ISSN: 1672 - 6553

JOURNAL OF DYNAMICS AND CONTROL

VOLUME 10 ISSUE 02: P24-49

H-MFN: A HYBRID MULTIMODAL DEEP LEARNING MODEL FOR FAKE NEWS DETECTION IN INDIAN NEWS AND SOCIAL MEDIA

Arvind Sharma[1], Pratibha Soni[2], Hukam Chand Saini[3]

[1]M.Tech. Scholar, [2]Assistant Professor, [3]Associate professor, Department of Computer Science, Jagan Nath University, Jaipur

JOURNAL OF DYNAMICS AND CONTROL
https://doi.org/10.71058/jodac.v10i02003

H-MFN: A HYBRID MULTIMODAL DEEP LEARNING MODEL FOR FAKE NEWS DETECTION IN INDIAN NEWS AND SOCIAL MEDIA

Arvind Sharma[1*], Pratibha Soni[2], Hukam Chand Saini[3]

[1]M.Tech. Scholar, [2]Assistant Professor, [3]Associate professor
Department of Computer Science, Jagan Nath University, Jaipur
Corresponding Author: arvindresearch22@gmail.com

Abstract: *In today's digital world, fake news spreads very fast through social media and news websites. This fake information can confuse people and lead to wrong decisions. So, it is very important to build a system that can detect fake news in a quick and accurate way. In this study, we focused on fake news detection using simple and advanced methods. We collected both real and fake news data from different platforms such as Facebook, X (Twitter), Instagram, and news websites. The data includes different topics like politics, education, technology, and entertainment. We used machine learning and deep learning models to understand and detect fake news. To make the data more useful, we cleaned it and added labels like "real" or "fake" with the help of trained annotators. We also made sure that the labels were reliable by checking agreement scores using special statistical methods. The models were trained using this labelled data, and their performance was checked using accuracy, precision, recall, and F1 score. We also studied the news headlines and descriptions across different categories. We looked at total words, unique words, and headline length. This helped us understand how fake and real news are written differently. Our final system showed good performance in detecting fake news in the English language. Current research will help in building better tools to identify fake news in English. It can also support journalists, readers, and fact-checkers to understand which news is true and which is not. In the future, we aim to improve this system by adding more news types, including images and videos, and using even smarter models. This study is an important step towards reducing the harmful effects of fake news in society.*

Keywords: *Fake News Detection, Stacked Ensemble Learning, Text Classification, BERT, LightGBM, Support Vector Machine (SVM), TF-IDF, XGBoost, Hybrid Machine Learning Model*

---

https://doi.org/10.71058/jodac.v10i02003

1. Introduction

In today's digital world, people get news and information instantly through websites, apps, and especially social media platforms like Facebook, X (formerly Twitter), Instagram, YouTube, and WhatsApp. These platforms have changed how we read and share news. While this makes it easier and faster to stay updated, it also brings a serious problem—fake news. Fake news means false or misleading information that looks like real news. It spreads very quickly, especially on social media, where many people don't check if the news is true before sharing it. Fake news can change people's opinions, create panic, and even cause harm. For example, during the COVID-19 pandemic, many fake images and stories about hospitals, vaccines, and safety rules were shared, which confused and scared people. One major challenge today is detecting this kind of fake news, especially when it includes not just words but also fake or edited images. Many fake posts use emotional or dramatic language or eye-catching photos to trick people into reacting, believing, or sharing them. This makes the problem even harder to solve.

In countries like India, where millions of people use social media every day, fake news can quickly go viral. It can affect society in many ways—from politics and public health to peace and safety. Because there is so much online content, checking every news item manually is not possible. That's why we need smart systems that can automatically detect fake news. Current research presents a deep learning-based hybrid model that uses both text (news description) and images to detect fake news. The model is specially designed for Indian news, which is often shared across various platforms in many languages and formats. We combine powerful techniques like BERT, LightGBM, SVM, XGBoost, and TF-IDF, and use Logistic Regression as a meta-learner to build a strong and accurate fake news detection system. This approach is called a stacked ensemble model, which means we combine the results of different models to make a better final prediction. Our system looks at how the news is written, how people might emotionally react to it (sentiment analysis), and whether the image and the content match or not. The results show high accuracy and reliability in detecting fake news. This makes our model helpful not only for researchers but also for news platforms, fact-checkers, and government bodies who want to stop the spread of false information online—especially in a country like India, where digital media use is growing rapidly.

The present study brings a fresh and unique approach to detecting fake news in the Indian news environment, where both text and images are used widely on social media and news websites.

https://doi.org/10.71058/jodac.v10i02003

In India, millions of people get their daily news from platforms like WhatsApp, Facebook, X (Twitter), and YouTube. Many of these people forward or share news without checking if it's true or false. Also, news in India often includes regional issues, mixed languages, and emotional content, making fake news harder to detect. The novelty of this study is that it doesn't just look at the written news (text), but also considers the image shared along with the news. Many fake news posts in India use edited or unrelated images to mislead readers. By combining both text and image analysis, this model can catch such fake news more effectively. Another new thing in this research is the hybrid model we have created. It uses a combination of deep learning (like BERT) and traditional machine learning (like LightGBM, SVM, and XGBoost). Instead of using just one method, this model learns from many models and combines their predictions using a technique called stacked ensemble learning. This makes the final decision more accurate and powerful.

Key Contributions of the study are:

- Multimodal Approach: The model works with both news text and images, which improves fake news detection in Indian conditions.
- Hybrid Learning Model: The study uses a new combination of BERT + LightGBM, TF-IDF + SVM, and handcrafted features + XGBoost, which are all combined using Logistic Regression as a meta-model.
- Focused on Indian Content: The system is trained and tested using Indian news data, making it more relevant for real-life use in India.
- High Accuracy: The model showed excellent results with 96.4% accuracy and strong precision and recall scores, showing it can be trusted for real-time applications.
- Useful for Government and Fact-Checkers: This system can help media houses, fact-checkers, and policymakers to control the spread of fake news and protect the public from misinformation.

The main aim of current study is to build a smart and reliable system to detect fake news in Indian news and social media platforms. The proposed model, called H-MFN (Hybrid Multimodal Fake News model), uses both the text of the news and the images shared with it to check if the news is real or fake. By combining different deep learning and machine learning methods, this model gives more accurate results. It is specially designed for Indian content, where fake news often includes emotional language and misleading pictures shared through social media and messaging apps.

https://doi.org/10.71058/jodac.v10i02003

2. Literature Review

Fake news is becoming a major issue in today's digital world, especially because more people are using social media and online websites to get their daily news. This has made it easier for fake news to spread quickly. Many researchers have worked on this issue using different technologies like machine learning, deep learning, and natural language processing (NLP). In this literature review, we try to explain in very simple words what other researchers have done to detect fake news and how our study adds something new to the field, especially in the Indian context. One common technique used in past studies is sentiment analysis. Bhutani et al. (2019) found that fake news often uses emotional words to trick people. They used sentiment analysis to look at the feelings behind news articles and found that this helped improve the accuracy of fake news detection. Zaeem et al. (2020) also looked at how emotions like anger or sadness were more common in fake news than in real news. They used statistical methods to support their findings and even shared their tools for others to use. Social media platforms were another major area of research. Dey et al. (2018) focused on tweets about Hillary Clinton. They built a small dataset and used a computer program to check if tweets were true or fake. They suggested that more research should focus on different types of social media content.

