and the input, which is passed at present time. The input gate is applied to the current input, which is then combined with the forget gate applied to the previous memory cell state. This element-wise addition and multiplication process ensures the appropriate storage and removal of information in the memory cell state, as shown in formula 6. The hidden state of the LSTM network is computed by applying the output gate to the updated memory cell state. This produces a filtered hidden state that captures the relevant information from the sequence as shown in formula 7.

By incorporating memory cells and gating mechanisms, LSTM networks can effectively capture and utilise long-term dependencies in sequential data. This makes them highly suitable for applications such as NLP, recognition of speech, and analysis of time series data, where understanding and modelling sequential relationships are critical.

$$\vec{i}_t = \sigma_g(\vec{Z}_i x_t + \overrightarrow{U_i h_{t-1}} + \vec{y}_i) \tag{2}$$

$$\vec{f}_t = \sigma_g(\vec{Z}_f x_t + \overrightarrow{U_f h_{t-1}} + \vec{y}_f) \tag{3}$$

$$\vec{g}_t = \sigma_g(\vec{Z}_g x_t + \overrightarrow{U_g h_{t-1}} + \vec{y}_g) \tag{4}$$

$$\vec{O}_t = \sigma_g(\vec{Z}_o x_t + \overrightarrow{U_o h_{t-1}} + \vec{y}_o) \tag{5}$$

$$\vec{C}_t = \vec{f}_t \odot \overrightarrow{C_{t-1}} + \vec{i}_t \odot \vec{g}_t \tag{6}$$

$$\vec{h}_t = \vec{O}_t \odot \sigma_h(\vec{C}_t) \tag{7}$$

5.5.2 Bi-LSTM

Bi-LSTM networks are an advanced variation of the LSTM architecture that enables the incorporation of information from both previous and upcoming contexts in sequence modelling. Bi-LSTMs differ from traditional LSTMs in their bidirectional processing nature, which allows them to simultaneously capture dependencies in both directions, the front and reverse (Verma et al. 2024). In Bi-LSTM networks, the input sequence is processed by two separate LSTM layers: the layers responsible for forward propagation and the layers responsible for backward propagation, as shown in Fig. 5. The sequence from beginning till the end is traversed by the forward layer, and the backward layer does this in the opposite direction. At each time step, the hidden states and memory cells of both layers are concatenated, producing a fused representation that encodes information from both directions as shown in formula 13.

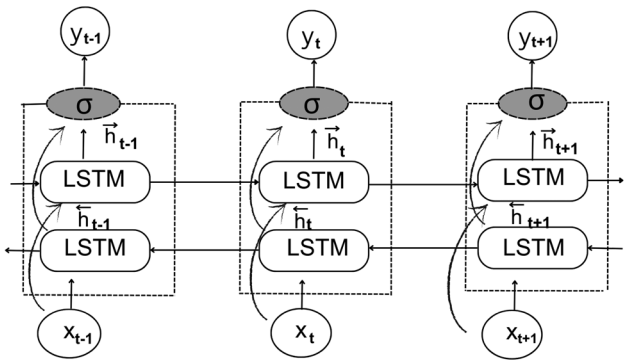

Fig. 5 The figure represents an unfolded model of a bi-LSTM cell with one step prior and one step ahead processing

By leveraging bidirectional processing, Bi-LSTMs possess an enhanced capacity to capture contextual dependencies that extend in both temporal directions. This capability is particularly advantageous in tasks where understanding the entire sequence, including future contexts, is essential, such as sentiment analysis, fake news detection, and named entity recognition.

The incorporation of bidirectional information introduces additional computational complexity and parameter requirements compared to traditional LSTMs [62]. Each layer processes the sequence independently, resulting in a more intricate architecture. However, the benefits of modelling bidirectional dependencies often outweigh the associated computational costs.

In summary, Bi-LSTM networks extend the LSTM architecture by introducing bidirectional processing. In this network, the information is processed in the same manner as in LSTM with the help of formulas 2, 3, 4, 5, 6, 7 in the forward direction and with the formulas 8, 9, 10, 11, 12, 13 in the backward direction. This empowers the networks to effectively capture bidirectional dependencies, making them well-suited for tasks that demand a comprehensive understanding of sequential data in both temporal directions. This is the primary reason for making this model to be a part of our ensemble technique.

