VISVESVARAYA TECHNOLOGICAL UNIVERSITY,

Jnana Sangama, Belgaum-590018

MINI PROJECT REPORT ON

" FAKE NEWS DETECTOR"

Submitted in partial fulfilment of the requirements for the award of the degree of
Bachelor of Engineering
In
Information Science & Engineering

SUBMITTED BY:

SIMRANJEET SINGH (3GN21IS042)
VRUSHAB PATIL (3GN21IS058)
AKHILESH BIRADAR (3GN21IS001)
NATRAJ HALE (3GN21IS024)

UNDER THE GUIDANCE OF
Dr. Masrath Begum

GURU NANAK DEV ENGINEERING COLLEGE,

MAILOOR ROAD, BIDAR, KARNATAKA-585403

Department Of Information Science And Engineering

GURU NANAK DEV ENGINEERING COLLEGE BIDAR,

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

CERTIFICATE

This is certified that the mini-project entitled **"Fake News Detector"** is a bonafide work carried out by **SIMRANJEET SINGH (3GN21IS042) , VRUSHAB PATIL (3GN21IS058) , AKHILESH BIRADAR (3GN21IS001), NATRAJ HALE (3GN21AI055)** in partial fulfillment of the requirements for third year of Engineering in **INFORMATION SCIENCE by VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI** during the year **2023-2024.** The mini project report has been approved as it satisfies the academic requirements prescribed for the Bachelor of Engineering Degree.

Guide	HOD	Principal
(Dr.Masrath Begum)	(**Dr**.**Masrath Begum)**	**(Dr. Suresh R. Reddy)**

Name of the Examiner Signature with Date

1.

2.

ACKNOWLEDGMENT

We would like to express our deep sense of gratitude to our principal **Dr. Suresh Reddy** Guru Nanak Dev Engineering College, Bidar for his inspiration and academic support by providing good facilities to complete the project.

We sincerely thanks our guide Dr. Masrath Begum Head of INFORMATION SCIENCE AND ENGINEERING department for their invaluable guidance, insightful advice, and constant encouragement. Their expertise and feedback have been instrumental in the successful completion of this project.
We thank all the staff members for supporting us in completion of the project, also we are thankful to our parents and all our dear friends who have directly or indirectly helped us.

Lastly, we would like to acknowledge the open-source community and the developers of the tools and libraries that made this project possible. Their contributions to the field of AI are immensely appreciated.

Finally, we express gratefulness to all those who knowingly or unknowingly helps us in successful completion of this project.

Simranjeet Singh
Vrushab Patil
Akhilesh Biradar
Natraj Hale

ABSTRACT

It is important to ensure the integrity of information, as well as that accurate and reliable news reaches further into society than fake counterparts. Abstract in this report summaries the necessary components of a working fake news detection system focused on finding and limiting false information spread. This system is based on new technologies like digitalized databases combined with machine learning algorithms to improve identification and authentication of news sources. Automated content analysis and fact-checking processes decrease the time needed to verify information, allowing us provision continuous monitoring of news feeds. It also has a very powerful verification system to assure the accuracy and authenticity of such news content. These tools also provide deeper analytical capabilities to trace misinformation patterns and predicted fake news outbreaks, as well helping in minimizing its expected aftermath. Real-time communication channels help news agencies, social media tailored and fact-checking organizations coordinate with the them stakeholders. If media organizations and platforms follow these holistic fake news detection steps, they will be able to fulfil the role of providing accurate information that people can trust in their civic life, which will ultimately protect society from misinformation.

TABLE OF CONTENTS

FIGURE LIST

CHAPTER-1

INTRODUCTION

The investigation of fake news has increased the moon and popularity of the public, especially scientists, around the world. A lot of research has been done on the effects of fake news and people's reactions to it. Fake news can be content that is untrue and designed to persuade readers to believe something that is not true. Today, there are many social media platforms and sharing applications that allow users to share information with millions of people with a single click. The real problem is that when people begin to accept this fact, their news becomes not only a "fake news" but has a new perspective. The problem starts when the public begins to believe fake news without checking its accuracy and true incidents. There are various tools or websites that can inform the public about the news and its facts. Now-a-days, fake news creates different problems, from negative reviews to news media and government campaigns in some media outlets. Fake news and lack of trust in the media has become a growing problem that has greatly affected our lives. It is clear that deliberately spreading false stories is "fake news", but recent discussions on social media have changed this definition. Facebook, Instagram, Twitter, WhatsApp and many other social media sites were at the center of many criticisms following the news. When users came across these articles, they used a feature on the site that spreads fake news. They also announced to the public that they were working on the difference between these sentences from a technical perspective. Of course, this is not an easy task. Since fake news exists at both ends of the spectrum, a given algorithm must be politically neutral and also give equal weight to legitimate news at both ends of the spectrum. But in order to solve this problem we need to understand what fake news is. Next, we need to learn how techniques in machine learning can help us detect fake news.

1.1 OBJECTIVES

The main purpose of this study is to detect fake news, which is a problem of classifying classical texts, with a simple method. Standards need to be developed to distinguish "real" news from "fake" news. By using machine learning algorithms with the help of Python programming language, data containing two types of news will be classified as fake news or real news.

- Accuracy: Build a model that can accurately classify news from fake news with high accuracy (e.g., >90% accuracy).
- Model Interpretation: Build a model to understand why a particular article is classified as fake or real and gain insight into the decision-making process.
- Handling Conflicting Information: Create a model that can handle conflicting information where one category (e.g. real news) dominates another (e.g. fake news).
- Real-time efficiency: Create systems that can process and distribute news in real time, allowing fake news to be quickly detected and alerted.
- Scalability: Create a model that can handle multiple media and scales to meet the needs of multiple applications (such as social media, news media, etc.).
- Continuous learning: Create a system that learns from feedback, adapts to changes and ideas, and improves performance over time.
- Fair decision-making: Make the system fair, equitable and transparent and avoid bias or discrimination.
- Human-computer interaction: Consider the use of human-computer interaction to analyze decision-making patterns and provide feedback for improvement.

CHAPTER-2

LITERATURE SURVEY

This study examined different techniques and methods previously used to detect fake news in order to identify the most effective and most fraudulent ones. The research examines a wide range of research. The problem of fake news is complex because individuals and organizations use different methods.

However, this research suggests four main methods:

1. Factual analysis
2. Word analysis
3. Visual analysis
4. Emotional analysis.

As discussed in this study, sentiment analysis uses neural networks to detect fake news; this method is also available in other literature, regardless of the use of specific tools and resources. Machine learning is a field of text analysis where fake news is an important part of data science that can distinguish fake news from real news.

It requires neural networks, word embeddings for data files, and commonly used techniques such as TF-IDF, Fast Text, Bag of Words (BOW), and Word2Vec. In addition, NLP library techniques are also shown in some studies.

