# MULTIMODAL MULTIMEDIA INTEGRITY VERIFICATION: DETECTION ONOMA

## (FAKE NEWS DETECTION)

Project ID: 24-25J-137

Project Proposal Report

**IT21267086 : Hasara L.A.N**

Supervisor: Mr. Kavinga Yapa Abeywardena

Co Supervisor: Dr. Harinda Fernando

BSc. (Hons) in Information Technology Specializing in Cybersecurity

Department of Computer Systems Engineering

Faculty of Computing

Sri Lanka Institute of Information Technology

Sri Lanka

August 2024

# DECLARATION

I declare that this is my own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

| NAME | STUDENT ID | SIGNATURE | DATE |
|---|---|---|---|
| Hasara L. A. N | IT21267086 | Hasara | 22/08/2024 |

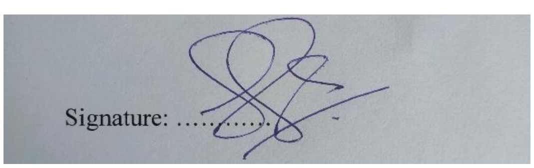
Name of supervisor: Mr. Kavinga Yapa Abeywardena

Name of co-supervisor: Dr. Harinda Fernando

The above candidate has carried out research for the bachelor's degree dissertation under my supervision.

Signature: …………………………….                                        Date: 22.08.2024

Supervisor: Mr. Kavinga Yapa Abeywardena




Co Supervisor: Dr. Harinda Fernando                                        Date: 22.08.2024

# ABSTRACT

Misinformation is fast spreading in the digital era and is an emerging global problem in the trust of people, politics, and society. The proliferation of misinformation over several languages and platforms imposes the need for a powerful and flexible detection system. This research will focus on the development of an advanced fake news detection system in the context of the broader Multimodal Multimedia Integrity Verification – Detection ONOMA effort. Addressing the limitations of current approaches, this study proposes a state-of-art NLP technique, multilingual embeddings like mBERT and utilize hybrid model of Convolutional Neural Networks and Bi-directional Long Short-Term Memory techniques to guarantee the effective capture of local and sequential features in textual content across languages, with a focus mainly on the English and Chinese languages. Apart from the content analysis, the system integrates real-time source credibility assessments by making calls to third-party APIs that estimate historical accuracy and domain reputation. Combining sophisticated content analysis with dynamic credibility assessment in a dual approach ensures a more resilient and accurate detection process. The continuous learning mechanism at the core of the system allows it to remain effective over time by adapting to emerging trends in misinformation. This research involves comprehensive data collection and preprocessing, extraction of features, development of the model, evaluation, and finally deployment of the system. This will result in a flexible and scalable solution to the challenge of fake news worldwide, which will have significant ramifications for government agencies, media companies, and cybersecurity sectors by offering a comprehensive framework to protect the integrity of digital information.

*Key Words – fake news detection, misinformation, CNN-BiLSTM, mBERT, source credibility, continuous learning, digital trust, ethical AI, Multimodal Multimedia Integrity Verification*

# ACKNOWLEDGMENT

I would like to sincerely express my gratitude to supervisors, Mr. Kavinga Yapa Abeywardena and Dr. Harinda Fernando, Department of Computer Systems Engineering at the Sri Lanka Institute of Information Technology, for their constant support and guidance regarding the development of this research proposal. Their knowledge and constructive feedback have been invaluable in shaping the direction of this project and in addressing possible challenges.

I would also like to thank the panel members for their invaluable advice. Indeed, their input was great in pointing out essential improvements to be made and in refining the proposal for an effective and feasible study. Lastly, my appreciation goes to the coordinators and the design and analysis lecturer of this course for all-round support during the whole proposal phase. Their guidance was very fundamental in laying a strong foundation of the project.

# TABLE OF CONTENTS

# LIST OF FIGURES



# LIST OF TABLES



# LIST OF APPENDICES

# LIST OF ABBREVIATIONS

| Abbreviation | Definition |
|---|---|
| **NLP** | Natural Language Processing |
| **CNN** | Convolutional Neural Network |
| **Bi-LSTM** | Bi-directional Long Short-Term Memory |
| **mBERT** | Multilingual Bidirectional Encoder Representations from Transformers |
| **GPU** | Graphics Processing Unit |
| **API** | Application Programming Interface |
| **ML** | Machine Learning |
| **DL** | Deep Learning |
| **AI** | Artificial Intelligence |
| **POS** | Part-Of-Speech |
| **AWS** | Amazon Web Services |
| **SQL** | Structured Query Language |
| **NoSQL** | Not Only Structured Query Language |

Table 1. List of Abbreviations

# 1. INTRODUCTION

## 1.1 Research Background

The dissemination of false information and misinformation popularly known as "fake news" has increased to become a serious global challenge in the digital age. Information that is deceptive can be quickly picked up by social media, news websites, and other digital channels, further affecting public opinion and eroding institutional trust right down to political outcomes. It is the rising sophistication of fake news powered by deepfakes and AI that makes the requirement for dependable detection systems very important [1].

Fake news is not a novel concept, the internet and social media have hugely amplified it into a pervasive problem. Traditional methods of fact-checking and control over editorials prove ineffective against the very fast and very extensive spread of misinformation [2]. Thus, to detect fake news, automated algorithms using ML and NLP have been developed. Most of the existing approaches, however, are usually scoped very narrowly, normally either content-based or source credibility-based, not both. This limits their efficacy against the multifaceted nature of fake news.

This challenge is further compounded by the multilingual nature of the global information ecosystem. Fake news goes across linguistic borders, affecting diverse linguistic groups around the world. Most of the algorithms that have so far been developed for fake news detection are tailored for a single language, usually English, greatly limiting its scope of applicability across the globe [3]. There is a growing need for systems that can detect fake news in multiple languages, especially in regions where linguistic barriers make access to information quite a challenge.

This research proposes the development of a fully functional multilingual fake news detection system to solve these challenges. In this respect, advanced NLP techniques and state-of-the-art machine learning models, which include BERT and deep learning architectures, will be fronted as an accurate and robust solution for fake news detection. This system combines content analysis and source credibility assessment in multiple languages to ensure an efficient and effective means of detecting fake news at a global scale [4].

The study also highlights how important it is for the system to continuously learn and adapt to be able to keep up with new developing patterns of fake news. This adaptability secures its long-term efficacy and makes it a useful weapon against fake news in real-world situations where the disinformation environment is ever-changing.

## 1.2 Research Scope

The current work focuses on developing a comprehensive multilingual fake news detection system that addresses the challenges of identifying false information across different languages and aspects. This project aims to build such a system using advanced Natural Language Processing (NLP) techniques to evaluate both the content of news articles and the credibility of their sources. By cross-checking information from textual material, the system seeks to create a more reliable and accurate method for identifying false news.

This paper emphasizes the ability of speaking many languages throughout. In this way, when created, the system will be able to process and evaluate material in at least two different languages, further extending its usefulness into all kinds of different linguistic situations. This is a very relevant characteristic for the globalized society that we live in today, since fake news might often cross language borders and affect several populations simultaneously. A multilingual approach will help the system offer a solution that is more inclusive and more successful in solving the problem of false news on a worldwide scale.

In this regard, it will make use of state-of-the-art machine learning models for the analysis of sophisticated language structures and the recognition of subtle patterns in the data, such as multilingual BERT [5]. This will be very instrumental in enhancing the capacity of the system to detect false information. This research will also indulge in setting up ways for learning to make sure that the system is able to fit into new and evolving patterns of false news over several periods. The ability to react to such changes will be paramount in ensuring that the system continues serving optimally in this ever-changing landscape of digital disinformation.

