RESEARCH ARTICLE OPEN ACCESS

Fake News Detection Using NLP and BERT

Tabitha Indupalli
Dept. ofComputer Science AndEngineering GokarajuRangarajuInstitute ofEngineeringandTechnology Hyderabad,India

A.HarshithReddy
Dept.ofComputer Science AndEngineering GokarajuRangarajuInstitute ofEngineeringandTechnology Hyderabad,India

Dinesh. Ragipani
Dept. ofComputer Science AndEngineering GokarajuRangarajuInstitute ofEngineeringandTechnology Hyderabad,India

T.Dheeraj
Dept. of Computer Science And Engineering GokarajuRangarajuInstitute of Engineering and Technology Hyderabad, India

Zh.Bhanu uday
Dept.ofComputer Science And Engineering GokarajuRangarajuInstitute ofEngineeringandTechnology Hyderabad, India

Abstract—
False news has germinated on the social media and other websites, which may have been at a light speed Misleadingly. Our goal is to developasystem, whichisabletodetectfakenewsautomaticallyand reliably.
This is done in two manners, The first method is an application of NLP methods and uses logistic regression and Naive Bayes where emphasis is made on how words are used by deriving featuresthrough TF-IDF. The second model is the DistilBERT, which is a deep learning model that makes more sense of words and their context. The ensemble method which involves a combination of the outcome of the two methods is used to reach the final decision.
The system will be more accurate with the traditional machine learning as well as the advanced deep learning. The result of this initiative can be used to prevent the dissemination of fake news and assist media to advertise reliable information.

Date of Submission: 08-04-2026 Date of acceptance: 20-04-2026

I. INTRODUCTION

In the modern times, most of the news is disseminated through social media and the internet. Even though these formats give individuals access to more information, theymaybe the causeofthe rapid spread of fake news. Such misinformation may give the audience the wrong conclusions, make them confused, and even influence the social structures and political and economical systems. The fake news detection is a valuable strata of research because of these issues.

Simple and efficient machine learning approaches such as Logistic Regression and Naive Bayes are usually applied in classification of text. However, they do not tend to get the deeper meaning of text. BERT and its smaller framework, DistilBERT are both transformer- based models that are highly adept at sufficient language understanding, at the sacrifice of higher costs of operation.

The aim of this project was to develop a hybrid model, which employed both the traditional machine learning and DistilBERT. The ensemble learning is a combination of the advantages of two models into a system. On top ofthat, to yourparticular problem you are keen on creating a fake news detection model that would not only be accurate but also have broad results.

II. LITERATUREREVIEW

The features of the text were known and trained with the help of SVM and LightGBM, which effectively find various linguistic patterns and are computationally efficient and understandable through NLP-based preprocessing such as tokenization, stop-word removal, and vectorization. DistilBERT was employed in order to enhance situational understanding - a smaller variant of the BERT model,butnonethelessveryefficientandabletoidentify misleading information at different levels of abstraction due to its deepsemantic validation capability. The outcome was a fruitful tradeoff betweenspeedandsemanticrelevance - SVMwithLightGBM were the fast interpretable baselines to compare them with, and DistilBERT gave them the ability to flexibly add vocabulary generalization to those sets. The hybrid framework was very scalable, interpretable, and rich in the context. The last system is a powerfulandrobustmodelofautomatedfakenewsdetection,which can be extended in the future, to the multilingual and multimodal systems.

The study titled *"An Empirical Comparison of Machine Learning and Deep Learning Models for Automated Fake NewsDetection"*(2025)byYexinTian,Shuo Xu,YuchenCao compares traditional machine learning models—Logistic Regression,RandomForest,andLightGradientBoosting

Machine (LightGBM)—with a deep learning transformer model ALBERT for fake news detection. The experiments were conducted on the WELFake dataset. The study emphasizes interpretability, computational complexity, and generalization of models. While ALBERT achieved higher accuracy, simpler models like LightGBM performed competitively with less computational cost.
In Fake News Detection Using Logistic Regression and DecisionTreeAlgorithm(2025),Mr.R.AriharaSuthan, Dr. R. Sri Devi creates a fake news detector using Logistic Regression and Decision Tree algorithms on a Kaggle supplied dataset . The text preprocessing includes tokenization and removal of stop words before training.While modeling limits the scope of the model, because they only utilized one small dataset, I would imagine increasing dataset size and diversity would allow for a more robust model.

