Department of Information Technology
NBA Accredited

A.P. Shah Institute of Technology

G.B.Road,Kasarvadavli, Thane(W), Mumbai-400615

UNIVERSITY OF MUMBAI

Academic Year 2020-2021

A Project Presentation on

ML Based Web Framework for Fake News Detection

Submitted in partial fulfillment of the degree of
Bachelor of Engineering(Sem-8)
in
INFORMATION TECHNOLOGY

By
Jaagrut Shah(19204002)
Jigar Desai (19204003)
Yash Jain (19204013)

Under the Guidance of
Prof. Apeksha Mohite
Prof. Geetanjali Kalme

1.Project Conception and Initiation

1.1 Abstract

- Fake news has been a drag ever since the web boomed. The news or social media network’s that allows us to gather information about incident’s happening over this world can be contaminated with fake news to run a particular political or personal agenda. Combating this fake news is vital because the world's view is formed by information. Verifying each news one by one by a person's being is totally unfeasible.
- This project attempts to expedite the method of identification of fake news by proposing a system which will try to reliably classify and predict if a news can be fake or not. The arduous task of detection of fake news are often made trivial with the usage of the proper models with the proper tools. To automate the analysis of such data, the world of Sentiment Analysis has emerged.

1.1 Abstract

- Sentiment Analysis may be a problem of text-based analysis, but there are some challenges that make it difficult as compared to traditional text-based analysis, This clearly states that there is a need for an attempt to work towards these problems and it has opened up several opportunities for future research for handling negations, hidden sentiments identification, slangs, polysemy.
- However, the growing scale of knowledge demands automatic data analysis techniques like analyzing text using Natural Language Processing. In this paper, an in-depth survey on different techniques utilized in Fake News Prediction & Sentiment Analysis is administered to know the extent of labor.

1.2 Objectives

- To show the news relevancy and analysis to attain accuracy in anticipating real and dependable news.
- To work on this issue, a layered model is proposed, which fine-tunes the information insight received from the data at each phase before attempting a prediction.
- To use a variety of Machine Learning approaches, achieve demonstratable success in the prediction of fake news and posts.
- To eliminate the propagation of false information on social media that may mislead users.
- To be able to give more and more accurate news on the screen.

1.3 Literature Review

Sr. No.	Authors	Paper Titles	Findings
1	Alim Al Ayub Ahmed, Ayman Aljabouh, Praveen Kumar Donepudi, Myung Suh Choi	Detection of online fake news using N-gram analysis and machine learning Technique.	ML models including Knearest neighbor (KNN), support vector machine (SVM), logistic regression (LR), linear support vector machine (LSVM), decision tree (DT), and stochastic gradient descent (SGD), achieving the highest accuracy (92%) with SVM and logistic regression.
2	Shu et al Sadia Afroz, Michael Brennan, and Rachel Greenstadt	Fake news detection on social media	Achieved better accuracies with different models by combining textual features with auxiliary information such as user social engagements on social Media.

1.4 Problem Definition

- Problem Identified:
  - The Easy Access and exponential growth of the information available on social media networks has made it intricate to distinguish between false and true information.
  - The easy dissemination of information by way of sharing has added to exponential growth of its falsification.
- Solution Proposed:
  - With The help of ML algorithms, and the data sets we would try to eliminatethe fake news which is being spread and trying to alter the emotions of people reading news.

1.5 Scope

- The fake news challenge is perilous and is spreading rapidly like a wildfire as it becomes easier for information to reach the mass in various flavours.
- Fake news can have a huge impact in politics and thereafter on the people like a domino effect. With the help of Machine Learning & Artificial Intelligence, we can control and limit the spread of such misinformation more quickly and efficiently as compared to manual efforts.
- There are various Up’s and Down’s on the market or in people’s lives after they read the news that was Falsely claimed to be True. This project will assure them, that the news is True or False, this approach will Solve many problems like riots.