The quality of the data that feeds the learning model is important, especially for the detection of fake news, where the accuracy of the model depends on the accuracy of the data. Both datasets and current information are frequently used. Some studies use data from the Fake News Challenge-1 (FNC-1) published in 2016, while others use a web crawler to extract more articles. Kaggle datasets DS1 and DS2 were also used to train the model.

Social media like Facebook and Twitter are all about rumors because most of the posts are claims rather than facts. Web scraping is usually done by the Beautiful Soup library, which is used to extract information from websites. In general, datasets are preferred over existing data because they are reliable and verifiable to ensure the accuracy and validity of the working model.

It involves cleaning the data by removing pre-written text, pauses, tags, and other

irrelevant content. Converting text into valuable features is crucial and this is where word embedding technology comes into play. Techniques include TF-IDF and BOW, which vectorize and rank words according to their frequency within and outside the system. Word2Vec and Fast Text are also used for word embedding, predicting words, and providing a sub word embedding to improve accuracy.

Fake news classification is a binary classification problem and a predictive analysis method. The system must meet these conditions in order to produce lies and fake news. Machine learning, especially in neural networks, has become an important method to achieve this task.

Traditional machine learning algorithms such as support vector machine (SVM) and logistic regression are generally the algorithms of choice for binary classification. Naive Bayes models perform better when combined with prediction.

However, the difficulty of traditional algorithms to cope with changing data and their inability to provide the required accuracy has led to the search for deep learning and NLP methods. There is a benefit to learning from experience and improving performance over time.

Dense Neural Networks (DNN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) are among the deep learning models for detecting fake news. This model gives good results compared to the old model.

The features of LSTM (Long Short-Term Memory) and RNN with bidirectional LSTM make it a popular choice for paper mining. CNN, although not accurate alone, works very well and gives better results when combined with LSTM (tools). Many ways to create fake news.

It highlights the importance of machine learning, the impact of word embedding techniques on accuracy, and the performance of neural network model

CHAPTER-3

REQUIREMENTS AND ANALYSIS

The fake news discovery extend requires comprehensive information collection, counting news articles labeled as fake or genuine from dependable sources, as well as metadata such as distribution date, creator, source validity, and social media engagement measurements. Information preprocessing is fundamental, including content cleaning to expel HTML labels, extraordinary characters, and stop words from news articles, and highlight extraction to determine significant highlights like word recurrence, estimation scores, and meaningfulness files. Normalization of numerical highlights guarantees consistency over the dataset. A well-designed database is vital, with normalization to decrease repetition and anticipate issues like erase special cases, guaranteeing information judgment and clear, reliable comes about.

4.1 Functional Requirements:

4.1.1 Data Collection:

- Collect news articles labeled as fake or genuine from reliable sources.
- Include metadata such as publication date, author, source credibility, and social media engagement metrics.

4.1.2. Data Preprocessing:

- Text cleaning to remove HTML tags, special characters, and stop words.
- Feature extraction to determine significant features like word frequency, sentiment scores, and readability indices.
- Normalization of numerical features to ensure consistency across the dataset.

4.1.3 Database Management:

- Well-designed database with normalization to reduce redundancy.
- Ensure data integrity and support efficient data retrieval and processing.

4.1.4. Machine Learning:

- Use of machine learning algorithms such as Decision Trees, Gradient Boosting Classifiers, Logistic Regression, and Random Forest.
- Training classifiers using extracted features to distinguish between fake and real news.

4.1.5. User Interface:

- Develop a user-friendly interface providing a dashboard for users to input news articles and receive classification results.
- Include visualizations showing the confidence level of predictions and key features affecting the decision.

4.1.6. Web Server:

- Deployment on a web server to handle a large number of concurrent requests and ensure accessibility and scalability.

4.2 Non-Functional Requirements

4.2.1. Performance:

- The system should handle data preprocessing and machine learning tasks efficiently without bottlenecks.
- The web server should manage concurrent requests effectively.

4.2.2. Scalability:

- Ability to scale resources based on load, especially during high traffic periods or when processing large data volumes.

4.2.3. Reliability:

- Ensure high availability of the system to provide continuous service without significant downtime.
- Maintain data integrity and consistency across the system.

4.2.4. Usability:

- The interface should be intuitive and easy to use, providing clear feedback and visualizations to users.
- Users should be able to easily input news articles and interpret classification results.

4.2.5. Security:

- Protect data from unauthorized access and ensure secure handling of user inputs and outputs.
- Implement measures to prevent data breaches and maintain user privacy.

4.2.6. Maintainability:

- The system should be modular and well-documented to facilitate updates and maintenance.
- Ensure that the codebase is organized and follows best practices for easy debugging and enhancements.

CHAPTER- 4

DESIGN

- ARCHITECTURE
- ACTIVITY DIAGRAM
- USE CASE DIAGRAM
- CONTROL FLOW DIAGRAM
- DATA FLOW DIAGRAM

4.1 ARCHITECTURE OF FAKE NEWS DETECTOR

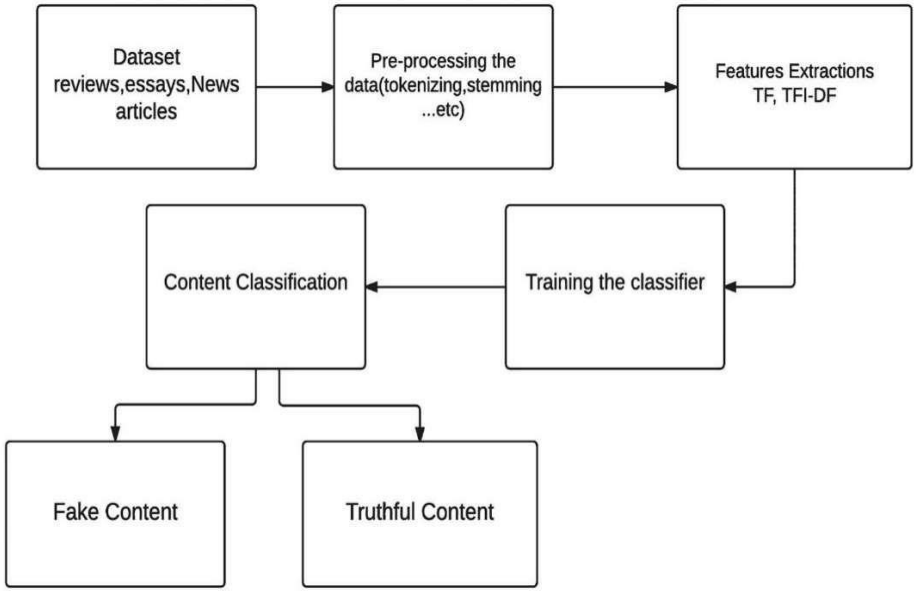


Fig 4.1

This diagram represents the architecture of a fake news detection system.