## 1.3 Research Area

### 1.3.1 Multilingual Fake News detection

The objective of the research area is to develop a system that could detect fake news across several languages, addressing this global challenge of misinformation that has no borders but is related to linguistic diversity. The intended outcome will be a very flexible and accurate fake news detection system that will work effectively in various linguistic settings and thus limit the number of pieces of information that spread misinformation across the world. This includes tapping into the deepest NLP techniques, especially multilingual models like mBERT, in processing and analyzing content in multiple languages to make sure that it has audiences all over the world.

### 1.3.2 NLP and machine learning

This is an area of research in which methods of NLP and machine learning models are applied to spot fake news through content analysis. The line of research mainly aims at increasing the accuracy and adaptiveness of fake news detection systems using state-of-the-art NLP models with robust machine learning algorithms. It implements content analysis through CNN-BiLSTM and deep learning architectures that treat multilingual text to provide a strong base for the system's capability to detect misinformation in various linguistic contexts.

### 1.3.3 Content Analysis

In content analysis, this study seeks to develop better methods that would find the perfect examination of the contents in a news article to pick any clues pointing at misinformation. Here, it is trying to set out a full-fledged framework that can point out semantic, syntactic, and stylistic features pointing to fake news. This is realized through convolutional neural networks in capturing local text features and bidirectional long short-term memory networks in understanding sequential dependencies of the text, hence culminating into a detailed, nuanced content analysis score that contributes to the overall detection process.

### 1.3.4 Source Credibility Assessment

This research area in real-time source credibility assessment will be combined with fake news detection for enhancing the system's accuracy. It purports to take the accuracy of detection a step further by checking the reliability of the news source itself in addition to the content. This is done with the help of fetching credibility scores returned from third-party APIs, which consider factors such as historical accuracy and domain reputation, to provide an overall, dynamic, and updated set on which the likelihood of a news source's reliability should be based in terms of its being fake.

### 1.3.5 Continuous Learning and Adaptation to Change

There is a need for research in continuous learning and adaptation so that the detection system remains relevant and effective over some time. To that end, mechanisms will need to be devised through which the model learns from new data and reacts to new patterns of fake news. Several methods are being studied for building flexibility in the system and for long-term efficiency, which include online learning and incorporation of user feedback into training.

### 1.3.6 Cybersecurity and Information Integrity

It contributes to the broad realm of cybersecurity by focusing on protecting digital information integrity. The objective is to offer state-of-the-art protection tools that avoid the proliferation of false information and increase the security of information, as well as confidence among the public. Developing all-inclusive solutions geared toward maintaining the integrity of digital communications, hence securing robust protection for this very pervasive problem of misinformation in the digital age, through AI, machine learning, and NLP techniques.

### 1.3.7 Ethical Considerations in AI-Driven Fake News Detection

This research area therefore has a focus on the ways in which ethical implications arising from the use of AI in fake news detection could be addressed to make the system responsible, minimize bias and false positives, and possible impacts on freedom of expression. It means designing a detection system that is infused with ethical risk assessments and transparency mechanisms to make sure it performs not only effectively but also according to ethical standards, thus promoting fairness and accountability in combating misinformation.

## 1.4 Background and Literature Survey

### 1.4.1 Overview of Fake News and its Effects

The term "fake news" designates the spread of incorrect or inaccurate information that is presented as news, often with the aim to mislead the public or achieve some other form of influence over the population. If that is not enough, the invention of digital platforms and social media completely accelerated the spread of false news, turning it into a serious issue in this information age. Some researchers have pointed to the social effects of fake news, which, according to this study [6], is deemed to have the potential to swing elections, incite social violence, and undermine trust in institutions. Considering the high frequency of false news usage, academics engage with a good number of methods geared toward detection and curtailment of its impact, especially through the application of artificial intelligence and machine learning [7].

### 1.4.2 Currently Existing Techniques of fake news detection

Traditional methods for fake news detection have been content-based; that is, the actual text of news items is thoroughly checked for false or inaccurate information. Earlier work has used a wide variety of techniques, from lexical classifying to sentiment analysis to stylistic features, which separate news items as either true or false [8]. The method of linguistic analysis and used differences in writing style as pointers or red flags for false information. These methods, however, have problems dealing with the complex language used in high-end fake news articles [9]. This has led to the continued development of evermore advanced machine learning models.

### 1.4.3 Developments of NLP for fake news identification

Developments in Natural Language Processing have changed fake news identification by allowing a deeper analysis of textual data. In recent times, other NLP techniques like sentiment analysis, POS tagging, and NER have been applied in attempts to build better false news detection systems. Deep learning has also recently boosted this area of research, with models such as BERT [10]. In this respect, the performance of false news detection algorithms was significantly improved by the realization of contextual textual relationships in the pre-trained language model BERT. For example, research conducted just how well BERT performed in the identification of small language cues indicative of fake news [11], thereby making it a very powerful tool in the fight against online disinformation.

### 1.4.4 Multilingual Recognition of False news

Despite all the progress in the English-language false news detection, there comes a growing realization of the need for multilingual algorithms that work in diverse linguistic contexts. Indeed, mechanisms are required to detect this misinformation in many languages to prevent disinformation from running free in non-English-speaking regions, as this information often spreads globally [12]. A study indicated that detecting false news in multiple languages is difficult because of linguistic variation and the unavailability of annotated datasets. Recently, some research has started to probe into cross-lingual transfer learning [13], fine-tuning models trained on one language for use in another. This field is still little known, hence requires further investigation to come up with reliable multilingual false news detection systems.

### 1.4.5 Limitations and Challenges to Overcome

The growth in fake news detection has been huge, but there are still several challenges. First, one of the most important limitations is the detection of fake news across different languages and cultural contexts. In this respect, it is foreseeable that linguistic and stylistic variations will act as a barrier to model performance [14]. The second challenge regards the availability and quality of annotated datasets for such tasks, considering non-English languages. Furthermore, the methods of disinformation are fast changing, and updating detection systems in a continuous manner could be resource-intensive and technically challenging [15]. Meeting these challenges is thus inescapable to build robust and adaptable fake news detection systems.

### 1.4.6 Ethical Dimensions in Fake News Identification

The second major concern is the major ethical questions thrown up by technologies for detecting false news [16]. The results of false positives, mislabeling genuine news as fraudulent, can be serious, comprising the strangulation of free expression and the defamation of respectable journalism. AI and machine learning in the detection of false news must also [17], therefore, be transparent and explainable to ensure that the decisions of these systems are understandable and reasonable [18]. Ensuring ethical AI practices are aligned within the design and deployment of false news detection systems will prevent unexpected consequences [19].

## 1.5 Significance of the Research

### 1.5.1 Addressing a global challenge of fake news

The research addresses a global concern: fake news. In this era of rapid diffusion of digital media, today's misinformation disseminates quicker than ever. Thus, the paper is focused on multilingual detection to identify the disinformation in different languages and avoid its global consequences. This approach is quite comprehensive for retaining the integrity of information globally by combining the analysis of content with credibility assessment of sources in real time.

### 1.5.2 Innovation in Natural Language Processing and Source Credibility

The research enhances the state of the art in NLP and source credibility rating by integrating these features into a novel fake news detection system. This approach fills a critical gap in existing methods for increased accuracy and reliability of detection and contributes effectively to the technological field.

