TRIBHUVAN UNIVERSITY

FACULTY OF HUMANITIES AND SOCIAL SCIENCES

LALITPUR ENGINEERING COLLEGE

TRUELENS: FAKE NEWS DETECTOR

BY

SIRJAN SHRESTHA (LEC077BCA06)

A PROJECT PROPOSAL

SUBMITTED TO THE DEPARTMENT OF COMPUTER APPLICATION

IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR

THE DEGREE OF BACHELORS IN COMPUTER APPLICATION

DEPARTMENT OF COMPUTER APPLICATION

LALITPUR, NEPAL

JUNE, 2024

Tribhuvan University

Faculty of Humanities and Social Sciences

TRUELENS: FAKE NEWS DETECTOR

Submitted to

Department of Computer Application

Lalitpur Engineering College

In partial fulfillment of the requirement for the degree of Bachelors in Computer Application

Submitted by

Sirjan Shrestha (LEC077BCA06)

JUNE, 2024

Under the Supervision of

Er. Bibat Thokar

COPYRIGHT ©

DECLARATION

I declare that the work hereby submitted for Bachelors in Computer Application at the Department of Computer Application , Lalitpur Engineering College entitled "**TrueLens: Fake News Detector**" is my own work and has not been previously submitted by me at any university for any academic award. I authorize the Department of Computer Application , Lalitpur Engineering College to lend this project work to other institutions or individuals for the purpose of scholarly research.

Sirjan Shrestha (LEC077BCA06)

June, 2024

RECOMMENDATION

The undersigned certify that they have read and recommend to the Department of Computer Application for acceptance, a project work entitled "**TrueLens: Fake News Detector**", submitted by **Sirjan Shrestha (LEC077BCA06)** in partial fulfillment of the requirement for the award of the degree of "**Bachelors in Computer Application**".

Project Supervisor

Er. Bibat Thokar

Lecturer

Department of Computer Application , Lalitpur Engineering College

---

BCA Program Coordinator

Er. Bibat Thokar

Lecturer

Department of Computer Application , Lalitpur Engineering College

June, 2024

DEPARTMENTAL ACCEPTANCE

The project work entitled "**TrueLens: Fake News Detector**", submitted by **Sirjan Shrestha (LEC077BCA06)**in partial fulfillment of the requirement for the award of the degree of "**Bachelors of Computer Application**" has been accepted as a genuine record of work independently carried out by the student in the department.

Er.Bibat Thokar
BCA Coordinator
Department of Computer Application ,
Lalitpur Engineering College ,
Faculty of Humanities and Social Sciences ,
Tribhuvan University, Nepal.

June, 2024

ACKNOWLEDGMENT

This project work would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First of all, I would like to express my sincere gratitude to my supervisor, **Er. Bibat Thokar**, of **Lalitpur Engineering College** for providing invaluable guidance, insightful comments, meticulous suggestions, and encouragement throughout the duration of this project work. My sincere thanks also goes to the BCA coordinator, **Er. Bibat Thokar** , for coordinating the project works, providing astute criticism, and having inexhaustible patience.

Furthermore, we would like to extend our gratitude to the entire faculty of the Department of Computer Application . Their dedication to fostering creativity, critical thinking, and technical proficiency has been useful in our project's development. The support and guidance received from our teachers have empowered us to transform our vision into a reality.

I am also grateful to my classmates and friends for offering me advice and moral support. To my family, thank you for encouraging me in all of my pursuits and inspiring me to follow my dreams. I am especially grateful to my parents, who supported me emotionally, believed in me and wanted the best for me.