$$\overleftarrow{i}_t = \sigma(g(\overleftarrow{Z}_i x_t + \overleftarrow{U_i h_{t-1}} + \overleftarrow{y}_i)) \tag{8}$$

$$\overleftarrow{f}_t = \sigma(g(\overleftarrow{Z}_f x_t + \overleftarrow{U_f h_{t-1}} + \overleftarrow{y}_f)) \tag{9}$$

$$\overleftarrow{C}_t = \overleftarrow{f}_t \odot \overleftarrow{C_{t-1}} + \overleftarrow{i}_t \odot \sigma_c(\overleftarrow{Z}_c x_t + \overleftarrow{U_c h_{t-1}} + \overleftarrow{y}_c) \tag{10}$$

$$\overleftarrow{O}_t = \sigma(g(\overleftarrow{Z}_o x_t + \overleftarrow{U_o h_{t-1}} + \overleftarrow{y}_o)) \tag{11}$$

$$\overleftarrow{h}_t = \overleftarrow{O}_t \odot \sigma_h(\overleftarrow{C}_t) \tag{12}$$

$$y_t = \sigma_g(W_y(\overrightarrow{h_t}, \overleftarrow{h_t}) + b_y) \quad (13)$$

5.6 CNN

We have employed CNN as a part of our ensemble model and for convolving the features found with the help of bi-LSTM model in the ensemble. CNNs have transcended the boundaries of intelligent vision tasks and have emerged as a formidable tool for NLP tasks, particularly in the realm of text classification. Leveraging their prowess in capturing local patterns and semantic features, CNNs have become a potent force in NLP classification applications such as sentiment analysis, text categorisation, and document classification.

When employed in NLP classification, CNNs operate on text inputs that are typically encoded as sequences of words or characters. Prior to feeding the text into the CNN architecture, it is transformed into word embeddings or character embeddings, serving as the foundation for subsequent computations. Convolutional layers are then applied to discern and extract pertinent local patterns within the input embeddings. Through the utilisation of filters, which traverse the sequence, salient features and n-grams of significance are identified, allowing for effective feature detection.

The feature maps coming as an output from convolutional layers are downsampled by using pooling layers which reduce dimensionality. The pooling operations serve the purpose of capturing the most prominent features or aggregating information across the sequence, respectively. The major ones being maxpooling operations and average pooling. Following the pooling layers, the resultant features undergo flattening and are fed into fully connected layers or dense layers. These layers are responsible for learning high-level representations and modelling intricate interactions between the extracted features, ultimately enabling accurate classification of the input text. In our case, CNNs offer numerous advantages by possessing the ability to capture both local and compositional features inherent in the text, thereby effectively assimilating word order and contextual information. Additionally, CNNs excel in their capacity to accommodate inputs of variable length, while simultaneously having the capability to autonomously learn relevant features, thereby obviating the need for explicit feature engineering.

5.7 ScrutNet: ensemble model

Our novel ensemble model for detecting fake news combines the power of Bi-LSTM and CNN layers to effectively capture both sequential and spatial features from textual data. The architecture of the model is shown in Fig. 6. The input to the model is transformed into vectors with the help of the GloVe. Now, this vectorised form is input to a 1-dimensional convolutional and a bi-LSTM layer at the same time. Further, we see that convolutional layers are accompanied by maxpooling layers, with some dropout layers in between to maintain the regularisation.

The CNN layer is crucial in extracting local and compositional features from the input text. This helps for capturing the meaningful information such as word associations, n-gram features, and syntactic patterns, which can be indicative of the presence of fake news. The Bi-LSTM layer, with its bidirectional nature, allows capturing the information related to context and also the dependencies that are seen for the long term by considering both the previous time and upcoming words in the input sequence. This enables the model to grasp the intricate patterns and relationships within the text, which are crucial for identifying misleading or fabricated information.

The outputs from the Bi-LSTM and CNN layers are then combined in a feature slicing layer, which aggregates the information learnt from both architectures. Here the features from both sides of the ensemble are made compatible to be fed in the proceeding convolutional layer. This fusion process facilitates the integration of the complementary strengths of Bi-LSTM and CNN, enhancing the overall performance of the ensemble model. It effectively captures both sequential dependencies and spatial features, allowing for a more accurate and robust identification of deceptive content.