### 1.5.3 Multilingual and Cross-Cultural Applicability

A significant contribution of this study includes the fact that this system would be able to apply multilingually and across cultures. To ascertain the detection of fake news within different linguistic and cultural contexts, the test is carried out in various languages such as English and Chinese. This addresses the need for effective global solutions in a connected world.

### 1.5.4 Continuous Learning and Adaptability

This feature would help to continuously learn from it and adapt to the ever-evolving tactics of disinformation, making sure that it remains effective over time, thus providing an effective solution that is future-proof against the ever-changing landscape that is digital disinformation.

### 1.5.5 Academic and Research Contributions

It provides an important contribution to the academic community by uniquely filling important gaps in multilingual fake news detection literature. The new framework integrating content analysis and source credibility thus provides the base for future studies. Second, while exploring the ethical dimensions of AI in fake news detection, the study adds important insight into how to develop such technologies responsibly.

# 2 RESEARCH GAP

Several areas where gaps remain in existing literature, thereby informing the intentions of this study. One glaring void is that of efficient multilingual false news detection systems. Although much work has gone into developing detecting algorithms for English language material, the technologies often fail to work in other languages, thus leaving people who do not speak English open to false information. Our work looks to fill this gap by developing a multilingual detection system that will be able to inspect information in at least two languages, hence making the methods of false news detection more expansively applicable around the globe.

Another notable gap in the literature is the integration of content analysis with real-time source credibility assessment. Most of the systems address only textual analysis and overlook the need for integrating content analysis with source credibility assessment in real time. This limitation becomes particularly a problem as fake news can originate from sources that were historically unreliable and keep changing. This study fills this gap by integrating within its framework an API-based source credibility score which is retrieved in real-time from third-party services.

On the other hand, there is little that is known about how easily models on fake news detection can be adapted. Most systems are based on static models trained on historic data. Although these algorithms may prove useful in detecting forms of fake news that have been previously identified, they are likely to underperform on novel and emerging forms of disinformation. The detection systems must be agile and capable of learning from new information at any time, since fake news is constantly evolving. This study closes this gap by including continuous learning methods in detection systems, so they can keep on learning new threats and hence get stronger over time.

Finally, there is a clear disconnection on the moral implications of fake news detection. Though well researched are techniques for identifying such fake news, less attention has been given to the ethical and social implications of these tools. Some of the important issues that need to be dealt with are possible false positives, transparency of AI-driven decisions, and its impact on free expression. This study will also focus particularly upon these ethical concerns so that the system works responsibly and in a fair manner apart from being successful.

This will, then, fill numerous gaps in the existing knowledge regarding the identification of misleading news. In these ways, this work focuses on multilingualism, comprehensive content and source credibility analysis, lifelong learning, and ethics to build a more resilient, flexible, and internationally applicable solution to the problem of misleading news. The findings will provide credence to information in this digital age and future efforts against the flow of false information.

| Research paper / Feature | A Hierarchical Framework for Fake News Detection Using Semantic Analysis and Source Reliability [20] | Content Features and Machine Learning Based Effective Fake News Detection [21] | Content-Based Fake News Detection with Machine and Deep Learning: A Systematic Review [22] | Multilingual Fake News Detection with Deep Learning [23] | SA-Bi-LSTM: Self Attention with Bi-Directional LSTM-Based Intelligent Model for Accurate Fake News Detection to Ensured Information Integrity on Social Media Platforms [24] | My Approach |
|---|---|---|---|---|---|---|
| **Multilingual** | | | | ✓ | | ✓ |
| **Real-time API integration** | | | | | | ✓ |
| **Source credibility Assessment** | ✓ | | | | | ✓ |
| **Hybrid (CNN+ BiLSTM)** | | | | | ✓ | ✓ |
| **Content Analysis** | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| **Continuous Learning Methods** | | | ✓ | | | ✓ |

Table 2. Gap Table

# 3 RESEARCH PROBLEM

In a fast-changing, digital, and connected world, one of the central questions this research seeks to answer is how to ensure more effective and timely detection and containment of fake news across languages in a very virulent setting. Although machine learning-based fake news detection systems have made significant strides, most existing solutions are still relatively narrow in scope. They primarily focus on analysis in a single language and often fail to account for the multilingual nature of the global information landscape. This leaves a large portion of the non-English speaking population vulnerable to well-funded disinformation campaigns. Most of the current approaches overlook this major element, source reliability. Most of the fake news has its origin from sources with low credibility; hence, not assessing the credibility of the source can never guarantee efficiency in detection. Thus, source reliability forms an important parameter to be considered to make the process of detection completer and more holistic.

After all, a great many static approaches to detection often miss quick responses to novel forms of disinformation, since the nature of false news is dynamic in tactics and format. Therefore, this research problem will focus on developing a global, flexible, and multilingual fake news detection system that is basically based on and NLP techniques with real-time source credibility assessment, with a view to detecting fake news more accurately and mitigating them in various real-life scenarios. This problem about information integrity is very important to be solved in this digital era, and the effectiveness of fake news detection initiatives being taken across the world needs to be improved.

# 4 RESEARCH QUESTIONS

- To what extent, and how accurately, can a multilingual approach using advanced NLP techniques and source credibility assessment improve the detection of false news across multiple languages?

- What are the important challenges and considerations in developing a multilingual fake news detection system, which considers at least two languages for the effectual processing and analysis of content, together with source credibility assessment done in real time?
- How could continuous learning mechanisms and real-time source credibility assessment be integrated into false news detection systems so that they can adapt to new and emerging misinformation patterns?
- Which ethical concerns should one raise and follow while developing and deploying an AI-driven fake news detection system to prevent its potential biases and false positives?

# 5 OBJECTIVES

## 5.1 Main Objective

To develop a comprehensive and robust platform capable of accurately detecting and verifying the authenticity of multimedia content across multiple modalities, including deepfake audio, deepfake video, fake news, and tampered images. The system will employ cutting-edge deep learning techniques to enhance the reliability and integrity of digital media, thereby mitigating the spread of misinformation and digital manipulation.

Reference to the multimodal multimedia integrity Verification system diagram [Appendices 1].

## 5.2 Specific Objective

To design and implement a novel multilingual fake news detection that assess both content and source credibility assessment, using integration of deep text analysis with advanced NLP and real-time API-driven source credibility assessment for enhancing detection accuracy across multiple languages.

## 5.3 Sub Objectives

- Collect and preprocess datasets of real and fake news in both English and Chinese languages to create a unified format for further multilingual analysis.
- Develop and integrate state-of-the-art models of NLP, such as multilingual BERT, with deep learning in multiple languages and real-time credibility assessment of sources for the improvement in detection accuracy.
- Designing a continuous learning mechanism for the false news detection system itself, so it can, at runtime, adapt to new patterns of misinformation.
- Ethical risks testing and identification during the development stage of the detection system to reduce the risk of false positives and guarantee transparency and impartiality.
- Thoroughly test and evaluate the system in multiple real-world scenarios to ensure its effectiveness and robustness in detecting misinformation from a wide variety of languages and sources.

# 6 RESEARCH METHODOLOGY

In this work, a mixed-methods approach will be applied to the development and evaluation of the proposed false news detection system. It helps bridge the quantitative and qualitative methodologies. The methodology avails the opportunity for an in-depth analysis with respect to the technical components of the system and the contextual factors that influence its effectiveness. The different phases of the study will be data collection, feature extraction and analysis, model development, training and evaluation, deployment, and monitoring. The design of the system shall be such that it is technically robust and contextually relevant; it should work in the real world, with a focus on this very end-goal being placed at each phase.

