WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Identification of Fake News in Hindi Language

Pooja Pawar[1], Gaurav Mohite[2], Vaishnavi Zurunge[3], Mayur More[4], Mr. Rajkumar Panchal[5]

[1]*Student, Department of Information Technology, VPKBIET, Pune, Maharashtra, India,* pawarpooja1704@gmail.com
[2]*Student, Department of Information Technology, VPKBIET, Pune, Maharashtra, India,* gauravmohite1j@gmail.com
[3]*Student, Department of Information Technology, VPKBIET, Pune, Maharashtra, India,* zurungevaishnavi07@gmail.com
[4]*Student, Department of Information Technology, VPKBIET, Pune, Maharashtra, India,* mayurmore6163@gmail.com
[5]*Assistant Professor, Department of Information Technology, VPKBIET, Pune, Maharashtra, India,* rajkumar.panchal@vpkbiet.org

ABSTRACT-

With the recent social media boom, the spread of fake news has become a great concern for everybody. It has been used to manipulate public opinions, influence the election, incite hatred and riots. The credibility and trust in the news media are at an all-time low. It is becoming increasingly difficult to determine which news is real and which is fake. Various machine learning methods have been used to separate real news from fake ones. This research presents a comprehensive approach to fake news detection specifically tailored for the Hindi language. Now there is some confusion present in the authenticity of the correctness. There are some of the aspects that has to be kept in mind considering the fact that fake news detection is not only a simple web interface but also a quite complex thing that includes a lot of backend work. Our approach combines deep learning techniques with machine learning algorithms to create an effective and accurate fake news detection system.

Keywords - Fake news, Hindi language, machine learning, deep learning.

1. INTRODUCTION

Fake news is one of the biggest discouragements in our digitally connected world. Fake news spreads at lightning-fast speed impacting millions of people in the form of clickbait, trigrams everyday. Therefore, noticing fake news becomes a vital problem attracting huge research efforts. The rapid expansion of digital communication has revolutionized the way information is disseminated and consumed. However, in this era of rapid information exchange, a new and concerning phenomenon has emerged: the proliferation of fake news. Fake news, defined as intentionally misleading or fabricated information presented as factual news, poses a serious threat to the integrity of information and public discourse. This challenge is particularly pronounced in languages other than English, where linguistic nuances and cultural contexts can complicate the detection process. Fake news and lack of trust in the media are growing problems with huge ramifications in our society. Some of them now use the term to dismiss the facts counter to their preferred viewpoints. So the proposed system can be helpful for the detection of news into fake and real. In this report, we present a survey on the state of the art pertaining to the type of fake news and solutions that are being proposed. The research in this field has been going on for a long time and in the Indian context, the ill effects of spreading fake news are far from what anyone might think. Unlike in the context of other countries, WhatsApp is the prime distributor of fake news as compared to other social networking sites like Facebook and Twitter. Due to the increase of internet users in India, which has increased 137million (in 2012) to over 600 million (in 2019) facing unique challenges day by day.

Fig. Hindi Fake News.

Following are the types of fake news:-

1. For entertainment purpose.

2. Use a fake image or title for irrelevant content.

3. Misinterpreted information

4. Completely baseless content.

5. Rumors spread by blind followers.

These are such fake news that is easily available on social media. Though most fake news is not defective, they are used only for entertainment purposes but the readers do not understand the fact of this news and they change themselves according to the theme of the news. So it is very difficult for readers to understand the motto of news whether it is released for entertainment purposes or any other purpose. That's why it is very necessary to develop a such model which can easily indicate the motto of news so that readers will not get distracted.

2. HISTORY AND RESEARCH STATUS

The concept of fake news dates back centuries, but its detection relied heavily on human judgment and traditional journalistic practices. Manual fact-checking and editorial oversight were the primary means of identifying and mitigating misinformation. With the rise of the internet and digital media in the late 20th century, the dissemination of information became faster and more widespread. Early attempts to address digital misinformation involved manual fact-checking by journalists and media organizations. The advent of Web 2.0 and social media platforms like Facebook, Twitter, and YouTube in the mid-2000s marked a turning point. The rapid spread of user-generated content made manual detection insufficient, prompting the need for automated solutions. Initial automated systems were largely rule-based, relying on predefined patterns and keywords to identify potential misinformation. These systems were limited in scope and often generated high false-positive rates.