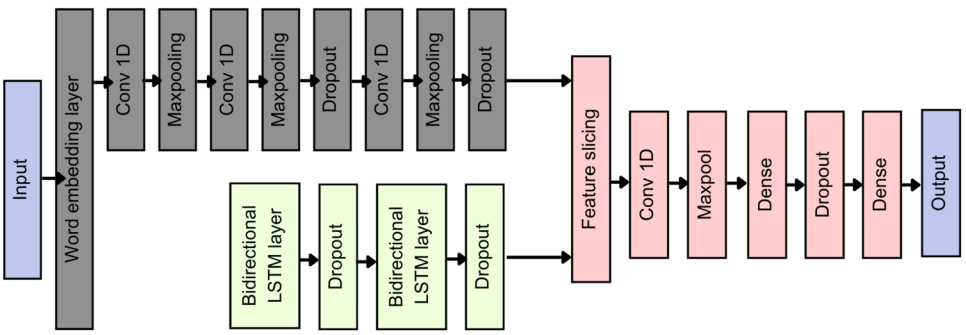

Fig. 6 Image showing the architecture of the ensemble model developed in this work

To ensure optimal performance, we conducted a systematic hyperparameter tuning process for ScrutNet. Key hyperparameters, including the number of Bi-LSTM units, CNN filter sizes, kernel sizes, and dropout rates, were optimized using a grid search approach. The tuning was guided by cross-validation on the training set to balance accuracy and generalization. For instance, the best performance was achieved with 128 Bi-LSTM units, 64 CNN filters, a kernel size of 3, and a dropout rate of 0.3, minimizing overfitting while maintaining high accuracy. The influence of hyperparameter variations on performance was carefully analyzed, demonstrating that the selected configuration enables robust performance across different dataset partitions, thereby ensuring reliability in broader applications.

6 Results and discussion

This section consists of the experimental results obtained while doing this research work. The training and validation accuracy and loss plots are given in Figs. 7 and 8. The training accuracy reaches 99.81% with a loss of 9.66% in 34^{th} epoch, while the validation accuracy is 99.68% with a loss of 11.20%. This trained model was evaluated on test data with the help of evaluation metrics, employed to gauge the capabilities of classification algorithms. We conducted experiments with the BERT model, incorporating Word-2Vec embeddings to enhance its performance. Our findings revealed that the BERT model exhibited significant overfitting on the dataset with 99.81% train accuracy and 96.23% test accuracy. Furthermore, employing LIME for interpretability analysis indicated that BERT struggles to capture the nuanced patterns in small datasets, posing a potential limitation for language-specific fake news detection tasks.

Relying solely on classification accuracy falls short in making this determination. Hence, additional assessment metrics become indispensable for a comprehensive evaluation. The confusion matrix, which shows the correct and incorrect classifications of the test samples by the algorithm, is shown in Fig. 9. With the help of this confusion matrix, we see that out of total 4160 samples, ScrutNet is able to detect 4142 samples correctly, which comprise 2121 fake class and 2021 real class news articles. Also, the ensemble accounts for 18 misclassifications. The following are the metrics obtained for our final trained ScrutNet model.

6.1 Accuracy

Accuracy measures the overall correctly classified cases from the total number of cases available for testing. In other words, it indicates the proportion of correctly classified news articles, including fake and real, over the total number of news articles. This is calculated according to formula 14 and comes out to be 99.56%.

$$Accuracy = \frac{(TP + TN)}{TP + FP + TN + FN} \tag{14}$$

6.2 Precision

Precision, also called positive predicted value, tells us about the capability of our trained network to correctly identify the real news articles out of all the test articles given to it. It is calculated according to the formula 15 and comes out to be 99.46%.

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

6.3 Recall

Recall is also known as the sensitivity or true positive rate (TPR) of the trained algorithm. This metric measures that out of all the genuine news articles, how many classifiers are able to detect. It focuses on the coverage of real news articles and is calculated according to the formula 16, which comes out to be 99.65%.

$$Recall = \frac{TP}{TP + FN} \tag{16}$$

6.4 F1-score

The F1 score is determined by taking the harmonic mean of the precision and recall metrics. This is a balanced metric for the classifier and has its value from 0 to 1. This is calculated according to formula 17 and comes out to be 99.49%.

$$\text{F1 Score} = \frac{2 * Precision * Recall}{Precision + Recall} \tag{17}$$

6.5 Specificity

Specificity is also known as true negative rate (TNR). This metric measures the total number of correctly predicted fake articles by the classifier out of the actual number of fake articles in the testing corpus. It is the complement of the false positive rate, and it is calculated according to formula 18, which comes out to be 99.44%.

$$Specificity = \frac{TN}{TN + FP} \tag{18}$$

Table 3 This tables shows the performance of various popular baseline architectures on same data along with our proposed method ScrutNet