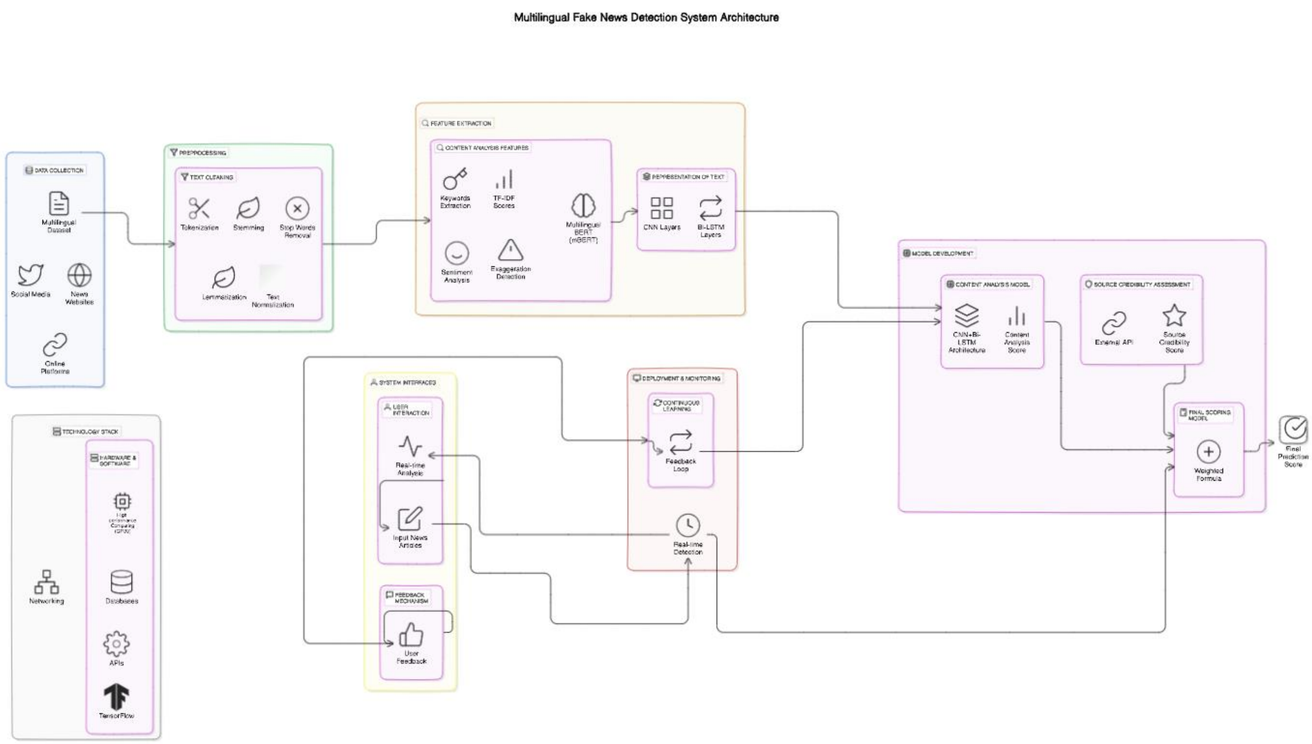


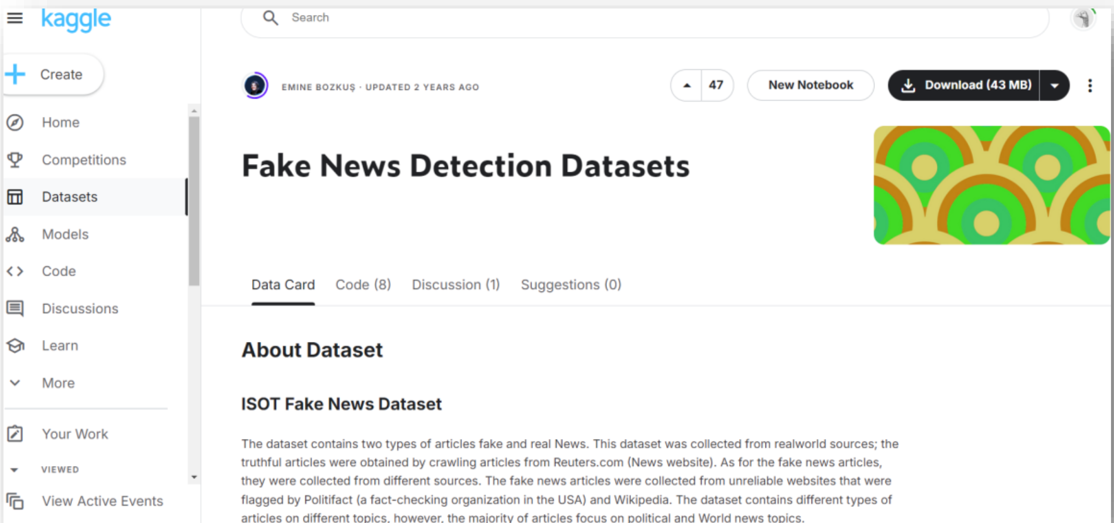

*Figure 1: Multilingual Fake news Detection System Architecture*

## 6.1 Methodology: Data Collection to Deployment Process

### 6.1.1 Data Collection and Preprocessing

The first step in developing a multilingual fake news detection system is the collection and preprocessing of a diversified dataset in both English and Chinese. It shall focus on textual data, originating from multiple platforms such as news websites and social media to ensure that it contains complete information on real and fake news. These preprocessing steps on this data include tokenization, removal of stop words, stemming, and lemmatization, which help remove irrelevant information from the text. All these preprocessing steps are to clean and normalize the text for subsequent feature extraction and model training.

#### 6.1.1.1 Selection of Datasets




Link to Dataset - https://www.kaggle.com/datasets/emineyetm/fake-news-detection-datasets

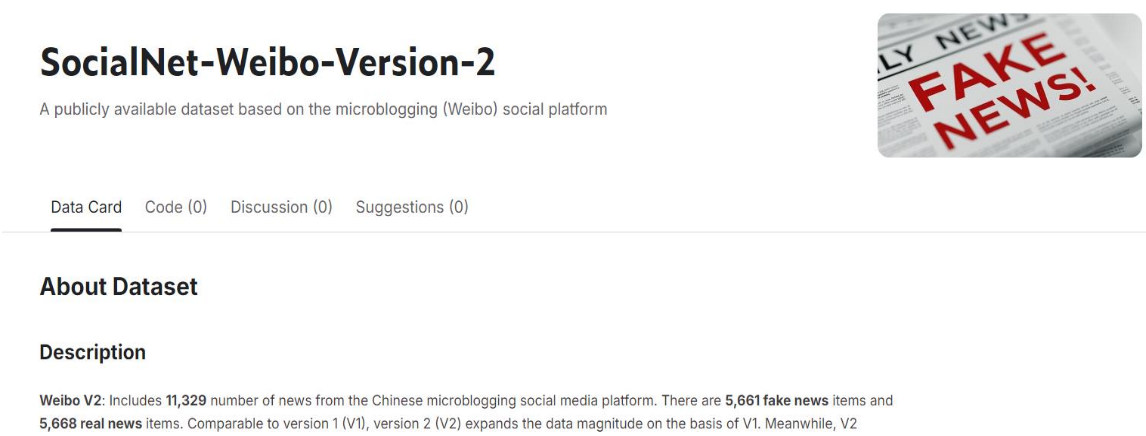



*Figure 2: Datasets*

Link to dataset - https://www.kaggle.com/datasets/yangzhou32/socialnet-weibo-version-2

### 6.1.2 Feature Extraction and Analysis

After preprocessing, feature extraction and analysis will be done. Advanced NLP techniques, such as those offered by multilingual BERT, will be applied to extract linguistic features like syntax, semantics, and contextual patterns from the text. These features will have a role in differentiating between real and fake news. In the establishment of source credibility, an API-based approach will be used. It will call out to a third-party service for the credibility scores in real-time, focusing on parameters such as historical accuracy and domain reputation. The score, as a percentage, will provide an up-to-date measure of a news source's trustworthiness.

