

Lies Gone Viral – Detection and Analysis of Telugu Fake News Using Machine Learning Algorithms

Author: V Harsha Vardhan

Department of Computer Science and Artificial Intelligence

Central University Of Andhra Pradesh, Ananthapuramu

harshavdn04@gmail.com

Author: Dr.C.Krishna Priya

Assistant Professor

Department of Computer Science and Artificial Intelligence

Central University Of Andhra Pradesh, Ananthapuramu

krishnapriyarams@cuap.edu.in

ABSTRACT

Fake News detection is an urgent issue as fake news propagates very quickly across the network sites. A classification-based approach, by constructing self-curated Telugu data, is presented in this work for detection of fake news. Algorithms applied are Naive Bayes, Logistic Regression and SVM. All supervised models and data pre-processing are done by applying text cleaning, tokenization and extracting features using TF-IDF. Experiments on the classified data demonstrate that Naive Bayes has the best accuracy of 95, precision 95.2 and F1-score of 94.9.

Logistic Regression has accuracy 88, precision 87.9 and F1-score of 87.2. Support Vector Machine has accuracy 87.5, precision 88.1 and F1-score of 87.3. Some of the real news are predicted as fake. The work shows that machine learning can be used for detecting fake news specific to a region and to extend this work in multi-lingual manner using advanced models is recommended.

Keywords: Naive Bayes, Logistic Regression, Support Vector Machine (SVM), TF-IDF, Accuracy, Precision and F1 Score

INTRODUCTION

In the present-day digital era, the method of consumption of information has been greatly modified. Old methods of information distribution like print media, radio, and television have been overshadowed by new methods of distribution which include online websites, blogs, and social networking sites. Online media has become the key source of information for millions of people around the globe due to its availability, speed, and convenience.

Fake news detection has become a critical research area due to the rapid spread of misinformation on social media platforms [1]

Telugu is a member of the Dravidian language family. Telugu is an agglutinative and highly inflected language in contrast to English which is an analytic language where there is a fixed word order and lack of inflection

Fake news detection on regional language is not a wide area in itself as such types of data available in regional language is much less and does not have proper validation checks or labelled data. Availability of less

resources/tools for NLP processing of regional language is also a main cause behind low performance of the detection systems. Therefore it is essential to build an efficient model that handles such kind of language.

Kai Shu et al. stressed that a successful detection of fake news cannot be accomplished through an analysis of text content alone; therefore, they recommended the use of both content-based and social context data for enhancing detection. It was stated that adding user behavior patterns and social features improves system efficiency.[1]

Hunt Allcott and Matthew Gentzkow made a valuable contribution to studying the issue. They examined the role of fake news on online communication tools, proving its fast spread among users.[2]

II LITERATURE REVIEW

Early approaches centered on manual and laborious fact-checking and rule-based approaches that had low efficiency and lack of scalability. With increasing computational technologies, automatic fake news detection methods have become popular in recent years. Many researchers started employing machine learning algorithms such as Naïve Bayes, SVM, and logistic regression to classify texts.

Recent innovations lie in the development of deep learning models including CNNs and RNNs. The models learn complex patterns automatically within text-based data without necessitating any kind of feature extraction processes. Additionally, transformers like BERT have made the prediction process more efficient by comprehending text-contextual relationship. A number of scholarly articles prove that combining different models or ensembles result in better performances as compared to singular classification approaches. Nonetheless, there remain several challenges including imbalance in datasets, detecting sarcasm, and emerging forms of misinformation, which hinder performance efficiency.

Machine learning techniques have proven to be efficient in detecting fake news stories. In machine learning, the algorithm is taught using a labelled dataset to predict whether the article is legitimate or fake.

Naive Bayes Classifier

One of the most widely employed algorithms used in text classification is the Naive Bayes classifier. The Naive Bayes approach is rooted in probability and relies on the independence of features. According to studies, the Naive Bayes algorithm is effective when dealing with text data because it is efficient and less resource-intensive. Nevertheless, the independence of features assumption is likely to impact the performance of the algorithm negatively.