The 2010s saw the introduction of machine learning (ML) techniques for fake news detection. ML models could learn from data and improve over time, providing a more robust solution compared to rule-based systems. Advances in NLP enabled more sophisticated text analysis, allowing detection systems to understand context, sentiment, and the subtle nuances of language. Techniques such as sentiment analysis and topic modeling became integral to fake news detection. The creation of annotated datasets like the Fake News Challenge (FNC) and LIAR benchmark provided a foundation for training and evaluating ML models. These datasets facilitated the development and comparison of different approaches. The advent of deep learning brought significant improvements in fake news detection. Neural networks, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), were used to capture complex patterns in text data. The introduction of transformer-based models, such as BERT (Bidirectional Encoder Representations from Transformers) and its variants, revolutionized NLP. These models, pre-trained on large corpora, could be fine-tuned for specific tasks like fake news detection, achieving state-of-the-art performance.

With the ever-increasing volume of online content, real-time fake news detection systems became essential. These systems use streaming data analysis and real-time processing to quickly identify and respond to misinformation. Alongside technological solutions, efforts to engage and educate users gained importance. Digital literacy campaigns and tools like browser extensions and fact-checking apps empower users to identify and avoid fake news. Collaboration between academia, industry, and government bodies has been crucial in advancing fake news detection. Projects like the Partnership on AI and initiatives by organizations like the International Fact-Checking Network (IFCN) promote research and the development of ethical standards. The future of fake news detection lies in improving the accuracy and scalability of detection systems, enhancing multilingual capabilities, and developing more sophisticated multimodal analysis techniques. Continued research into explainable AI and user-friendly tools will be vital in the ongoing battle against misinformation.

3. LITEARTURE SURVEY

In paper [1], the authors proposed the fake news detection by using BERT model to classify the news into fake and real. They have experimented using BERT and different machine learning and deep learning methods and compared those techniques : LSTM, tf-idf which are good at detecting fake news. Dataset used here, have been taken from the GitHub repository of jelwin13afc/Fake News Detection To the best of our knowledge, this paper is the first to use the BERT model with a Hindi data set. The machine learning model employed in this study was the BERT model. The accuracy of the BERT model we employed in this study is 52%. When it was contrasted with other models like the tf-idf model and LSTM (Long-Short Term Memory), found that the accuracy provided by the LSTM model in comparison to our model's efficiency is 92.36% of previous studies are accurate, while 52% of BERT model. This study demonstrates how several fundamental machine learning and deep learning algorithms can be merged together.

In paper [2], authors proposed a fake news detection system by using m-BERT to classify news. The information used in the research came from the Cornell University Hostility Detection Dataset and BBC-Boomlive. This study will concentrate on several machine learning models and how well they can detect fake news in Hindi. They have been taken into account Naive Bayes Classifier and Multi-Layer Perceptron as these are the usual methods for text-based processing since the problem at hand is text categorization. Transformers have also been covered; they are useful for classifying texts because of their self-attention mechanism and improved comprehension of word features. The fine tuned mBERT model is explained in order to make comparisons and determine outcomes. Since the information that was obtained was so tiny, so used a variety of data augmentation techniques to expand the dataset and improve prediction. The core components of the suggested model are model architecture, tokenization, text preprocessing, and data augmentation. The mBERT model outperformed the other models in most performance metrics when compared to them, with the exception of F1 score before data augmentation, where Naive Bayes performed better with a score of 0.94, and precision and recall after data augmentation, where MLP performed better with a precision score of 0.97 and Naive Bayes with a recall score of 0.96. The addition of transformers, which are helpful for text categorization because to their self-attention mechanism and a better understanding of word characteristics, is largely responsible for the improved performance of the mBERT model.

In paper [3], the authors proposed the fake news detection system by using the algorithm tf-idf and passive aggressive classifier algorithm. The fake news is divided into non hostile, defamatory, offensive, and phony categories, and after working with the provided input sample, the user is provided with the final result. The ultimate accuracy of the suggested model is approximately 67.45%, which is higher than the BERT model's accuracy of 52% but lower than the LSTM model's acuracy of 94%. This is due to the dataset being compiled from many sources and own labeling efforts. The dataset's nature is significantly diverse, and because of this, the model doesn't receive enough training, which leads to a gradual decline in accuracy.