	LSTM		Bi-LSTM		CNN		ScrutNet	
Fold No	Train Acc	Test Acc	Train Acc	Test Acc	Train Acc	Test Acc	Train Acc	Test Acc
1	96.71%	94.92%	98.73%	97.93%	96.87%	96.01%	99.57%	99.24%
2	97.40%	95.73%	98.17%	97.55%	98.73%	97.29%	99.72%	99.55%
3	97.19%	95.08%	98.80%	97.16%	97.60%	97.17%	99.53%	99.40%
4	96.82%	95.27%	98.56%	97.23%	97.13%	96.77%	99.71%	99.59%
5	97.45%	95.31%	97.90%	97.48%	97.47%	97.08%	99.79%	99.64%

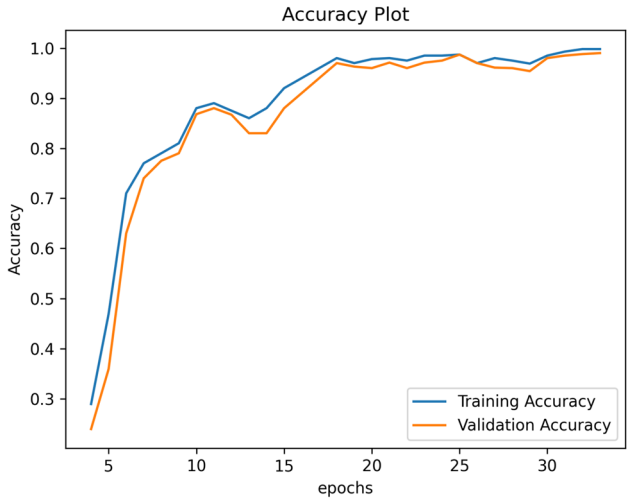

Fig. 7 The image represents the accuracy plotted over multiple epochs while training and testing the ScrutNet

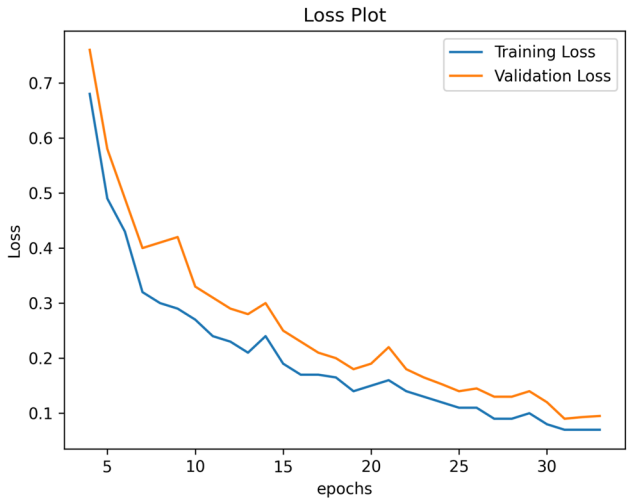

Fig. 8 The image represents the loss plotted over multiple epochs while training and testing the ScrutNet

We performed training of several other architectures, including LSTM, Bi-LSTM, CNN, and ScrutNet. This is done to contrast the capabilities of our novel model as

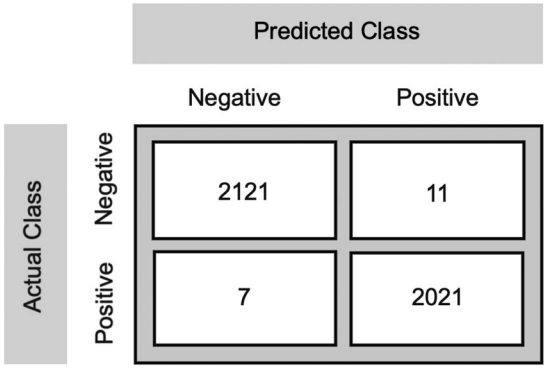

Fig. 9 The image represents the confusion matrix of the evaluation of ScrutNet

compared to baseline architectures. The training process is carried on with a nested cross-validation approach, as shown in Fig. 10. The nested CV allows for more generalised models to be trained and more true insights as the test partition is not seen by the model in any fold of data. This method is better than holdout validation and K-fold CV. The results are summarised in Table 3. As noticed from the given table, the proposed methodology clearly outperforms the baseline architectures with a margin of 2–3% accuracy. Not only this happens with the training data, but also on the test data, the model outperforms the baseline architectures, which shows the generalisation power of our approach.