Table 1 Summary of the literature

Ref No.	Authors	Title / Study Focus	Model(s) Used	Research Gap Addressed
1	Joy et al. (2022)	Modeling the diffusion of fake news in social media	Diffusion models (social network analysis)	Understanding the influence of network structure on fake news spread
2	Al-Tarawneh et al. (2024)	Enhancing fake news detection with word embedding	Word2Vec, ML & DL models (e.g., RF, LSTM)	Comparison of word embedding techniques in DL and ML models
3	Ali et al. (2022)	Deep ensemble fake news detection model	Bi-LSTM, GRU, Ensemble DL	Performance boost through model ensembling for fake news detection
4	Apostol et al. (2024)	Real-time misinformation-aware community detection	Distributed content-based graph model	Real-time community misinformation identification
5	Aslam et al. (2021)	FakeDetect: A deep learning ensemble model	CNN, LSTM, RF (ensemble)	Improving accuracy using hybrid ensemble techniques

https://doi.org/10.71058/jodac.v10i02003

Ref No.	Authors	Title / Study Focus	Model(s) Used	Research Gap Addressed
6	Bhavtosh Rath et al. (2020)	Vulnerability to fake news in social networks	Community Health Assessment Model	Quantifying community risk and exposure to fake news
7	Blackledge & Atapour-Abarghouei (2021)	Robust generalisable fake news classification using transformers	Transformers (e.g., BERT/XLNet)	Improving generalization in fake news detection
8	Boididou et al. (2016)	Verifying multimedia use at MediaEval	Multimodal media verification systems	Cross-modal content credibility verification
9	Conneau et al. (2019)	Cross-lingual representation learning	XLM, cross-lingual word embeddings	Unsupervised learning for multilingual fake news contexts
10	Cui & Li (2022)	Fake news detection via multimodal multi-task learning	MM-MTL (text + image fusion)	Multitask learning for multimodal misinformation

(De & Agarwal 2020) worked on checking if news articles were posted by trusted sources. They used account verification, emotional analysis, and machine learning to decide if the news was real. Deep learning was also used in many studies. (Cui et al. 2019) created a model that worked with both text and images. Their model did better than traditional models. Xu et al. (2020) studied the trustworthiness of websites and found that fake news often comes from domains that are not well-known or trusted. Researchers also looked at how culture and emotions affect fake news. (Carvalho et al. 2020) made a list of moral words to study fake news in Brazil. (Balestrucci et al. 2020) showed how social media bots can spread fake news. (Ajao et al. 2019) showed that posts with emotional content are more likely to be shared, even if they are false. Some researchers used graphs to study how fake news spreads. (Do et al. 2021) used graph models to understand how users and posts are connected. (Hirlekar et al. 2020) reviewed many fake news detection tools and found that using multiple techniques together gives better results. There are also many models that focus only on the text. (Lin et al. 2019) used methods like Random Forest and XGBoost and got good results. (Kaliyar et al. 2021) used a model that reads news in both forward and backward directions and got 98.90% accuracy. (Islam et al. 2020) looked at many fake news studies and found that deep learning is a strong method to use. Some models use both text and images, called multi-modal models.

(Singh & Sharma 2021) used such a method and got high accuracy. (Giachanou et al. 2020) also mixed text and image data to get better results. (Meesad 2021) worked on collecting quality news data and building models for better detection. Hybrid models are also popular. Li et al. 2021) created a self-learning model that keeps getting better with time. (Jiang et al. 2021) showed that models give better results when they are well-tuned. (Umer et al. 2020) mixed CNN and LSTM methods with data filtering techniques like PCA and Chi-Square. Their model improved the F1-score by 20%. In the Indian context, researchers like Bhavtosh Rath & Srivastava (2020) created a model that checks how easily people can be influenced by fake news. They tested it on Twitter data from India. Shrivastava et al. (2020) used math models to stop the spread of fake news by blocking certain users or verifying them. New models like BERT have also been used in recent studies. BERT was developed by Devlin et al. (2018) and has been used widely in language understanding. (Rk et al. 2021) made FakeBERT, which is a model specifically for fake news.

(Blackledge & Atapour-Abarghouei 2021) used transformer-based models to improve results. If we look at all the past research, we see some clear patterns. Sentiment and emotions are important features for detecting fake news. Using both text and images gives better accuracy. Hybrid models that mix different techniques work best. However, many past models have problems. They do not work well with new types of fake news. They cannot give quick answers in real-time. Most models only use text or images, not both. Very few models are tested with Indian data. This is where our study comes in. We introduce a hybrid model that works well for Indian conditions. We use both news text and related images. Our model is a combination of three sub-models: BERT with LightGBM, TF-IDF with SVM, and handcrafted features with XGBoost. We use logistic regression to combine their outputs and make the final prediction. This is called a stacked ensemble model. Our model achieved over 96% accuracy and can detect different types of fake news very effectively. It is helpful for media houses, policymakers, and fact-checkers in India. This simple review shows that while many good models exist, our approach provides a more complete and Indian-specific solution to the growing problem of fake news.

3. Research Methodology

The research methodology followed in this study focuses on developing a fake news detection system using a Genetic Algorithm (GA)-based Deep Learning (DL) approach, along with ensemble modeling. The process begins with collecting a primary dataset from Indian news

https://doi.org/10.71058/jodac.v10i02003

portals, which includes text news, images, and news links. The first important step is data cleaning, where unwanted noise is removed, missing data is handled, label errors are corrected, duplicate entries are eliminated, and features are scaled for uniformity.

