(Fake News Detection)

Submitted in partial fulfillment of the requirements

of the Mini-Project 1/2 for Final Year of

Bachelors of Engineering

by

Ansh Tiwari\_201P019

Ali Mo Salman\_201P048

Faizan\_201P001

Suchit sawant\_201P040

Guide:

Prof. Ramaya Ma’am



Department of Computer Engineering

Rizvi College of Engineering



University of Mumbai

2022-2023

**CERTIFICATE**

This is to certify that the mini-project entitled **“Fake News Detection”** is a bonafide work of “Ansh Tiwari (201P060), Suchit (201P014), Ali Mo Salman (201P048), Shikalgar Faizan (201P001)”submitted to the University of Mumbai in partial fulfillment of the requirement for the Mini-Project for the final year of the Bachelor of Engineeringin **“Computer Engineering”**.

Prof. Ramaya Ma’a**m**

**Guide**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_

Prof. Shiburaj Pappu Dr. Varsha Shah

**Head of Department** **Principal**

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited or from whom proper permission has not been taken when needed.

-----------------------------------------

(Signature)

-----------------------------------------

(Ansh Tiwari\_201P019)

(Suchit sawant\_201P0140)

(Ali Mo Salman\_201P048)

(Shikalgar Faizan\_201P001)

Date:

**ABSTRACT**

This project is designed to be used by FAKE NEWS DETECTION Recently, fake news has been incurring many problems in our society. As a result, many researchers have been working on identifying fake news. Most of the fake news detection systems utilize the linguistic features of the news.

However, they have difficulty sensing highly ambiguous fake news which can be detected only after identifying the meaning and latest related information. Our system receives a proposition and searches the semantically related articles from Fact DB to verify whether the given proposition is true or not by comparing the proposition with the related articles DB. To achieve this, we utilize a deep learning model, Bidirectional Multi-Perspective Matching for Natural Language Sentences (BiMPM), which has demonstrated a good performance for the sentence matching task.

Index

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sr. No | Title | | | Page No |
| 1. | Introduction | | | 3 |
| 2. | Review and Literature | | | 2 |
|  | 2.1. | Paper 1 | | 3 |
|  | 2.2. | Paper 2 | | 4 |
| 3. | Theory, Methodology, and Algorithm | | | 5 |
|  | 3.1 | Section | | 6 |
|  |  | 3.1.1. | Subsection | 7 |
| 4. | Results and Discussions | | | 8 |
| 5. | Conclusion | | | 9 |
| 6. | References | | | 10 |
|  | Appendix | | | 11 |
|  | Acknowledgment | | | 12 |
|  | Publication | | | 13 |

**Chapter 1**

**Introduction**

The term “Fake News” was a lot less unheard of and not prevalent a couple of decades ago but in this digital era of social media, it has surfaced as a huge monster. Fake news, information

bubbles, news manipulation and the lack of trust in the media are growing problems within our society. However, to start addressing this problem, an in-depth understanding of fake news and its origins is required. Only then one can look into the different techniques and fields of machine learning (ML), natural language processing (NLP), and artificial intelligence (AI) that could help us fight this situation.“Fake news”, natural language processing (NLP) and artificial intelligence (AI) that could help us fight this situation.“Fake news” has been used in a multitude of ways in the last half a year and multiple definitions have been given. For instance, the New York Times defines it as “a made-up story to deceive”.Measuring fake news or even defining it properly could very quickly become a subjective matter, rather than an objective metric. In its purest form, fake news is completely made up, and manipulated to resemble credible journalism and attract maximum attention and, with it, advertising revenue [2]. Despite all these shortcomings, several entities have tried to categorize fake news in different manners

* 1. **Relevant Issues /Problem Identification**

Social media for news consumption is a double-edged sword. On the one hand, its low cost, easy access, and rapid dissemination of information lead people to seek out and consume news from social media. On the other hand, it enables the wide spread of “fake news”, i.e., low quality news with intentionally false information. The extensive spread of fake news has the potential for extremely negative impacts on individuals and society. Therefore, fake news detection on social media has recently become an emerging research that is attracting tremendous attention. Fake news detection on social media presents unique characteristics and challenges that make existing detection algorithms from traditional news media ineffective or not applicable. First, fake news isintentionally written to mislead readers to believe false information, which makes it difficult and nontrivial to detect based on news content; therefore, we need to include auxiliary information, such as user social engagements on social media, to help make a determination. Second, exploiting this auxiliary information is challenging in and of itself as users’ social engagements with fake news produce data that is big, incomplete, unstructured, and noisy

* 1. **Task Identification**

The approaches used to identify fake news. In order to do this, a systematic literature review

was done. This section presents the search terms that were used, the selection criteria and the

source selection.Our job now was to build an AI logic that can run on any modern Internet browser, unlike common AI facial manipulators which need a lot of system resources to run and can hardly ever run on a mobile phone’s web browser...