In paper [4], artificial intelligence techniques make up the suggested answer. Because detecting false news is difficult and requires dealing with texts, feature engineering and natural language processing approaches are employed. For classifying news stories, various machine learning and deep learning algorithms are applied. A dataset named "Hindi Fake and True Dataset" was created. News articles from various Hindi news channel websites which are BBC-Hindi1 and NDTV2. This dataset is tested on different machine learning and deep learning algorithms. The different algorithms used are Naive Bayes, logistic regression and LSTM. Naive Bayes classifier is performed on TF-IDF vectorizer. It gives 88.23% accuracy. Then logistic regression classifier is used on TF IDF vectorizer. This gives slightly optimized performance than earlier. The accuracy obtained is 89.15%. Thirdly, LSTM model is used which gives the best result of 92.36%. For this task of classifying fake news, LSTM algorithm gives the best accuracy of 92.36%.

In paper [5], three stages made up the methodology used in this investigation. This study used specific data filtering and cleaning procedures to extract in the first step (pre processing) the raw dataset's semantic characteristics. These comprised of stopword filter to sort the data into categories by deleting adverbs of place. In addition, HTML was used in this investigation tags and technology to filter out non English characters the data and eliminate contaminants that weren't helpful in classification. The following stage (extracting numerical data) methods were employed to convert the semantic characteristics into feature vectors, with classifiers being used as the last stage. Classifiers using machine learning and deep learning to divide the components of the dataset. Each of these techniques stood alone used on the identical dataset. The average accuracy of machine learning classifiers using tf-idf technology was 85.87%. The accuracy of machine learning using (count, n-gram and char vector) was obtained 87.34% for count vector, 83.6% for n gram level vector and 91.87% for character level vector. Accuracy of deep learning classifiers was obtained 93.52% for simple neural network, 90.12% for RNN + LSTM and approximately 100% for CNN + LSTM.

In paper [6], here compositionally incompetent method called web scrapping was also introduced, which gave us insight into how we can update our dataset on a regular basis to check the veracity of the recently updated Facebook posts. Bayes classifier was specifically used for fake news detection; we tested the difference in accuracy by taking different lengths of articles for detecting the fake news. The dataset used to evaluate the model's effectiveness was generated by GitHub and contains 11,000 news articles classified as authentic or fraudulent. It comprises 4 columns and 6335 rows. The four columns are label, text, title, and index. This dataset contains news categories for business, science and technology, entertainment, and health. The fact that this dataset was examined by journalists before being classified as "REAL" or "FAKE" lends credibility to its validity. The obtained AUC scores without n-grams was 0.806 for title and 0.912 for text. The obtained AUC score with n-gram was 0.807 for title and 0.931 for text.

In paper [7], proposed multilingual capsule network learning-based model with multilingual embeddings integrated with semantics infusion. There are several layers in the proposed architecture for effective multilanguage fake news detection. The semantic infusion is used for extra lexical semantics. Two parallel architectures for source and destination languages make up our paradigm. A Capsule Network and BiLSTM are utilized to extract contextual attributes and hierarchically positioned relationships, respectively. The effectiveness of our proposed model was demonstrated by experiments on TALLIP datasets, which showed that the classification performance had been significantly improved and outperformed state-of-the-art algorithms by about 3.97% for English to English, 1.41% for English to Hindi, 5.47% for English to Indonesian, 2.18% for English to Swahili, and 2.88% for English to Vietnamese languages. According to the experimental results on these two well-known datasets, accuracy increased by 7.8% on the ISOT dataset, 3.1% on the validation set, and 1% on the test set.

In paper [8], the classification model has been BERT with an LSTM layer. In order to group the news titles. To do this, a classification model called BERT with an LSTM layer is used. It acquires contextualized word representations by using a large number of unlabelled text corpora. Due to its intricate structure and significant nonlinear representation learning capacity The NLP tests were successful for BERT. With memorization and finding the essential information's pattern, LSTMs effectively improved performance. The proposed model achieved maximum accuracy of 88.75% as compared to other models. The proposed model got an increment of a minimum of 1.35% and a maximum of 17.55% accuracy for baseline models in the PolitiFact dataset. Also, an increment of a minimum of 0.3% and a maximum of 10.5% is seen in accuracy when compared to baseline models in the GossipCop dataset7(a-b) shows a pictorial representation of the evaluated metrics over PolitiFact and GossipCop respectively.