Sirjan Shrestha (LEC077BCA06)

June, 2024

ABSTRACT

In an age where misinformation proliferates rapidly across digital platforms, discerning the credibility of news is crucial. This project, ”TrueLens: A Fake News Detection System,” addresses this challenge by leveraging advanced machine learning techniques integrated with user-friendly web technologies. The backend is developed using Python and Django with the Django Rest Framework (DRF), facilitating robust API management and data handling, while the machine learning model, powered by TensorFlow, employs sophisticated natural language processing (NLP) to analyze and classify news articles. The frontend, also developed with Django, provides an intuitive interface for users to input or paste news content for verification, and the system's data is efficiently managed with PostgreSQL. The project uses Git and GitHub for version control, ensuring seamless collaboration and continuous integration. TrueLens not only detects fake news but also serves as an educational tool, enhancing users' understanding of misinformation patterns and fostering critical evaluation skills. By making powerful AI accessible through a web interface, TrueLens contributes significantly to the fight against misinformation, offering a practical solution for individuals and organizations committed to maintaining the integrity of information in the digital age.

Keywords: *Fake News, Machine Learning, Natural Language Processing (NLP)*

TABLE OF CONTENTS

COPYRIGHT ..... **iii**

DECLARATION ..... **iv**

RECOMMENDATION ..... **v**

DEPARTMENTAL ACCEPTANCE ..... **vi**

ACKNOWLEDGMENT ..... **vii**

ABSTRACT ..... **viii**

TABLE OF CONTENTS ..... **ix**

LIST OF FIGURES ..... **xi**

LIST OF ABBREVIATIONS ..... **xii**

1 INTRODUCTION **1**

1.1 Introduction 1

1.2 Problem Statement 1

1.3 Objectives 1

1.4 Scope 2

1.5 Report Organisation 2

2 BACKGROUND AND LITERATURE REVIEW **3**

2.1 Background Study 3

2.2 Limitation 3

2.3 Literature Review 3

3 METHODOLOGY **5**

3.1 Iterative Approach 5

3.2 Requirement Analysis 5

3.3 Feasibility Analysis 5

3.3.1 Financial Feasibility 6

3.3.2 Operational Feasibility 6

3.3.3 Technical Feasibility 6

3.4 System Design 7

3.4.1 Architecture Design 7
3.4.2 Data Modelling(ER-Diagram) 8
3.4.3 Activity Diagram 9
3.4.4 DFD 10
3.4.5 Use Case Diagram 11
4 IMPLEMENTATION 12
4.1 Tools Used 12
5 CONCLUSION AND EXPECTED OUTCOMES 15
5.1 Conclusion 15
5.2 Expected Outcome 15
APPENDIX A
A.1 Project Schedule 16
REFERENCES 17

LIST OF FIGURES

Figure 3.1 Main Architecture of System 7
Figure 3.2 ER Diagram of System Data 8
Figure 3.3 Activity Diagram 9
Figure 3.4 Data Flow Diagram (Context Level) 10
Figure 3.5 Use Case Diagram 11
Figure A.1 Gantt Chart of Schedule 16

LIST OF ABBREVIATIONS

TL	TrueLens
FN	Fake News
ML	Machine Learning
AI	Artificial Intelligence
DB	Database
DFD	Data Flow Diagram
FE	Frontend
ER	Entity-Relationship
HTML	Hypertext Markup Language
IT	Information Technology
JS	JavaScript
API	Application Programming Interface

1 INTRODUCTION

1.1 Introduction

In today's digital era, the rapid spread of fake news poses a serious threat to public trust and societal well-being. TrueLens: Fake News Detection System offers a practical solution by using advanced machine learning techniques to identify and flag misleading information. Powered by Python and TensorFlow, TrueLens integrates seamlessly with Django to provide a user-friendly interface where individuals can verify the authenticity of news articles. With efficient data management via PostgreSQL and robust version control through GitHub, TrueLens stands as a crucial tool in the fight against misinformation, promoting informed decision-making and enhancing the integrity of online content.

1.2 Problem Statement

In the digital age, the rampant spread of fake news undermines public trust and distorts reality. The volume and sophistication of misinformation on digital platforms make it difficult for individuals to discern fact from fiction. Traditional verification methods are often inadequate, leading to widespread misinformation with significant societal and political impacts. There is a lack of effective, user-friendly tools for quick and accurate news verification, which exacerbates the issue.