1. **Dataset (reviews, essays, news articles):**

- This is an information resource that includes many articles such as reviews, essays, and news. These articles are the raw material for detecting fake news.

2. **Pre-processing the data (tokenizing, stemming, etc.):**

- The text needs to be pre-processed before the text data is analyzed. This step involves several tasks:
  - **Tokenizing:** Breaking down the text into individual words or tokens.
  - **Stemming**: Reducing words to their root form (e.g., "running" becomes "run").
  - Other possible steps might include removing stop words, lowercasing, and handling punctuation.

3. **Features Extractions (TF, TF-IDF):**

- In this step, important features are extracted from the pre-processed text. Common methods include:
  - **Term Frequency (TF):** Counts how often a word appears in a document.
  - **Term Frequency-Inverse Document Frequency (TF-IDF):** Assess the importance of a word based on its frequency in the document relative to its frequency in the entire document.

4. **Training the classifier:**

- Uses the extracted features to train a machine learning classifier. This involves providing the distributor with a collection of fake and real content examples and learning to distinguish between them.

5. **Content Classification:**

- Once trained, the classifier can be used to classify new, unseen items. A classifier analyzes the features of new content and decides whether it is fake or real.

6. **Fake Content / Truthful Content:**

- The result of the content classification process is that the content is divided into two groups: fake content and real content.

4.2 ACTIVITY DIAGRAM

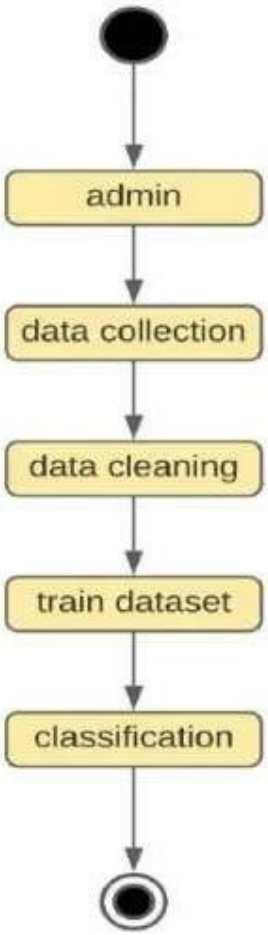


Fig 4.2

This diagram represents an activity diagram for a fake news detection system.

1. **Admin**:

- The process begins with the admin, who is responsible for overseeing all activities aimed at detecting fake news.

2. **Data Collection**:

- This step involves data collection, which includes information gathered from various sources such as news, media reports, reviews, and notes. The collected data will be the raw material of the system.

3. **Data Cleaning**:

- After data is collected, it must be cleaned to remove noise or irrelevant data. Data cleaning involves removing duplicates, handling missing values, correcting errors, and modelling the data.

4. **Train Dataset**:

- After the data is cleaned, it is used for training the dataset. This involves splitting the data into training and testing and then using the training process to train machine learning models. The training process helps the model learn the patterns and characteristics that distinguish fake news from real news.

5. **Classification**:

- In the final step, the training model is used to classify new, unseen objects. This model analyses the features of the new product to determine whether it is fake or real.

4.3 USE CASE DIAGRAM

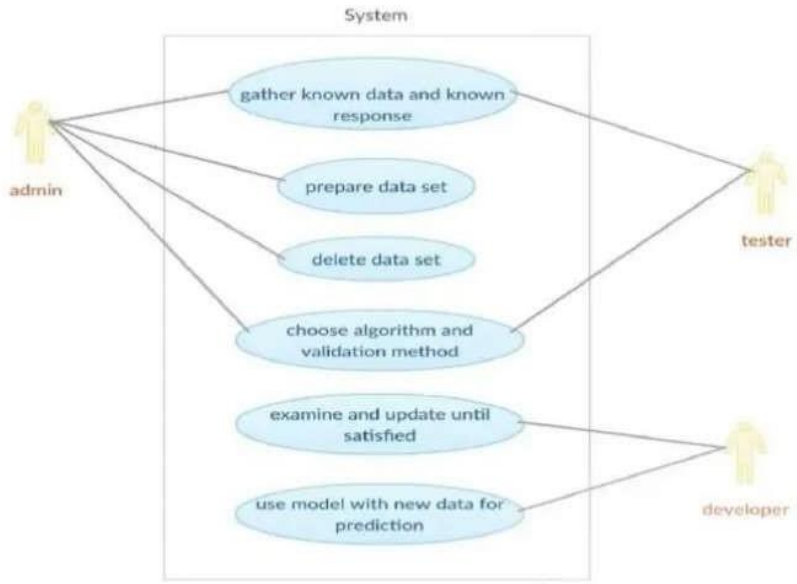


Fig 4.3

This diagram represents a use case diagram for a fake news detection system.

1. **Admin:**

- **Gather known data and known response:** Collects data and its corresponding labels (true/false).
- **Prepare data set:** Organizes and preprocesses the data for analysis.
- **Delete data set:** Removes unwanted or outdated data.

2. **Tester:**

- **Choose algorithm and validation method:** Selects the machine learning algorithm and validation techniques.
- Examine and update until satisfied: Tests and refines the model to improve performance.

3. **Developer:**

- Use model with new data for prediction: Deploys the model to predict the authenticity of new data.

4.4 CONTROL FLOW DIAGRAM

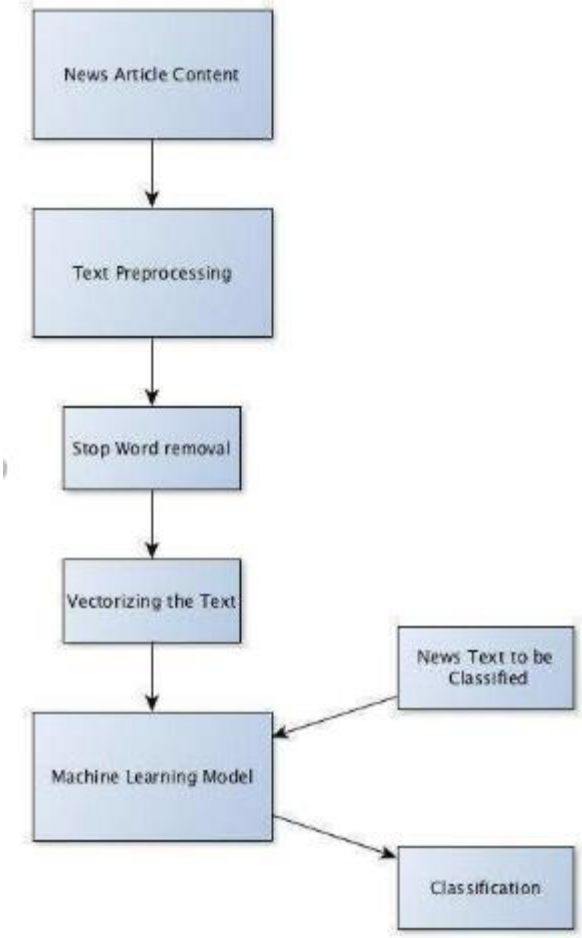


Fig 4.4

This diagram represents a control flow diagram for a fake news detection system.