4. DATASET

The Hindi News Fake and True dataset is a meticulously curated collection of news articles in the Hindi language, classified into two categories: fake and true. This dataset is designed to aid researchers in developing and evaluating machine learning models for fake news detection in Hindi, providing a valuable resource for advancing natural language processing (NLP) in regional languages. The articles were sourced from various Hindi news websites, social media platforms, and fact-checking portals. Each article was manually labeled based on thorough verification against credible sources and fact-checking agencies. A team of bilingual (Hindi-English) annotators ensured the accuracy and reliability of the labels. The dataset is tailored specifically for the Hindi language, addressing a critical gap in the availability of regional language datasets for fake news detection. The articles encompass a broad range of subjects, including politics, health, technology, entertainment, and more, providing a rich context for model training. Efforts have been made to balance the number of fake and true articles to ensure effective model training and evaluation. The dataset is available for download and can be accessed via the provided link (“https://github.com/Siddhartha15/Hindi-Fake-News-Detection/tree/main/Data”). Researchers and practitioners are encouraged to use the dataset in compliance with ethical guidelines and cite the source in their publications.

Dataset	Contains
Total Articles	7111
Number of Columns	2(Articles, Labels)
Total number of true news articles	6613
Total number of fake news articles	498

5. MODEL PIPELINE

The process begins with the raw text itself. This could be from any source, whether a website, social media post, or even an email. First, the system cleanses the text. This might involve breaking it down into smaller units like words or even characters, a process called tokenization. Normalization follows, which could include converting all letters to lowercase, removing punctuation, or transforming words to their root form (stemming). This initial step ensures the model focuses on the core meaning of the text and avoids getting hung up on stylistic variations.

Next, each word in the cleaned text is transformed into a numerical code. Imagine this as assigning a unique fingerprint to each word based on its meaning and connection to other words. This is achieved through word embedding techniques like Word2Vec. By converting words into numerical vectors, the system creates a format that the machine learning model can understand and manipulate. With the text transformed into a sequence of numbers, it's fed into a Convolutional Neural Network (CNN) layer. Think of the CNN as a filter, meticulously scanning the sequence to identify patterns. These patterns could be frequent sequences of words, like "breaking news" or "shocking discovery," often used to grab attention in sensationalized headlines. The CNN might also identify patterns in how different types of words (nouns, verbs, adjectives) are used together, or flag the presence of overly emotional language, which can be a red flag for fake news. After the CNN extracts these features, a pooling layer steps in to simplify the data. This helps improve processing speed and prevents overfitting, a situation where the model memorizes specific patterns from the training data but struggles to analyze new information.

Now, the processed data enters one or more Long Short-Term Memory (LSTM) layers. LSTMs excel at handling sequential data like text. Unlike the CNN, which focuses on localized patterns, LSTMs can capture long-term dependencies between words. For instance, an LSTM might learn that the phrase "according to a recent study" often precedes fabricated claims in fake news articles. This ability to grasp the broader context within a sentence is crucial for accurate fake news detection.

Finally, the system reaches the output stage. The information extracted by the LSTM is fed into a final layer, a dense layer, which combines it all to generate a single probability score. This score represents the likelihood of the news article being fake. A score closer to 1 indicates a high probability of deception, while a score closer to 0 suggests the article is likely legitimate.