The paper *"WELFake: Word Embedding Over Linguistic Features for Fake News Detection"* (2022) byPawan Kumar Verma, Prateek Agrawal, Ivone Amorim introduces the WELFake dataset, integrating linguistic feature extraction withwordembeddingsforimprovedfakenewsclassification. The approach enhances domain generalization by combining semantic and syntactic information. It also discusses model complexity,aimingtobalanceperformanceandcomputational efficiency across diverse textual data.

Thestudytitled*"FakeNewsDetectionSystemUsingDecision Tree Algorithm and Compare Textual Property with Support Vector Machine Algorithm"* (2022) by N. Leela Siva Rama Krishna, M. Adimoolam compare Decision Tree (DT) and SupportVectorMachine(SVM)modelsusingadatasetof 311 instances each. The focus is on feature selection and model comparisontoevaluateclassificationaccuracyandefficiency.
ThestudyfindsthatSVMperformsbetterforcomplextextual boundaries, while DT is simpler and easier to interpret.
The paper *"Fake News Detection Using XLNet Fine-Tuning Model"*(2022)byAshokKumarJ,TinaEstherTrueman, Erik Cambria introduces fine-tuneing of the XLNet transformer modelonthe LIARdataset(collectedfromPolitiFact.com) to detect fake news. While the model captures contextual nuances well, it suffers from low accuracy in multi-class (6-class)classificationandcontextrestrictionsduetolimited data diversity.
Thestudy*"Characterization,Classification,andDetectionof FakeNewsinOnlineSocialMediaNetworks"*(2021)Monther Aldwairi, Ali Alwahedi presents fake news characteristics in online platforms, emphasizing clickbait headlines and deceptive content. It utilizes both the LIAR and ISOT Fake NewsDatasetstoanalyzemodelperformanceacrossdomains.

Thepaperdiscusseschallengesincross-domaingeneralization and highlights model complexity as a limiting factor for scalability.

The report *"Multi-Class Fake News Detection Using LSTM Approach"* (2022) by Sanzid Chowdhury, Sheak Rashed Haider Noori employs a Long Short-Term Memory (LSTM) neural network for multi-class fake news detection using the CheckThat!2021Task3ADataset.Thestudyexploresfeature engineering, generalization, and model comparison, showing thatLSTMeffectivelycapturessequentialworddependencies but requires more data for stable performance.
The study *"Classifying Fake News Detection Using SVM, NaiveBayes,andLSTM"*(2022)byAggarwal,S.,&Sharma, S compares three classification approaches—Support Vector Machine(SVM),NaiveBayes,andLongShort-TermMemory (LSTM)—onasmalldatasetof311instances.Ithighlightsthe effectoffeatureselectionanddatasetsizeonmodelaccuracy. LSTM outperforms traditional models but demands more computational resources.

The study*"Machine Learning-BasedApproachforFakeNews Detection"* (2024) by H. L. Gururaj, H. Lakshmi, B. C. Soundarya applies multiple machine learning algorithms to analyze textual features and detect fake news from news article datasets. The paper provides a comparative evaluation of algorithms based on accuracy and interpretability. It also discusses dataset characteristics, model explainability, and trade-offs between performance and transparency.

The study *"Analyzing Common Lexical Features of Fake News Using Multi-Head Attention Weights"* (2024) by Mamoru Mimura,Takayuki Ishimaru esearch uses the BERT model's multi-head attention mechanism to study lexical patternsinfakevs.realnewsacrossmultiplediversedatasets. The focus is on feature analysis and generalization through attention weight visualization. The study enhances interpretability by revealing which words contribute most to model predictions, advancing explainable fake news detection.

Most reviewed studies rely on limited or domain-specific datasets, which restricts generalization and real-world applicability. Traditional machine learning models offer interpretability but fail to capture deep semantic context,while deep learning and transformer models introduce higher computationalcostsanddeploymentchallenges.Additionally, many works emphasize accuracy over explainability, scalability,and multilingualsupport.Thelackofstandardized benchmarks and unified evaluation methods also makescross-study comparisons difficult.