Additionally, there is a significant educational gap, with many users unaware of how to critically evaluate fake news. Current solutions either lack real-time detection capabilities or are too complex for general use. This highlights the need for a robust system that not only detects fake news using advanced machine learning but also educates users about misinformation. TrueLens aims to fill this gap by providing an accessible, efficient, and educational platform that empowers users to combat false information and make informed decisions.

1.3 Objectives

- To empower users with a user-friendly tool for swiftly and accurately verifying news authenticity, thereby combating the spread of misinformation effectively.

1.4 Scope

- Implement advanced machine learning algorithms to accurately detect and classify fake news articles based on linguistic and statistical analysis.
- Develop a user-friendly web interface that allows users to easily submit news articles for verification and receive clear, understandable results regarding their authenticity.
- Provide educational resources within the platform to enhance users' understanding of fake news, promoting critical thinking and media literacy skills among the general public.

1.5 Report Organisation

The material in this project report is organised into seven chapters. After this introductory chapter introduces the problem topic this research tries to address, chapter 2 contains the literature review of vital and relevant publications, pointing toward a notable research gap. Chapter 3 describes the methodology for the implementation of this project. Chapter 4 provides an overview of what has been accomplished. Chapter 5 contains some crucial discussions on the used model and methods. Chapter 6 mentions pathways for future research direction for the same problem or in the same domain. Chapter 7 concludes the project shortly, mentioning the accomplishment and comparing it with the main objectives.

2 BACKGROUND AND LITERATURE REVIEW

2.1 Background Study

In today's digital age, the rapid spread of fake news on social media and other online platforms has become a critical issue, influencing public opinion and eroding trust in genuine news sources. Traditional fact-checking methods are often too slow to combat the swift spread of misinformation. Machine learning offers a powerful solution by analyzing text for patterns that indicate falsehoods, but many existing tools are either complex or lack real-time capabilities. TrueLens: Fake News Detection System addresses these challenges by providing a user-friendly platform that uses advanced machine learning to quickly verify the authenticity of news articles, thereby helping users navigate the information landscape more effectively and promoting informed decision-making.

2.2 Limitation

- The system may struggle with detecting fake news if its training data is biased or lacks diversity, limiting its effectiveness in identifying new or less common misinformation.

- The system can misinterpret nuanced language or context, potentially leading to inaccuracies in detecting sarcasm, subtle biases, or contextually misleading information.

2.3 Literature Review

The exponential growth of digital media and social networks has led to an unprecedented increase in the spread of fake news, significantly affecting public opinion and trust in credible news sources. It highlight how the rapid dissemination of fake news, particularly during the 2016 U.S. Presidential Election, influenced voter behavior and public perceptions, emphasizing the need for robust mechanisms to detect and mitigate the spread of misinformation. Their study sheds light on the economic and social repercussions of fake news, which necessitates the development of effective detection systems to maintain information integrity in the digital era.[1]

Machine learning has emerged as a crucial tool in the detection of fake news, offering the ability to analyze large datasets and identify patterns that human analysis might miss. Conroy, Rubin, and Chen (2015) classify fake news detection methodologies into three main categories: linguistic, network-based, and hybrid approaches. Linguistic approaches focus on analyzing the language and writing style of news articles to identify deceptive content, while network-based methods examine the dissemination patterns and relationships between news articles and their sources. Hybrid approaches combine linguistic and network-based techniques to enhance detection accuracy, leveraging the strengths of both methodologies to combat misinformation effectively.[2]

Despite these advancements, challenges remain in developing effective fake news detection systems. Shu, Wang, and Liu (2017) discuss the difficulties associated with the evolving nature of fake news, which requires continuous updates to detection models to keep up with new types of misinformation. They also emphasize the need for diverse and high-quality datasets to train and test detection algorithms effectively. Zhou and Zafarani (2018) highlight the scarcity of comprehensive labeled datasets for fake news detection, which limits the ability to develop and validate robust models. The TrueLens project aims to address these challenges by integrating advanced machine learning techniques with a user-friendly web interface, providing an accessible and scalable solution for fake news detection.[3]