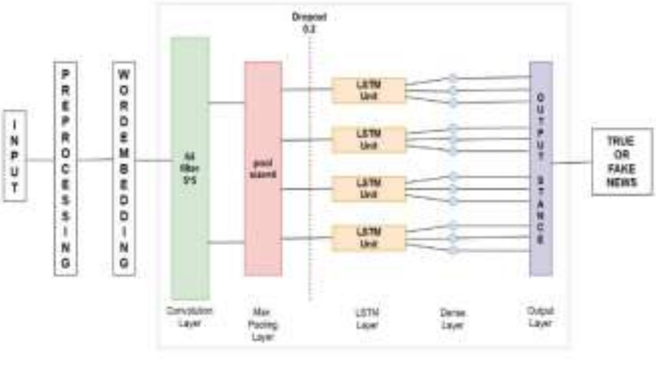


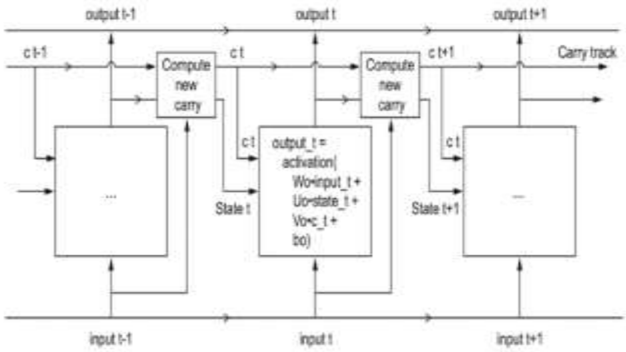
Fig. 1: Flow of Proposed system

6. Proposed Approach

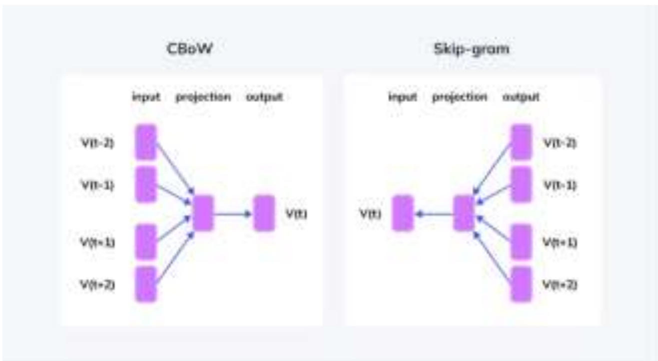
It takes a lot of trial-and-error utilizing machine learning algorithms on a variety of datasets to do research on fake news identification. Innovative methods must gain a thorough grasp of the characteristics of fraudulent news and the methods by which it travels throughout the globe. By putting out a model based on cutting-edge methods that demonstrate the significance of deep learning models for the false news detection challenge, the current work makes a contribution in this field. More specifically, it presents a convolutional neural network (CNN) and recurrent neural network (RNN) with the combination of Word-2-Vec embedding model that improved the suggested fake news detection model's performance. Explore various computational approaches to developing a robust fake news detection system specifically for the Hindi language. We delve into advanced machine learning and deep learning techniques, including Hybrid CNN-RNN, RNN, CNN, Logistic Regression, and Multinomial Naive Bayes. Each of these methodologies offers distinct advantages in processing and analysing textual data, enabling the detection of fake news with varying degrees of accuracy and efficiency. By leveraging these techniques, we aim to create a comprehensive framework that can effectively identify and mitigate the spread of fake news in Hindi, contributing to a more informed and discerning digital community.

I. *Hybrid CNN-RNN:*

A hybrid CNN-RNN model leverages the strengths of both Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to effectively handle complex data with both spatial and temporal dimensions. This approach begins by taking input data, such as sequences of images, frames, or text. For text, CNNs first capture local features and patterns by processing word embeddings through convolutional layers. For images or frames, the input data passes through CNN layers that apply convolutional filters to extract spatial features and perform pooling operations to reduce dimensionality. These spatial features are then flattened into a 1D feature vector, suitable for sequential processing. The flattened features are fed into RNN layers, composed of Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) units, which capture temporal dependencies and contextual information by processing the sequence of feature vectors over time. Each RNN unit takes the current input and the previous state to compute the current state and output, passing this information to the next time step. Finally, the output from the RNN layers is passed through dense (fully connected) layers to generate the final output, which could be classification labels, probabilities, or continuous values, depending on the application. This combination allows the model to learn both syntactic and sequential patterns, making it particularly useful for longer texts and complex sequential data. However, it is computationally intensive and requires substantial data for effective training.