* 1. **Known Risks**

A major risk that we were well aware of was related to the running environment of the AI.

As we know, Python has dominated most of the AI related operation field in modern times.

Unfortunately, none of the existing web-browsers can interpret pure Python codes.

The initiation and spread of fake news presents significant risks from many different

perspectives, including from a national security standpoint. A good example of this is

deliberately misleading news that attempt to influence an individual’s perception about another

individual or election results. In politically divided environments, such as those being witnessed

in the US and Europe, people tend to gravitate towards news from sources that are congenial to

their belief or political taste.

**Chapter 2: Literature Survey**

**2.1 SPACY V3.JS: MACHINE LEARNING FOR THE WEB AND BEYOND**

**Author:** R. Alekya1 , Neelima Devi Boddeti2 , K. Salomi Monica3 , Dr.R. Prabha4

The enabling of portability of models between Python and JavaScript is what did the trick for us in the first place.The spread of fake news is unstoppable with the adoption of different social networks. On Twitter, Facebook, Reddit, people take advantage of fake news to spread rumours, win political benefits and click rates.Detecting fake news is critical for a healthy society, and there are multiple different approaches to detect fake news. From a machine learning standpoint, fake news detection is a binary classification problem; hence we can use traditional classification methods or state-of-the-art Neural Networks to deal with this problem.The library is part of the SpaCy v3 ecosystem, providing a set of APIs that are compatible with those in Python, allowing models to be ported between the Python and JavaScript ecosystems.

**2.1.1 The JavaScript environment**

One of the challenges of JS is that it runs in different environments. Computation can happen

client-side in a browser, server-side, most notably as part of the Node.js framework, and more

recently on the desktop. Machine Learning for the Web and Beyond via frameworks like

Electron. Although the majority of our work to date has been tuning it for client-side

development in a web browser.

2.1.2 Performance.

A second key challenge, specific to the browser environment, is performance. JS is an interpreted language so it does not typically match the speed of a compiled language like C++ or Java for

numerical computation. Unlike Python which can bind to C++ libraries, browsers do not expose this capability. For security reasons, browser applications don’t have direct access to the GPU, which is typically where numerical computation happens for modern deep learning systems.

To address these performance issues, a few new JS standards are emerging. Most modern

browsers also support WebGL (Kronos, 2011), an API that exposes OpenGL to JS. OpenGL is a cross-language, cross-platform API for rendering 2D and 3D vector graphics, enabling games

and other high-performance rendering tasks directly in a webpage.

**2.2 Implementation of an Fake News Detection Based on Web and SpaCy v3**

**Author:** K. Siva Sankar1, G. Joga Rao2, B. Ravi3, B. Alekhya4, P. Sampathkumar5, D. Karthik Ganesh6, Rakesh. E7

Health has prime importance in our day-to-day life. Sound health is necessary to do the daily work properly.

This project aims at developing a system which gives body temperature and heart rate using LM35 and pulse sensor respectively. These sensors are interfaced with controller node MCU. Wireless data transmission done by node MCU through wifi module. ESP8266 is used for wireless data transmission on IoT platform. Data visualization is done on Blynk App. So that record of data can be stored over period of time .This data stored on web server so that it can seen to who logged.

**Chapter \_**

**Results and Discussions**

**Chapter \_05**

**Conclusions**

IDF. In the future, the models can be applied on different datasets with more documents, and hybridization between more than one classifier can enhance the accuracy of the classification

As a conclusion, In the 21st century, the majority of the tasks are done online. Newspapers that were earlier preferred as hard copies are now being substituted by applications like Facebook, and Twitter, and news articles to be read online. WhatsApp’s forwards are also a major source. The growing problem of fake news only makes things more complicated and tries to change or hamper the opinion and attitudes of people toward use of digital technology. When a person is deceived by the real news two possible things happen- People start believing that their perceptions about a particular topic are true as assumed. Thus, to curb the phenomenon, we have developed our Fake News Detection system that takes input from

the user and classify it to be true or fake. The model is trained using an appropriate dataset and performance evaluation is also done using various performance measures. the model with the highest accuracy is used to classify the news headlines or articles. As evident above for static search, our best model came out to be Logistic Regression with an accuracy of 65%. Hence we then used grid search parameter optimization to increase the performance of logistic regression which then gave us an accuracy of 75%. Hence we can say that if a user feeds a particular news article or its headline in our model, there are 75% chance that it will be classified to its true nature. The user can check the news article or keywords online; he can also check the authenticity of the website. The accuracy for the dynamic system is 93% and it increases with every iteration. We intend to build our dataset which will be kept up to

date according to the latest news. All the live news and latest data will be kept in a database using Web Crawler and an online database.