1. **Start:** The process begins.
2. **Input News:** News articles or content are input into the system
3. **Preprocess Data:** The input data is cleaned and pre-processed to remove noise and irrelevant information,

4. **Extract Features:** Important features from the data are extracted for analysis.
5. **Train ML Model:** A machine learning model is trained using the processed data and extracted features.
6. **Predict Fake/Real News:** The trained model is used to predict whether the news is fake or real.
7. **Is it Fake News?:** The prediction result is checked.
8. **Yes:** If the news is predicted as fake, proceed to the next step.
9. **No:** If the news is not fake, proceed to the end.
10. **Display Result:** The result (whether the news is fake or real) is displayed.
11. **End:** The process ends.

4.5 DATA FLOW DIAGRAM

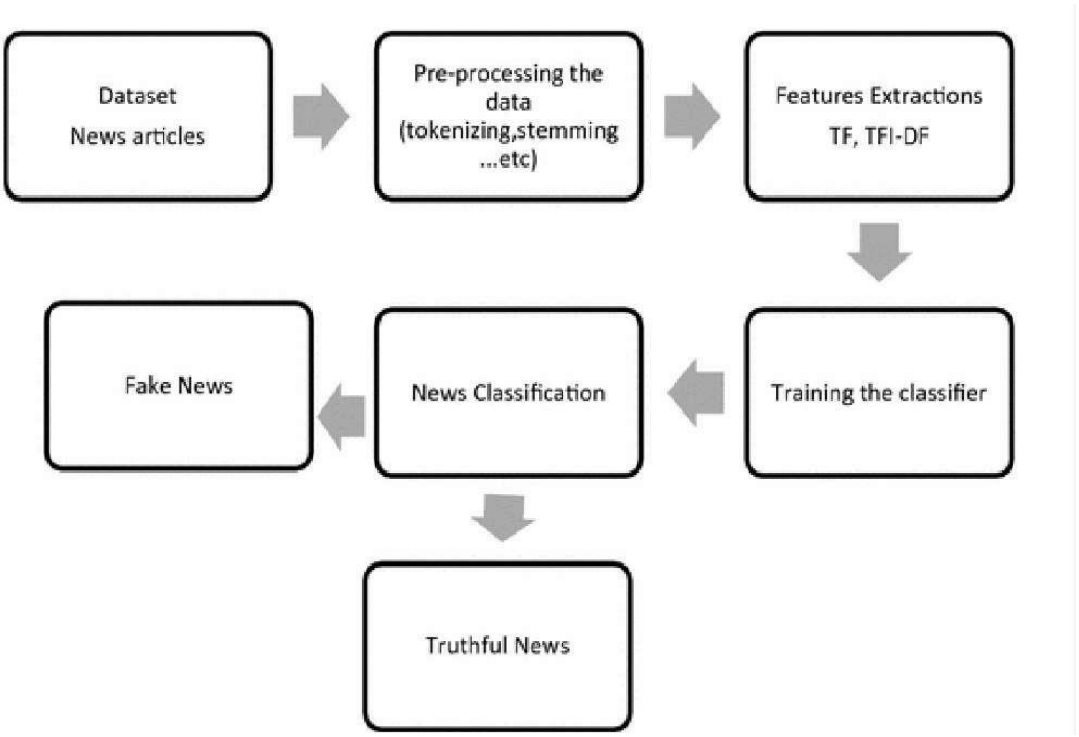


Fig 4.5

This diagram represents a data flow diagram for a fake news detection system.

1. **Dataset News articles:** The process begins with a collection of news articles.

2. **Pre-processing the data (tokenizing, stemming, etc.):** The data undergoes preprocessing, which includes tokenizing, stemming, and other techniques to clean and prepare the text for analysis.

3. **Features Extractions (TF, TF-IDF):** Important features are extracted from the pre-processed data using methods like Term Frequency (TF) and Term Frequency-Inverse Document Frequency (TF-IDF).

4. **Training the classifier:** The extracted features are used to train a machine learning classifier.

5. **News Classification**: The trained classifier is used to classify news articles.

6. **Fake News**: If the article is classified as fake news, it is labelled accordingly.

7. **Truthful News:** If the article is classified as truthful news, it is labelled accordingly.

CHAPTER - 5

IMPLEMENTATTION

Fake news detection: -

Information in the search for fake news is created in a way that is useful and helpful to users. This includes a comprehensive view of the frequency of specific words, emotional score, credibility of the site, etc. that is derived from raw data. Raw data includes events related to social and user interactions, including text messages, metadata such as publication date and author details, and user's interactions such as likes, shares, and comments. This raw data is extracted, pre-processed and transformed into usable data required for effective machine learning models.

A great database is significant for overseeing information productively, avoiding clashes, and keeping up the keenness of critical data. For a fake news location venture, a well-structured database guarantees that information is organized and effortlessly open, encouraging solid comes about. Normalization may be a key prepare in database administration that decreases excess and avoids issues like erase exemptions, where profitable information can be misplaced when a record is erased. By normalizing tables, the database gives clear and reliable comes about, which is imperative for preparing precise and vigorous machine learning models.

Algorithms that have been used are as follows: -

5.1 TF-IDF Vectorizer:

TF-IDF is a statistical measure that calculates the significance of a word is to a document in a collection of documents. This is calculated by multiplying two metrics : number of times a word appears in a document and the inverse document frequency of the word across a set of documents.

Term Frequency (TF) is he number of times a word appears during a document divided by total number of words within the document. Every document has a different term frequency.

Inverse Data Frequency (IDF) is the log of the number of documents divided by the number

of documents that contain the word w. Inverse data frequency determines the weightage of rare words across all documents in the corpus.

```
x= data['text']
y= data['class']
```

```
x_train, x_test, y_train, y_test = train_test_split(x,y, test_size=0.25, random_state=42)
```

```
from sklearn.feature_extraction.text import TfidfVectorizer

vectorization = TfidfVectorizer()
xv_train = vectorization.fit_transform(x_train)
xv_test = vectorization.transform(x_test)
```

Fig 5.1 TF-IDF vectorizer implementation

5.2. LOGISTIC REGRESSION

Logistic regression is a statistical model used for binary classification. It predicts the probability of a binary outcome based on one or more input features.

It uses a logistic function to map predicted values to probabilities. The model estimates coefficients for input features to maximize the likelihood of observing the given data.

