GOVERNMENT COLLEGE OF TECHNOLOGY

COIMBATORE

(An Autonomous Institution Affiliated to Anna University)

COLLEGE CODE-7177

PROJECT TITLE: FAKE NEWS DETECTION USING NLP

TEAM MEMBERS:

SUBASH S -71772114146

ATCHAYA A -71772114104

ARUNTHATHI R -71772114103

RITHISHA P -71772114L01

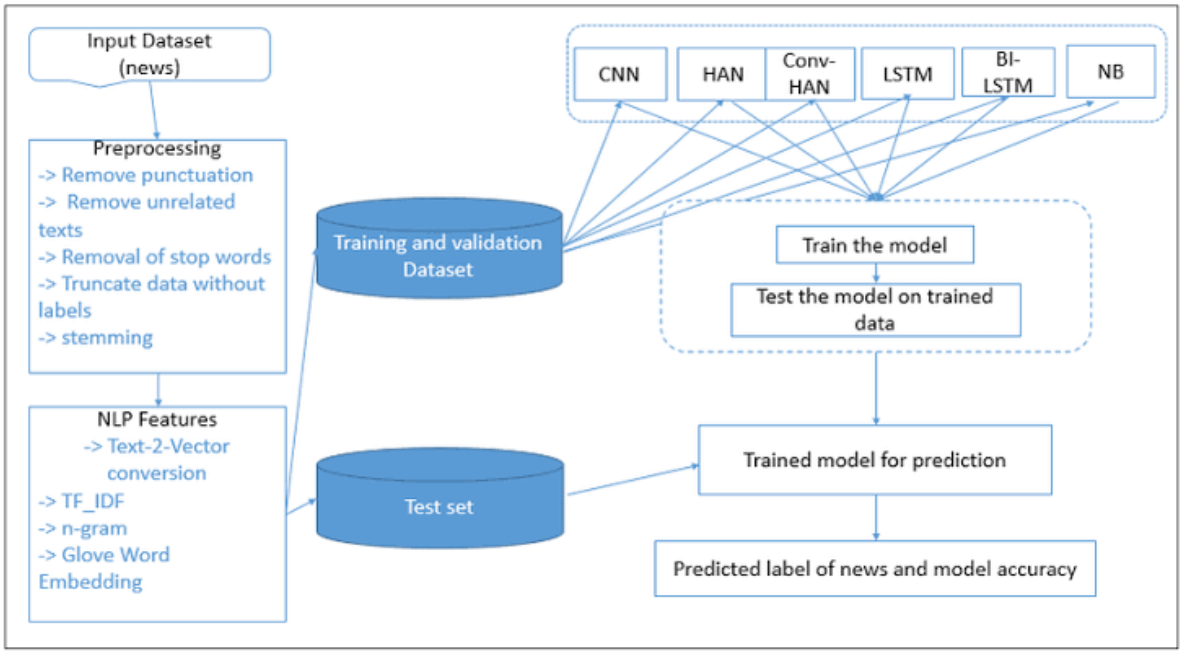
PROBLEM STATEMENT:

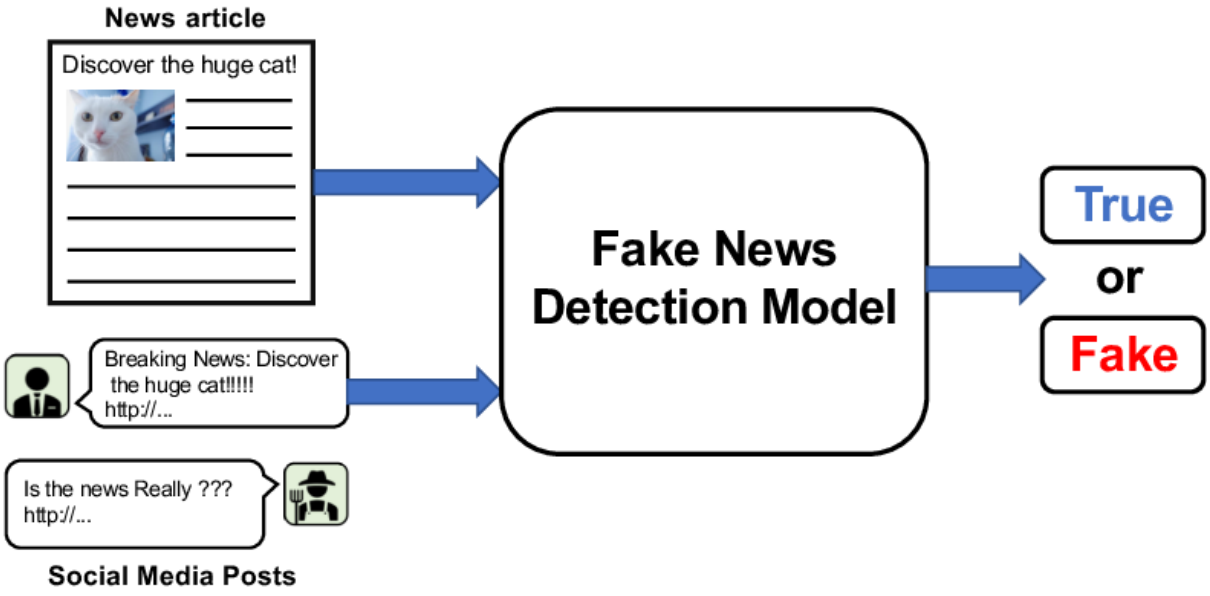
It involves developing algorithms and models to accurately identify and classify misinformation in news articles or social media posts. The goal is to create a system that can distinguish between reliable and fake news sources based on various factors such as credibility, source reputation, fact-checking, and content analysis. The challenge lies in designing effective algorithms that can analyse large amounts of data and detect patterns or indicators of fake news with high accuracy.