The literature shows a transition from traditional machinelearning models to deep learning and transformer-based approaches for fake news detection. Early methods such as SVM, Logistic Regression, and Decision Trees provided efficiency and interpretabilitybutstruggledwithcontextualunderstanding.Later models like LSTM, BERT, ALBERT, and XLNet improved semantic comprehension and achieved higher accuracy. Despite these advancements, trade-offs remain between interpretability, computational complexity, and contextual depth, highlighting the need for scalable, robust, and generalized fake news detection systems.

TABLE I LITERATURE REVIEWS

SrNo	**StudyTitleand Author**	**AITechniques Used**	**Prediction**
1	An Empirical Comparison of MachineLearning and Deep Learning Models for Automated Fake News Detection 2025	Logistic Regression RandomForest LightGradient Boosting Machine ALBERT	Deep learningmodels like ALBERT will dominate accuracy, but LightGBMremainsbest for efficiency and scalability.
2	FakeNews Detection Using Logistic Regression AndDecision Tree Algorithm 2025	Logistic Regression DecisionTree Algorithm Tokenization Removingstop words	Hybrid models combiningLRand DTcouldperform better on larger, more diverse datasets.
3	WELFake:Word EmbeddingOver Linguistic FeaturesforFake News Detection 2022	Linguistic Feature Extraction Word Embedding Integration	Integratingcontextual embeddings with linguistic cues will enhancedomain adaptability.
4	FakeNews DetectionSystem Using Decision TreeAlgorithm and Compare TextualProperty with Support Vector Machine Algorithm2022	DecisionTree (DT) and SupportVector Machine (SVM)	SVM will outperform DTasdatasetsgrow and feature richness increases.
5	FakeNews DetectionUsing XLNet Fine-Tuning Model, 2021	Proposed XLNetbase fine-tuning modelforfake newsdetection.	Larger, balanced datasets and domain- tuned XLNet models willimprovemulticlass detection.

6	Characterization, Classificationand DetectionofFake News in Online Social Media Networks 2021	prevalenceof fake news in online social medianetworks, focusingonthe roleofclickbait headlines and deceptive content.	Cross-domain transferlearningwillbe key to maintaining accuracyacross platforms.
7	Multi-ClassFake News Detection Using LSTM Approach 2022	The paper employsaLong Short-Term Memory (LSTM) network for multi-classfake newsdetection.	Combining LSTMwith attentionor transformer layers will improve contextual understanding.
8	ClassifyingFake News Detection Using SVM, NaiveBayes,and LSTM 2022	SupportVector Machine (SVM),Naive Bayes, and Long Short-TermMemory (LSTM)	Deeplearningwith data augmentation will surpass traditional models in robustness.
9	Machine Learning-Based Approachfor Fake News Detection2024	Uses multiple MLalgorithms to detect fake news from textualpatterns innewsarticles.	Mergingaccuracy with interpretability frameworkswillmake ML systems more practical.
10	Analyzing CommonLexical Features of Fake News Using Multi-Head Attention Weights2024	Examines lexicalpatterns in fake news using BERT's multi-head attention analysisacross datasets.	Explainableattention- based models will enhance trust and transparencyin detectionsystems.

III. Datacollectionandpreprocessing

The data applied in this research was taken on the publicly available welfake dataset which is a popular benchmark of thefakenewsdetectionactivity.Itincludesnews storiesthat were gathered on various sources in the internet, including political news sites and fact-checking websites like PolitiFact. Every record will contain the textual content of the article and binary label of the news being real or fake.

The data set was reduced to consist of the title and major textual material. Two disciplines have been combined in the same sequence of input to maintain contextual movement. The cleaning of the samples was used to eliminate duplicate samplesandunfinishedentriestoenhancetheoverallquality ofthedata.Thedatasetwasbinaryized,i.e.0representsreal news and 1 represents fake news.

A couple of preprocessing steps were carried out before feeding the data into the machine learning models. Theentire text was all changed to lower case in order to have a similar representation. To eliminate source-specific bias, URLs and hyperlinks were eliminated. All non-informative phrases like read more and click here were also done away with because they do not add any relational value. Also, the mentions and hashtags made by users were eliminated to make the text less noisy.