Logistic regression can predict the likelihood of a news article being fake based on textual features and metadata. It's simple and interpretable, making it useful for understanding key factors associated with fake news

```
[76]: from sklearn.linear_model import LogisticRegression

      LR = LogisticRegression()
      LR.fit(xv_train, y_train)

[76]: ▾ LogisticRegression
      LogisticRegression()

[77]: LR.score(xv_train, y_train)

[77]: 0.991116788971746

[27]: pred_lr= LR.predict(xv_test)

[28]: LR.score(xv_test, y_test)

[28]: 0.9849376114081997

[29]: print(classification_report(y_test, pred_lr))

              precision    recall  f1-score   support

           0       0.99      0.98      0.99      5853
           1       0.98      0.99      0.98      5367

    accuracy                           0.98     11220
   macro avg       0.98      0.98      0.98     11220
weighted avg       0.98      0.98      0.98     11220
```

Fig 5.2 Logistic Regression Implementation

5.3.DECISION TREE:

A decision tree is a supervised learning algorithm used for classification and regression tasks. It splits the data into subsets based on the value of input

features, forming a tree-like structure of decisions.

Starting from the root node, the algorithm selects the best feature that splits the data into purest subsets, continues this process recursively, and forms a tree where each leaf represents a class label (for classification) or a continuous value (for regression) Decision trees can classify news as fake or real by learning from labeled examples. They identify patterns and keywords that are commonly associated

```
[78]: from sklearn.tree import DecisionTreeClassifier

      DT = DecisionTreeClassifier()
      DT.fit(xv_train, y_train)

[78]: ▾ DecisionTreeClassifier
      DecisionTreeClassifier()

[79]: DT.score(xv_train, y_train)

[79]: 0.9999702902641195

[31]: pred_dt= DT.predict(xv_test)

[32]: DT.score(xv_test, y_test)

[32]: 0.9961675579322639

[33]: print(classification_report(y_test, pred_lr))

                  precision    recall  f1-score   support

               0       0.99      0.98      0.99      5853
               1       0.98      0.99      0.98      5367

        accuracy                           0.98     11220
       macro avg       0.98      0.98      0.98     11220
    weighted avg       0.98      0.98      0.98     11220
```

with fake or real news.

Fig 5.3 Decision tree implementation

5.4.GRADIENT BOOSTING CLASSIFIER

Gradient Boosting is an ensemble learning technique that builds multiple decision trees sequentially, where each tree tries to correct the errors of the previous one.

It combines weak learners i.e. shallow trees into a strong learner by optimizing the loss function, using gradient descent. Each new tree is added to reduce the residual errors of the combined trees.

It enhances the accuracy of classification by focusing on the misclassified instances from previous trees, leading to a more robust model that effectively distinguishes fake news

```
[80]: from sklearn.ensemble import GradientBoostingClassifier
      GB = GradientBoostingClassifier(random_state =0)
      GB.fit(xv_train, y_train)

[80]: GradientBoostingClassifier
      GradientBoostingClassifier(random_state=0)

[81]: GB.score(xv_train, y_train)

[81]: 0.9972072848272379

[35]: pred_gb = GB.predict(xv_test)

[36]: GB.score(xv_test, y_test)

[36]: 0.9953654188948307

[37]: print(classification_report(y_test, pred_lr))
              precision    recall  f1-score   support

           0       0.99      0.98      0.99      5853
           1       0.98      0.99      0.98      5367

    accuracy                           0.98     11220
   macro avg       0.98      0.98      0.98     11220
weighted avg       0.98      0.98      0.98     11220
```

Fig 5.4 gradient Boosting Classifier Implementation

5.5 RANDOM FOREST

Random Forest is an ensemble learning method that constructs multiple decision trees during training and merges their outputs to improve accuracy and control overfitting.

It builds a "forest" of decision trees using randomly selected subsets of data and features. The final prediction is made by averaging the predictions of all trees for regression or taking a majority vote for classification. By manipulating the power of multiple decision trees, Random Forest can fully classify news articles as fake or real. It reduces the risk of overfitting and improves generalization.

```
[82]: from sklearn.ensemble import RandomForestClassifier
      RF = RandomForestClassifier()
      RF.fit(xv_train, y_train)

[82]: RandomForestClassifier
      RandomForestClassifier()

[83]: RF.score(xv_train, y_train)

[83]: 0.9999702902641195

[39]: pred_rf= RF.predict(xv_test)

[40]: RF.score(xv_test, y_test)

[40]: 0.9867201426024955

[41]: print(classification_report(y_test, pred_lr))
              precision    recall  f1-score   support

           0       0.99      0.98      0.99      5853
           1       0.98      0.99      0.98      5367

    accuracy                           0.98     11220
   macro avg       0.98      0.98      0.98     11220
weighted avg       0.98      0.98      0.98     11220
```

Fig 5.5 Random Forest implementation

These algorithms are trained on a dataset containing labeled news articles (fake or real).

Text features such as word frequency, n-grams, and metadata (author, source) are extracted.

The algorithm uses these features to distinguish fake news from real news.

Once trained, the model can predict the likelihood that a new, unknown news article is fake, aiding automated fake news detection systems.

CHAPTER-6

RESULTS

The project on Fake News Detection has been executed effectively utilizing Python programming language. Machine learning calculations have been utilized to confirm the information precision, and subsequently the comes about illustrated a fabulous exactness rate.

Algorithms such as Logistic Regression, Decision tree classifier, Gradient boosting classifier and Random Forest strategies were utilized to analyze and classify the news articles effectively.

After actualizing the previously mentioned forms, a manual testing stage was conducted wherein the client was required to input the information. The comes about reliably given exact data almost the genuineness of the news, recognizing whether it was fake or genuine. In this specific testing occurrence, the news was precisely recognized as fake.

Further, the system's vigor was tried with different datasets, guaranteeing its capability to handle assorted and expansive volumes of news articles. The client interface was outlined to be instinctive, empowering clients to effectively input news articles and get real-time genuineness results.