PROBLEM SOLUTION :

- To solve the problem of fake news detection, various approaches can be taken.
- One solution involves using natural language processing (NLP) techniques to analyse the content of news articles and identify patterns that indicate misinformation.
- This can include analysing the language used, fact-checking claims, and identifying biased or misleading information.
- Additionally, machine learning algorithms can be trained on labelled datasets to classify news articles as either reliable or fake based on features such as source credibility, writing style, and social media engagement.
- Regular updates and improvements to the algorithms are crucial to keep up with evolving techniques used by those spreading fake news.
- To tackle the problem of fake news detection is by leveraging the power of crowdsourcing.
- Platforms can engage users to report and flag suspicious or misleading content, which can then be reviewed by fact-checkers or community moderators.
- This collaborative approach helps in identifying and verifying the accuracy of news articles through collective efforts.
- Additionally, educating users about media literacy and critical thinking skills can empower them to identify and question sources of fake news, thereby reducing its impact.

PROCESS:

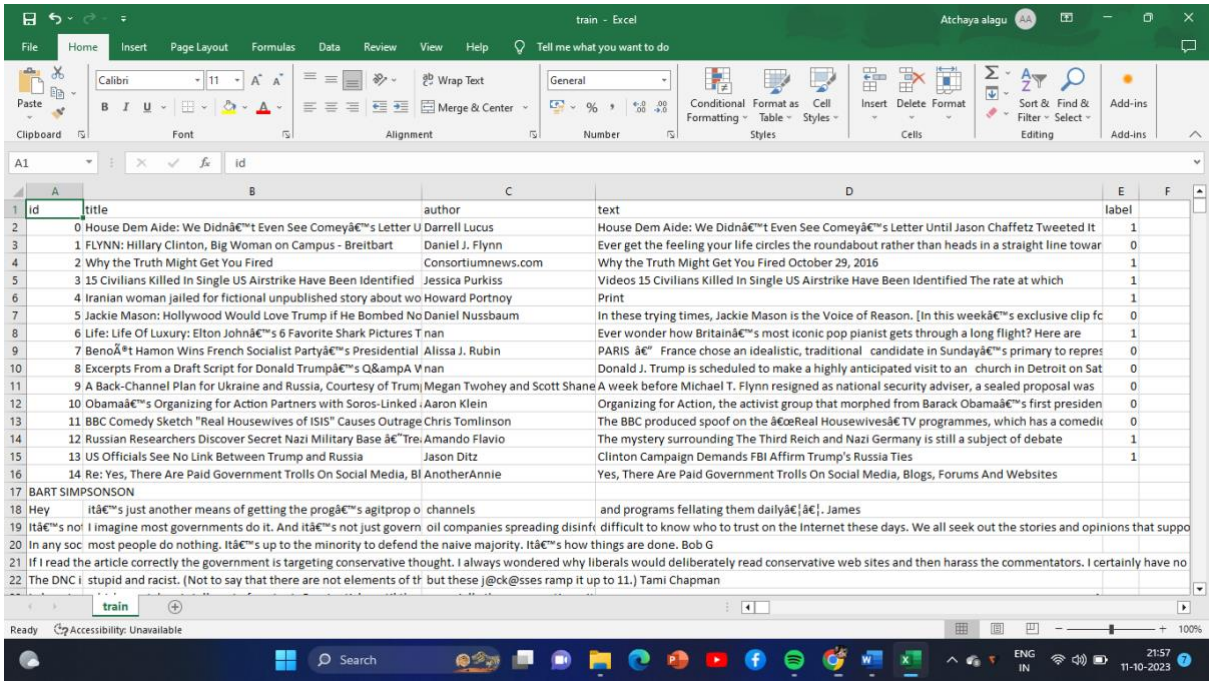




DATASET:

The dataset is taken from

https://www.kaggle.com/c/fake-news/data



id	title	author	text	label
0	House Dem Aide: We Didnâ€™t Even See Comeyâ€™s Letter U	Darrell Lucus	House Dem Aide: We Didnâ€™t Even See Comeyâ€™s Letter Until Jason Chaffetz Tweeted It	1
1	FLYNN: Hillary Clinton, Big Woman on Campus - Breitbart	Daniel J. Flynn	Ever get the feeling your life circles the roundabout rather than heads in a straight line towar	0
2	Why the Truth Might Get You Fired	Consortiumnews.com	Why the Truth Might Get You Fired October 29, 2016	1
3	15 Civilians Killed In Single US Airstrike Have Been Identified	Jessica Purkiss	Videos 15 Civilians Killed In Single US Airstrike Have Been Identified The rate at which	1
4	Iranian woman jailed for fictional unpublished story about wo	Howard Portnoy	Print	1
5	Jackie Mason: Hollywood Would Love Trump if He Bombed No	Daniel Nussbaum	In these trying times, Jackie Mason is the Voice of Reason. [In this weekâ€™s exclusive clip fc	0
6	Life: Life Of Luxury: Elton Johnâ€™s 6 Favorite Shark Pictures T	nan	Ever wonder how Britainâ€™s most iconic pop pianist gets through a long flight? Here are	1
7	BenoÃ®t Hamon Wins French Socialist Partyâ€™s Presidential	Alissa J. Rubin	PARIS â€" France chose an idealistic, traditional candidate in Sundayâ€™s primary to repres	0
8	Excerpts From a Draft Script for Donald Trumpâ€™s Q&ampA V	nan	Donald J. Trump is scheduled to make a highly anticipated visit to an church in Detroit on Sat	0
9	A Back-Channel Plan for Ukraine and Russia, Courtesy of Trum	Megan Twohey and Scott Shane	A week before Michael T. Flynn resigned as national security adviser, a sealed proposal was	0
10	Obamaâ€™s Organizing for Action Partners with Soros-Linked	Aaron Klein	Organizing for Action, the activist group that morphed from Barack Obamaâ€™s first presiden	0
11	BBC Comedy Sketch "Real Housewives of ISIS" Causes Outrage	Chris Tomlinson	The BBC produced spoof on the â€œReal Housewivesâ€ TV programmes, which has a comedic	0
12	Russian Researchers Discover Secret Nazi Military Base â€"Tre	Amando Flavio	The mystery surrounding The Third Reich and Nazi Germany is still a subject of debate	1
13	US Officials See No Link Between Trump and Russia	Jason Ditz	Clinton Campaign Demands FBI Affirm Trump's Russia Ties	1
14	Re: Yes, There Are Paid Government Trolls On Social Media, Bl	AnotherAnnie	Yes, There Are Paid Government Trolls On Social Media, Blogs, Forums And Websites	
BART SIMPSONSON				
Hey	itâ€™s just another means of getting the progâ€™s agitprop o	channels	and programs fellating them dailyâ€¦â€¦. James	
Itâ€™s not	I imagine most governments do it. And itâ€™s not just govern	oil companies spreading disinfo	difficult to know who to trust on the Internet these days. We all seek out the stories and opinions that suppo	
In any soc	most people do nothing. Itâ€™s up to the minority to defend the naive majority. Itâ€™s how things are done. Bob G			
If I read the	article correctly the government is targeting conservative thought. I always wondered why liberals would deliberately read conservative web sites and then harass the commentators. I certainly have no			
The DNC i	stupid and racist. (Not to say that there are not elements of th	but these j@ck@sses ramp it up to 11.) Tami Chapman		

SOFTWARE USED:

Python – 3.X

LIBRARIES USED:

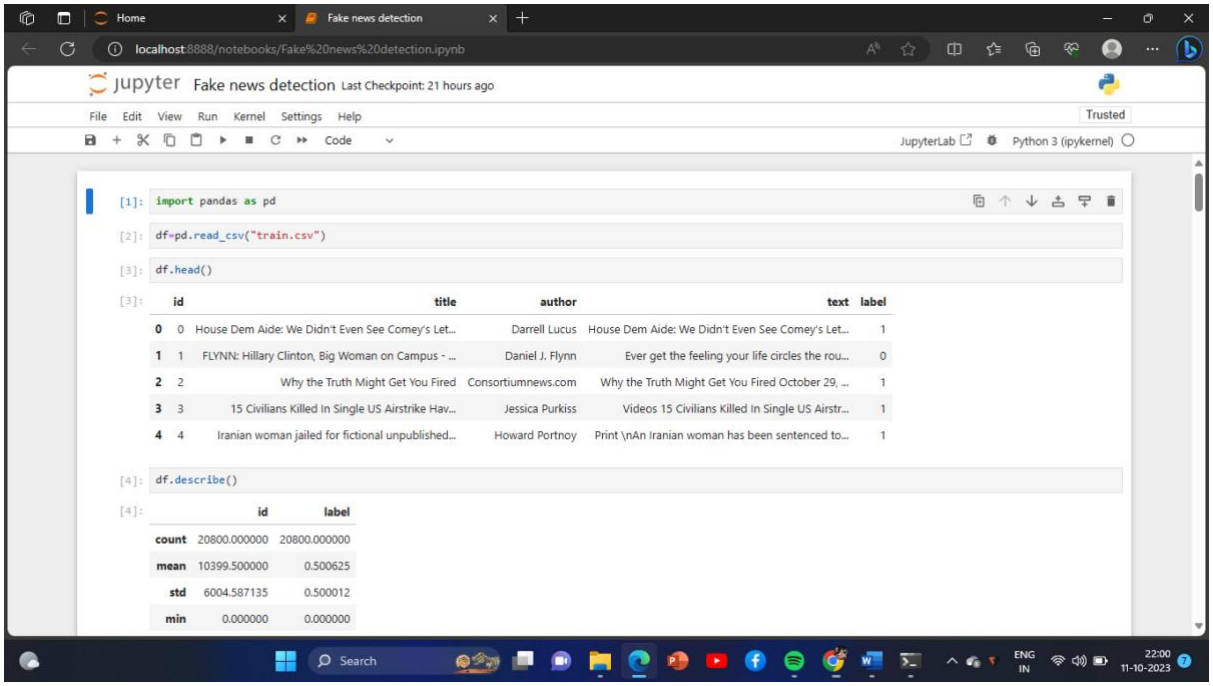
- Pandas
- Scikit – learn
- Nltk (Natural Language Toolkit)

IMPORTING LIBRARIES:

```
import pandas as pd

import numpy as np

import nltk

from nltk.corpus import stopwords

from nltk.stem import PorterStemmer
```

```
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.model_selection import train_test_split
```

JUPYTER NOTEBOOK:



PROGRAM:

```
port_stem = PorterStemmer()
vectorization = TfidfVectorizer()

vector_form = pickle.load(open('vector.pkl', 'rb'))
load_model = pickle.load(open('model.pkl', 'rb'))

def stemming(content):
    con=re.sub('[^a-zA-Z]', ' ', content)
```

```
    con=con.lower()
    con=con.split()
    con=[port_stem.stem(word) for word in con if not word in
stopwords.words('english')]
    con=' '.join(con)
    return con


def fake_news(news):
    news=stemming(news)
    input_data=[news]
    vector_form1=vector_form.transform(input_data)
    prediction = load_model.predict(vector_form1)
    return prediction




if __name__ == '__main__':
    st.title('Fake News Classification app ')
    st.subheader("Input the News content below")
    sentence = st.text_area("Enter your news content here",
"",height=200)
    predict_btt = st.button("predict")
    if predict_btt:
        prediction_class=fake_news(sentence)
        print(prediction_class)
```

```
if prediction_class == [0]:
    st.success('Reliable')
if prediction_class == [1]:
    st.warning('Unreliable')
```

UP-TO-DATE:

- Get data from BuzzFeed news and politifact.
- Extend the dataset using data from social media.
- Detect communities around fake news posts and real news posts.
- Analyse our results to detect unusual user behaviour.

CONCLUSION:

After finishing all these process run this file and make api server with local host then paste the news . The compilation process will check whether the entered news is reliable or not.