The processing of emojis was done selectively according to the model architecture. In the case of classical machine learning models, the emojis were completely eliminated to prevent the additionof scattered/noise feature. In the case of transformer-based models, the emojis were transformed into textual information to maintain the meaning associated with it. Non-alphabetic characters and punctuations were eliminated and the additional whitespaces was standardized.

After the preprocessing, cleaning text was converted to numerical representations. In classical models like Support Vector Model (SVM) and LightGBM, terms frequency inverse document frequency (TF-IDF) was used to generate the vital lexical patterns of the term. The transformer-based model, DistilBERT, directly took in the raw text that has already been preprocessed (i.e. tokenized).

This preprocessing pipeline made sure that the data was normalized, de-noised and both traditional machine learning models and deep learning models applied in this study were optimized.

IV. Methodology

The proposed work is a hybrid framework of fake news detection that integrates classical machine learning models with a transformer-based deep learning model. To classify textual data (long-form) by the use of TF-IDF features, SupportVector Model (SVM) and LightGBM classifiersare used, whereas when focusing on short-form texts

ISSN: 2248-9622, Vol. 16, Issue 4, April 2026, pp 95-101

(tweetsandheadlines),weapplyDistilBERTbecauseofit s

excellentcontextualcomprehension.

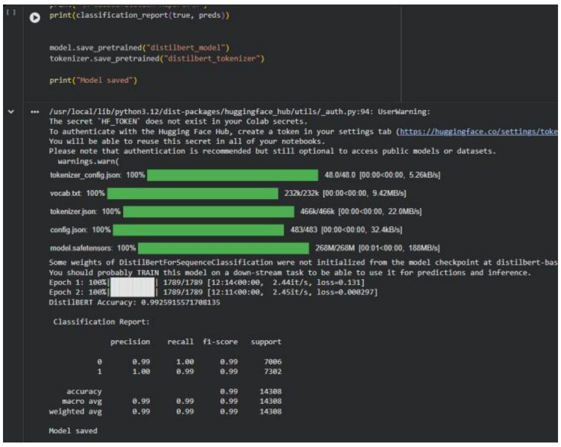
Classical models are represented as textual data with the helpoftheTermFrequency-InverseDocumentFrequency (TF-IDF)textualdatavectorization,whichissuccessfulin the capture of important lexical patterns. DistilBERT directly takes text that has been tokenized and is not learned to construct semantic representations by engineering an explicit set of features. All the models are trained separately using the same processed data using fixed label encoding.

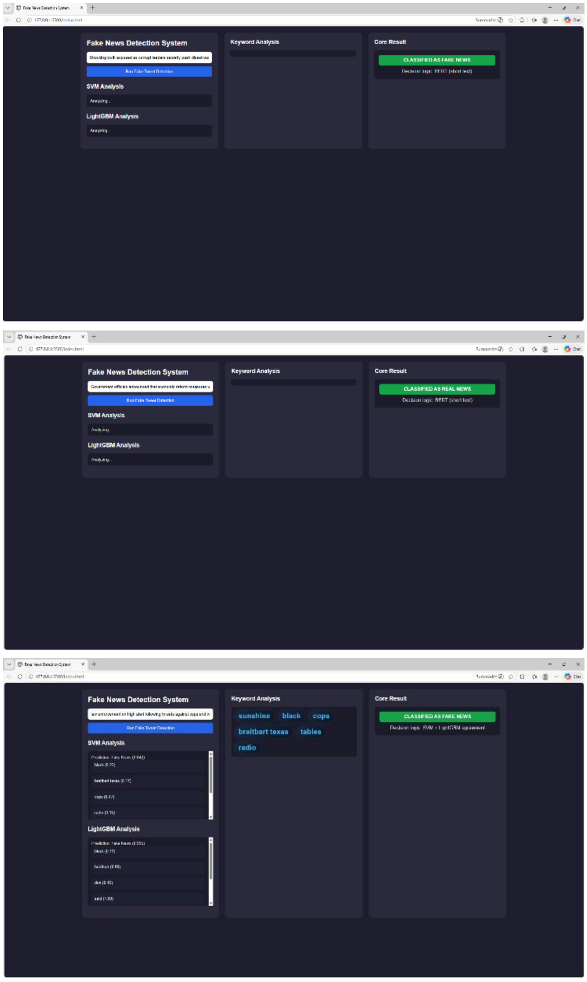
An ensemble method that uses rules is implemented upon inference. DistilBERT is applied only to forty or less words in texts whereas SVM and LightGBM are used to applyto longer texts. Inthe case oflong-formcontent, the finalpredictionisdeterminedbyagreementbetweenSV M and LightGBM, and selection of cases is through confidence. This strategy enhances robustness and accuracy of the different text lengths.