1.6 Technology stack

- Software Requirements:
- Python libraries
- Bootstrap
- Framework: Flask
- Machine-Learning Classification models
- OS Requirements: Windows XP or above
- JavaScript supported browsers

- Hardware Requirements:
- Minimum 2GB Ram
- Pentium / Intel i3 Processors or above

1.7 Benefits for Society

- In today's world, there is a great deal of information available, yet 80 percent of it is noise and only 20% of it is valuable. Our idea will serve as a forum for them to classify the news they receive as "genuine" or "fake," as well as the emotions portrayed in the news.

2. Project Design

2.1 Proposed System

The model to be made for predicting the relevance of a News can be approached through the proposed steps to be followed. They are listed below as:

- Collecting Data from datasets
- Data cleaning and Pre-processing
- Feature Extraction using TFIDF
- Providing data to Models
- Result Analysis

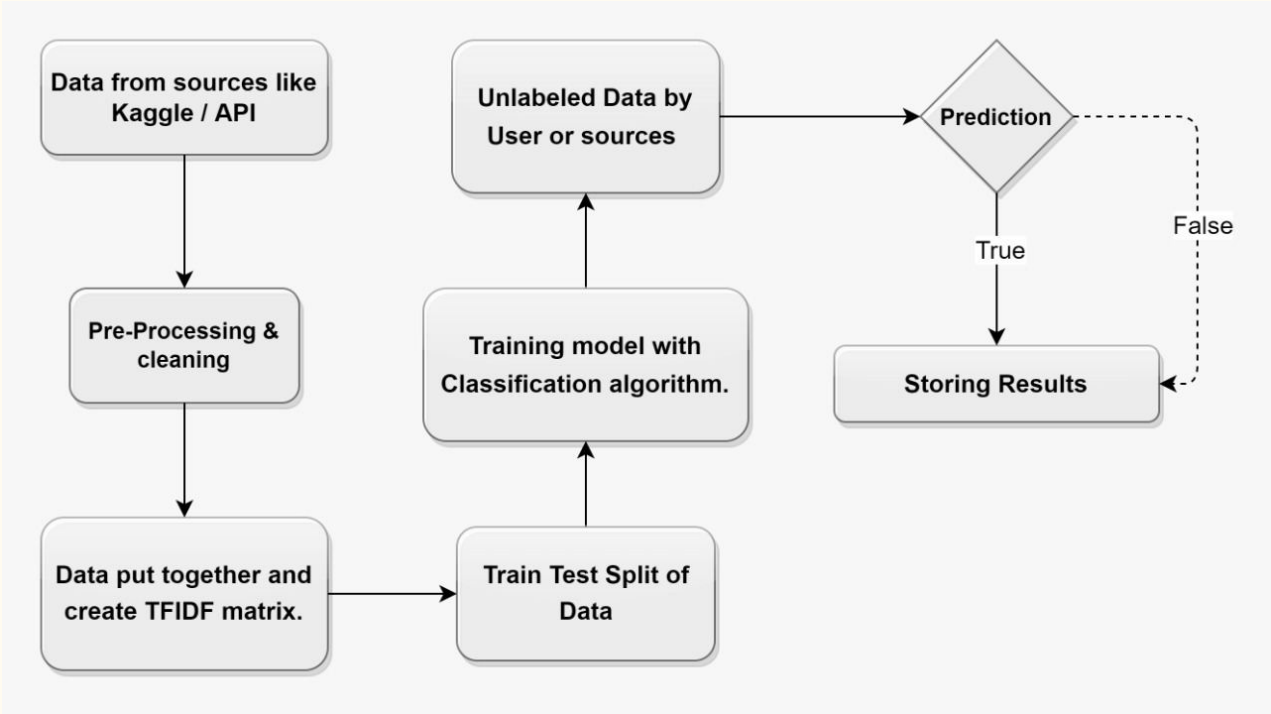
The Model is chosen based on the best accuracy providing one and the user passed data is passed through the model to predict the result. Also, the Sentiments are similarly classified into Positive, Negative and Neutral. The analysis of sentiments depends upon the:

- Methods Of NLP

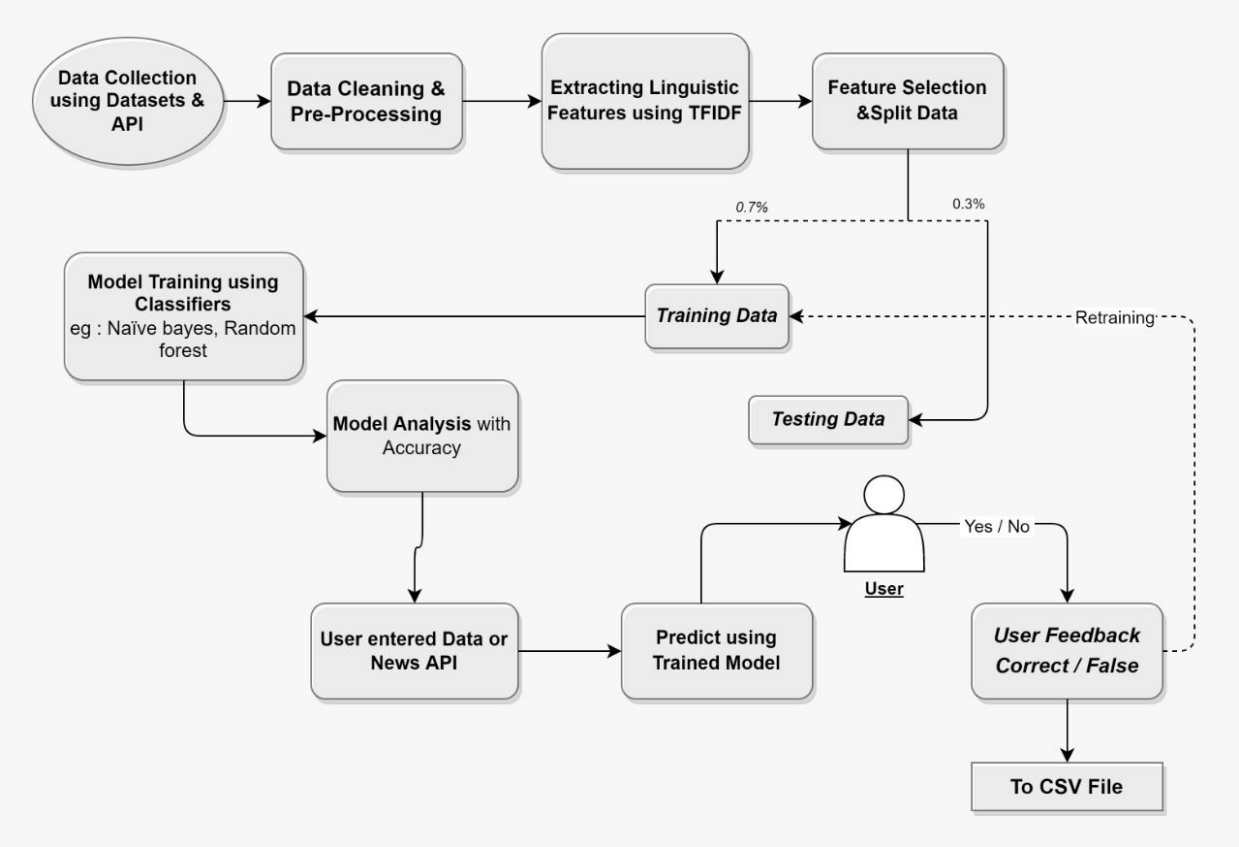
Similarly, the prediction of Fake or Not depends upon the features extracted (i.e., no. of words along with their inverse frequency converted to numeric using vectorization) using TFIDF from the combined text total obtained by either user of twitter API. The classification here depends on:

- Features present in Text.
- Probability calculation

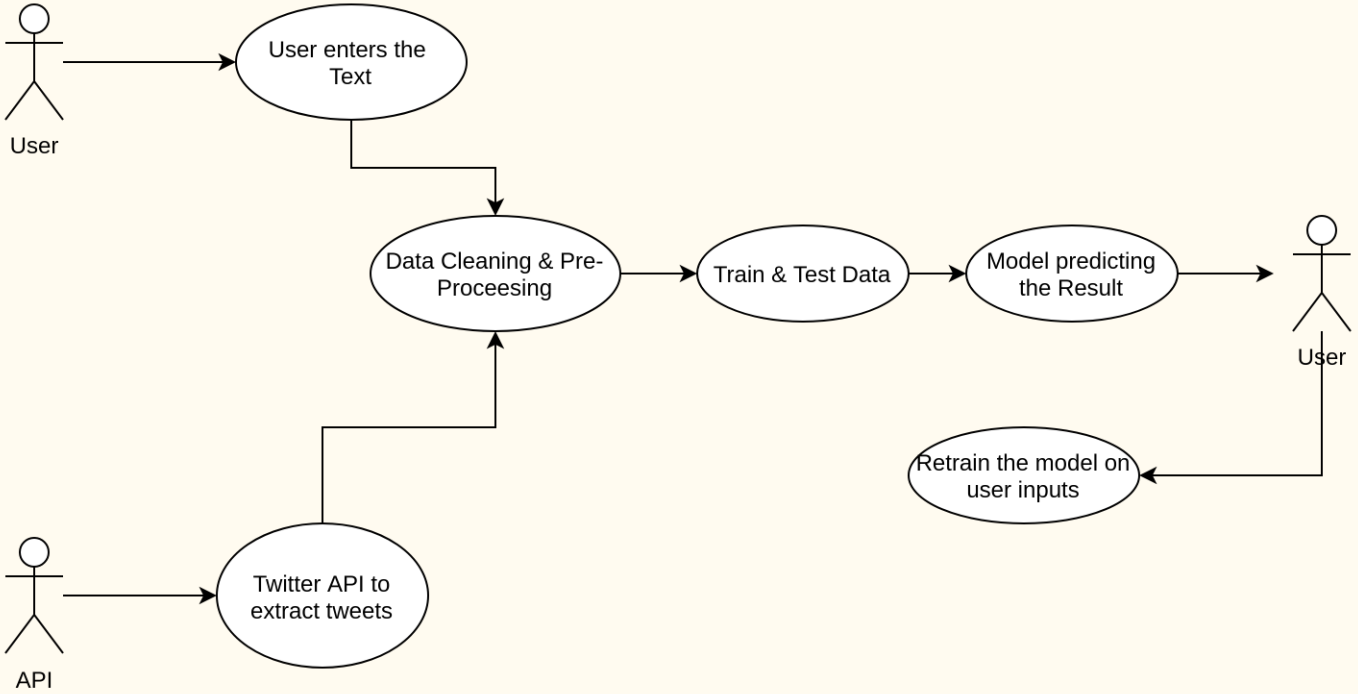
2.2 Design(Block Diagram)