II. *Word-2-Vec:*



In the context of fake news detection for Hindi news, Word2Vec embeddings are instrumental in transforming text data into numerical vectors that encapsulate the semantic meaning of words. The process begins with the collection of a substantial corpus of Hindi news articles, encompassing both genuine and fake news. This text data undergoes preprocessing steps such as tokenization into individual words, removal of stop words, and other text cleaning procedures specific to the Hindi language. A Word2Vec model is then trained on this corpus using either the Continuous Bag of Words (CBOW) or Skip-gram approach. CBOW predicts a word based on its surrounding context, while Skip-gram predicts context words from a given word. This training results in word embeddings, where each Hindi word is represented by a dense vector in a multi-dimensional space, capturing its semantic relationships with other words. Once the Word2Vec model is trained, each news article is converted into a sequence of word vectors. These vectors are aggregated to form a single vector representation for each article, using methods like averaging the word vectors or applying attention mechanisms to emphasize more significant words. These aggregated vectors serve as input features for a machine learning classifier, which could be a Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) network, or traditional classifiers like Support Vector Machines (SVM) or Logistic Regression. The classifier is trained on a labelled dataset of real and fake Hindi news articles, learning to identify patterns and features that distinguish fake news. By leveraging Word2Vec, the classifier benefits from a rich, semantic understanding of the text, enabling it to make more informed distinctions between real and fake news. This process not only enhances the detection accuracy but also efficiently handles improving the high-dimensional text data by converting it into meaningful, lower-dimensional vectors. Thus, Word2Vec embeddings play a crucial role in robustness and effectiveness of fake news detection systems for Hindi news.

III. *RNN:*

RNNs alone can also be used for fake news detection, as they are designed to handle sequential data by maintaining a hidden state that captures information about previous words in the sequence. By processing word embeddings through LSTM or GRU layers, RNNs capture temporal dependencies and context, making them suitable for tasks that involve understanding the sequence of words. Although RNNs are good at handling sequential data, they can suffer from vanishing gradient problems and are resource-intensive.

IV. *CNN:*

CNNs, although primarily used for image processing, can also be effective for text classification tasks by capturing local patterns in the text. In this approach, word embeddings are processed through convolutional layers with various filter sizes to capture different n-gram features. The output is then passed through max-pooling layers to down-sample the feature maps, followed by flattening and dense layers for classification. While CNNs are less computationally intensive than RNNs and are effective at capturing local dependencies, they might not be as adept at capturing long-range dependencies.

V. *Logistic Regression:*

Logistic Regression offers a simpler and more interpretable approach. Here, the text is vectorized using methods like TF-IDF or word embeddings, and important features are selected. The logistic regression model is then trained to predict the probability of a news article being fake or real based on these features. Despite its simplicity and speed, logistic regression is limited to linear relationships and may not capture complex patterns in the data.

VI. *Multinomial Naïve Bayes:*

Multinomial Naive Bayes is a probabilistic classifier based on Bayes' Theorem, assuming feature independence given the class label. The text is vectorized using methods like TF-IDF, and the model calculates the probability of a document belonging to each class based on word frequency. Although it is simple, fast, and effective for large-scale text classification, the assumption of feature independence is often unrealistic for text

7. SYSTEM PERFROMANCE MESUREMENT

Accuracy (Acc)**:**

Accuracy is a metric used to evaluate the performance of a classification model. It represents the proportion of correctly classified instances (both true positives and true negatives) out of the total instances in the dataset. In other words, accuracy measures how often the classifier makes the correct prediction.

The formula for calculating accuracy is:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision (Pre)**:**

Precision is a metric used to evaluate the accuracy of positive predictions made by a classification model. It represents the proportion of true positive instances out of all instances that were predicted as positive.

$$\text{Precision} = \frac{TP}{TP+FP}$$

Recall (Rc)**:**

Recall, also known as sensitivity or true positive rate, is a metric used to evaluate the ability of a classification model to identify all relevant instances of the positive class. It represents the proportion of true positive instances out of all actual positive instances in the dataset.

$$\text{Recall} = \frac{TP}{TP+FN}$$

FAKE NEWS DETECTION MODEL	ACCURACY	PRECISION	RECALL	F1-SCORE
HYBRID CNN-RNN	0.931	0.928	0.945	0.936
RNN	0.912	0.926	0.917	0.922
CNN	0.670	0.680	0.780	0.720
LOGISTIC REGRESSION	0.683	0.686	0.783	0.730
MULTINOMIAL NAÏVE BAYES	0.760	0.740	0.880	0.800

