PROJECT-REPORT

DEPLOYING A FAKE NEWS DETECTION MODEL USING NLP ON HEROKU

NAME : L.SATYAJIT

REG NO : 18BCD7143

Guided by : Dr. Karthikeyan S

VIT-AP UNIVERSITY

Outline:

Abstract

- Social media has become a tool to spread false information with the help of its large complex network. The consequences of such misinformation could be very severe.

- Many Social Media companies want to curb these false information and block the urls of these fake news spreaders.

- Nowadays with the emergence of Natural Language Processing techniques we can analyze huge volumes of text to figure out patterns to help us detect the authenticity of a certain news article.

Introduction

- A purposely misleading story is called a "Fake News".

- They can be used for various negative ramifications in a variety of domains

- I will attempt to use artificial intelligence technologies particularly machine learning techniques and natural language processing (NLP) to automatically detect fake news and prevent its viral spread using the huge volume of news data available online

Modules

- Data Collection

- Text Preprocessing

- Prediction

- Storing Results
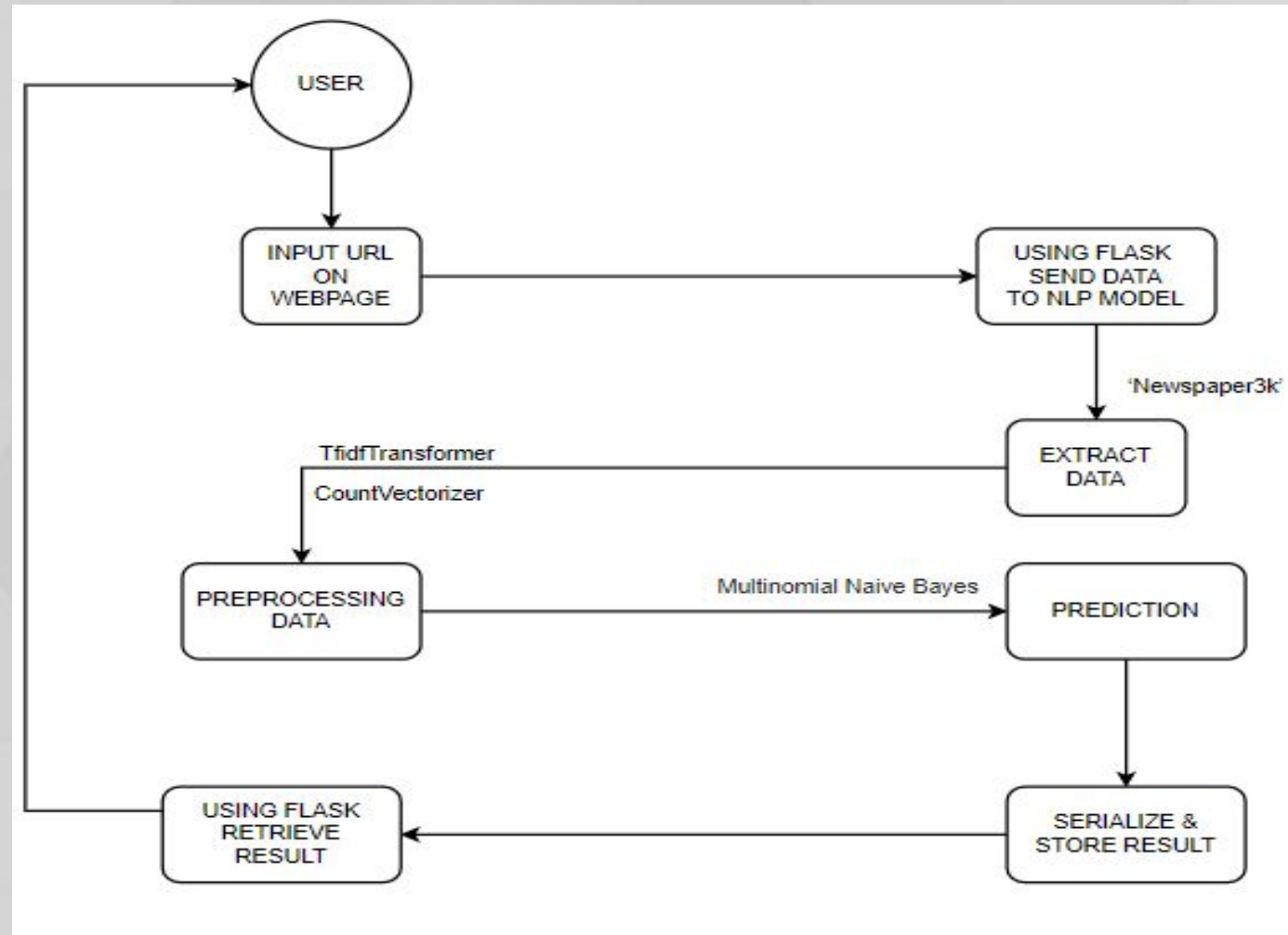
- Showing Outcome

Software Requirements

- Model - **SAAS**
- Front End: Html,CSS
- Back End: Python
- Frameworks/API used: FLASK
- Name of the Cloud tool: Heroku CLI,PyCaret
- Name of the Cloud Service: Heroku