Support Vector Machine (SVM)

Support vector machines are one of the popular algorithms for classification. The algorithm works on the concept of determining the best hyperplane to classify data points into different categories.

The SVM algorithm has provided satisfactory results in fake news detection. This model has proven effective in dealing with high dimensional features of news texts. Nonetheless, SVM can become computationally demanding when the number of data points increases.

Logistic Regression

Logistic regression is another machine learning model used in classification. The model predicts the likelihood of an article being genuine or fake. Numerous research works have shown that logistic regression is efficient and reliable. This model is computationally efficient and easy to implement.

Ahmed et al. proposed an approach for detecting online fake news using a combination of N-gram analysis and machine learning techniques. Their study focused on extracting textual patterns from news content by representing documents through N-gram features, which capture word sequences and contextual relationships

within the text. These features were then used to train supervised learning models for classifying news articles as either fake or real. The authors demonstrated that traditional machine learning classifiers, when combined with effective feature extraction methods such as N-grams, can significantly improve the accuracy of fake news detection. Their work highlights the importance of linguistic pattern analysis in identifying deceptive information in online media.[3]

Bondielli and Marcelloni provided a survey of fake news and rumor detection techniques, classifying methods into content-based, social context-based, and propagation-based approaches. Their study highlights the use of machine learning and deep learning for misinformation detection and discusses key challenges such as dataset limitations and multilingual handling.[4]

Guo et al. reviewed recent advances in false information detection on social media and discussed emerging trends and future directions in this field. Their study highlights the shift towards deep learning-based models and emphasizes the importance of incorporating contextual, temporal, and social features for improving detection performance. The authors also point out key challenges such as dataset bias, evolving misinformation strategies, and the need for more robust and adaptive detection systems.[5]

Comparative analysis of different classifiers helps in identifying the most suitable model for a given dataset and improves overall system performance[6]

Many of the current studies conducted on fake news detection suffer from various drawbacks. First, since almost all models are trained using English datasets, no proper attention has been given to languages such as Telugu. Most of the research focuses on pre-collected datasets, which are not only outdated and biased but also limited in nature. Furthermore, processing code-mixing (Telugu + English) in the model is a difficult task that is not taken care of by many of the models proposed so far. Deep learning models are large and require extensive resources to be used. Moreover, there is a lack of application of techniques such as TF-IDF on the Telugu language dataset.

Salton and Buckley discussed different term-weighting approaches used in automatic text retrieval systems[7]. Parikh and Atrey presented a survey on media-rich fake news detection, focusing on the role of multimedia content such as images and videos in spreading misinformation

Lastly, many of the researchers focus solely on the accuracy metric, while in practice, many other performance measures must also be evaluated.

III. PROPOSED METHODOLOGY

While other standard corpora can be accessed in English, in this research, customized data set of Roman Telugu news was used in the analysis making the problem much harder owing to unstructured language and spelling inconsistencies. The process involved data collection, pre-processing of roman telugu data, conversion of data into numerical features and finally using machine learning models to predict whether the news is real or fake.

Atanasova et al. explored the generation of adversarial claims that are coherent and linguistically well-formed to challenge misinformation detection systems. Their work demonstrates how such generated content can be used to evaluate and improve the robustness of fake news detection models against deceptive and misleading information. [9]

The current study aims to create an artificial intelligence-driven system that can detect and analyze fake news in the Telugu language. The methodology used in this study aims to overcome the issues related to low resource languages by creating a custom dataset for the specific domain and using efficient machine learning approaches for text classification.