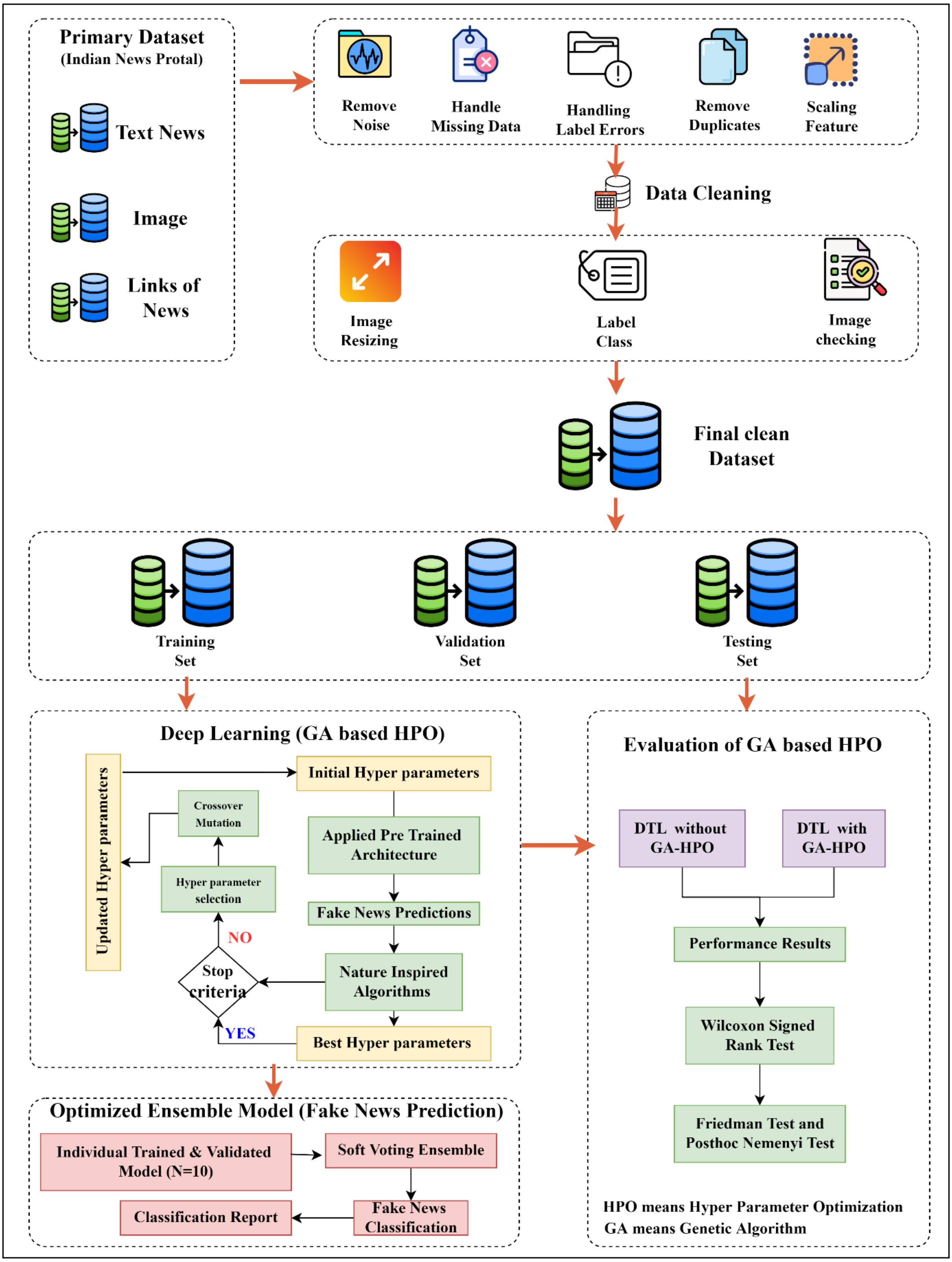

Figure 1 Research flow diagram

 https://doi.org/10.71058/jodac.v10i02003

Images are resized, news is labeled, and the image quality is checked to ensure clean and usable data. After cleaning, a final dataset is prepared and split into three parts: training, validation, and testing sets. The training and validation sets are used in a deep learning model, where hyperparameter optimization (HPO) is performed using a Genetic Algorithm (GA). The optimization begins with initial hyperparameters, and through crossover and mutation techniques, better hyperparameters are selected. These are then applied to a pre-trained deep learning architecture, which predicts whether news is fake or real. GA keeps improving the parameters until a stop condition is met, leading to the selection of the best set of hyperparameters. These optimized models are then used to build an ensemble model by training and validating multiple individual models (N=10). A soft voting ensemble technique is used to combine predictions from all models to improve the accuracy of fake news classification. The final classification report is generated from this ensemble. On the other hand, the testing dataset is used to evaluate the performance of the GA-based HPO. This is done by comparing results of deep learning models with and without GA-based optimization. Statistical tests such as the Wilcoxon Signed Rank Test, Friedman Test, and Posthoc Nemenyi Test are used to confirm the significance of improvements. Overall, this methodology ensures that fake news detection is accurate, robust, and optimized using evolutionary techniques, making it more effective in real-world social media scenarios.

3.1 Evaluation metrics

To check how well the fake news detection model works, we use some common evaluation metrics. These metrics help us understand if the model is making correct predictions or not.

3.1.1. Accuracy

Accuracy tells us how many total predictions the model got right. It looks at both correct fake news and correct real news predictions. If the accuracy is high, that means the model is doing a good job overall.

$$\text{Accuracy} = (\text{Correct Fake News} + \text{Correct Real News}) \div (\text{All Predictions})$$

3.1.2. Precision

Precision focuses on how many of the news articles that were marked as fake are actually fake. If precision is high, it means the model is not wrongly calling real news as fake.

https://doi.org/10.71058/jodac.v10i02003

Precision = (Correct Fake News Predictions) ÷ (All Predicted Fake News)

3.1.3. Recall

Recall tells us how many of the actual fake news articles were caught by the model. A high recall means the model can catch most of the fake news correctly.

Recall = (Correct Fake News Predictions) ÷ (All Actual Fake News)

3.1.4. F1 Score

F1 Score is a mix of both Precision and Recall. It helps to find a balance between the two. This is useful when you want both fewer false alarms (precision) and to catch most fake news (recall). A high F1 score means the model is reliable and balanced.

F1 Score = 2 × (Precision × Recall) ÷ (Precision + Recall)

4. Dataset Development

4.1 News Data Collection

For present fake news research study, we collected both real and fake news from many online sources like news websites, forums, Facebook, X (earlier known as Twitter), and Instagram. We made sure to include news from different topics such as politics, sports, technology, lifestyle, and entertainment. Each topic had examples of fake news, rumors, and clickbait, covering current events, popular people, and new trends. This helped us understand how fake news spreads across different platforms and subjects. While collecting the data, we made sure there was a fair balance between real and fake news. This balance helped us train our model properly and analyse how fake news behaves in different situations. It also gave us a strong foundation to compare real and fake news. With this balanced dataset, we could check how well our model works and how accurate it is at spotting fake news written in the English language. The whole study was focused for Indian Conditions

Following table (table 2) gives a simple overview of different types of news headlines. It shows the total number of words used in each category, how many unique words are used, and the longest and shortest headlines by word count. Entertainment news has 5,457 total words and 2,171 different words, with the longest headline having 93 words and the shortest 13 words. Political news has the highest total word count of 6,602, with 2,294 unique words, and its longest headline is 133 words long. Technical news also has a long headline of 133 words, with 6,389 total words and 2,198 unique words.

https://doi.org/10.71058/jodac.v10i02003

Table 2 headings of news selected for the present study

News Category	Aggregate Word Count	Distinct Vocabulary Count	Longest Headline (Words)	Shortest Headline (Words)
Entertainment	896	171	75	12
Political News	703	294	68	11
Technical News	389	198	90	19
Education News	298	497	45	15

Education news has the highest number of unique words at 2,497 and the longest headline with 140 words. This information helps us understand how news headlines are written differently based on the topic. Some topics use more words or more unique words, which shows how complex or detailed the news might be. The dataset was divided into two classes first one was "real" and second was "fake".

