|  |  |  |
| --- | --- | --- |
| **Tech Saksham**  Final Project Report  **Track Name** |  |  |

**DETECTION OF FAKE NEWS”**

**MIT ARTS AND SCIENCE FOR WOMAN”**

|  |  |
| --- | --- |
| **ROLL NO** | **NAME** |
| CB20S262227 | ABINAYA R |
| CB20S262241 | SNEHA D |
| CB20S262243 | VANISHREE P |
| CB20S262244 | VISHNUPRIYA K |
|  |  |

|  |  |
| --- | --- |
|  | MAYANK SHRIVASTAVA |
|  | Trainer Name |
|  | Master Trainer |

**ABSTRACT**

**Do you trust all the news you hear from social media?**

**All news are not real, right?**

**How will you detect fake news?**

**The answer is Python. By practicing this advanced python project of detecting fake news, you will easily make a difference between real and fake news.**

**Before moving ahead in this machine learning project, get aware of the terms related to it like fake news, tfidfvectorizer, PassiveAggressive Classifier.**

**Also, I like to add that DataFlair has published a series of machine learning Projects where you will get interesting and open-source advanced ml projects. Do check, and then share your experience through comments**

**INDEX**

|  |  |  |
| --- | --- | --- |
| **Sr. No.** | **Table of Contents** | **Page No.** |
| 1 | Chapter 1: Introduction |  |
| 2 | Chapter 2: Services and Tools Required |  |
| 3 | Chapter 3: Project Architecture |  |
| 4 | Chapter 4: Architecture Blocks Detail Working |  |
| 5 | Conclusion |  |
| 6 | References |  |
| 7 | Code |  |

**CHAPTER 1**

**INTRODUCTION**

**Today, we learned to detect fake news with Python. We took a political dataset, implemented a TfidfVectorizer, initialized a PassiveAggressiveClassifier, and fit our model. We ended up obtaining an accuracy of 92.82% in magnitude.**

**Hope you enjoyed the fake news detection python project. Keep visiting DataFlair for more interesting python, data science, and machine learning projects.**

**Fake News**

**A type of yellow journalism, fake news encapsulates pieces of news that may be hoaxes and is generally spread through social media and other online media. This is often done to further or impose certain ideas and is often achieved with political agendas. Such news items may contain false and/or exaggerated claims, and may end up being viralized by algorithms, and users may**

**CHAPTER 2**

**SERVICES AND TOOLS REQUIRED**

**2.1 Services Used**

**2.1.1 Liberty Profile**

**2.2 Tools and Softwares used**

**2.2.1 Jupiter node**

**2.2.2 python**

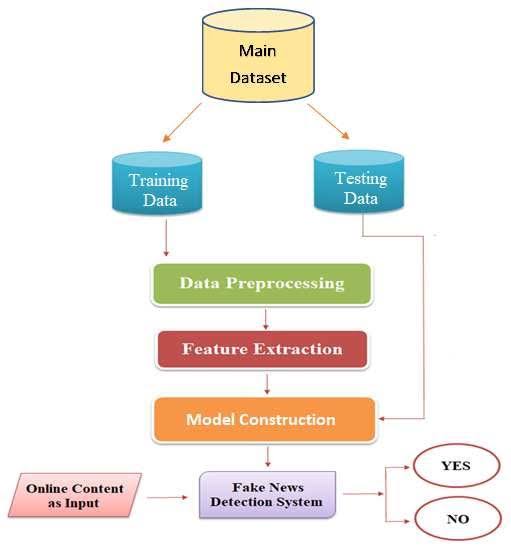
**2.2.3 Cloud Foundry**

**CHAPTER 3**

**PROJECT ARCHITECTURE**

**3.1 Architecture**

|  |
| --- |
|  |



**CHAPTER 4**

**ARCHITECTURE BLOCKS DETAIL WORKING**

14 A Comparative Study of Computational Fake News Detection on Social MediaFig. 4: A General Fake News Detection Architecture.•Feature Extraction: This components aims to create a formal mathematicalstructure in order to represent news content and related auxiliary data thatis helpful to construct fake news detection model. The feature extractioncomponent performs important NLP functions including (i). bag-of-words(it is a method of counting the number of times a word appears in a docu-ment which is helpful to compare documents and assess their similarities forthe required applications including search, classifying documents and topicmodelling [72]), (ii). n-grams (it is an n-item contiguous sequence createdfrom a given text sample which is helpful to extract text corpus features [73]for building fake news detection model), (iii). TF-IDF weighting (TF-IDF is an acronym for Term-Frequency-Inverse Document Frequency whichis a powerful technique of determining the topic of an article based on thewords it includes that measures relevance, but not frequency [74] and it hashelped to create popular and valuable tools such as Google Search [75]),(iv). word2Vec (it is a two-layer neural network that ”vectorizes” words

**CONCLUSION**

Through utilization of different kinds of Machine Learning Algorithms, this paper is aimed to exploit different aspects of dataset which has not been deeply considered in literature and to find a good way of detection of the fake and automated accounts. In this paper we have presented a Machine Learning pipeline for detecting fake accounts in online social networks. Rather than making a prediction using one single algorithm, our system uses three different classification algorithms to determine whether or not an account in the provided dataset is a fake account or not. Our evaluation using Support Vector Machine, Random Forest and Neural Networks showed strong performance, and the comparison of the accuracy of prediction seemed to be higher using Support Vector Machine for the given dataset. The Accuracy of detecting fake accounts is found to be higher using Random Forest Algorithm followed by Neural Networks Algorithm for a given dataset. As a future work,[5] recurrent neural networks can be utilized for the time series user data for a better detection of fake accounts and the algorithms can be applied to various social online platforms such as Instagram, LinkedIn and Twitter to detect the fake accounts.

**REFERENCES**

**CODE**

**Please Provide Code through Git Hub Repo Link**

**import numpy as np**

**import pandas as pd**

**import itertools**

**from sklearn.model\_selection import train\_test\_split**

**from sklearn.feature\_extraction.text import TfidfVectorizer**

**from sklearn.linear\_model import PassiveAggressiveClassifier**

**from sklearn.metrics import accuracy\_score, confusion\_matrix**