**Chapter \_06**

**References**

ASME standard

Website links,

[1] https://www.studocu.com/in/document/indraprastha-institute-of-information-technology-delhi/computer-sciences-and-applied-mathematics/final-report-fake-news-detection/39608349 [2] https://www.geeksforgeeks.org/fake-news-detection-using-machine-learning/

[3] https://www.scribd.com/document/428453346/Fake-News-Detection-Project-Report

IEEE standard

Journal Paper,

[1] Prabha R, Ramaraj N. An improved multipath MANET routing using link estimation and swarm intelligence. EURASIP Journal on Wireless Communications and Networking. 2019;2020(1):1–9.

chromeextension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.ijert.org/research/potatoleaf-disease-detection-using-deep-learning-IJERTV11IS110029.pdf [2] Farooque Azam, Dr. Avinash Sharma: A Sentiment Analysis of Movies Reviews using Machine Learning’s Classification Algorithms, International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 10 | Oct 2020, e-ISSN: 2395-0056 p-ISSN: 2395-0072

[3] Yuan Tian, Chunjiang Zhao, Shenglian Lu and Xinyu Guo. "SVMbased Multiple Classifier System for Recognition of Wheat Leaf Diseases" Proceedings of 2010 Conference on Dependable Computing (CDC’2010) November 20-22, 2010, Yichang, China [4] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. ALRahamneh, “Fast and Accurate Detection and Classification of Plant Diseases” IJCA, vol. 17(1), pp. 31-38, March 2011, IEEE-2010

**Acknowledgments**

I am profoundly grateful to Prof. Ramaya Ma’am for his expert guidance and continuous encouragement throughout to see that this project reaches its target.

I would like to express deepest appreciation towards Dr. Varsha Shah, Principal RCOE, Mumbai and Prof Shiburaj Pappu HOD Computer Department whose invaluable guidance supported me in this project.

At last I must express my sincere heartfelt gratitude to all the staff members of Computer Engineering Department who helped us directly or indirectly during this course of work.

(Ansh Tiwari\_201P019)

(Suchit\_201P040)

(Ali Mo Salman\_201P048)

(Shikalgar Faizan\_201P001)

**MINI-PROJECT**

**ASSESSMENT SHEET**

**Termwork:** 25 marks

Group Members

Student 1 : Ansh Tiwari

Student 2 Ali Mo Salman

Student 3 : Suchit

Student 4 : Shikalgar Faizan

Guide Name: Prof Ramaya maam

**Attendance Percentage**

|  |  |
| --- | --- |
| **Student** | **Semester Attendance %** |
| Ansh Tiwari |  |
| Ali Mo Salman |  |
| Suchit |  |
| Shikalgar Faizan |  |

Attendance to TW Conversion

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| >=90% | <90% & >=80% | <80% & >=70% | <70% & >=60% | <60% |
| 5 | 4 | 3 | 2 | 1 |

**Project Review Performance:**

Rubrics used: Quality of survey/ need identification, Clarity of Problem definition based on need, Innovativeness in solutions, Feasibility of proposed problem solutions and selection of best solution, Cost effectiveness, Full functioning of working model as per stated requirements, Effective use of skill sets, Effective use of standard engineering norms.

|  |  |
| --- | --- |
| **Student** | **Average Points of Rubrics received after Review** |
| Ansh Tiwari |  |
| Ali Mo Salman |  |
| Suchit |  |
| Shikalgar Faizan |  |

Review RUBRICS to TW Conversion

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| >=18 | <18 & >=10 | <10 & >=5 | <5 & >=3 | <3 |
| 5 | 4 | 3 | 2 | 1 |

**Rubrics for Report:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Criteria** | **1**  **Unsatisfactory** | **2**  **Average** | **3**  **Good** | **Assessed by Guide**  **(1 to 3)** |
| **Content** | Insufficient content | Some topics or part missing | All necessary topics covered. |  |
| **References** | No research papers referred | Few research papers referred but no IEEE/ scopus indexed paper referred | Scopus / IEEE / reputed paper referred |  |
| **Representation** | No alignment, No caption in figures and tables and no citation | Citation missing but alignment and caption proper | Citation to references present along with captions and alignment of content. |  |
| **Abidance to Template** | Not at all | Some what | Good |  |
|  |  |  | **Total** |  |

Report Rubrics to TW Conversion

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| >=10 | <10 & >=8 | <8 & >=6 | <6 & >=4 | <4 |
| 5 | 4 | 3 | 2 | 1 |

**Final Term work Calculation**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Distribution** | **Student 1**  **Obtained** | **Student 2**  **Obtained** | **Student 3**  **Obtained** | **Student 4**  **Obtained** | **Outoff** |
| Attendance  (To be filled by Project Coordinator) |  |  |  |  | 5 |
| Project Review Performance  (To be filled by Project Coordinator) |  |  |  |  | 5 |
| Report  (To be filled by Guide) |  |  |  |  | 5 |
| CIE by Guide (Weekly)  (To be filled by Guide) |  |  |  |  | 10 |
| **Total Term work** |  |  |  |  | **25** |

**H.o.D. Computer Project Coordinator Project Guide**