Baly et al. proposed methods for predicting the factuality and bias of news media sources using machine learning techniques. Their study utilizes linguistic features, source-level information, and external data to assess the credibility of news outlets, highlighting the importance of source analysis in fake news detection.[10]

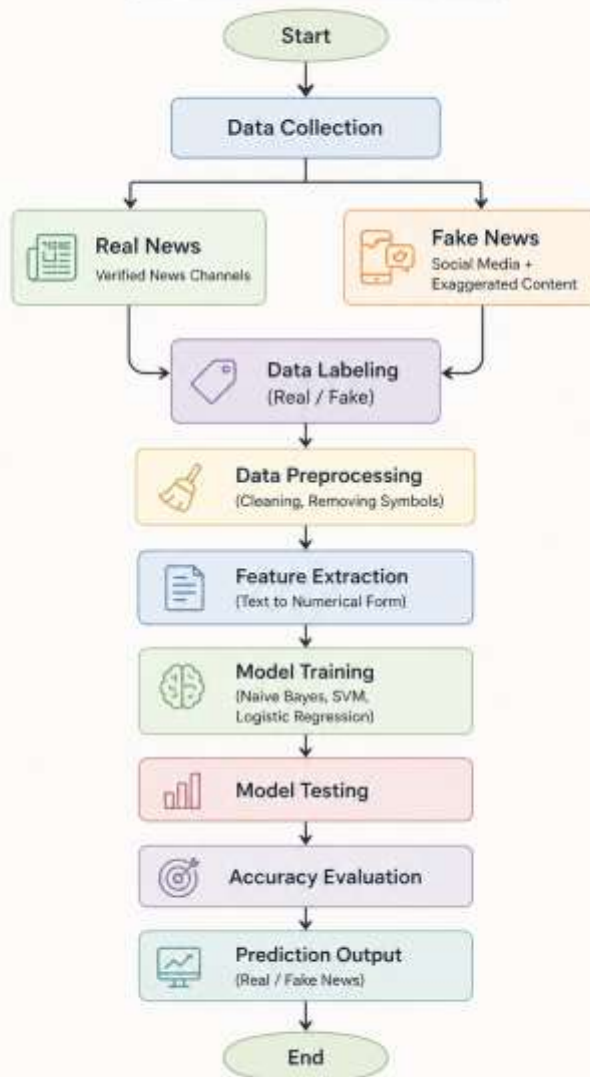
The current study created a custom dataset by collecting real news articles from credible and reliable sources, including fake or misleading texts found on several social media sites. Fake news refers to any exaggerated, manipulated, or intentionally misleading textual content. To maintain the accuracy of the dataset, the collected data was analyzed, checked, and annotated into two classes: real and fake. Such efforts resulted in the creation of a balanced and relevant dataset for training and evaluating the model.