Deep learning has significantly advanced fake news detection models. Ruchansky, Seo, and Liu (2017) present the ”CSI” model, a hybrid deep learning framework integrating content analysis, social context, and user behavior to detect fake news more accurately. Their approach combines diverse data sources and analytical techniques, improving detection compared to traditional methods by capturing the dynamic spread of misinformation and its social influences.[4]

3 METHODOLOGY

3.1 Iterative Approach

Iterative development is a methodology where a project is broken down into smaller, manageable parts or iterations, each delivering a piece of functionality that is tested and refined before moving on to the next iteration. Here are the steps for iterative development, particularly applicable to developing a fake news detection system like TrueLens:

- Planning and Requirements Gathering
- Planning and Design
- Development and Implementation
- Testing and Quality Assurance
- Review and Evaluation
- Deployment and Release
- Feedback and iteration
- Documentation and Maintenance

3.2 Requirement Analysis

3.3 Feasibility Analysis

Feasibility analysis is a systematic assessment of the practicality, viability, and potential success of a proposed project or initiative. It evaluates various aspects including technical, operational, economic, legal, and schedule feasibility to determine whether the project can be realistically implemented and achieve its intended objectives. Feasibility analysis helps stakeholders make informed decisions by identifying risks, constraints, and opportunities associated with the project, ultimately guiding resource allocation and planning to maximize the likelihood of successful outcomes.

3.3.1 Financial Feasibility

For the TrueLens project, our toolkit includes Django for robust backend development, Python for scripting and backend logic, PostgreSQL for efficient database management, and a combination of HTML, CSS, and JavaScript for dynamic frontend implementation. Additionally, Figma will be used for designing user interfaces. Our main financial allocation will be dedicated to securing reliable server hosting to ensure optimal performance and seamless accessibility of the system.

3.3.2 Operational Feasibility

Operational feasibility for the TrueLens project involves assessing user acceptance among stakeholders like media consumers, journalists, and fact-checkers. It includes providing adequate training and support, ensuring compliance with data protection regulations and ethical standards, and evaluating scalability to meet growing demands effectively. These factors ensure the system is practical, user-friendly, and capable of maintaining trust while combating misinformation.

3.3.3 Technical Feasibility

Technical feasibility for the TrueLens project involves evaluating several critical factors to ensure its successful implementation. This includes assessing the availability of expertise in machine learning (ML) algorithms and frameworks such as PyTorch or TensorFlow, essential for developing accurate fake news detection models. Integration with Django for backend development must be evaluated to ensure seamless compatibility and optimal performance, especially concerning scalability requirements to handle large datasets and real-time processing effectively. Additionally, robust infrastructure planning is necessary to support the system's technological needs and future growth.

3.4 System Design

3.4.1 Architecture Design

The following diagram shows diagram of our Architecture. Mainly shows what are the functions can be accessed after starting our application.