To evaluate the effectiveness of the proposed ScrutNet model trained on GloVe embeddings, we conducted an ablation study by replacing GloVe with alternative feature representation techniques, including Word2Vec, TF-IDF, Bag-of-Words, FastText, and ELMo. The results, summarized in Table 4, demonstrate that ScrutNet with GloVe achieved the highest performance, with a training accuracy of 99.81% and test accuracy of 99.68%, alongside the lowest training and test loss of 9.66% and 11.20%, respectively. Among the alternatives, FastText and Word2Vec delivered competitive results, albeit slightly lower than GloVe, while traditional methods such as TF-IDF and Bag-of-Words exhibited significantly reduced accuracy and higher loss values. These findings underscore the importance of leveraging context-rich embeddings like GloVe to maximize the model's ability

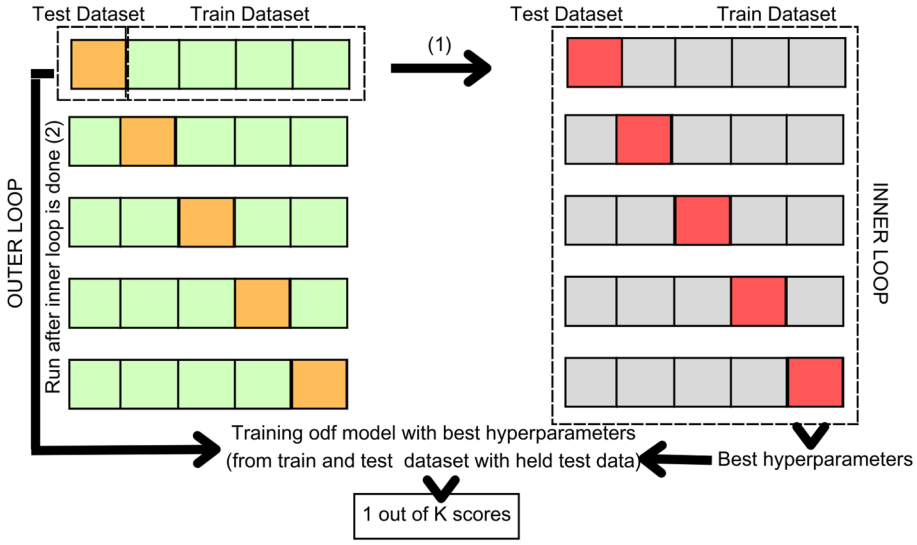

Fig. 10 The image represents the nested cross-validation approach used for training the models

Table 4 Ablation study results for scrutnet with different feature representations on fake news detection data

Method	Training accuracy (%)	Test accuracy (%)	Training loss (%)	Test loss (%)
Proposed (GloVe)	99.81	99.68	9.66	11.20
Word2Vec	98.95	98.42	12.30	13.75
TF-IDF	97.80	96.85	14.85	16.20
Bag-of-Words	96.95	95.60	16.50	18.35
FastText	99.10	98.65	11.40	12.85
ELMo	98.25	97.85	13.20	14.55

to capture intricate patterns in language, which is crucial for fake news detection.

7 Conclusion

In this study, we developed a pioneering deep ensemble network explicitly designed to detect fake news articles. By integrating the unique strengths of Bi-LSTM and CNN architectures, our approach achieved an exceptional test accuracy of 99.56%, setting a new benchmark across various domains.

The synergy of Bi-LSTM and CNN within our framework endowed it with the capability to effectively capture both sequential features and the intricate structural characteristics present in news articles. The Bi-LSTM component excelled at understanding long-range dependencies and context awareness, while the CNN component adeptly identified localised patterns and hierarchical representations. Rigorous experimentation and evaluation conducted on a substantial dataset confirmed the unparalleled performance of our proposed deep ensemble network, ScrutNet.

Compared to the current leading approaches for fake news detection leveraging ML and deep learning, our methodology outperformed others in terms of accuracy. A summarised comparative analysis of previous prominent works with our work is presented in Table 5. Moreover, our approach addressed the challenge of generalisation by employing an ensemble technique that harnessed the collective insights of multiple models trained on distinct subsets of the data. This strategic integration not only elevated the system's overall performance but also improved precision to 99.46%. Our model offers a promising avenue for deployment in real-world applications, such as social media platforms and news aggregators, where addressing the spread of fake news is critically important.