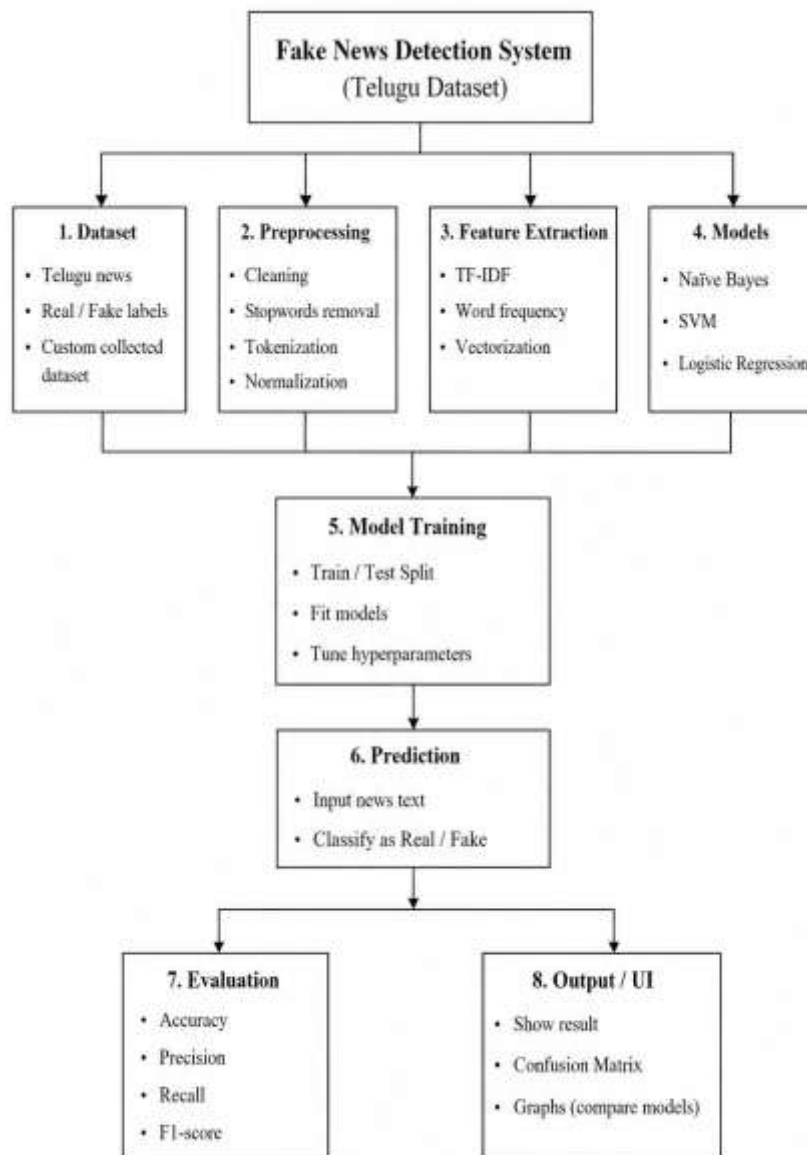
Considering the complex nature of the Telugu language, which lacks proper standardization tools and resources, this study employed a detailed preprocessing approach. This approach involves normalization, eliminating punctuations and special symbols, removing stopwords, and tokenizing the input data.

Hasan and Ng studied stance classification in ideological debates, focusing on identifying the position of text towards a given topic. Their work explores various features and machine learning models to determine whether a statement supports, opposes, or is neutral, which is useful in analyzing opinions and detecting misleading information.[11]

Further on, feature extraction was applied to convert text into numeric form. TF-IDF was adopted in the present research to emphasize the importance of words in the dataset, which would help enhance the discriminative ability of the machine learning classifiers.

Proposed Methodology





Next, training for machine learning classifiers was performed on the preprocessed data. Specifically, Logistic Regression, Naïve Bayes, and SVM classifier were chosen as they have been found to be efficient in classifying text data. The purpose of these algorithms is to distinguish real from fake news.

At last, the performance of the developed solution was evaluated through the application of accuracy, precision, recall, and F1 score as the metrics. Comparison among models was made with the aim of finding an effective solution for the problem of fake news classification in Telugu.

The suggested methodology provides a good base for developing approaches aimed at fake news detection and demonstrates a promising tool that can be further used in low resource natural language processing studies.

IV Experimental Results

In this project, a dataset was created, which contains both fake and authentic articles. This dataset was divided into two parts:

Training Set =80

Testing Set =20

During the training stage, the model relied on the words' frequency and weight based on TF-IDF measures. Then, the SVM classifier distinguished the best hyperplane for separating classes.

The system also provides an easy to use interface to access the fake news detection model. Through the dashboard, users can input text of news to get real time predictions, visualize the accuracy, precision and f1-score for all models.