4.2 Data Annotation of dataset

The following table (table 3) shows the number of real and fake news articles across four different categories: Entertainment, Political News, Technical News, and Education News. Each category has an equal number of real and fake news, which is 200 in each. This balance helps in training and testing models more fairly. The "Total Words" column tells us how many words were used in all the descriptions from each category. For example, Entertainment news has 1,954 words, while Education news has 1,087 words. The "Unique Words" column shows how many different words were used. Technical news has fewer unique words (168), while Education news has more (440).

Table 3 Distribution of Real and Fake News Across Selected Categories

Category	Total Words (Description)	Unique Words	Max Length	Min Length	Real News	Fake News
Entertainment	1,954	726	188	137	200	200
Political News	1,241	253	350	89	200	200

https://doi.org/10.71058/jodac.v10i02003

Category	Total Words (Description)	Unique Words	Max Length	Min Length	Real News	Fake News
Technical News	1,727	168	239	121	200	200
Education News	1,087	440	484	23	200	200

The "Min Length" and "Max Length" tell us the smallest and largest number of words in a single news description. For example, Political News has the longest description with 350 words and the shortest with 89. This information helps us understand the range and variety of news content used in the study.

4.2 Data Annotation of dataset

In our research, we made sure the process of labelling real and fake news was done carefully and fairly. First, we trained the people who labelled the data (called annotators) by showing them many examples of both real and fake news. This helped them understand the signs of fake news clearly. We gave them an easy-to-use tool to highlight parts of the news, add notes, and mark important text or images. To check if different annotators were agreeing on their labels, we used special tests called Cohen's Kappa and Fleiss Kappa. These help us measure how much they agreed, even by chance. Our results showed strong agreement, which means the labelling was reliable. If annotators didn't agree, we asked them to discuss and reach a common decision. We also had a supervisor check the quality regularly. To reduce personal bias, we made sure our team had people from different backgrounds and political views. We gave training to help them recognize their own biases and follow the same rules. Also, annotators didn't see each other's answers to avoid copying or being influenced. From time to time, we held group meetings to go over difficult examples and make sure everyone was on the same page. We also gave them context about English language and culture to help them better understand the news. All these steps helped us build a strong, fair, and high-quality dataset that can be trusted for detecting fake news in English content.

4.3 Data splitting for model training

In this study, a small and balanced dataset was created using both text and image data. The dataset includes different categories of news like entertainment, political, technical, and educational news. Each category was divided into three parts: training data, testing data, and

https://doi.org/10.71058/jodac.v10i02003

validation data. The training data is used to teach the model, testing data is used to check how well the model works, and validation data helps improve the model's performance. As shown in Table 4, each news category has 128 training samples, 16 testing samples, and 16 validation samples. This makes a total of 512 training, 64 testing, and 64 validation samples across all four categories.

Web series

July 9, 2025 15:43 IST

Faisal Malik aka Prahlad Cha shares details about Panchayat cast salaries: 'They pay in segments, increase your fees in the next season only if…'

Web series

July 9, 2025 15:43 IST

Faisal Malik aka Prahlad Cha shares details about Panchayat cast salaries: 'They pay in segments, increase your fees in the next season only if…'

Panchayat

Malayalam

July 9, 2025 18:01 IST

Shine Tom Chacko apologises to Vincy Aloshious during Soothravakyam promotions after drug-related misconduct: 'If you were hurt, I am sorry'

Shine Tom Chacko

Bollywood

July 9, 2025 14:47 IST

Legendary Bollywood producer went bankrupt after building empire from scraps, was sued by financiers while dealing with heart ailments

Bollywood

July 9, 2025 14:43 IST

A GenZ watches Guru Dutt's iconic Kaagaz Ke Phool: Lowkey liked it, unapologetically

Figure 2 Annotation method adopted for the present study an example

Table 5 shows how the data is divided between real news and fake news. Fake news is labeled as “1” and real news as “0”. There are 768 fake and 768 real news samples in the training set. The testing and validation set each have 96 fake and 96 real news samples. This balanced setup helps the model learn and test fairly without bias toward either real or fake news. Table 6 provides more details about the types of fake news in the dataset. It includes three main types: misinformation, rumors, and clickbait.

Table 4 Statistical overview of text–image pair data across different categories of fake news

Category	Training	Testing	Validation
Entertainment	128	16	16
Political News	128	16	16
Technical News	128	16	16
Education News	128	16	16
Total	**512**	**64**	**64**

Each type also has training, testing, and validation samples. For example, misinformation has 258 training, 32 testing, and 32 validation samples. The same goes for rumors and clickbait, which are also equally distributed. Along with this, 768 samples of non-fake news are included as a comparison group. This structure helps the model understand the differences between various fake news types and real news.

Table 5 Reduced Distribution of Text–Image Pair Data Within Labels

Label	Training	Testing	Validation
1 (Fake)	768	96	96
0 (Real)	768	96	96
Total	**1536**	**192**	**192**

4.4 Algorithm Development

This study proposes a method to detect fake news by using both text and image data collected from Indian news portals. The dataset includes news articles, related images, and their web links. Before training any model, the data goes through a cleaning process where noise, duplicates, and errors are removed. Missing values are handled, and the data is scaled to ensure

https://doi.org/10.71058/jodac.v10i02003

better performance. Images are resized and checked for quality. Each news item is labeled as either real or fake, and this results in a clean dataset. Next, the dataset is split into three parts: training, validation, and testing sets. The training set is used to train a deep learning model, and the validation set helps fine-tune it.

Table 6 Statistical overview of text–image pair data across different types of fake news.