### 6.1.3 Model Development

This research investigates a model that combines content analysis with source credibility evaluation in a multilingual framework. The architecture proposed is a hybrid CNN-BiLSTM model:

CNN (Convolutional Neural Networks): It captures local text features such as key phrases and word combinations.
BiLSTM (Bidirectional Long Short-Term Memory): It puts the content in sequential dependencies, which helps understand the flow of ideas and narrative structure.

A Content Analysis Score is generated by such hybrid architecture that analyzes the textual content. At the same time, Source Credibility Score is computed in real-time according to API’s output. These two scores are then merged to be used for final prediction score, which means it can check information across different modalities to basically tell whether they are trustworthy – textual contents and source credibility. The multilingual dataset is utilized to train the model so that it can effectively process English and Chinese data at once.

### 6.1.4 Model Training and Evaluation

The preprocessed dataset will then be trained to identify patterns associated with fake news in both English and Chinese. During training, many evaluation metrics will be calculated to benchmark model performance, including accuracy, precision, recall, and the F1-score. The training dataset and an independent validation set will be used to evaluate the model's robustness and its ability to generalize across different datasets. Besides, extensive testing will include cross-validation on data the model has not seen, and the checking of its effectiveness in real-world scenarios.

### 6.1.5 Deployment and Monitoring

To make the fake news detection system effective over time, a continuous learning and adaption methodology will be implemented. For this purpose, the training dataset must be updated periodically with new examples of fake news, and fine-tuning performed on the model regarding user feedback and other measures of performance. Online learning techniques can be exploited to enable the model to learn from new data in real-time, thus ensuring it is able to adapt to new patterns of fake news. This technique in continuous learning will help ensure that the relevance and precision of the system are not overtaken by time by the development of strategies aimed at spreading disinformation.

## 6.2 Software Architecture of the Research

The adapted Software Development Life Cycle has formed the foundation for which the software architecture of the study has been based. This architecture ensures that the development will proceed in a systematic and iterative manner within phases, influencing both the software and the hardware. It consists of planning, analysis, design, implementation, and maintenance phases [25]. Each of these phases contributes to careful testing, refinement, and execution. Different phases are duly explained below [26].

- **Planning**: This is the initial phase where all the requirements that need to be fulfilled to achieve the objectives of the project are gathered and documented. It involves the setting of clear goals, definition of scope for the project, and identification of the resources which would be needed for its proper execution.
- **Analysis**: This stage involves a thorough examination of the requirements of the entire project. The process followed includes problem identification, root cause analysis, collection of data, and its evaluation. From here, the detailed system specifications can be determined and insight into the user's needs is obtained.

- **Design**: In this stage, the insights gained are used in developing the architecture of the system and a user interface. It revolves around the making of a detailed blueprint of the components, flows, and interactions that go to make up the system. Designs should support all requirements of both the user and the technical system.
- **Implementation**: This is the phase of actual coding and development of the system. The software will be built following the design specifications, and integration with testing goes on simultaneously to make sure that every component works correctly. The result should be a prototype that can work or a final product ready to be deployed.
- **Maintenance**: The system enters the maintenance phase, and thus continues with constant monitoring and updating to rectify issues or bring improvements to it. This phase of the life cycle is held responsible for keeping the system running, efficient, and relevant to the changing requirements or new challenges.

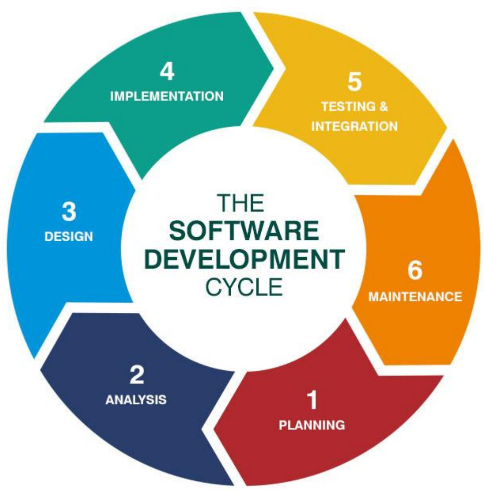



Figure 3. SDLC Methodology Life Cycle

## 6.3 Defining and Analyzing Requirements

Requirements gathering and elicitation is an extremely crucial step in the process to ensure that the newly built fake news detection system meets all requirements and expectations of the stakeholders and users for whom the system is targeted. Various requirements are gathered and understood on a systematic approach in this phase, which would form the basis for design, development, and deployment.

### 6.3.1 Identifying Stakeholders

The entire process of requirement gathering initiates with the identification of the stakeholders of the system. The stakeholders include the public and news organizations, social media platforms, fact-checking groups, government agencies, etc. The requirements and views of all such groups need to be taken into consideration. In-depth understanding of the specific needs and expectations that these stakeholders have of the fake news detection system can be gained through interviews, questionnaires, and workshops involving these stakeholders.

### 6.3.2 Requirements Gathering

In this phase, detailed requirements are gathered from interviews, questionnaires [Appendices 2], workshops, and consultations, which identify the stakeholders. That way, one will capture the goals, challenges, and features of the system as desired by it. Observation of current workflows and reviews of existing tools will explain areas for improvement and pain points. There will also be the addressing of legal and ethical considerations that would ensure compliance. All these ensure a comprehensive compilation of functional and non-functional requirements.

### 6.3.3 Documenting Requirements

Gathered requirements are then put down in writing to clearly direct development. Functional requirements include multilingual analysis and real-time credibility assessment, while non-functional ones put a focus on performance, security, and the user interface. Use cases and user stories explain how stakeholders will make use of entities; the data requirements describe any given sets of data required to implement the proposed system, while legal requirements describe compliance measures. Good documentation would lead to the entity understanding the scope of work.

### 6.3.4 Requirements Analysis

The next step is the feasibility, relevance, and impact analysis of the requirements. A feasibility study assesses any technical and financial constraints; a risk and impact analysis show possible challenges and their consequences for the project. The dependency mapping describes any interdependencies, guaranteeing a realistic timeline and deliverables. This phase refines the requirements considering the goals set to ensure successful implementation.

### 6.3.5 Requirements Prioritization

The requirements are ranked in order of importance to the stakeholder, their alignment with the goals of the project, and their technical feasibility. High-impact requirements, like source credibility in real-time, will be addressed first in this process. Resource availability and risk areas, such as ethical considerations, are accounted for in the selection of what should be done at the most preliminary stage. This will instrumentally help in the determination of a roadmap for the project, thereby guiding the development.