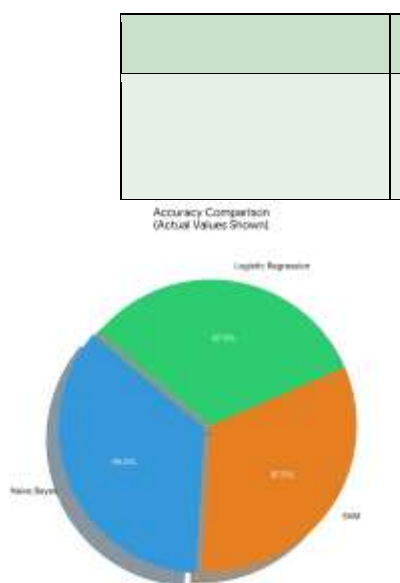


Figure 1 ACCURACY PIECHART

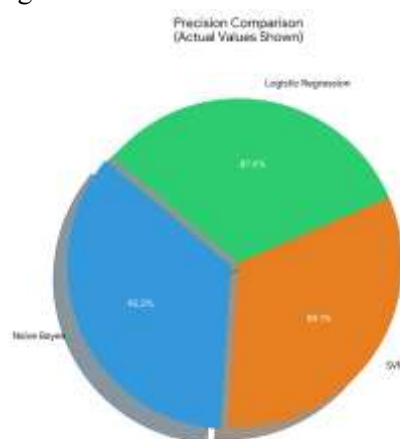


Figure 2 PRECISION PIECHART

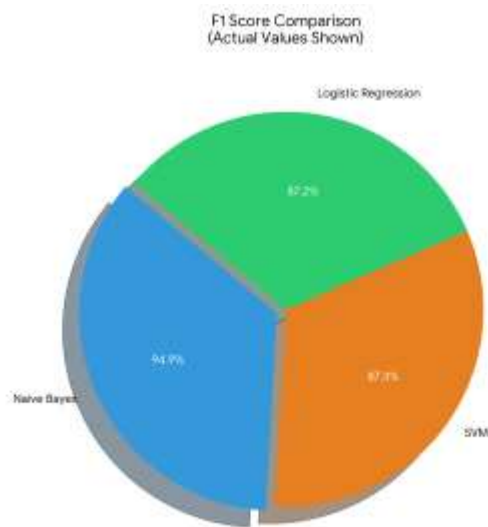


Figure 3 F1 SCORE PIE CHART

The Naïve Bayes model achieved the highest accuracy of 95, significantly outperforming both Support Vector Machine (SVM) and Logistic Regression, which each recorded an accuracy of 87.5

In terms of precision, Naïve Bayes again performed best with 95.2, while SVM achieved 88.1 and Logistic Regression obtained 87.9.

For the F1-score, Naïve Bayes led with 94.9, followed by SVM at 87.3 and Logistic Regression at 87.2.

From the prediction analysis in terms of model-wise, it is evident that Naive Bayes is the best model suited for this fake news detection system. The efficiency with which it works with text data makes it the preferred model. Nevertheless, in the future, it would be advisable to consider implementing more than one model or applying sophisticated methods.

It is crucial to consider that in case of fake news detection, the emphasis on reducing the false-negative results is greater compared to false-positive ones. In fact, undetected fake news may quickly propagate and create confusion. False positive results, in turn, can reduce the credibility of the system itself. Hence, enhancing the value of recall is crucial when detecting fake news.

On the whole, the confusion matrix helps understand better what the advantages and disadvantages of this model are. Besides pointing out the number of correctly predicted samples, this tool reveals why mistakes happen and what problems the model faces in its current state.