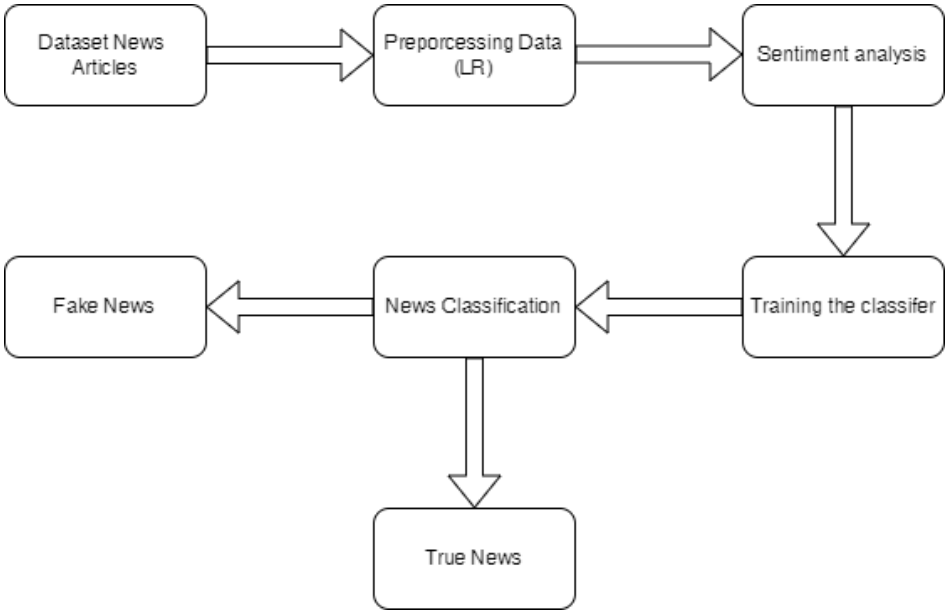

Figure 3.1: Main Architecture of System

3.4.2 Data Modelling(ER-Diagram)

ER Diagram is mainly used to design database schema. With the help of below er diagram we can easily design database in SQL.

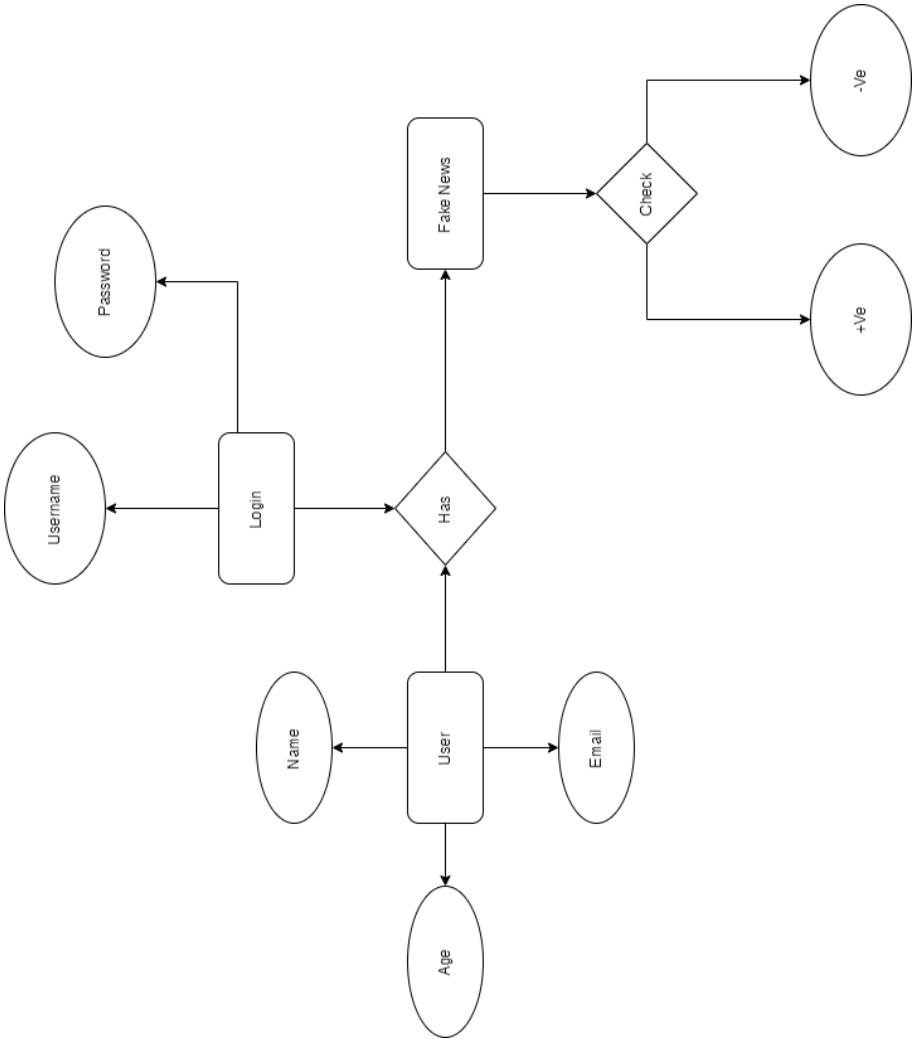

Figure 3.2: ER Diagram of System Data

3.4.3 Activity Diagram

An activity diagram visually presents a series of actions or flow of control in a system similar to a flowchart or a data flow diagram. This diagram showed how our program flow goes on.