2.3 Design(Flow Diagram)



2.4 Design(Use Case Diagram)



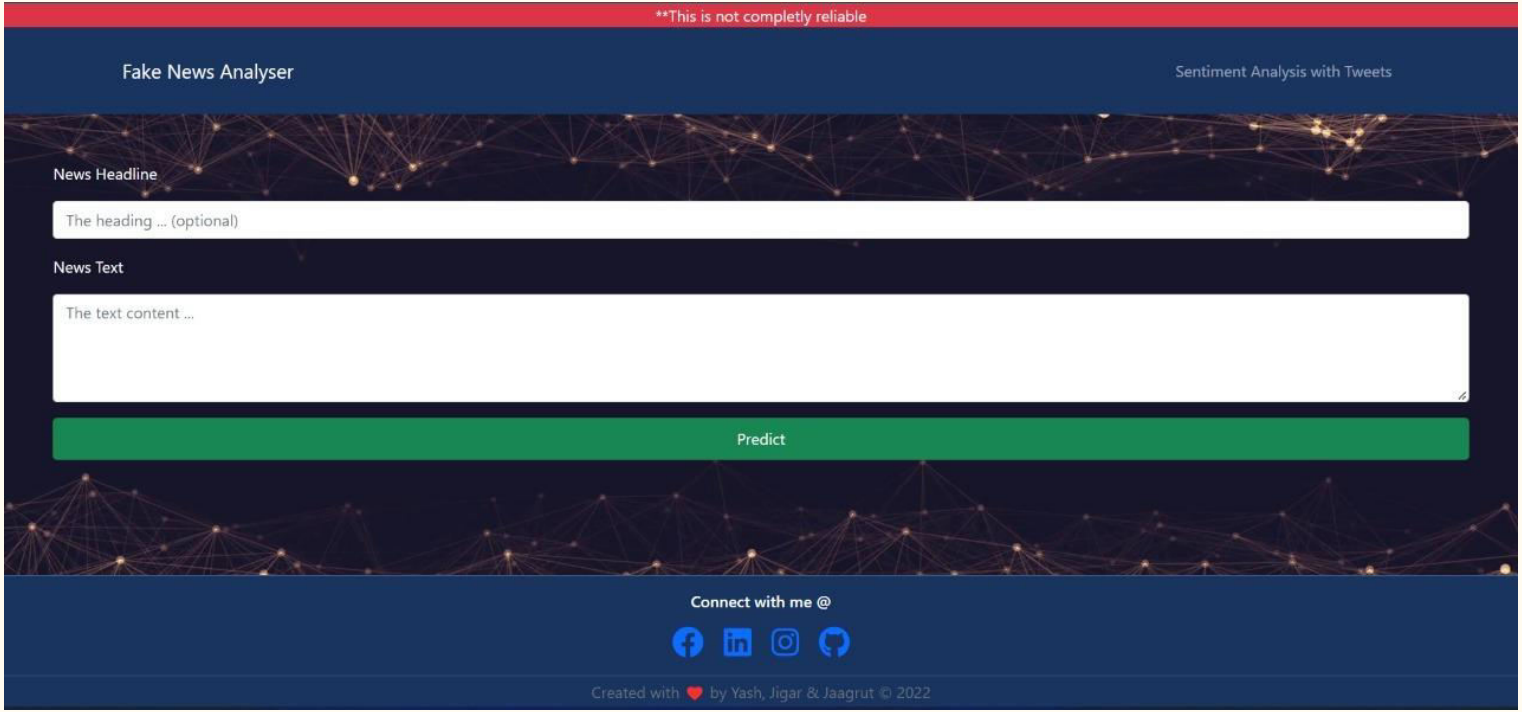
3. Implementation

3.1 Application Implementation

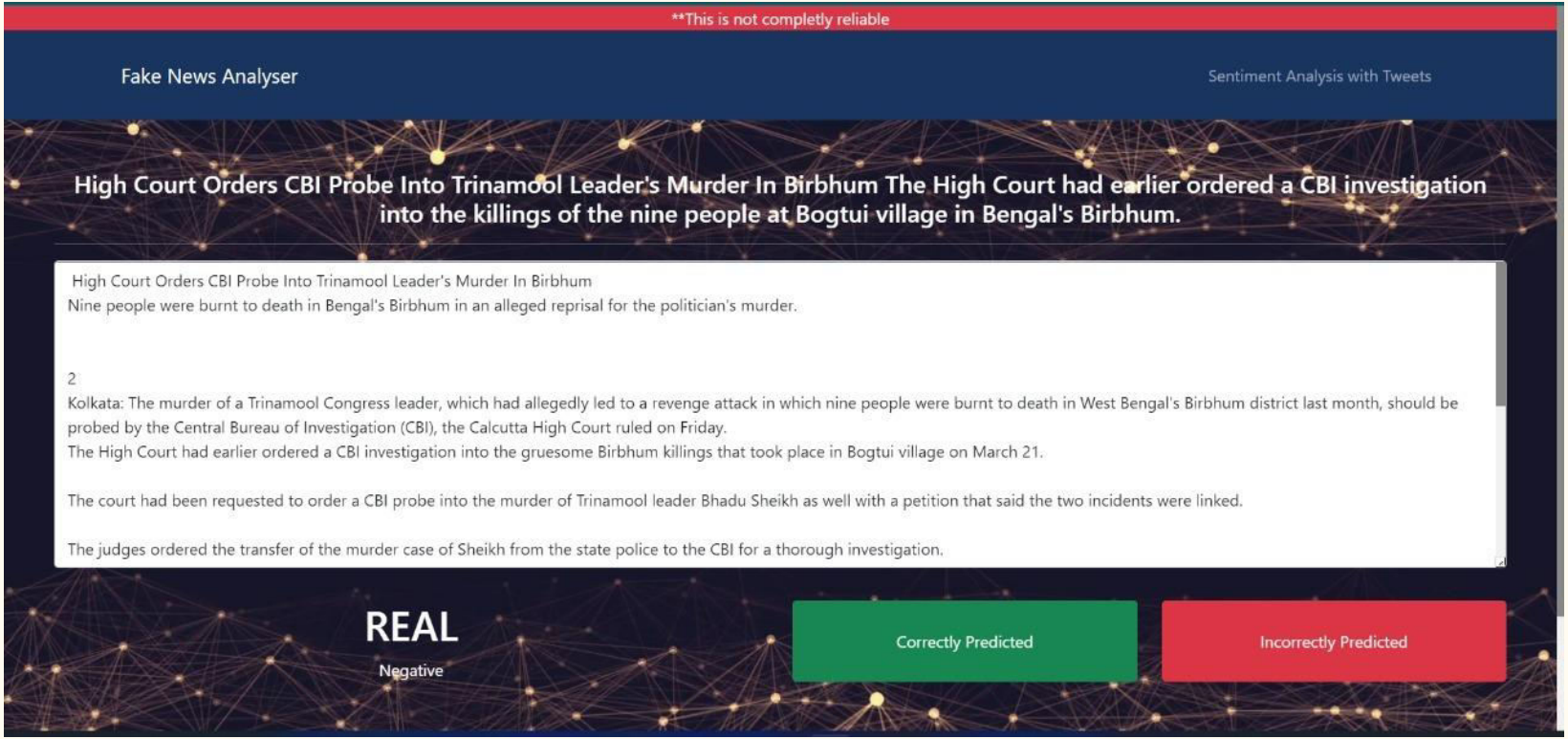
```python
@app.route('/results', methods=['POST'])
def result():

    result = request.form
    title = result['title']
    text = result['text']
    val = preprocess(title, text)

    pred = predict(val)                          #news prediction
    sentiment_res = sentimentPrediction(title)               #sentiments prediction

    return render_template('result.html', news_result=pred, sentiment_res=sentiment_res, title=title, text=text)

@app.route('/retrain', methods=['POST'])
def reTrain():

    retrain = request.form
    title = retrain['title']
    text = retrain['text']
    result = retrain['news_result']
    poll = retrain['poll']

    val = preprocess(title, text)

    #this function retrain the model if poll==correct || if poll==incorrect change the bit and then retrain
    result = 1 if result == 'REAL' else 0
    print("Before Retraining ", pipeline.predict_proba_one(val))

    if poll == 'correct':
        pipeline.learn_one(val, result)
        print("After Retraining ", pipeline.predict_proba_one(val))

    else:
        result = 0 if result == 1 else 1
        pipeline.learn_one(val, result)
        print("After Retraining ", pipeline.predict_proba_one(val))
```

4. Result

4.1 Result


**This is not completly reliable
Fake News Analyser
Sentiment Analysis with Tweets
News Headline
The heading ... (optional)
News Text
The text content ...
Predict
Connect with me @
Created with ❤ by Yash, Jigar & Jaagrut © 2022