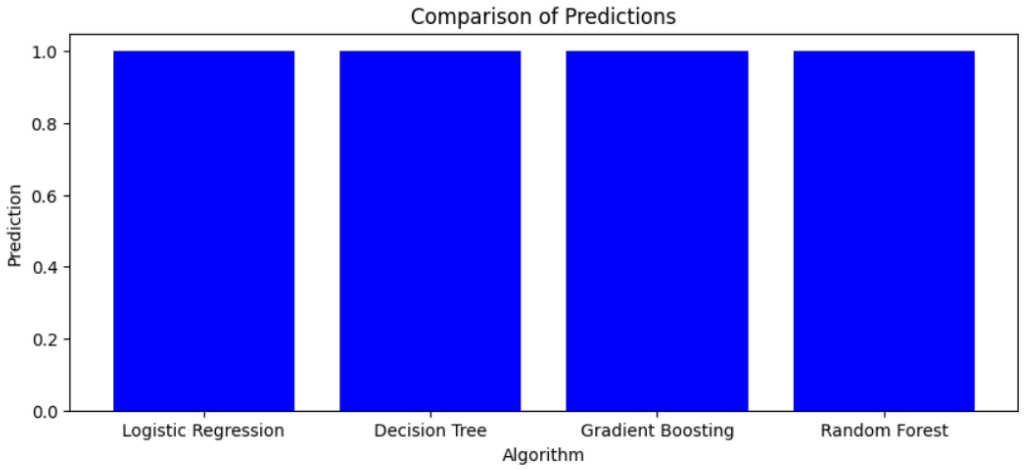
Hence, the project on the Fake News Detection has been actualized, executed, and confirmed effectively. The framework has demonstrated to be a solid instrument in combating the spread of deception, giving a significant asset for confirming news in today's information-rich society.

CHAPTER-8

SCREENSHOTS

For predicting that given news is true:

```
Enter a news article:  WASHINGTON (Reuters) - Alabama Secretary of State John Merrill said he will certify Democratic Senator-elect Doug Jones as winner
on Thursday despite opponent Roy Mooreâ€™s challenge, in a phone call on CNN. Moore, a conservative who had faced allegations of groping teenage girls wh
en he was in his 30s, filed a court challenge late on Wednesday to the outcome of a U.S. Senate election he unexpectedly lost.
```




```
The news is likely to be TRUE.
```

Fig. 8.1 predicting true news

For predicting news is fake:

```
[94]: news = str(input())
      manual_testing(news)
```

```
 On Friday, it was revealed that former Milwaukee Sheriff David Clarke, who was being considered for Homeland Security Secretary in Donald Trump s admin
istration, has an email scandal of his own.In January, there was a brief run-in on a plane between Clarke and fellow passenger Dan Black, who he later h
ad detained by the police for no reason whatsoever, except that maybe his feelings were hurt. Clarke messaged the police to stop Black after he deplane
d, and now, a search warrant has been executed by the FBI to see the exchanges.Clarke is calling it fake news even though copies of the search warrant a
re on the Internet. I am UNINTIMIDATED by lib media attempts to smear and discredit me with their FAKE NEWS reports designed to silence me,  the former
sheriff tweeted.  I will continue to poke them in the eye with a sharp stick and bitch slap these scum bags til they get it. I have been attacked by bet
ter people than them #MAGA I am UNINTIMIDATED by lib media attempts to smear and discredit me with their FAKE NEWS reports designed to silence me. I wil
l continue to poke them in the eye with a sharp stick and bitch slap these scum bags til they get it. I have been attacked by better people than them #M
AGA pic.twitter.com/XtZW5PdU2b  David A. Clarke, Jr. (@SheriffClarke) December 30, 2017He didn t stop there.BREAKING NEWS! When LYING LIB MEDIA makes up
FAKE NEWS to smear me, the ANTIDOTE is go right at them. Punch them in the nose & MAKE THEM TASTE THEIR OWN BLOOD. Nothing gets a bully like LYING LIB M
EDIA S attention better than to give them a taste of their own blood #neverbackdown pic.twitter.com/T2NY2psHCR  David A. Clarke, Jr. (@SheriffClarke) De
cember 30, 2017The internet called him out.This is your local newspaper and that search warrant isn t fake, and just because the chose not to file charg
es at the time doesn t mean they won t! Especially if you continue to lie. Months after decision not to charge Clarke, email search warrant filed http
s://t.co/zcbyc4Wp5b  KeithLeBlanc (@KeithLeBlanc63) December 30, 2017I just hope the rest of the Village People aren t implicated.  Kirk Ketchum (@kirkk
etchum) December 30, 2017Slaw, baked potatoes, or French fries? pic.twitter.com/fWfXsZupxy  ALT- Immigration  (@ALT_uscis) December 30, 2017pic.twitte
r.com/ymsOBLjfxU  Pendulum Swinger (@PendulumSwngr) December 30, 2017you called your police friends to stand up for you when someone made fun of your ha
t  Chris Jackson (@ChrisCJackson) December 30, 2017Is it me, with this masterful pshop of your hat, which I seem to never tire of. I think it s the stee
ly resolve in your one visible eye pic.twitter.com/dWr5k8ZEZV  Chris Mohney (@chrismohney) December 30, 2017Are you indicating with your fingers how man
y people died in your jail? I think you re a few fingers short, dipshit  Ike Barinholtz (@ikebarinholtz) December 30, 2017ROFL. Internet tough guy with
fake flair. pic.twitter.com/ulCFddhkdy  KellMeCrazy (@Kel_MoonFace) December 30, 2017You re so edgy, buddy.  Mrs. SMH (@MRSSMH2) December 30, 2017Is his
break over at Applebees?  Aaron (@feltrrr2) December 30, 2017Are you trying to earn your  still relevant  badge?  CircusRebel (@CircusDrew) December 30,
2017make sure to hydrate, drink lots of water. It s rumored that prisoners can be denied water by prison officials.  Robert Klinc (@RobertKlinc1) Decemb
er 30, 2017Terrill Thomas, the 38-year-old black man who died of thirst in Clarke s Milwaukee County Jail cell this April, was a victim of homicide. We
just thought we should point that out. It can t be repeated enough.Photo by Spencer Platt/Getty Images.


 LR Prediction:Fake News
Accuracy: 98.49%
DT Prediction:Fake News
Accuracy: 99.58%
GBC Prediction:Fake News
Accuracy: 99.54%
RFC Prediction:Fake News
Accuracy: 98.55%
```