As we advance, further research should focus on integrating domain-specific knowledge and external linguistic resources to enhance the system's robustness and accuracy. Investigating the use of hybrid architectures that blend deep learning with traditional machine learning techniques could offer new perspectives on improving generalizability. Moreover, expanding the dataset to include multilingual and cross-cultural news articles can enable the model to operate effectively across diverse contexts. Finally, the model's real-time implementation and scalability can pave the way for combating fake news in dynamic and large-scale environments. Our research signifies a significant breakthrough

Table 5 Comparative analysis between existing popular approaches and the proposed method

S. No	Dataset	Method	NLP Technique	Accuracy	Citation
1	Fake News	Passive Aggressive	TF-IDF	83.80%	(Mandical et al. 2020)
2	Fake News	CNN+LSTM	GloVe	94.71%	(Agarwal et al. 2020)
3	Fake News	CNN	Tensorflow Embedding Layer	96.0%	(Ajao et al. 2018)
4	Fake News	CNN	TF-IDF	98.30%	(Kaliyar 2018)
5	Fake News	Deep CNN	GloVe	98.36%	(Kaliyar et al. 2020)
6	Fake News	Bi-LSTM	GloVe	98.75%	(Bahad et al. 2019)
7	Fake News	fakeBERT	GloVe	98.90%	(Kaliyar et al. 2021)
8	Fake News	ScrutNet	GloVe	99.56%	N/A

in fake news detection, introducing an exceptional deep ensemble network that surpasses existing methods based on reported performance metrics.

Data Availability The dataset utilised in this research comprises a comprehensive collection of news articles acquired from Kaggle. In adherence to ethical and legal considerations, the specific dataset utilised cannot be directly disclosed due to the probable disclosure of misleading data or any metadata. However, researchers interested in accessing the fake news dataset can readily obtain it by navigating to https://www.kaggle.com/c/fake-news/data. Access to and availability of the dataset are contingent upon compliance with Kaggle's terms and conditions. For detailed information regarding the dataset's origins, methodology employed in its collection, and potential applications, we recommend consulting the original Kaggle dataset page and associated documentation.

Declarations

Conflict of interest The authors of this research work declare no conflict of interest.

References

Agarwal A, Mittal M, Pathak A, Goyal LM (2020) Fake news detection using a blend of neural networks: an application of deep learning. SN Comput Sci 1:1–9

Ahmad I, Ahmad I, Ahmad I, Ahmad I, Yousaf MM, Yousaf M, Yousaf S, Yousaf S, Ahmad MO, Ahmad MO (2020) Fake news detection using machine learning ensemble methods. Complexity. https://doi.org/10.1155/2020/8885861

Ahmed H, Traore I, Saad S (2018) Detecting opinion spams and fake news using text classification. Securi Priv 1(1):9

Ajao O, Bhowmik D, Zargari S (2019) Sentiment aware fake news detection on online social networks. In: ICASSP 2019-2019 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, pp 2507–2511

Ajao O, Bhowmik D, Zargari S (2018) Fake news identification on twitter with hybrid CNN and RNN models. In: Proceedings of the 9th international conference on social media and society, pp 226–230

Allcott H, Gentzkow M (2017) Social media and fake news in the 2016 election. J Econ Perspect 31(2):211–236

Alsaeedi A, Al-Sarem M (2020) Detecting rumors on social media based on a CNN deep learning technique. Arab J Sci Eng 45:10813–10844

Asghar MZ, Habib A, Habib A, Khan A, Ali R, Khattak A (2021) Exploring deep neural networks for rumor detection. J Ambient Intell Humaniz Comput 12:4315–4333

Bahad P, Saxena P, Kamal R (2019) Fake news detection using bi-directional LSTM-recurrent neural network. Proced Comput Sci 165:74–82

Bhutani B, Rastogi N, Sehgal P, Purwar A (2019) Fake news detection using sentiment analysis. In: 2019 twelfth international conference on contemporary computing (IC3), IEEE, pp 1–5

Essa E, Omar K, Alqahtani A (2023) Fake news detection based on a hybrid BERT and lightGBM models. Complex Intell Syst 9(6):6581–6592

Fernández-Reyes FC, Shinde S (2018) 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. Springer, pp 206–216

Ghanem B, Rosso P, Rangel F (2020) An emotional analysis of false information in social media and news articles. ACM Trans Internet Technol 20(2):1–18

Giachanou A, Rosso P, Crestani F (2019) Leveraging emotional signals for credibility detection. In: Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval, pp 877–880

Girgis S, Amer E, Gadallah M (2018) Deep learning algorithms for detecting fake news in online text. In: 2018 13th international conference on computer engineering and systems (ICCES), IEEE, pp 93–97

Granik M, Mesyura V (2017) Fake news detection using Naive Bayes classifier. In: 2017 IEEE first Ukraine conference on electrical and computer engineering (UKRCON), IEEE, pp 900–903