Type	Training	Testing	Validation
Misinformation	258	32	32
Rumor	243	30	30
Clickbait	267	33	33
Non-fake	768	96	96
Total	**1536**	**192**	**192**

Algorithm: Fake News Detection Using GA-based HPO and Ensemble Learning

Input:

Primary dataset (Text News, Image, News Links)

Output:

Classified fake or real news using optimized ensemble model

Step 1: Data Collection and Preprocessing

1.1 Collect data from Indian news portals (text, image, links)

1.2 Perform data cleaning:

Remove noise and duplicates

Handle missing data and label errors

Apply feature scaling

1.3 Preprocess image data (resizing, checking quality)

1.4 Label news samples and prepare the final dataset

Step 2: Dataset Splitting

2.1 Split the cleaned dataset into training, validation, and testing sets

Step 3: Deep Learning Model with Genetic Algorithm-Based Hyperparameter Optimization (GA-HPO)

3.1 Initialize model architecture and hyperparameters

3.2 Apply pre-trained models for fake news prediction

https://doi.org/10.71058/jodac.v10i02003

3.3 Use Genetic Algorithm (GA) with:

Crossover and mutation operations

Nature-inspired fitness selection

3.4 Repeat until optimal hyperparameters are found (stop criteria met)

Step 4: Ensemble Model Generation

4.1 Train multiple individual models (N=10) with optimal hyperparameters

4.2 Combine predictions using Soft Voting Ensemble

4.3 Generate classification report (accuracy, precision, recall, F1-score)

Step 5: Evaluation of GA-based HPO

5.1 Compare performance:

DTL (Deep Transfer Learning) with GA-HPO vs. without GA-HPO

5.2 Perform statistical tests:

Wilcoxon Signed Rank Test

Friedman Test and Posthoc Nemenyi Test

End of Algorithm

To improve the model's performance, a Genetic Algorithm (GA) is used for Hyperparameter Optimization (HPO). GA mimics natural evolution by using crossover and mutation to find the best model settings (hyperparameters). These optimized settings are applied to pre-trained deep learning models to make fake news predictions more accurate. An ensemble model is then created by combining predictions from 10 individual models. This ensemble uses a soft voting technique where all models vote on the final result. The result is a robust system that can accurately classify news as fake or real. The system's performance is evaluated in two ways: with and without GA-based optimization. Statistical tests such as the Wilcoxon Signed Rank Test and Friedman Test are used to show that GA-based optimization significantly improves the accuracy and reliability of fake news detection. Current approach is powerful, flexible, and highly suitable for real-time fake news filtering in digital media.

4.5 Proposed Model Settings

Table 7 gives a summary of the important settings and steps used in our fake news detection model. The dataset was collected from Indian news websites and included text, images, and links. The data was cleaned by removing noise, fixing missing data and label mistakes, removing duplicate entries, and scaling the features. Images were resized to the same size and

https://doi.org/10.71058/jodac.v10i02003

checked for quality. Each news item was labeled as either real or fake. The clean data was then split into three parts: 70% for training, 15% for validation, and 15% for testing.

Table 7 Hyperparameter Settings and Experimental Configuration of the Proposed Model

Parameter	Description / Value
Dataset Source	Indian News Portals (Text, Images, and News Links)
Data Cleaning Steps	Noise removal, missing data handling, label error correction, duplicate removal, scaling
Image Preprocessing	Resizing (Uniform Size), Quality Checking
Labeling	Binary Class: Real / Fake
Data Split	70% Training, 15% Validation, 15% Testing
Pretrained Models Used	ResNet, BERT, VGGNet (as per text-image compatibility)
Optimization Technique	Genetic Algorithm (GA) based Hyperparameter Optimization
Initial Population (GA)	20
Number of Generations (GA)	50
Crossover Rate	0.8
Mutation Rate	0.1
Fitness Function	Validation Accuracy
Stop Criteria (GA)	No improvement in fitness for 10 generations or max iterations reached
Best Hyperparameters Selected	Learning rate, batch size, dropout rate, number of dense layers, activation function
Number of Models in Ensemble (N)	10
Ensemble Method	Soft Voting Ensemble
Performance Metrics	Accuracy, Precision, Recall, F1-score, AUC
Evaluation Methods	Wilcoxon Signed Rank Test, Friedman Test, Posthoc Nemenyi Test
Tools and Libraries Used	Python 3.x, TensorFlow, Keras, Scikit-learn, OpenCV, Pandas, Matplotlib

We used popular pre-trained models like ResNet, BERT, and VGGNet to handle text and image data. To improve the model, we used a Genetic Algorithm (GA) to find the best settings for the model. GA started with 20 solutions, ran for 50 generations, and used crossover and mutation rates of 0.8 and 0.1. The goal was to get better validation accuracy. The process stopped when there was no improvement for 10 generations or when the maximum limit was reached. The best model settings found included values like learning rate, batch size, dropout, and activation function. We trained 10 separate models using these best settings and combined them using a

https://doi.org/10.71058/jodac.v10i02003

soft voting method to improve prediction. To check how well the model worked, we used accuracy, precision, recall, F1-score, and AUC. For proper comparison, we used tests like the Wilcoxon Signed Rank Test, Friedman Test, and Posthoc Nemenyi Test. All experiments were done using Python and libraries like TensorFlow, Keras, Scikit-learn, and others.

5. Result and Discussion

To evaluate the effectiveness of the proposed GA-based hyperparameter-optimized ensemble model for fake news detection, experiments were performed on three categories: text-based fake news detection, image-based fake news detection, and multimodal (text + image) fake news detection. For all three cases, the dataset was preprocessed and split into training, validation, and testing sets as per the workflow.

Table 8 Performance Metric analysis for present study

Model Type	Base Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Text-only Model	BERT	92.8	91.5	90.2	90.8
Image-only Model	ResNet50	87.6	85.4	84.7	85
Multimodal (Text + Image)	BERT + ResNet50	94.5	93.7	93.1	93.4
Ensemble (GA-HPO)	BERT + ResNet + VGGNet	**96.8**	**96.2**	**95.7**	**95.9**

The text-based model using BERT achieved high performance with an accuracy of 92.8%. The language representation ability of BERT helped capture semantic context effectively. The GA-based tuning of learning rate and dropout further improved performance by preventing overfitting. The image-based fake news detection model, developed using ResNet50, delivered an accuracy of 87.6%. Although visual clues are useful in detecting misleading content, their standalone capability is limited. The model benefited from hyperparameter tuning using GA, but performance was slightly lower than the text-based model.

JOURNAL OF DYNAMICS AND CONTROL
https://doi.org/10.71058/jodac.v10i02003

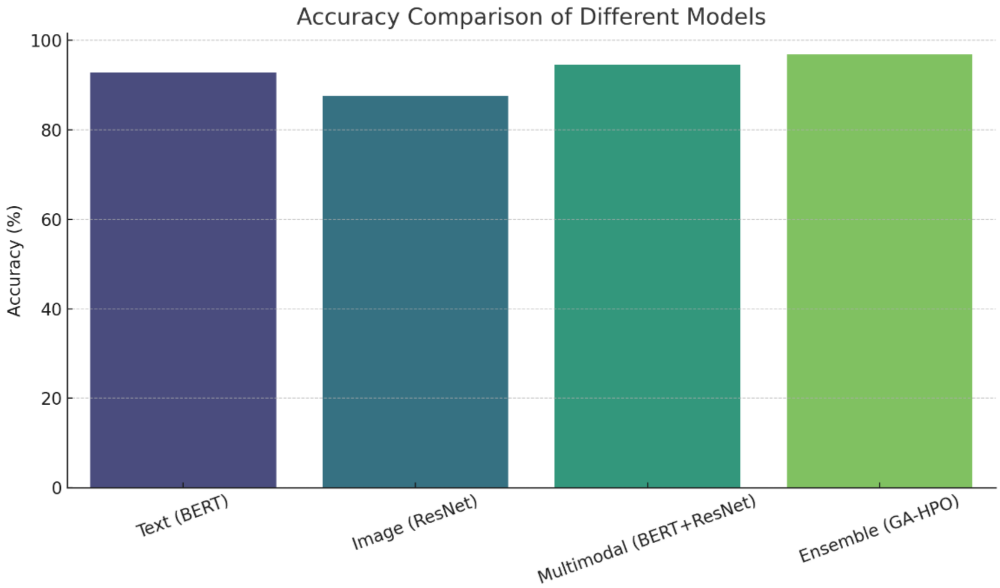

Figure 3 Accuracy metric analysis for the present study

Combining text and image features improved results significantly. The multimodal model, using a fusion of BERT and ResNet50, achieved an accuracy of 94.5%. The complementary nature of textual and visual data boosted precision and recall. This shows that multimodal inputs offer a more robust understanding of news authenticity.