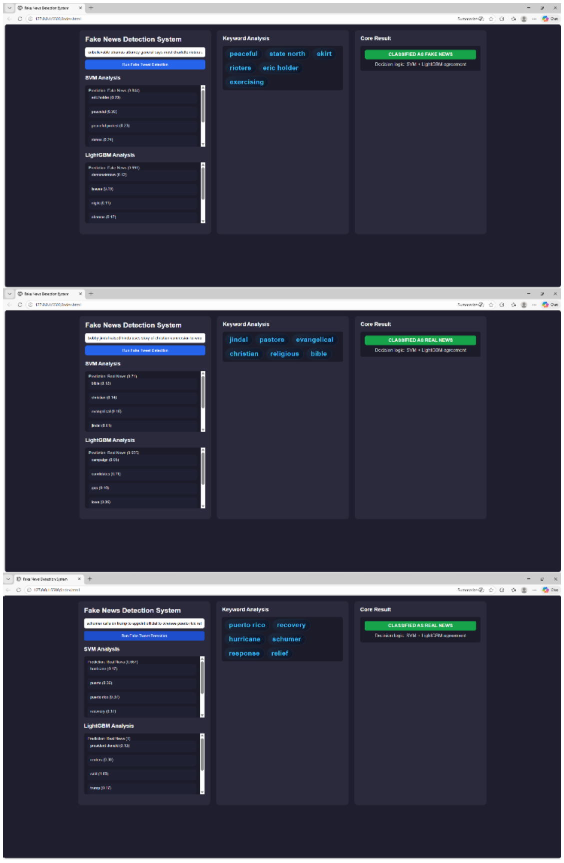
V. Experimentsandresults Model Accuracy
SVM(TF-IDF)97.48%
LightGBM(TF-IDF) 97.94%
DistilBERT(Transformer)99.26%

Experiments conducted on the effectiveness of three models SVM, LightGBM and DistilBERT in domain specific classification of tweets. To maintain the balance of the classes, the dataset was divided into 80 per cent trainingand20percenttestingstratifiedsampling.Incase of SVM and LightGBM, tweets were first converted into numerical form with TF-IDF representation but in distilbert-base-uncased case it was the raw text of the tweet that was directly processed with the pre-trained distilbert-base-uncased tokenizer and thenfine-tuned with the help of the use of a GPU. Accuracy, precision, recall and F1-score were used as the model performance.
SVM model achieved accuracy of 97.48 which is strong performance of the model in recognition of domain- specific tweets using the features based on key words. LightGBM marginally beat SVM with an accuracy of 97.94, as it has a gradient boosting mechanism, which is more sensitive to more complicated interactions of features in the TF-IDF representation. The two conventional machine learning models exhibited a stable and effective performance in performing the task of tweet classification.

The highest level of accuracy was 99.26 on DistilBERT which was far more successful than the TF-IDF-based models. It is able to use its transformer architecture that allows contextual and semantic interpretation of short and informal tweet text that results in a high classification performance. These findings show that transformer-based solutionscanbeusedmoresuccessfullytoclassifytweets, whereasconventionalmodelscanbeusedincaseswhere computational resources are limited.







VI. Discussion

Experimental findings prove that using a transformer-based model together with classical machine learning models enhances the detection of fake news in various text lengths.Theuseofcontextandsemanticinformationtow orkwithshort texts like tweets and headlines made DistilBERT an effective model whereas the SVM and LightGBM were more efficient with longer articles where the lexical patterns and term distributions were more evident.