4.2 Result


**This is not completly reliable
Fake News Analyser
Sentiment Analysis with Tweets
High Court Orders CBI Probe Into Trinamool Leader's Murder In Birbhum The High Court had earlier ordered a CBI investigation into the killings of the nine people at Bogtui village in Bengal's Birbhum.
High Court Orders CBI Probe Into Trinamool Leader's Murder In Birbhum
Nine people were burnt to death in Bengal's Birbhum in an alleged reprisal for the politician's murder.
2
Kolkata: The murder of a Trinamool Congress leader, which had allegedly led to a revenge attack in which nine people were burnt to death in West Bengal's Birbhum district last month, should be probed by the Central Bureau of Investigation (CBI), the Calcutta High Court ruled on Friday.
The High Court had earlier ordered a CBI investigation into the gruesome Birbhum killings that took place in Bogtui village on March 21.
The court had been requested to order a CBI probe into the murder of Trinamool leader Bhadu Sheikh as well with a petition that said the two incidents were linked.
The judges ordered the transfer of the murder case of Sheikh from the state police to the CBI for a thorough investigation.
REAL
Negative
Correctly Predicted
Incorrectly Predicted

4.3 Result

We give user an option to enter the text of news received, as title and the content. After they submit it through predict button the text is preprocessed and next they are passed through a pipeline where is passed to TFIDF for generating numeric metric of the text, as soon as it converted it is passed through the model which was trained earlier on the dataset. The model then provides he result as 1 or 0 where, 1 being Real and 0 as Not real. Which is then displayed to the user along with the news entered, and simultaneously in this process the sentiments are being analyzed and displayed on the same page. We have provided a feedback button on the page where the results are displayed so that the user get`s an option to provide his input over the news result. Use this input to retrain the model and save it to a CSV file.

5. Conclusion and Future Scope

Fake news is categorized as any kind of cooked-up story with an intention to deceive or to mislead. In this paper we are trying to present the solution for fake news detection task by using Machine Learning techniques. With this, endeavors are being made to automate the task of fake news detection. The most mainstream of such actions include blacklisting of sources and authors that are unreliable. Even though these tools are useful, but in order to produce a progressive complete end to end solution, we are required to represent for tougher cases where reliable sources and authors are responsible for releasing fake news. The outcomes of this project shows the capability of ML to be fruitful in this task. We have tried to build a model that helps in catching many intuitive indications of real and fake news as well as in the visualization of the classification decision.

References

- S. A. García, G. G. García, M. S. Prieto, A. J. M. Guerrero, and C. R. Jiménez, “The impact of termfake news on the scientific community scientific performance and mapping in web of science,” *Social Sciences*, vol. 9, no. 5, 2020.
- Ammara Habib, Muhammad Zubair Asghar, Adil Khan, Anam Habib, Aurangzeb Khan, “Falseinformation detection in online content and its role in decision making: a systematic literature review”,Springer-Verlag GmbH Austria, part of Springer Nature 2019.
- D. M. J. Lazer, M. A. Baum, Y. Benkler et al., “The science of fake news,” *Science*, vol. 359, no. 6380, pp. 1094–1096, 2018.
- Mykhailo Granik, Volodymyr Mesyura, “Fake News Detection using Naïve Bayes Classifier”, 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering.
- Shlok Gilda,” Evaluating Machine Learning Algorithms for Fake News Detection**”,** 2017 IEEE15th Student Conference on Research and Development.

Paper Publication

- 2nd International Conference on Advance Computing and Innovative Technologies in Engineering Conference
  Date: 28th to 29th April, 2022
  Submission Status : Need to reduce the plagiarism percentage.

- ICT4SD CONFERENCE Seventh International Conference on ICT for Sustainable Development
  Date: 26th April, 2022
  Submission Status : We must submit a new submission once we have reviewed the plagiarism percentage.
- International Conference on Educating the Millennials (ICEM)'22
  Date: 22th April, 2022
  Submission Status : The abstract has been approved, but the paper must be resubmitted for final review.

Thank You