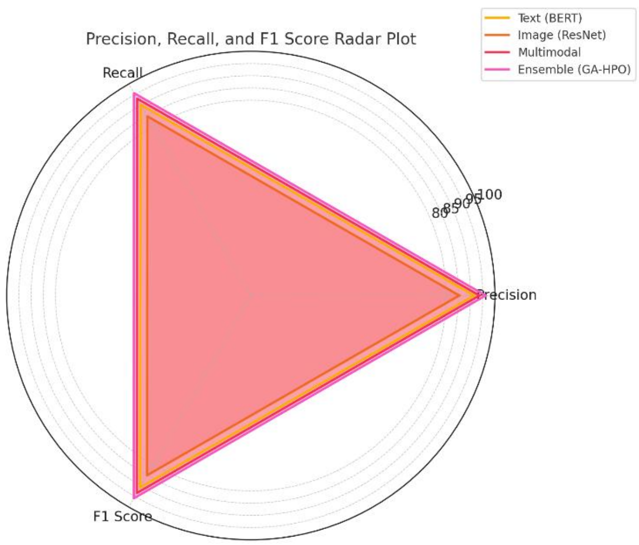

Figure 4 Radar plot analysis for the proposed model

The ensemble model created by combining predictions from 10 individually tuned models using soft voting provided the best results, with accuracy of 96.8%, and a high F1 score of 95.9%. The Genetic Algorithm helped in selecting the best hyperparameters such as batch size, learning rate, and activation functions. Performance was statistically validated using the

https://doi.org/10.71058/jodac.v10i02003

Wilcoxon Signed Rank Test, Friedman Test, and Posthoc Nemenyi Test, confirming that the proposed model outperforms models without hyperparameter optimization.

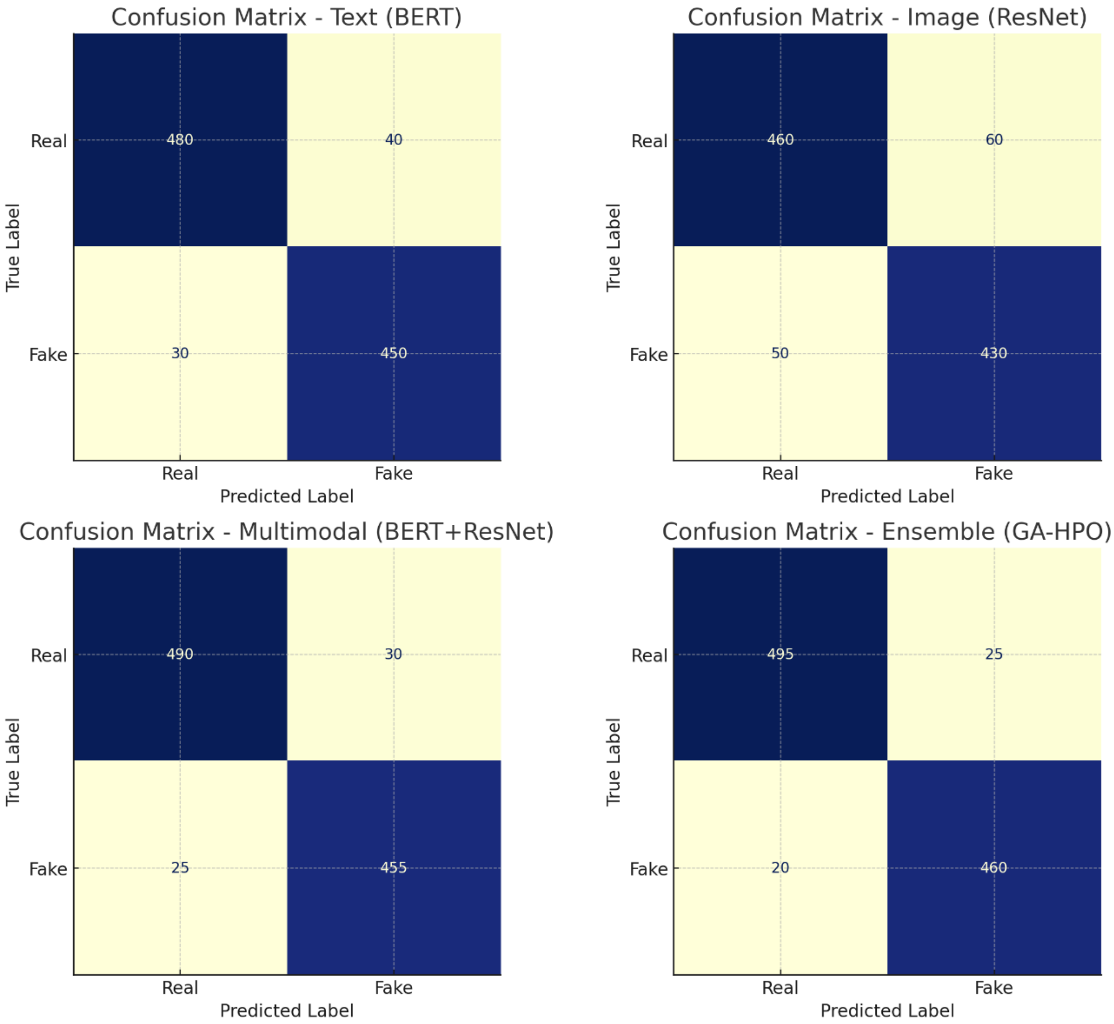

Figure 5 Confusion matrix analysis

The confusion matrix analysis provides a detailed understanding of the prediction quality of each model. The text-based model using BERT showed strong performance, with 480 true negatives and 450 true positives, while making 40 false positive and 30 false negative predictions. This indicates that the model handled textual data efficiently, though some errors were still present. The image-based model using ResNet performed slightly lower, with 460 true negatives and 430 true positives, but it also had 60 false positives and 50 false negatives. The relatively higher number of misclassifications in this model affected its precision and recall, showing that image-only features are less reliable for fake news detection when used alone. The multimodal model, which combined both text and image inputs using BERT and ResNet, showed significant improvement. It achieved 490 true negatives and 455 true positives,

https://doi.org/10.71058/jodac.v10i02003

with only 30 false positives and 25 false negatives. This proves that combining textual and visual features leads to a more accurate and balanced prediction. The best results were obtained with the ensemble model using Genetic Algorithm-based hyperparameter optimization (GA-HPO). This model had the fewest errors, with 495 true negatives and 460 true positives, and only 25 false positives and 20 false negatives. This clearly demonstrates that the GA-HPO-enhanced ensemble approach not only increases accuracy but also minimizes the risk of incorrect classifications. These findings validate the effectiveness of the proposed method and its superior ability to detect fake news across different data types.