Fig.8.2 Predicting the news is Fake

```
[52]: news = str(input())
      manual_testing(news,0)
```

```
 Donald Trump just couldn t wish all Americans a Happy New Year and leave it at that. Instead, he had to give a shout out to his enemies, haters and  th
e very dishonest fake news media.  The former reality show star had just one job to do and he couldn t do it. As our Country rapidly grows stronger and
smarter, I want to wish all of my friends, supporters, enemies, haters, and even the very dishonest Fake News Media, a Happy and Healthy New Year,  Pres
ident Angry Pants tweeted.  2018 will be a great year for America! As our Country rapidly grows stronger and smarter, I want to wish all of my friends,
supporters, enemies, haters, and even the very dishonest Fake News Media, a Happy and Healthy New Year. 2018 will be a great year for America!  Donald
J. Trump (@realDonaldTrump) December 31, 2017Trump s tweet went down about as welll as you d expect.What kind of president sends a New Year s greeting l
ike this despicable, petty, infantile gibberish? Only Trump! His lack of decency won t even allow him to rise above the gutter long enough to wish the A
merican citizens a happy new year!  Bishop Talbert Swan (@TalbertSwan) December 31, 2017no one likes you  Calvin (@calvinstowell) December 31, 2017Your
impeachment would make 2018 a great year for America, but I ll also accept regaining control of Congress.  Miranda Yaver (@mirandayaver) December 31, 20
17Do you hear yourself talk? When you have to include that many people that hate you you have to wonder? Why do the they all hate me?  Alan Sandoval (@A
lanSandoval13) December 31, 2017Who uses the word Haters in a New Years wish??  Marlene (@marlene399) December 31, 2017You can t just say happy new yea
r?  Koren pollitt (@Korencarpenter) December 31, 2017Here s Trump s New Year s Eve tweet from 2016.Happy New Year to all, including to my many enemies a
nd those who have fought me and lost so badly they just don t know what to do. Love!  Donald J. Trump (@realDonaldTrump) December 31, 2016This is nothin
g new for Trump. He s been doing this for years.Trump has directed messages to his  enemies  and  haters  for New Year s, Easter, Thanksgiving, and the
anniversary of 9/11. pic.twitter.com/4FPAe2KypA  Daniel Dale (@ddale8) December 31, 2017Trump s holiday tweets are clearly not presidential.How long did
he work at Hallmark before becoming President?  Steven Goodine (@SGoodine) December 31, 2017He s always been like this . . . the only difference is that
in the last few years, his filter has been breaking down.  Roy Schulze (@thbthttt) December 31, 2017Who, apart from a teenager uses the term haters?  We
ndy (@WendyWhistles) December 31, 2017he s a fucking 5 year old  Who Knows (@rainyday80) December 31, 2017So, to all the people who voted for this a hol
e thinking he would change once he got into power, you were wrong! 70-year-old men don t change and now he s a year older.Photo by Andrew Burton/Getty I
mages.


 LR Prediction:Fake News
Accuracy: 100.00%
DT Prediction:Fake News
Accuracy: 100.00%
GBC Prediction:Fake News
Accuracy: 100.00%
RFC Prediction:Fake News
Accuracy: 100.00%
```

Fig. 8.3 Predicting the news is fake with target value is fake or 0

```
[53]: news = str(input())
      manual_testing(news,1)
```

```
 Donald Trump just couldn t wish all Americans a Happy New Year and leave it at that. Instead, he had to give a shout out to his enemies, haters and  th
e very dishonest fake news media.  The former reality show star had just one job to do and he couldn t do it. As our Country rapidly grows stronger and
smarter, I want to wish all of my friends, supporters, enemies, haters, and even the very dishonest Fake News Media, a Happy and Healthy New Year,  Pres
ident Angry Pants tweeted.  2018 will be a great year for America! As our Country rapidly grows stronger and smarter, I want to wish all of my friends,
supporters, enemies, haters, and even the very dishonest Fake News Media, a Happy and Healthy New Year. 2018 will be a great year for America!  Donald
J. Trump (@realDonaldTrump) December 31, 2017Trump s tweet went down about as welll as you d expect.What kind of president sends a New Year s greeting l
ike this despicable, petty, infantile gibberish? Only Trump! His lack of decency won t even allow him to rise above the gutter long enough to wish the A
merican citizens a happy new year!  Bishop Talbert Swan (@TalbertSwan) December 31, 2017no one likes you  Calvin (@calvinstowell) December 31, 2017Your
impeachment would make 2018 a great year for America, but I ll also accept regaining control of Congress.  Miranda Yaver (@mirandayaver) December 31, 20
17Do you hear yourself talk? When you have to include that many people that hate you you have to wonder? Why do the they all hate me?  Alan Sandoval (@A
lanSandoval13) December 31, 2017Who uses the word Haters in a New Years wish??  Marlene (@marlene399) December 31, 2017You can t just say happy new yea
r?  Koren pollitt (@Korencarpenter) December 31, 2017Here s Trump s New Year s Eve tweet from 2016.Happy New Year to all, including to my many enemies a
nd those who have fought me and lost so badly they just don t know what to do. Love!  Donald J. Trump (@realDonaldTrump) December 31, 2016This is nothin
g new for Trump. He s been doing this for years.Trump has directed messages to his  enemies  and  haters  for New Year s, Easter, Thanksgiving, and the
anniversary of 9/11. pic.twitter.com/4FPAe2KypA  Daniel Dale (@ddale8) December 31, 2017Trump s holiday tweets are clearly not presidential.How long did
he work at Hallmark before becoming President?  Steven Goodine (@SGoodine) December 31, 2017He s always been like this . . . the only difference is that
in the last few years, his filter has been breaking down.  Roy Schulze (@thbthttt) December 31, 2017Who, apart from a teenager uses the term haters?  We
ndy (@WendyWhistles) December 31, 2017he s a fucking 5 year old  Who Knows (@rainyday80) December 31, 2017So, to all the people who voted for this a hol
e thinking he would change once he got into power, you were wrong! 70-year-old men don t change and now he s a year older.Photo by Andrew Burton/Getty I
mages.


 LR Prediction:Fake News
Accuracy: 0.00%
DT Prediction:Fake News
Accuracy: 0.00%
GBC Prediction:Fake News
Accuracy: 0.00%
RFC Prediction:Fake News
Accuracy: 0.00%
```