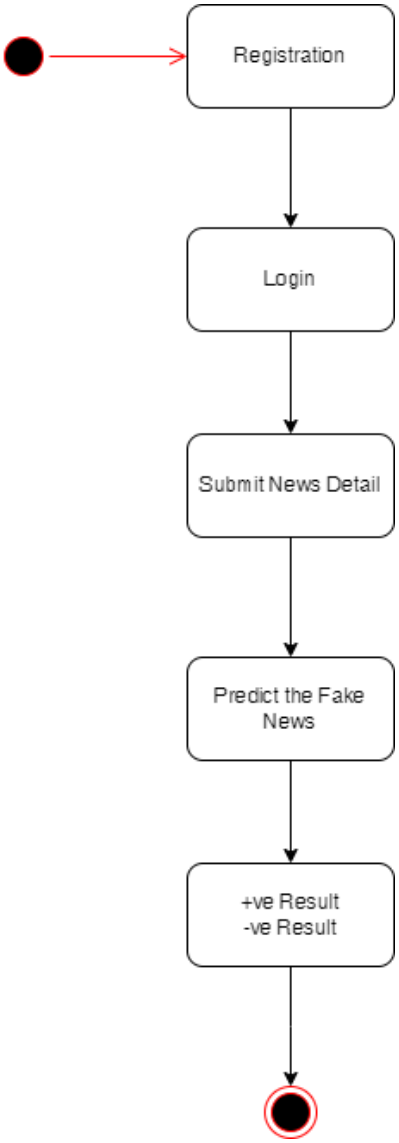

Figure 3.3: Activity Diagram

3.4.4 DFD

DFD or Data Flow Diagram is mainly used to show how data are being flowed in and out of our system. There are 3 levels of DFD i.e Context Level(Level 0),Level 1 and Level 2

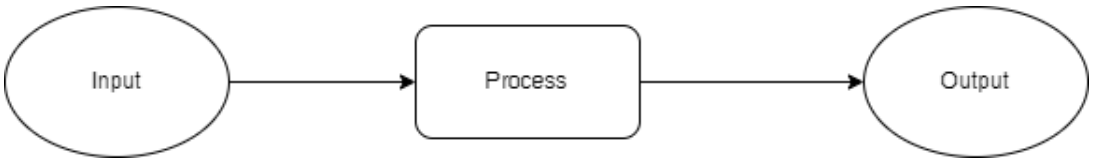

Figure 3.4: Data Flow Diagram (Context Level)

3.4.5 Use Case Diagram

A use case diagram, part of UML, visually represents interactions between actors and a system. Actors are external entities, while use cases depict specific functionalities. Relationships, such as association, generalization, include, and extend, illustrate connections between actors and use cases. The diagram helps in understanding system behavior, requirements, and scope.