### 6.3.6 Requirement Validation

Requirements are finally validated to ensure that they indeed reflect the needs of all stakeholders and are, in fact, feasible. Review sessions, prototyping, and proof-of-concept models offer stakeholders a chance to give feedback early in the process. With requirement traceability, each requirement can be traced through development. Initial testing identifies any gaps or issues. Validation ensures a project built upon clearly defined and agreed-upon requirements.

## 6.4 Selection of Tools and Technologies

- Selection of Datasets
    - Kaggle

- Feature Extraction and Analysis
    - mBERT (Multilingual BERT)
    - NLTK (Natural Language Toolkit)
    - Jieba
    - Scikit-learn

- Model Development
    - TensorFlow/PyTorch
    - Keras
    - mBERT

        - API Integration

- Model Training and Evaluation
    - TensorFlow/PyTorch
    - Scikit-learn
    - Cross-Validation Libraries
    - Google Colab/Kaggle Notebooks
    - GPU-enabled Hardware

- Deployment and Monitoring
    - Docker
    - Kubernetes
    - Flask/Django
    - Prometheus/Grafana
    - Continuous Integration/Continuous Deployment (CI/CD) Tools Such as Jenkins or GitLab
    - AWS

# 7 PROJECT REQUIREMENTS

## 7.1 Functional Requirements

- **Data Collection:**
    - Gather multilingual datasets containing real and fake news articles in both English and Chinese.
    - Ensure that data collected is sourced from various platforms such as news websites and social media.

- **Data Preprocessing:**

- Clean and normalize text data using tokenization, removal of stop words, stemming, and lemmatization.
- All data formats are standardized across languages so that uniform analysis can be carried out.

- **Feature Extraction and Analysis:**
  - Semantic, Syntactic, and Stylistic Features: Extract using advanced NLP techniques such as multilingual BERT.
  - Perform an API to run source credibility scores in real-time based on historical accuracy and domain reputation.

- **Model Development:**
  - Develop a hybrid CNN-BiLSTM model to perform content analysis, capturing both local text features and sequential dependencies.
  - Source credibility scores should be integrated with content analysis to derive the final detection score.

- **Model Training:**
  - Train the model based on multilingual datasets with techniques of supervised learning.
  - The model is to be evaluated based on metrics like accuracy, precision, recall, and F1-score for the evaluation of the model.

- **Detection:**
  - Design a model for real-time fake news detection in several languages by integrating both content analysis and source credibility.

- **System Deployment:**
  - Deploy the system in a real-world environment, ensuring the processing and analysis of new data in real-time.
  - Integrate continuous mechanisms for learning to keep the model up to date against new data and emerging patterns of fake news.

- **Monitoring and Feedback:**
    - Continue monitoring the performance and updating the model with feedback from users and new data.
    - Enhance its detection accuracy and reduce false positives through user feedback into the system.

## 7.2 User Requirements

- **Multilingual Capability:**
    - The system should support more than one language like English and Chinese for fake news detection.
    - Users should be able to input any news article in either of these two languages for checking.

- **Real-time Analysis:**
    - Users expect real-time analysis regarding the credibility of news content and sources.
    - The system shall update instant feedback regarding the probability that the news is fake.

- **User Feedback Integration:**
    - The system should make provisions for user feedback from the output of the detection results to enable updates of the model over time.

- **Ethical Transparency:**
    - The user wants to be able to understand AI-driven decisions, including why a news item has been flagged as fake. The system shall decrease bias and false positives to retain trustworthiness.

## 7.3 System Requirements

- **Hardware Requirements:**
  - High-performance computing resources with GPUs for handling deep learning tasks.
  - Sufficient storage capacity for large multilingual datasets and real-time data processing.

- **Software Requirements:**
  - Advanced NLP libraries like TensorFlow, PyTorch for model development
  - APIs for the source credibility assessment task in real-time
  - Database management systems for storing and retrieving datasets, model outputs, and user feedback.

- **Network Requirements:**
  - Reliable and fast internet connectivity shall be required for real-time data retrieval, making API calls
  - Secure channels of communication to protect user data and hence privacy

## 7.4 Non - Functional Requirements

- **Performance:**
  - The system is supposed to process and analyze huge volumes of data in real-time.
  - It should be accurate and reliable regarding the detection of fake news.
- **Scalability:**
  - It should be highly scalable in terms of large data loads and increasing user demands without degradation of performance.
- **Reliability:**

- The system should provide consistent and reliable results for fake news detection.
- It must be resistant in the case of several types of fake news and manipulations.

- **Usability:**
    - The system should offer intuitive, accessible user interfaces for all categories of users.
    - This enables easy interaction and fast understanding among people with different technical backgrounds.
- **Security:**
    - The system shall provide security and privacy of data by protecting sensitive information.
- **Adaptability:**
    - The system shall adapt to new patterns of fake news through in-built learning mechanisms continuously.
    - It should be able to take in user feedback for fine-tuning its detection capabilities and making them better over time.
- **Maintainability:**
    - The system should be relatively easy to maintain and update, so it allows enhancements and bug fixes regularly.
    - Documentation should be adequate to support development and maintenance efforts in the long term.

# 8 BUDGET AND BUDGET JUSTIFICATION

- There will be a need to buy large, multilingual datasets from various sources such as news websites, social media, and paid data providers. The cost of **Rs 3000 to Rs.10000** includes access fees to premium datasets, as well as potential licensing fees for using the data.
- This budget accounts for ongoing usage costs associated with calling external APIs of Rs. **10,000 to Rs.25000**. The cost will depend on the volume of API requests made.

# 9 GANTT CHART

Below Gantt chart illustrates a timeline for the project from June 2024 to April 2025 with its major phases and duration.