Table. System Performance Measurements

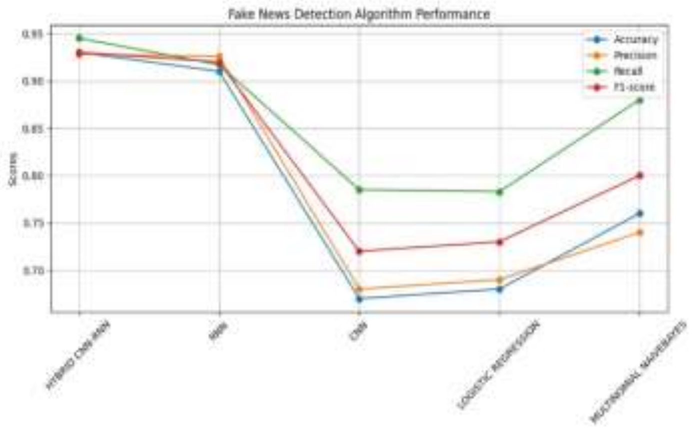


Table. Graph of Performance Measures

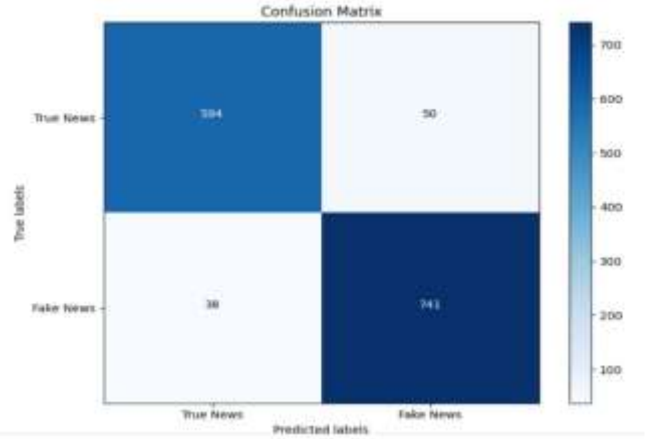


Table. Confusion Matrix

8.CONCLUSION

Manually classifying news stories calls for in-depth knowledge and proficiency in spotting irregularities. Because it takes a while to manually check a single item, we have discussed the problem of classifying bogus news articles using machine learning models and ensemble techniques.We need to figure out how to spot false news, or at the very least, be aware that not everything we read on social media is accurate. Therefore, critical thinking must be used constantly. By doing this, we may help the general public form wiser judgements and shield them from those who try to brainwash them.

9. REFERENCES

[1] Enjoy Maity, Ankush Tomar, Ruhi Peter, “Fake news detection system: In Hindi Dataset Using BERT”, International Research Journal Of Modernization In Engineering Technology And Science, vol.5, 5 May 2022.

[2] Kausthub Thekke Madathil, Meeraj mirji, Charan R and Anand Kumar M, “Fake news detection for hindi language”, The international conference and workshop on Agglutinative language technologies, 2021.

[3] Noor Mohd., Kireet Joshi, Rohan Raj, Dibyahash Bordoloi, “Fake news detection in hindi news using a passive classifier”, webology, vol.18, 2021.

[4] Sudhanshu Kumar, Thoudam Doren Singh, “Fake news detection on hindi news dataset”, Global transition proceedings, 2022.

[5] Awf Abdulrahman, Muhammet baykara, “Fake detection using machine learning and deep learning algorithms”, International conference on advanced science and engineering,IEEE 2020.

[6] Akshay jain, Amey Kasbe, “Fake news detection”, International students conference on electrical, electronics and computer science IEEE 2018.

[7] Rami Mohawesh, Sumbal Maqsood, Qutaibah Althebyan, “Multilingual deep learning framework for fake news detection using capsule neural network”, Journal of intelligent information systems 2023.

[8] Jamal Abdul Nasir, Osama Subhani Khan, Iraklis Varlamis, “Fake news detection: A hybrid CNN RNN based deep learning approach”, International journal of information management data insights, 2021