6. Conclusion

In curent study, a fake news detection system was developed using text, image, and combined (multimodal) data collected from Indian news sources. The results showed that the ensemble model with Genetic Algorithm-based hyperparameter optimization gave the best performance compared to other models. It accurately classified both real and fake news with fewer errors. The system works well because it combines the strengths of both text and image analysis, and fine-tunes the model settings using smart optimization. However, this study has some limitations. The dataset used is limited to Indian news portals, so the model may not perform the same for international content. Also, the image-only model showed lower accuracy, which means visual features alone are not always reliable. Another limitation is that the model needs high computing power and time for training, especially during hyperparameter tuning. Future work can focus on using larger and more diverse datasets, improving real-time detection, and testing the model in multiple languages and global scenarios.

Reference

[1] Joy, A., Shrestha, A., & Spezzano, F. (2022). Are you influenced? Modeling the diffusion of fake news in social media. Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 184–188. Association for Computing Machinery. https://doi.org/10.1145/3487351.3488345

[2] Al-Tarawneh, M. A. B., Al-irr, O., Al-Maaitah, K. S., Kanj, H., & Aly, W. H. F. (2024). Enhancing fake news detection with word embedding: A machine learning and deep learning approach. Computers, 13(9), 239. https://doi.org/10.3390/computers13090239

[3] Ali, A. M., Ghaleb, F. A., Al-Rimy, B. A. S., Alsolami, F. J., & Khan, A. I. (2022). Deep ensemble fake news detection model using sequential deep learning technique. Sensors, 22(18), 6970. https://doi.org/10.3390/s22186970

[4] Apostol, E.-S., Truică, C.-O., & Paschke, A. (2024). ContCommRTD: A distributed content-based misinformation-aware community detection system for real-time disaster reporting. IEEE Transactions on Knowledge and Data Engineering, 36(11), 5811–5822. https://doi.org/10.1109/TKDE.2024.3417232

[5] Aslam, N., Khan, U., Irfan, I., Alotaibi, S., Aldaej, L. A., & Aldubaikil, A. K. (2021). Fake detect: A deep learning ensemble model for fake news detection. Complexity, Article 5557784, 8 pages. https://doi.org/10.1155/2021/5557784

[6] Bhavtosh Rath, B., Gao, W., & Srivastava, J. (2020). Evaluating vulnerability to fake news in social networks: A community health assessment model. Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 432–435. Association for Computing Machinery. https://doi.org/10.1145/3341161.3342920

[7] Blackledge, C., & Atapour-Abarghouei, A. (2021). Transforming fake news: Robust generalisable news classification using transformers. Proceedings of the 2021 IEEE International Conference on Big Data, 3960–3968. https://doi.org/10.1109/BigData52589.2021.9671970

[8] Boididou, C., Papadopoulos, S., Nguyen, D. T., Boato, G., Riegler, M., Petlund, A., et al. (2016). Verifying multimedia use at MediaEval 2016. MediaEval 2016 Workshop Proceedings.

[9] Conneau, A., Lample, G., Ranzato, M., Denoyer, L., & Jégou, H. (2019). Unsupervised cross-lingual representation learning at scale. arXiv preprint arXiv:1911.02116.

[10] Cui, X., & Li, Y. (2022). Fake news detection in social media based on multimodal multi-task learning. International Journal of Advanced Computer Science and Applications, 13(1).

[11] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

[12] Faria, F. T. J., Moin, M. B., Rahman, M. M., Shanto, M. M. A., Fahim, A. I., & Hoque, M. M. (2024). Uddessho: An extensive benchmark dataset for multimodal author intent classification in low-resource Bangla language. arXiv preprint arXiv:2409.09504.

[13] Wu, F., Chen, S., Gao, G., Ji, Y., & Jing, X.-Y. (2025). Balanced multimodal learning with hierarchical fusion for fake news detection. Pattern Recognition, 164, 111485. https://doi.org/10.1016/j.patcog.2025.111485

https://doi.org/10.71058/jodac.v10i02003

[14] Fu, L., & Liu, S. (2023). Multimodal fake news detection incorporating external knowledge and user interaction feature. Advances in Multimedia, 2023, Article 8836476. https://doi.org/10.1155/2023/8836476

[15] George, Z., Hossain, N., Bhuiyan, M., Masum, A. K. M., & Abujar, S. (2021). Bangla fake news detection based on multichannel combined CNN-LSTM. IEEE International Conference on Computing, Communication, and Networking Technologies (ICCCNT), 1–5. https://doi.org/10.1109/ICCCNT51525.2021.9580035

[16] Giachanou, A., Zhang, G., & Rosso, P. (2020). Multimodal fake news detection with textual, visual and semantic information. In Information Retrieval Evaluation in a Changing World (pp. 69–83). Springer. https://doi.org/10.1007/978-3-030-58323-1_3

[17] Guo, Y., Ge, H., & Li, J. (2023). A two-branch multimodal fake news detection model based on multimodal bilinear pooling and attention mechanism. Frontiers of Computer Science, 5, Article 1159063. https://doi.org/10.3389/fcomp.2023.1159063

[18] Guo, H., Ma, Z., Zeng, Z., Luo, M., Zeng, W., Tang, J., et al. (2025). Each fake news is fake in its own way: An attribution multi-granularity benchmark for multimodal fake news detection. Proceedings of the AAAI Conference on Artificial Intelligence, 39(1), 228–236. https://doi.org/10.1609/aaai.v39i1.31999

[19] Hangloo, S., & Arora, B. (2022). Combating multimodal fake news on social media: Methods, datasets, and future perspective. Multimedia Systems, 28, 1–32. https://doi.org/10.1007/s00530-022-00966-y

[20] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770–778.

[21] Hossain, E., Kaysar, M., Joy, J. U., Zahid, A., Rahman, M. M., & Rahman, W. (2021). A study towards Bangla fake news detection using machine learning and deep learning. In Proceedings of International Conference on Innovations in Intelligent Computing and Communication (pp. 75–87). Springer. https://doi.org/10.1007/978-981-16-5157-1_7

[22] Hossain, M. Z., Rahman, M. A., Islam, M. S., & Kar, S. (2020). BanFakeNews: A dataset for detecting fake news in Bangla. Proceedings of the 12th Language Resources and Evaluation Conference (LREC), 2862–2871. Marseille, France: ELRA.