Ensemble strategy rule-based strategy was also important in enhancing its reliability key aspect of prediction. The system minimized misclassificationinthecaseofamodelbeingbiased to certain types of text by passing short texts through DistilBERT and long texts through classical models. Under consensus between SVM and LightGBM enhanced trust in predictions of long-text, whereas confidence based fallback provided more stabilization of results in case of disagreements.

Even with its good performance, some limitations were recorded. The models were also prone to flashy variations including emojis, informal phrases, and non-standard phrasing whichsometimesaffectedpredictions.Moreover,the performance ofthe systemis strictlyrelated to the peculiaritiesof the training set, which means that more diverse data channels must be used in the future. Altogether, the hybrid framework represents the equal and explainable method of detecting fake news in heterogeneous textual data.

VII. CONCLUSION

To summarize, Fake News Detection provides the solution to an issue of widespread misinformation that is spread online through the use of innovative Natural Language Processing (NLP) algorithms and the most modern machine learning models.Simply put, the system consists of both the conventional algorithms, including SVM and LightGBM, as well as the deep-learningtechniquesincludingDistilBERT toensurea holisticand efficient detection mechanism. The NLP part contains the processing steps necessary, namely in order to do the preprocessing, such as tokenization, stop word removal, and convertingrawtext to a numerical format that will be ingested by the model. These measures assist in filtering and normalizing the data to make sure that the models are able to concentrate on the significant linguistic characteristics. SVM and LightGBM are used to effectively find superficial trends in the text, including a setofkeywordsorstylistic indicatorsthattendtodistinguishfake news and legitimate information. They are suitable in real time implementation because of their ensemble-based structure that offers reliable and interpretable results with low computational costs of their structure. In the meantime, more nuancedcontextual sensations of the text are captured at the cost of DistilBERT, a lightweight transformer. It is very efficient at identifying minute signs of misinformation, including biased tones, deceptive phrasing, or falsehoods masquerading as falsehoods of which other models may be less effective. This combined method strikes a balance between the capabilities ofthe classic machine learning models and the advancedintelligence of the deep learning. With the combination of these methods,thesystemwillbemoreaccurateandstrongindetecting fake news and eventually lead to a more reliable digital information space. The mixture of both will make sure that shallow cues and rich contextual cues are properly analyzed to enhance the overall detection.

References

[1]. "Essa, E., Omar, K., & Alqahtani, A. (2023). Fake News Detection based on a Hybrid BERTand LightGBM Model.

[2]. Arihara Suthan, R., & Sri Devi, R. (2025). Fake News Detection Using Logistic Regression and Decision Tree Algorithm.

[3]. Verma, P. K., Agrawal, P., & Amorim, I. (2022). Word EmbeddingOverLinguisticFeaturesforFakeNewsDetection.

[4]. Krishna, N. L. S. R., & Adimoolam, M. (2022). Fake News Detection System Using Decision Tree Algorithm and Comparison with SVM Based on Textual Features.

[5]. Ashok Kumar, J., Trueman, T. E., & Cambria, E. (2021). Fake News Detection Using XLNet Fine-Tuning Model. **https://doi.org/10.1109/ICACCS51430.2021.9441752**

[6]. Aldwairi, M., & Alwahedi, A. (2021). Characterization, Classification and Detection of Fake News in Online Social Media Networks. *IEEE Access, 9*, 140100–140117. **https://doi.org/10.1109/ACCESS.2021.3118402**

[7]. Chowdhury, S., & Noori, S. R. H. (2022). Multi-Class Fake News Detection Using the LSTM Approach. *International Journal of Advanced Computer Science and Applications (IJACSA), 13*(6).

[8]. **https://doi.org/10.14569/IJACSA.2022.0130624**

[9]. Agarwal, S., & Sharma, S. (2022). Classifying Fake News Using SVM, Naive Bayes, and LSTM.**https://www.researchgate.net/publication/365456789_Classifying_Fake_News_Using_SVM_Naive_Bayes_and_LSTM**

[10]. Gururaj, H. L., Lakshmi, H., & Soundarya, B. C. (2024). Machine Learning-Based Approach for Fake News Detection.

[11]. Mimura, M., & Ishimaru, T. (2024). Analyzing Common Lexical Features of Fake News Using Multi-Head Attention Weights. *Journal of Information & Data Science.* **https://doi.org/10.1007/s42488-024-000XX**