Guleria V, Verma A, Dhenkawat R, Chaurasia U, Singh NP (2023) Brain tumor detection using texture based LBP feature on MRI images using feature selection technique. In: Artificial intelligence, blockchain, computing and security, vol 1, CRC Press, pp 30–36

Ibrishimova MD, Li KF (2020) A machine learning approach to fake news detection using knowledge verification and natural language processing. In: Advances in intelligent networking and collaborative systems: the 11th international conference on intelligent networking and collaborative systems (INCoS-2019), Springer, pp 223–234

Jadhav SS, Thepade SD (2019) Fake news identification and classification using DSSM and improved recurrent neural network classifier. Appl Artif Intell 33(12):1058–1068

Jain P, Jain P, Jain P, Sharma S, Sharma S, Sharma S, Monica N, Monica N, MÃ´nica M, Aggarwal PK, Aggarwal PK, Aggarwal P (2022) Classifying fake news detection using SVM, Naive Bayes and lSTM. Confluence https://doi.org/10.1109/confluence52989.2022.9734129

Jaiswal A, Verma H, Sachdeva N (2023) Swarm optimized fake news detection on social-media textual content using deep learning. 2023 international conference on advances in computing, communication and applied informatics (ACCAI) https://doi.org/10.1109/accai58221.2023.10201229

Jang SM, Kim JK (2018) Third person effects of fake news: fake news regulation and media literacy interventions. Comput Hum Behav 80:295–302

Jazi SY, Mirzaeinia A, Jazi SY (2024) Analyzing gender polarity in short social media texts with BERT: the role of emojis and emoticons. arXiv preprint arXiv:2406.09573

Kaliyar RK, Goswami A, Narang P (2021) FakeBERT: fake news detection in social media with a BERT-based deep learning approach. Multimed Tools Appl 80(8):11765–11788

Kaliyar RK, Goswami A, Narang P, Sinha S (2020) FNDNet-a deep convolutional neural network for fake news detection. Cogn Syst Res 61:32–44

Kaliyar RK (2018) Fake news detection using a deep neural network. In: 2018 4th International conference on computing communication and automation (ICCCA), IEEE, pp 1–7

Koru GK, Uluyol Ç (2024) Detection of Turkish fake news from tweets with BERT models. IEEE Access 12:14918–14931

Li Q, Hu Q, Lu Y, Yang Y, Cheng J (2020) Multi-level word features based on CNN for fake news detection in cultural communication. Pers Ubiquit Comput 24:259–272

Long Y, Lu Q, Xiang R, Li M, Huang C-R (2017) Fake news detection through multi-perspective speaker profiles. In: Proceedings of the eighth international joint conference on natural language processing (volume 2: short papers), pp 252–256

Madani M, Motameni H, Mohamadi H (2022) Fake news detection using deep learning integrating feature extraction, natural language processing, and statistical descriptors. Secur Priv 5(6):264

Mahir EM, Akhter S, Huq MR, et al (2019) Detecting fake news using machine learning and deep learning algorithms. In: 2019 7th international conference on smart computing and communications (ICSCC), IEEE, pp 1–5

Mandical RR, Mamatha N, Shivakumar N, Monica R, Krishna A (2020) Identification of fake news using machine learning. In: 2020 IEEE international conference on electronics, computing and communication technologies (CONECCT), IEEE, pp 1–6

Manzoor SI, Singla J, et al (2019) Fake news detection using machine learning approaches: a systematic review. In: 2019 3rd international conference on trends in electronics and informatics (ICOEI), IEEE, pp 230–234

Mishra S, Shukla P, Agarwal R (2022) Analyzing machine learning enabled fake news detection techniques for diversified datasets. Wirel Commun Mob Comput 2022:1–18

Mridha MF, Keya AJ, Hamid MA, Monowar MM, Rahman MS (2021) A comprehensive review on fake news detection with deep learning. IEEE Access 9:156151–156170

Nasir JA, Khan OS, Varlamis I (2021) Fake news detection: a hybrid CNN-RNN based deep learning approach. Int J Inf Manage Data Insights 1(1):100007

Ozbay FA, Alatas B (2020) Fake news detection within online social media using supervised artificial intelligence algorithms. Phys A 540:123174

Padnekar SM, Kumar GS, Deepak P (2020) BiLSTM-autoencoder architecture for stance prediction. In: 2020 international conference on data science and engineering (ICDSE), IEEE, pp 1–5

Park MJ, Chai S (2023) Constructing a user-centered fake news detection model by using classification algorithms in machine learning techniques. IEEE Access. https://doi.org/10.1109/access.2023.3294613