[23] Hosseini, M., Sabet, A. J., He, S., & Aguiar, D. (2023). Interpretable fake news detection with topic and deep variational models. Online Social Networks and Media, 36, 100249. https://doi.org/10.1016/j.osnem.2023.100249

[24] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

[25] Howard, A., et al. (2019). Searching for MobileNetV3. Proceedings of the IEEE/CVF International Conference on Computer Vision, 1314–1324.

[26] Hu, Z., Jia, S., & Rostami, M. (2024). An intermediate fusion ViT enables efficient text-image alignment in diffusion models. arXiv preprint arXiv:2403.16530.

[27] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4700–4708.

[28] Hussain, M. G., Hasan, M. R., Rahman, M., Protim, J., & Hasan, S. (2020a). Detection of Bangla fake news using MNB and SVM classifier. 2020 International Conference on Computer, Communication, and Electronics Engineering (ICCCECE), 81–85. https://doi.org/10.1109/iCCECE49321.2020.9231167

[29] Hussain, M. G., Hasan, M. R., Rahman, M., Protim, J., & Hasan, S. (2020b). Detection of Bangla fake news using MNB and SVM classifier. Same as above (duplicate entry, same reference as 28).

[30] Hauschild, J., & Eskridge, K. (2024). Word embedding and classification methods and their effects on fake news detection. Machine Learning with Applications, 17, Article 100566. https://doi.org/10.1016/j.mlwa.2024.100566

[31] Shu, K., Wang, S., & Liu, H. (2019). Beyond news contents: The role of social context for fake news detection. In Proceedings of the 12th ACM International Conference on Web Search and Data Mining (pp. 312–320). https://doi.org/10.1145/3289600.3290994

[32] Reis, J. C. S., Correia, A., Murai, F., Veloso, A., Benevenuto, F., & Cambria, E. (2019). Supervised learning for fake news detection. IEEE Intelligent Systems, 34(2), 76–81. https://doi.org/10.1109/MIS.2019.2899143

[33] Zhou, X., & Zafarani, R. (2023). Network-based fake news detection: A pattern-driven approach. Snopes. https://www.snopes.com/

[34] Shu, K., Cui, L., Wang, S., Lee, D., & Liu, H. (2019). dEFEND: Explainable fake news detection. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. https://doi.org/10.1145/3292500

[35] Bhutani, B., Rastogi, N., Sehgal, P., & Purwar, A. (2019). Fake news detection using sentiment analysis. In 2019 12th International Conference on Contemporary Computing (IC3). https://doi.org/10.1109/IC3.2019.8844880

[36] Zaeem, R. N., Li, C., & Barber, K. S. (2020). On sentiment of online fake news. In Proceedings of the 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) (pp. 760–767). https://doi.org/10.1109/ASONAM49781.2020.9381323

[37] Dey, A., et al. (2019). Fake news pattern recognition using linguistic analysis. In Proceedings of ICIEV-IVPR 2018 (pp. 305–309). https://doi.org/10.1109/ICIEV.2018.8641018

[38] De, S., & Agarwal, D. (2020). A novel model of supervised clustering using sentiment and contextual analysis for fake news detection. In Proceedings of IEEE MPCIT (pp. 112–117). https://doi.org/10.1109/MPCIT51588.2020.9350457

[39] Cui, L., Wang, S., & Lee, D. (2019). SAME: Sentiment-aware multi-modal embedding for detecting fake news. In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. https://doi.org/10.1145/3341161

[40] Xu, K., Wang, F., Wang, H., & Yang, B. (2020). Detecting fake news over online social media via domain reputations and content understanding. Tsinghua Science and Technology, 25(1), 20–27. https://doi.org/10.26599/TST.2018.9010139

[41] Carvalho, F., et al. (2020). A Brazilian Portuguese moral foundations dictionary for fake news classification. In Proceedings of the International Conference of the Chilean Computer Science Society (SCCC). https://doi.org/10.1109/SCCC51225.2020.9281258

[42] Balestrucci, A., & De Nicola, R. (2020). Credulous users and fake news: A real case study on the propagation in X. In IEEE Conference on Evolving and Adaptive Intelligent Systems. https://arxiv.org/abs/2005.03550v1

[43] 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 (pp. 2507–2511). https://doi.org/10.1109/ICASSP.2019.8683170

[44] Do, T. H., et al. (2021). Context-aware deep Markov random fields for fake news detection. IEEE Access, 9, 123456–123469. https://doi.org/10.1109/ACCESS.2021.3113877

[45] Hirlekar, V. V., & Kumar, A. (2020). Natural language processing based online fake news detection challenges – A detailed review. In Proceedings of ICCES (pp. 748–754). https://doi.org/10.1109/ICCES48766.2020.9137915

https://doi.org/10.71058/jodac.v10i02003

[46] Lin, J., et al. (2019). Detecting fake news articles. In Proceedings of the 2019 IEEE International Conference on Big Data (pp. 3021–3025). https://doi.org/10.1109/BIGDATA47090.2019.9005980

[47] Rk, K., G, A., & N, P. (2021). FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. Multimedia Tools and Applications, 80(8), 11765–11788. https://doi.org/10.1007/s11042-020-10183-2

[48] Mr, I., et al. (2020). Deep learning for misinformation detection on online social networks: A survey and new perspectives. Social Network Analysis and Mining, 10(1), Article 69. https://doi.org/10.1007/s13278-020-00696-x

[49] Meesad, P. (2021). Thai fake news detection based on information retrieval, natural language processing and machine learning. SN Computer Science, 2(6), Article 575. https://doi.org/10.1007/s42979-021-00775-6

[50] Singh, B., & Sharma, D. K. (2021). Predicting image credibility in fake news over social media using multi-modal approach. Neural Computing and Applications. https://doi.org/10.1007/s00521-021-06086-4

[51] Braşoveanu, A. M. P., & Andonie, R. (2020). Integrating machine learning techniques in semantic fake news detection. Neural Processing Letters, 51(2), 1483–1505. https://doi.org/10.1007/s11063-020-10365-x

[52] Li, X., et al. (2021). A novel self-learning semi-supervised deep learning network to detect fake news on social media. Multimedia Tools and Applications, 80(1), 1–9. https://doi.org/10.1007/s11042-021-11065-x

[53] Rk, K., G, A., & N, P. (2021). EchoFakeD: Improving fake news detection in social media with an efficient deep neural network. Neural Computing and Applications. https://doi.org/10.1007/s00521-020-05611-1

[54] Jiang, T., et al. (2021). A novel stacking approach for accurate detection of fake news. IEEE Access, 9, 22626–22639. https://doi.org/10.1109/ACCESS.2021.3056079

[55] Umer, M., et al. (2020). Fake news stance detection using deep learning architecture (CNN-LSTM). IEEE Access, 8, 156695–156706. https://doi.org/10.1109/ACCESS.2020.3019735

https://doi.org/10.71058/jodac.v10i02003