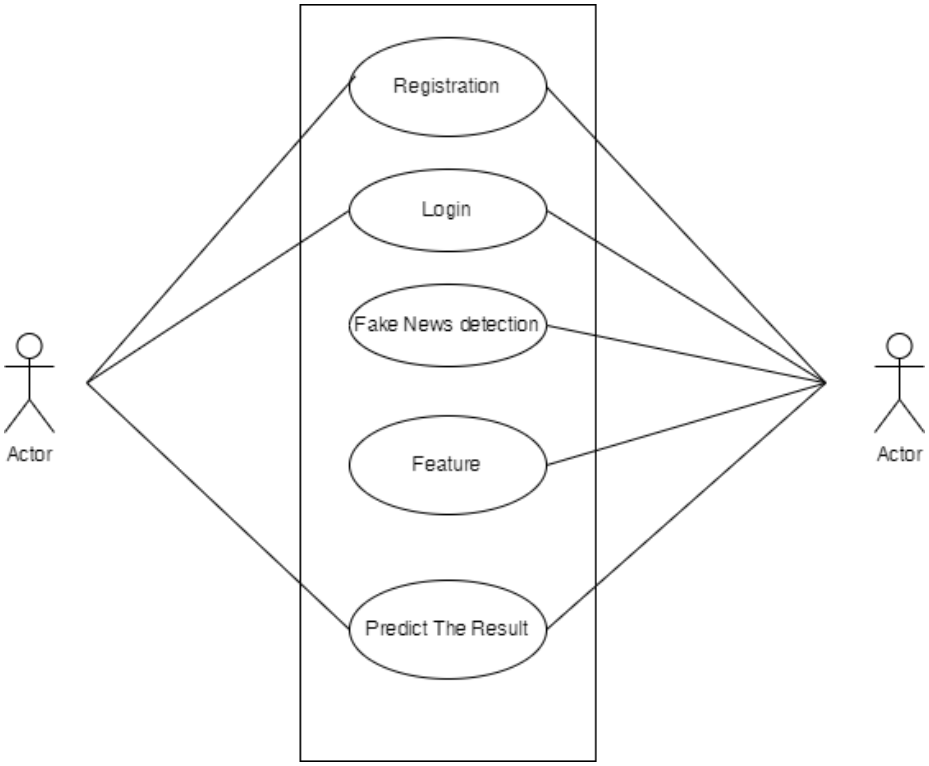

Figure 3.5: Use Case Diagram

4 IMPLEMENTATION

4.1 Tools Used

Python

Python is a versatile, high-level programming language known for its simplicity, readability, and extensive standard library. Developed in the late 1980s, Python's clear syntax supports multiple programming paradigms, including procedural and object-oriented styles. Its dynamic typing and interpreted nature facilitate rapid development and interactive debugging, making it popular across diverse domains such as web development, scientific computing, and artificial intelligence. Python's thriving community and robust ecosystem of libraries contribute to its widespread adoption and continued relevance in modern software development.

Django

Django is a high-level Python web framework designed for rapid development and clean, pragmatic design. Built on the principles of DRY (Don't Repeat Yourself) and convention over configuration, Django simplifies the creation of complex, database-driven websites by providing reusable components and a built-in admin interface. It encourages rapid development and clean, maintainable code through its "batteries-included" philosophy, offering features such as URL routing, template engine, ORM (Object-Relational Mapping), and authentication system out of the box. Django's scalability, security features, and extensive documentation make it a preferred choice for building robust web applications efficiently.

PostgreSQL

PostgreSQL, often referred to as Postgres, is an advanced open-source relational database management system (RDBMS) known for its reliability, robustness, and extensibility. Developed over three decades with a strong emphasis on standards compliance and SQL features, PostgreSQL supports a wide range of data types, indexing techniques, and advanced features like full-text search and JSON support. It offers ACID (Atomicity, Consistency, Isolation, Durability) compliance, ensuring data integrity and reliability even in high-transaction environments. PostgreSQL's active community contributes to its continuous improvement and extensive documen-

tation, making it a popular choice for applications requiring secure, scalable, and efficient data storage solutions.

Git/Github

Git is a distributed version control system (VCS) designed for managing software development projects and tracking changes to source code during development. Developed by Linus Torvalds in 2005, Git allows multiple developers to collaborate on projects simultaneously by maintaining a history of changes, facilitating code review, and enabling seamless integration of new features. Git operates locally on a developer's machine, providing branching and merging capabilities that support parallel development workflows and experimental features without affecting the main codebase. GitHub, on the other hand, is a web-based hosting service for Git repositories, offering additional features such as issue tracking, pull requests, and project management tools. It enhances collaboration and community engagement by providing a platform for developers to share, discuss, and contribute to open-source projects globally. Together, Git and GitHub form a powerful ecosystem that supports agile software development practices and fosters collaboration among developers worldwide.