Fig. 8.4 Predicting the news is fake with target value is true or 1

In this result, we are using the target values 0 and 1 for checking that it predict the accurate results or not. If it predict the accurate results then it gives the accuracy 100% , if not then it gives the accuracy 0%.

```
[55]: news = str(input())
      manual_testing(news,0)
```

```
 On Christmas day, Donald Trump announced that he would  be back to work  the following day, but he is golfing for the fourth day in a row. The former r
eality show star blasted former President Barack Obama for playing golf and now Trump is on track to outpace the number of golf games his predecessor pl
ayed.Updated my tracker of Trump s appearances at Trump properties.71 rounds of golf including today s. At this pace, he ll pass Obama s first-term tota
l by July 24 next year. https://t.co/Fg7VacxRtJ pic.twitter.com/5gEMcjQTbH  Philip Bump (@pbump) December 29, 2017 That makes what a Washington Post rep
orter discovered on Trump s website really weird, but everything about this administration is bizarre AF. The coding contained a reference to Obama and
golf:  Unlike Obama, we are working to fix the problem  and not on the golf course.  However, the coding wasn t done correctly.The website of Donald Tr
ump, who has spent several days in a row at the golf course, is coded to serve up the following message in the event of an internal server error: http
s://t.co/zrWpyMXRcz pic.twitter.com/wiQSQNNzw0  Christopher Ingraham (@_cingraham) December 28, 2017That snippet of code appears to be on all https://t.
co/dkhw0AlHB4 pages, which the footer says is paid for by the RNC? pic.twitter.com/oaZDT126B3  Christopher Ingraham (@_cingraham) December 28, 2017It s
also all over https://t.co/ayBlGmk65Z. As others have noted in this thread, this is weird code and it s not clear it would ever actually display, but wh
o knows.  Christopher Ingraham (@_cingraham) December 28, 2017After the coding was called out, the reference to Obama was deleted.UPDATE: The golf error
message has been removed from the Trump and GOP websites. They also fixed the javascript  =  vs  ==  problem. Still not clear when these messages would
actually display, since the actual 404 (and presumably 500) page displays a different message pic.twitter.com/Z7dmyQ5smy  Christopher Ingraham (@_cingra
ham) December 29, 2017That suggests someone at either RNC or the Trump admin is sensitive enough to Trump s golf problem to make this issue go away quic
kly once people noticed. You have no idea how much I d love to see the email exchange that led us here.  Christopher Ingraham (@_cingraham) December 29,
2017 The code was f-cked up.The best part about this is that they are using the  =  (assignment) operator which means that bit of code will never get ru
n. If you look a few lines up  errorCode  will always be  404          (@tw1trsux) December 28, 2017trump s coders can t code. Nobody is surprised.  Tim
Peterson (@timrpeterson) December 28, 2017Donald Trump is obsessed with Obama that his name was even in the coding of his website while he played golf a
gain.Photo by Joe Raedle/Getty Images.


 LR Prediction:Fake News
Accuracy: 100.00%
DT Prediction:Fake News
Accuracy: 100.00%
GBC Prediction:Fake News
Accuracy: 100.00%
RFC Prediction:Fake News
Accuracy: 100.00%
```

Fig 8.5 predicting the news is fake or real with target value

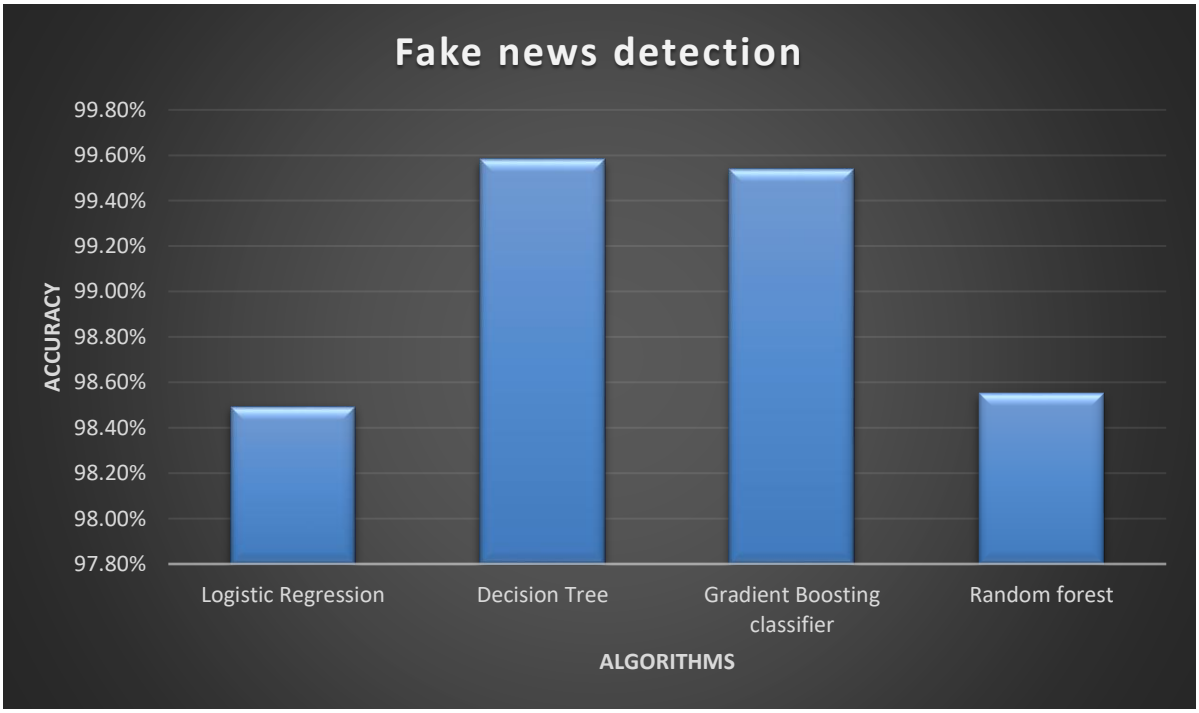


Fig 8.6 graphical representation of algorithm vs accuracy

CHAPTER-9

CONCLUSION AND FUTURE SCOPE

Based on the results, the study concludes that the fake news detection system is significantly better than manual verification methods. The results show that respondents prefer fake news detection systems over manual verification because they offer many advantages and benefits that lead to efficiency and efficiency. It can be concluded that as users' trust in the system increases, fake news detectors improve the accuracy and reliability of information by providing a more effective way to manage the verification process.

In the future, our fake news detection project may expand its scope in many directions. You may consider developing a mobile app version to improve accessibility and provide features such as real-time news checking and push notifications for questionable news. Integration with social media platforms can provide real-time analytics and identify potentially spreading fake news. Advanced analytics and reporting capabilities can help you optimize detection algorithms and provide detailed information about disinformation trends.

Integrating location services could allow users to find news verification for a specific area and geographically track the spread of fake news. Introducing sentiment analysis algorithms can further enhance the detection process by analyzing the tone and context of news articles. Blockchain technology can enhance the security and traceability of news verification transactions by ensuring the integrity of the verification process.

Community engagement features like social media and collaborative fact-checking can help build a sense of community among users in the fight against fake news. Machine learning algorithms can be used to predict misinformation trends and recommend targeted intervention strategies.

BIBILOGRAPHY

REFRENCES

- Bostrom, N., & Yudkowsky, E. (2011).

  The Ethics of Artificial Intelligence and Robotics. *Cambridge Handbook of Artificial Intelligence*, 316-334.

- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017).

  Fake News Detection on Social Media: A Data Mining Perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36.

- Ahmed, H., Traore, I., & Saad, S. (2017).

  Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques. *Proceedings of the International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments*, 127-138.

- Vosoughi, S., Roy, D., & Aral, S. (2018).

  The Spread of True and False News Online. *Science*, 359(6380), 1146- 1151.

- Zhou, X., & Zafrani, R. (2020).

  Detecting Fake News with Deep Learning: An Interdisciplinary Perspective. *ACM Computing Surveys (CSUR)*, 53(4), 1-36.

ARTICLES

- Industrial Engineering Journal
- International Research Journal of Modernization in Engineering Technology and Sciences

WEBSITES

- https://www.google.com/
- https://chatgpt.com/
- https://www.geeksforgeeks.org/
- https://www.youtube.com/