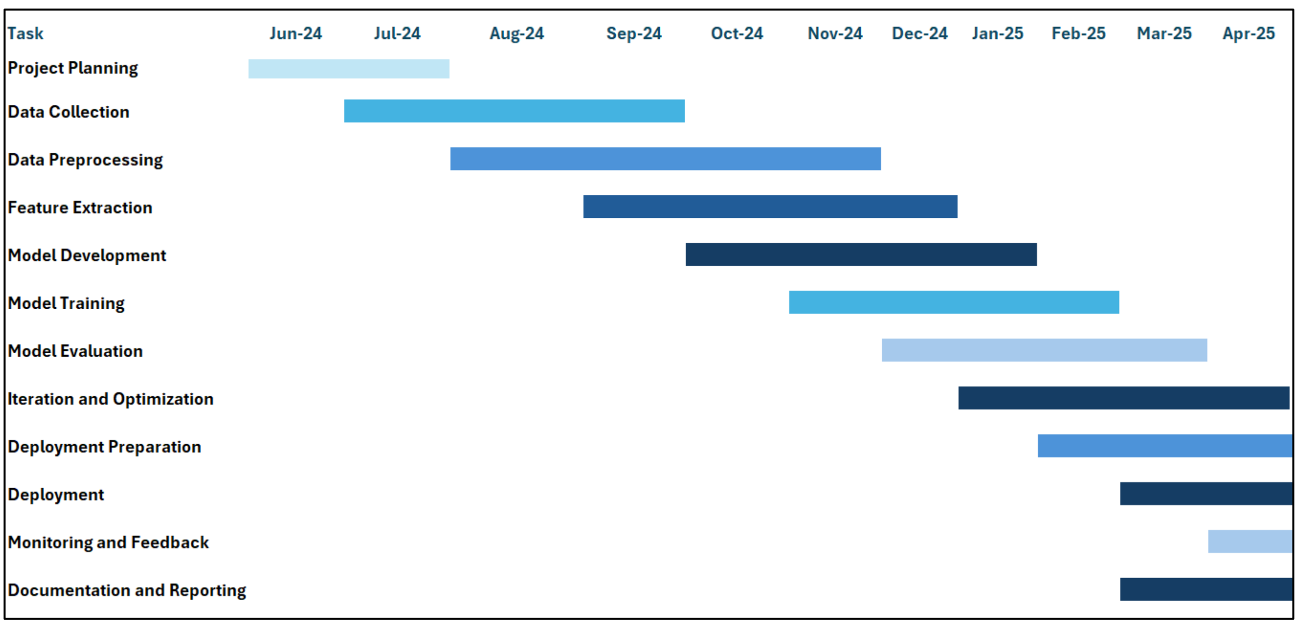



*Figure 4: Gantt Chart*

# 10 COMMERCIALIZATION ASEPECTS OF THE PRODUCT

The commercialization potential associated with this discovery is very large because the demand for effective countermeasures against misinformation and fake news is indeed spiraling upwards in various industries. A staunch, multilingual fake news detection system would have a good deal of significance for business from government agencies to social media platforms, media companies, and cybersecurity firms, in a nutshell, at a time when the integrity of information means everything.

A major commercialization avenue is licensing the technology to media entities and social media networks. These are organizations leading in the fight against disinformation; hence, they require the capability of detecting false material in different languages. By implementing the proposed method in their existing content management and distribution systems, these organizations will be better placed to ensure that the quality of information distributed is improved by filtering, in advance, fake news from going out to the public. The system can also be packaged as a SaaS solution, which will enable businesses to sign up for it and easily integrate with current infrastructure.

The cybersecurity sector is another wide-open opportunity for commercialization. With increasing misinformation being used as a tool for cyberattacks through sophisticated strategies involving social engineering, phishing, and disinformation campaigns, the need for advanced solutions to safeguard clients is what cybersecurity companies are seeking. The threat intelligence platform would be inquisitive about using the enhanced fake news detection system to analyze content from various sources for the identification and analysis of threats and their neutralization in real-time. For any business, financial institution, or government organization, the ability to do so would be imperative for the preservation of public faith in its respective operation.

There is also great potential for collaborations with government agencies and NGOs in projects involving media literacy and awareness campaigns. Such a system would, therefore, help to reduce the spread of misinformation during times such as elections, health epidemics, and social movements. This research would greatly save democratic processes and public health by providing tools for the identification and flagging of fake news in several languages. It can also be customized to answer the needs of specific markets or sectors. For instance, it would first identify highly linguistically diversified languages and, within a sector, for instance, in health and finance, fine-tune for disinformation. This gives more reach and effectiveness in other areas.

The continuous learning ability may prove to be a very powerful selling point, as it will offer customers a solution that grows with the changing landscape of misinformation. This way, the system guarantees relevance and effectiveness over time for its clients. Moreover, the ethical framework already built into the product is what gives the assurance of fairness and transparency; therefore, this places the product in a different category compared to others available in the market, thus drawing in businesses committed to ethical AI practices. Inclusion of academic institutions and research organizations would further develop and enhance the technology of the commercializing plan. This would ensure that the system is at the leading edge of its field of new discoveries and developments.

# CONCLUSION

This research attempts to tackle the increasing problem of fake news detection amidst fast-paced information transfer and interconnectedness in almost every part of the world today. In this regard, a multilingual fake news detection system has been proposed and regarded as one great leap forward in integrating advanced natural language processing and real-time analysis of source reliability assessment. This goes beyond just developing an overall adaptable, comprehensive system that can look at data across different languages like English and Chinese. It will help facilitate cross-examination amongst various sources on different occasions so that they can spot out fake news items as they arise within their circumstances. The significance of this investigation lies not only in addressing current limitations faced by existing detection methods which are often devoid of multilingualism aspects or fail to include evaluation of source trustworthiness but also establish a strong basis for fighting against distortion of news in this era of globalization.

To keep lagging on new and developing types of misinformation, the continuous learning mechanism of this system is very important for its efficiency over time. As a result, the system can easily respond to the quickly evolving world of digital disinformation. Besides, this research emphasizes ethical aspects in designing the system so that fairness and transparency are guaranteed. In a society that is overly obsessed with the effects of AI and machine learning technologies whether positive or negative, implementation of such systems is likely to be trusted as people believe that they will save them from potential harms.

This research caters for immense commercialization opportunities for the model even as it shows possible utilization among media companies, social media platforms and cybersecurity firms. With its flexible and ethical construction, this system becomes more marketable as it can accommodate customized versions developed for certain sectors or regions. Overall, this research aims to bridge the knowledge gap in fake news detection by providing a technologically advanced and ethically sound solution, contributing to the preservation of information integrity in the digital age. Therefore, it emerges as an effective instrument for managing and countering disinformation thus serving both individuals and institutions well.

# REFERENCES


[1] E. Aïmeur, S. Amri, and G. Brassard, "Fake news, disinformation and misinformation in social media: a review," *Soc. Netw. Anal. Min.*, vol. 13, no. 1, Dec. 2023, doi: 10.1007/s13278-023-01028-5.

[2] E. Aïmeur, S. Amri, and G. Brassard, "Fake news, disinformation and misinformation in social media: a review," *Soc. Netw. Anal. Min.*, vol. 13, no. 1, p. 30, 2023, doi: 10.1007/s13278-023-01028-5.

[3] N. Ahuja and S. Kumar, "Mul-FaD: attention based detection of multiLingual fake news.," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 3, pp. 2481–2491, 2023, doi: 10.1007/s12652-022-04499-0.

[4] K. M. Karaoğlan, "Novel approaches for fake news detection based on attention-based deep

multiple-instance learning using contextualized neural language models," *Neurocomputing*, vol. 602, p. 128263, 2024, doi: https://doi.org/10.1016/j.neucom.2024.128263.

[5] J. Benítez-Andrades, J. Alija-Perez, M.-E. Vidal, R. Pastor Vargas, and M. García-Ordás, "Traditional Machine Learning Models and Bidirectional Encoder Representations From Transformer (BERT)-Based Automatic Classification of Tweets About Eating Disorders: Algorithm Development and Validation Study," *JMIR Med. Informatics*, vol. 10, p. e34492, Feb. 2022, doi: 10.2196/34492.

[6] J. P. Baptista and A. Gradim, "Understanding Fake News Consumption: A Review," *Social Sciences*, vol. 9, no. 10. 2020, doi: 10.3390/socsci9100185.

[7] D. Lan and T. Tung, "Exploring fake news awareness and trust in the age of social media among university student TikTok users," *Cogent Soc. Sci.*, vol. 10, Jan. 2024, doi: 10.1080/23311886.2024.2302216.

[8] B. Hu, Z. Mao, and Y. Zhang, "An overview of fake news detection: From a new perspective," *Fundam. Res.*, 2024, doi: https://doi.org/10.1016/j.fmre.2024.01.017.

[9] T. Mehedi, A. Rahman, A. Kabir, and M. R. Islam, "A Review of Methodologies for Fake News Analysis," *IEEE Access*, vol. PP, p. 1, Jan. 2023, doi: 10.1109/ACCESS.2023.3294989.

[10] N. Oliveira, P. Pisa, M. Andreoni, D. Medeiros, and D. Menezes, "Identifying Fake News on Social Networks Based on Natural Language Processing: Trends and Challenges," *Information*, vol. 12, p. 38, Jan. 2021, doi: 10.3390/info12010038.