PyTorch

PyTorch is an open-source machine learning library developed primarily by Facebook's AI Research lab (FAIR). It is widely used for building and training deep learning models, particularly neural networks, due to its dynamic computational graph and GPU acceleration capabilities. PyTorch emphasizes flexibility and ease of use, allowing developers to quickly prototype and experiment with different network architectures and algorithms. Key features include tensor computations with support for automatic differentiation, making it straightforward to implement and optimize complex neural networks. PyTorch is favored by researchers and practitioners in fields such as computer vision, natural language processing, and reinforcement learning for its intuitive API, extensive documentation, and active community support, making it a leading choice for cutting-edge machine learning projects. PyTorch is renowned for its seamless integration with Python.

JavaScript

JavaScript is a versatile programming language primarily used for creating dynamic and interactive web pages. Initially developed by Netscape in 1995, JavaScript has evolved into a fundamental tool for front-end web development, allowing developers to manipulate webpage content, handle events, and interact with users in real-time. It is widely supported by all modern web browsers and can be integrated with HTML and CSS to enhance user interfaces and create responsive web applications. JavaScript's asynchronous capabilities and extensive ecosystem of libraries and frameworks (such as React, Angular, and Vue.js) further contribute to its widespread use in both client-side and server-side development, making it essential for building modern, interactive web experiences.

5 CONCLUSION AND EXPECTED OUTCOMES

5.1 Conclusion

The TrueLens project endeavors to develop a system for detecting fake news, aiming to enhance media integrity and combat misinformation effectively. By leveraging advanced machine learning techniques, the project seeks to create a reliable platform capable of distinguishing between genuine and fabricated news articles. The system will provide users with a trustworthy tool to verify the authenticity of information, thereby promoting informed decision-making and more reliable digital environment.

5.2 Expected Outcome

TrueLens project is the development of a comprehensive and effective fake news detection system designed to tackle the pervasive issue of misinformation in digital media. By harnessing the power of advanced machine learning algorithms, the system aims to achieve high levels of accuracy and reliability in distinguishing between genuine and fabricated news articles. This capability will empower users to make informed decisions about the information they consume, thereby promoting media literacy and critical thinking. The system's user-friendly interface will facilitate easy input of news articles and provide clear, actionable insights into the credibility of each piece of information. Ultimately, TrueLens seeks to contribute to a more trustworthy digital landscape by providing a tool that enhances transparency, integrity, and reliability in news dissemination and consumption.

APPENDIX A

A.1 Project Schedule

Below is the Gantt chart of our project Schedule. We have planned to perfrom these specific tasks between these time frames.

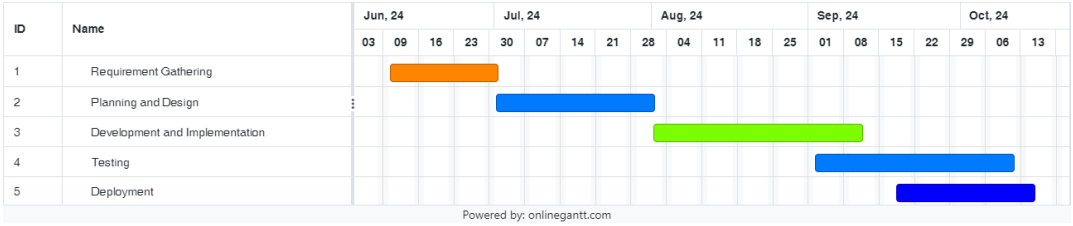

Figure A.1: Gantt Chart of Schedule

REFERENCES

[1] Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2):211–236, 2017.

[2] Nicole K Conroy, Victoria L Rubin, and Yimin Chen. Automatic deception detection: Methods for finding fake news. In *Proceedings of the Association for Information Science and Technology*, volume 52, pages 1–4. Association for Information Science and Technology, 2015.

[3] Xinyi Zhou and Reza Zafarani. Fake news: A survey of research, detection methods, and opportunities. *arXiv preprint arXiv:1812.00315*, 2018.

[4] Natali Ruchansky, Sungyong Seo, and Yan Liu. Csi: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 797–806. ACM, 2017.