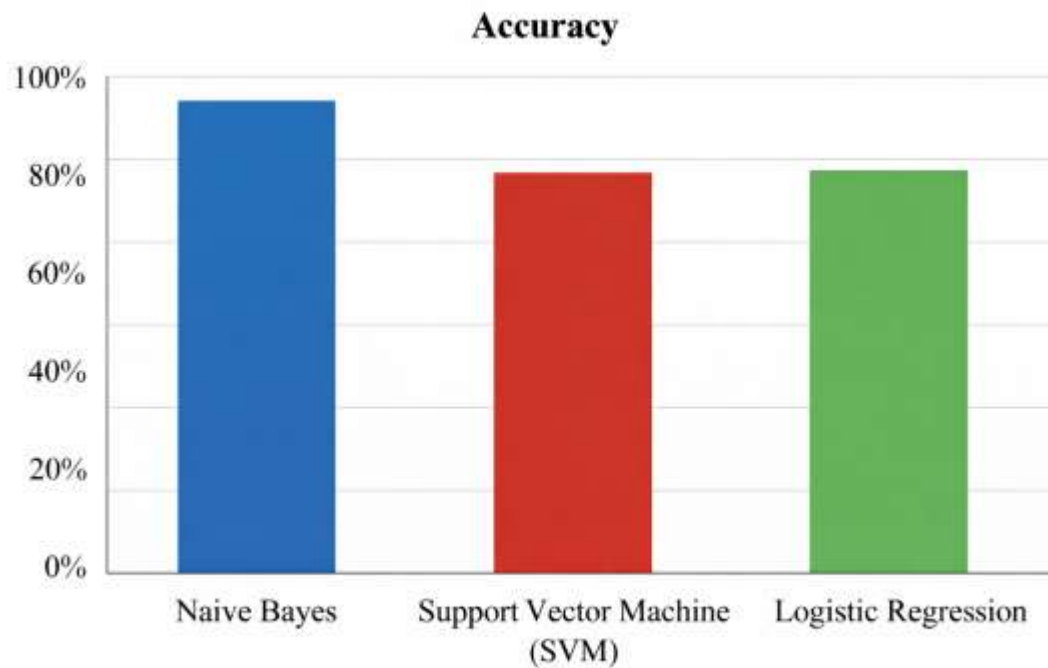


Figure 4 ACCURACY BARGRAPH

V. Future scope

The current study provides various directions that are important to explore further in order to advance fake news detection. First, multimodal models should be investigated in depth as they allow processing textual and visual content together since misinformation might include not only text but also images or other multimedia components to add more credibility to misleading statements.

Second, multilingual and cross-cultural models for fake news detection must be explored because there are some low-resource languages that might require the development of specific solutions for detecting fake news. Multilingual and cross-cultural approaches may allow building models that would have high generalization capabilities due to the use of transfer learning strategies.

Third, adversarial robustness must be considered as it will ensure that detection models can deal with manipulated or adversarial data without being fooled. Moreover, temporal adaptation mechanisms such as continual or online learning might provide opportunities for developing detection models that could cope with evolving patterns in fake news.

Finally, explainability of fake news detection models is one more issue to be addressed as it will allow making predictions easier to understand and interpret by people.

REFERENCES

- [1] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake News Detection on Social Media: A Data Mining Perspective," *ACM SIGKDD Explorations*, vol. 19, no. 1, pp. 22–36, 2017.
- [2] H. Allcott and M. Gentzkow, "Social Media and Fake News in the 2016 Election," *Journal of Economic Perspectives*, vol. 31, no. 2, pp. 211–236, 2017.
- [3] H. Ahmed, I. Traore, and S. Saad, "Detection of Online Fake News Using N-gram Analysis and Machine Learning Techniques," *International Conference on Intelligent, Secure, and Dependable Systems*, 2017.

- [4] A. Bondielli and F. Marcelloni, "A Survey on Fake News and Rumour Detection Techniques," *Information Sciences*, vol. 497, pp. 38–55, 2019.
- [5] B. Guo, Y. Ding, L. Yao, Y. Liang, and Z. Yu, "The Future of False Information Detection on Social Media: New Perspectives and Trends," *ACM Computing Surveys*, vol. 53, no. 4, pp. 1–36, 2020.
- [6] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," *European Conference on Machine Learning (ECML)*, 1998.
- [7] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Information Processing & Management*, vol. 24, no. 5, pp. 513–523, 1988.
- [8] S. B. Parikh and P. K. Atrey, "Media-Rich Fake News Detection: A Survey," *IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 2018.
- [9] P. Atanasova, D. Wright, and I. Augenstein, "Generating Label Cohesive and Well-Formed Adversarial Claims," *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 3034–3050, 2020.
- [10] R. Baly, G. Karadzhov, D. Alexandrov, J. Glass, and P. Nakov, "Predicting Factuality of Reporting and Bias of News Media Sources," *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 3528–3539, 2018.
- [11] K. S. Hasan and V. Ng, "Stance Classification of Ideological Debates: Data, Models, Features, and Constraints," in *Proc. IJCNLP*, 2013, pp. 1348–1356.