[11] S. Hamed, M. Aziz, and M. R. Yaakub, "Fake News Detection Model on Social Media by Leveraging Sentiment Analysis of News Content and Emotion Analysis of Users' Comments," *Sensors*, vol. 23, pp. 1–27, Feb. 2023, doi: 10.3390/s23041748.

[12] D. Dementieva, M. Kuimov, and A. Panchenko, "Multiverse: Multilingual Evidence for Fake News Detection," *J. Imaging*, vol. 9, p. 77, Mar. 2023, doi: 10.3390/jimaging9040077.

[13] S. Majumder and D. Das, "Studies Towards Language Independent Fake News Detection," *Proc. 18th Int. Conf. Nat. Lang. Process.*, pp. 439–446, 2021, [Online]. Available: https://aclanthology.org/2021.icon-main.53.

[14] A. Giachanou, G. Zhang, and P. Rosso, *Multimodal Multi-image Fake News Detection*. 2020.

[15] M. Sudhakar and K. P. Kaliyamurthie, “Detection of fake news from social media using support vector machine learning algorithms,” *Meas. Sensors*, vol. 32, p. 101028, 2024, doi: https://doi.org/10.1016/j.measen.2024.101028.

[16] S. F. Ahmed *et al.*, “Deep learning modelling techniques: current progress, applications, advantages, and challenges,” *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 13521–13617, 2023, doi: 10.1007/s10462-023-10466-8.

[17] L. Hu, S. Wei, Z. Zhao, and B. Wu, “Deep learning for fake news detection: A comprehensive survey,” *AI Open*, vol. 3, pp. 133–155, 2022, doi: https://doi.org/10.1016/j.aiopen.2022.09.001.

[18] F. C. C. Santos, “Artificial Intelligence in Automated Detection of Disinformation: A Thematic Analysis,” *Journal. Media*, vol. 4, no. 2, pp. 679–687, 2023, doi: 10.3390/journalmedia4020043.

[19] M. Choraś *et al.*, “Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study,” *Appl. Soft Comput.*, vol. 101, p. 107050, Dec. 2020, doi: 10.1016/j.asoc.2020.107050.

[20] Y. Cardinale, I. Dongo, G. Robayo, D. Cabeza, A. Aguilera and S. Medina, "T-CREo: A Twitter Credibility Analysis Framework," in IEEE Access, vol. 9, pp. 32498-32516, 2021, doi: 10.1109/ACCESS.2021.3060623.

[21] Nicola Capuano, Giuseppe Fenza, Vincenzo Loia, and Francesco David Nota. 2023. Content-Based Fake News Detection With Machine and Deep Learning: a Systematic Review. Neurocomput. 530, C (Apr 2023), 91–103. https://doi.org/10.1016/j.neucom.2023.02.005

[22] S. E. Vadakkethil Somanathan Pillai and A. S. R, "A Hierarchical Framework for Fake News Detection Using Semantic Analysis and Source Reliability," 2024 2nd International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2024, pp. 910-915, doi: 10.1109/ICDT61202.2024.10489413.

[23] Chalehchaleh, R., Farahbakhsh, R., Crespi, N. (2024). Multilingual Fake News Detection:

A Study on Various Models and Training Scenarios. In: Arai, K. (eds) Intelligent Systems and Applications. IntelliSys 2024. Lecture Notes in Networks and Systems, vol 1066. Springer, Cham. https://doi.org/10.1007/978-3-031-66428-1_5

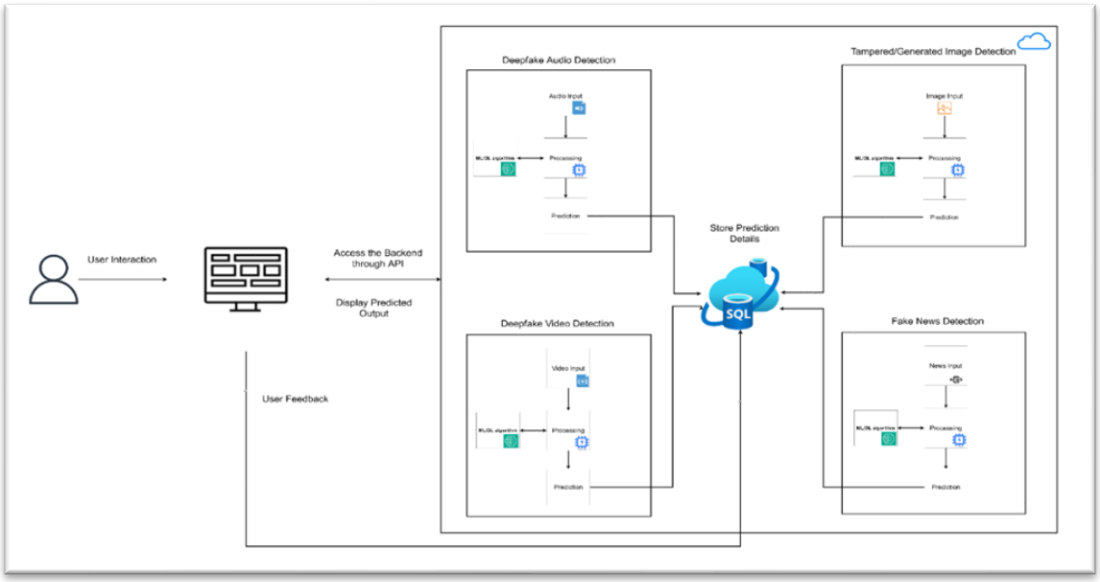
[24] Mohawesh, R., Maqsood, S. & Althebyan, Q. Multilingual deep learning framework for fake news detection using capsule neural network. J Intell Inf Syst 60, 655–671 (2023). https://doi.org/10.1007/s10844-023-00788-y

[25] M. A. Alonso, D. Vilares, C. Gómez-Rodríguez, and J. Vilares, "Sentiment Analysis for Fake News Detection," *Electronics*, vol. 10, no. 11. 2021, doi: 10.3390/electronics10111348.

[26] S. Swanlund, V. Wilson, V. Kapoor, M. Gabbard, and C. Srinath, "Software Development Life Cycle," *J. Soc. Clin. Data Manag.*, vol. 4, Apr. 2024, doi: 10.47912/jscdm.343.

# APPENDICES

## Multimodal Multimedia Integrity Verification System Diagram




*Appendices 1: Multimodal Multimedia Integrity Verification System Diagram*

## Research Survey

# Assessing the Need for Advanced Tools in Detecting Digital Misinformation: A Focus on Deep Fake Audio, Video, Image Tampering, and Fake News

Hello everyone!

We're a team of final-year undergraduates at SLIIT working on an exciting research project focused on the growing threat of digital misinformation. Our goal is to develop a cutting-edge tool capable of detecting deep fake audio, deep fake video, tampered images, and fake news. 🎯

In a world where digital content is easily manipulated, it's more important than ever to ensure the authenticity of what we see, hear, and share online. With your help, we aim to create a reliable and accessible solution that empowers people to identify fake content with confidence. 🌐 🔍

Your insights are invaluable! We'd greatly appreciate it if you could take a few minutes to complete our survey and pass it along to anyone who might be interested, including tech enthusiasts, media professionals, and anyone concerned about online misinformation. Every response brings us closer to making the digital world a safer place! 🙏

https://docs.google.com/forms/d/e/1FAIpQLSdciC0U-wv59aZ21bniHXB0SwR31dJOwCuTiqKR1TS31UsKLg/viewform?usp=sf_link

*Appendices 2: Research Survey*