Methodology

- Model deployed on Heroku cloud platform using github repository

- Initially use 'Newspaper3k' library to extract the news article from the news url provided by user

- Preprocess the extracted data using TfidfTransformer and CountVectorizer

- Predicting whether the data is real or fake using the Multinomial Naive Bayes model

- Serializing and storing the predicted outcome

- Using Flask and Html to display the predicted outcome

Flow Chart

Work Progress

- Data Collection - Completed

- Developing NLP Model - Completed

- Flask API - Completed

- Deployment on Heroku - Completed

	title	text	label
8476	You Can Smell Hillaryâ€™s Fear	Daniel	FAKE
10294	Watch The Exact Moment Paul Ryan Com	Google	FAKE
3608	Kerry to go to Paris in gesture of sympath	U.S.	REAL
10142	Bernie supporters on Twitter erupt in ang	â€"	FAKE
875	The Battle of New York: Why This Primar	It's	REAL
6903	Tehran, USA		FAKE
7341	Girl Horrified At What She Watches Boyfr	Share	FAKE
95	â€˜Britainâ€™s Schindlerâ€™ Dies at 106	A Czech st	REAL
4869	Fact check: Trump and Clinton at the 'con	Hillary	REAL
2909	Iran reportedly makes new push for urani	Iranian	REAL
1357	With all three Clintons in Iowa, a glimpse	CEDAR	REAL
988	Donald Trumpâ€™s Shockingly Weak Dele	Donald	REAL
7041	Strong Solar Storm, Tech Risks Today | SO	Click Here	FAKE

```
C:\Users\CaptainSatyu7\PycharmProjects\FakeNews\venv\Scr
              precision    recall   f1-score    support

      FAKE        0.95       0.70      0.81        627
      REAL        0.77       0.97      0.86        640


  accuracy                             0.84       1267
 macro avg        0.86       0.83      0.83       1267
weighted avg      0.86       0.84      0.83       1267


[[441 186]
 [ 22 618]]


Process finished with exit code 0
```

Demo

Demo

News Classifier

https://www.manchestereveningnews.co.uk/sport/football/transfer-

Predict

News Classifier

Enter the news url

Predict

This News is "FAKE"

Demo

News Articles Used :
https://www.manchestereveningnews.co.uk/sport/football/transfer-news/cristiano-ronaldo-man-utd-transfer-20376729 **(FAKE)**
https://zapatopi.net/treeoctopus/ **(FAKE)**
https://www.washingtonpost.com/politics/2021/04/13/trump-rnc-audio-excerpts/ **(REAL)**
https://www.wionews.com/world/florida-governor-tightens-voter-law-in-appeal-to-trump-base-383370 **(REAL)**

Cloud Site Link:
https://spoofnewsdetection.herokuapp.com/predict
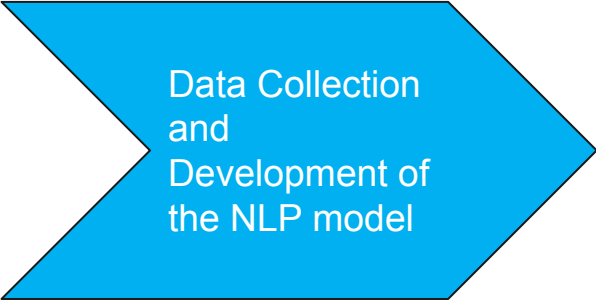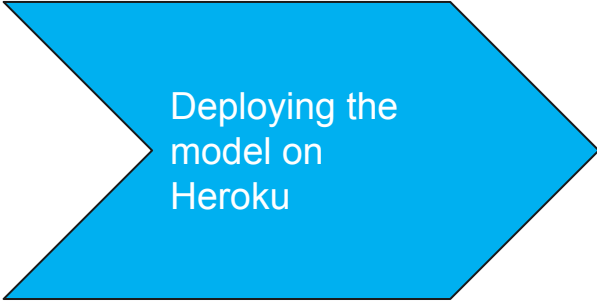
Timeline

April

Data Collection and Development of the NLP model

May

Building the Web API using Flask and Html

June

Deploying the model on Heroku

References

- Dataset- https://www.kaggle.com/hassanamin/textdb3

- https://flask.palletsprojects.com/en/1.0.x/api/#flask.Flask

- https://newspaper.readthedocs.io/en/latest/

- https://web.stanford.edu/~jurafsky/slp3/slides/7_NB.pdf

- https://devcenter.heroku.com/articles/heroku-cli

- https://www.hindawi.com/journals/complexity/2020/8885861/

- https://www.kaggle.com/learn/natural-language-processing

- https://codepen.io/frytyler/pen/EGdtg

Thank You