Pennycook G, Rand DG (2021) The psychology of fake news. Trends Cogn Sci 25(5):388–402

Qawasmeh E, Tawalbeh M, Abdullah M (2019) Automatic identification of fake news using deep learning. In: 2019 sixth international conference on social networks analysis, management and security (SNAMS), IEEE, pp 383–388

Rasool T, Butt WH, Shaukat A, Akram MU (2019) Multi-label fake news detection using multi-layered supervised learning. In: Proceedings of the 2019 11th international conference on computer and automation engineering, pp 73–77

Reddy H, Raj N, Gala M, Basava A (2020) Text-mining-based fake news detection using ensemble methods. Int J Autom Comput 17(2):210–221

Rusli A, Young JC, Iswari NMS (2020) Identifying fake news in Indonesian via supervised binary text classification. In: 2020 IEEE international conference on industry 4.0, artificial intelligence, and communications technology (IAICT), IEEE, pp 86–90

Sahoo SR, Gupta BB (2021) Multiple features based approach for automatic fake news detection on social networks using deep learning. Appl Soft Comput 100:106983

Shaik S (2023) A review and analysis on fake news detection based on artificial intelligence and data science. Tuijin Jishu J Propuls Technol. https://doi.org/10.52783/tjjpt.v44.i3.2107

Singh S, Verma A, Guleria V, Yadav S, Singh NP (2023) Deep learning-based networks to detect leaf disease in maize and corn. In: 2023 international conference on IoT, communication and automation technology (ICICAT), IEEE, pp 1–6

Singhania S, Fernandez N, Rao S (2017) 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, Springer, pp 572–581

Thota A, Tilak P, Ahluwalia S, Lohia N (2018) Fake news detection: a deep learning approach. SMU Data Sci Rev 1(3):10

Tsai CM (2023) Stylometric fake news detection based on natural language processing using named entity recognition: in-domain and cross-domain analysis. Electronics. https://doi.org/10.3390/electronics12173676

Ujgare NS, Singh NP, Verma PK, Patil M, Verma A. Non-invasive blood group prediction using optimized EfficientNet architecture: a systematic approach

Umer M, Imtiaz Z, Ullah S, Mehmood A, Choi GS, On B-W (2020) Fake news stance detection using deep learning architecture (CNN-LSTM). IEEE Access 8:156695–156706

Ushashree P, Naik A, Gurav S, Kumar A, Chethan SR, Madhumala BS (2023) Fake news detection using neural network. 2023. In: IEEE international conference on integrated circuits and communication

systems (ICICACS) https://doi.org/10.1109/icicacs57338.2023.10100208
Vardhan H, Verma A, Singh NP (2023) An ensemble learning approach for large scale birds species classification. In: Artificial intelligence, blockchain, computing and security, vol 1, CRC Press, pp 3–8
Verma PK, Agrawal P, Madaan V, Prodan R (2023) MCred: multimodal message credibility for fake news detection using BERT and CNN. J Ambient Intell Humaniz Comput 14(8):10617–10629
Verma A, Amin SB, Naeem M, Saha M (2022) Detecting COVID-19 from chest computed tomography scans using AI-driven android application. Comput Biol Med 143:105298
Verma A, Gupta N, Bhatele P, Khanna P (2023) JMCD dataset for brain tumor detection and analysis using explainable deep learning. SN Comput Sci 4(6):840
Verma A, Rahi R, Singh NP (2023) Novel ALBP and OLBP features for gender prediction from offline handwriting. Int J Inf Technol 15(3):1453–1464
Verma A, Singh N, Khanna V, Singh BP, Singh NP (2024) Automated tongue contour extraction from ultrasound sequences using signal enhancing neural network and energy minimized spline. Multimed Tools Appl 83(19):57511–57530
Veronica P-R, Bennett K, Alexandra L, Rada M (2018) Automatic detection of fake news. In: International conference on computational linguistics (COLING)
Yousefpanah K, Ebadi M, Sabzekar S, Zakaria NH, Osman NA, Ahmadian A (2024) An emerging network for COVID-19 CT-scan classification using an ensemble deep transfer learning model. Acta Tropica 107277
Zhang X, Ghorbani AA (2020) An overview of online fake news: characterization, detection, and discussion. Inf Process Manage 57(2):102025
Zhang Z, Lv Q, Jia X, Yun W, Miao G, Mao Z, Wu G (2024) GBCA: graph convolution network and BERT combined with co-attention for fake news detection. Pattern Recogn Lett 180:26–32