# Reality Check

Submitted in partial fulfillment of the requirements of the degree

**BACHELOR OF ENGINEERING IN** **COMPUTER ENGINEERING**

By

**Somya Jain 30**

**Saniya Dangat 14**

**Simran Karamchandani 37**

**Tanisha Pandit 46**

**Prof.** Richard Joseph



# Vivekanand Education Society’s Institute of Technology,

**An Autonomous Institute affiliated to University of Mumbai**

# HAMC, Collector’s Colony, Chembur,

**Mumbai-400074**

**University of Mumbai (AY 2024-25)**

# CERTIFICATE

This is to certify that the Mini Project entitled **“ Reality Check ”** is a bonafide work of **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of **“Bachelor of Engineering”** in **“Computer Engineering” .**

## ( Prof. Richard Joseph )

Mentor

## (Prof. ) (Prof. )

Head of Department Principal

# Mini Project Approval

This Mini Project entitled “Reality Check**”** by **Somya Jain (30), Simran Karamchandani (37), Tanisha Pandit (46), Saniya Dangat(14)** is approved for the degree of **Bachelor of Engineering** in **Computer Engineering.**

**Examiners**

**1………………………………………**

(Internal Examiner Name & Sign)

## 2…………………………………………

(External Examiner name & Sign)

Date: Place:

# Contents

## Abstract 5

## Acknowledgments 6

## List of Abbreviations 7

## List of Figures 7

## List of Tables 7

## Introduction 8

* 1. Introduction
  2. Motivation
  3. Problem Statement & Objectives
  4. Organization of the Report

1. **Literature Survey 12**
   1. Survey of Existing System
   2. Limitation Existing system or Research gap
   3. Mini Project Contribution

## Proposed System 16

* 1. Introduction
  2. Architectural Framework / Conceptual Design
  3. Algorithm and Process Design
  4. Methodology Applied
  5. Hardware & Software Specifications

**3.6** Experiment and Results for Validation and Verification

**3.7**  Result Analysis and Discussion

**3.8** Conclusion and Future work.

## References 27

**Abstract:**

Misinformation can cause public confusion, political polarization, and potential harm to public health and safety. This underscores the need for accurate identification of false news. Although fact-checking websites provide essential services, their manual processes restrict their scalability and coverage, and the data they rely on to fact-check user queries is often outdated. While numerous models currently exist, our model is designed to significantly enhance the accuracy of fact-checking while keeping it up to date. This research proposes the development of an automated, web-based fact-checking platform to address this challenge. Our solution, which is a large language model (LLM), leverages a specialized, continuously updated dataset sourced from credible RSS feeds and web scraping techniques to enhance fact-checking accuracy within specific domains. The platform will feature a user-friendly interface, enabling users to submit claims and receive detailed verification results, including explanations and sources. We will evaluate and refine various models to address domain-specific challenges, aiming for a highly accurate fact-checker

**Acknowledgement:**

We would like to express our sincere gratitude to Vivekanand Education Society’s Institute of Technology for their support throughout our project. Their assistance in gathering project-related information was invaluable.

We are delighted to extend our heartfelt appreciation to Mr. Richard Joseph, our TE Mini-Project Mentor, for his generous assistance and valuable advice in developing the project synopsis. His guidance and suggestions were instrumental in our project's development.

Our profound thanks go to Dr. (Mrs.) Nupur Giri, the Head of the Computer Department, and Dr. (Mrs.) J.M. Nair, our Principal, for granting us the opportunity to undertake this project. Their support was essential in successfully completing both the project synopsis and project review.

We would like to express our deep gratitude to all the teaching and non-teaching staff for their unwavering encouragement, support, and selfless assistance throughout the project. Their contributions were indispensable to our project's success.

**List of Abbreviations:**

LLM : Large Language Model

RSS: Really Simple Syndication

RAG: Retrieval Augmented Generation

## List of Figures:

Figure 1: Architectural Framework

Figure 2: Domain Specific Fine Tuning

Figure 3: RAG Implementation

Figure 4: Semi Structured Dataset

Figure 5: Input Query - 1

Figure 6: Output for input query - 1

Figure 7: Improvised Dataset

Figure 8: Input Query 2

Figure 9: Output for input query - 2

Figure 10: Input Query 3

Figure 11: Output for input query- 3

## List of Tables:

Table no.1 Comparisons of LLM models

Table no.2 Literature Survey

## Introduction

**1.1 Introduction:**

Fake news is a serious problem that can mislead the public, fuel animosity, and distort reality. The emergence of advanced language models like ChatGPT has exacerbated this issue due to their tendency to hallucinate; thus, the aim is to reduce these hallucinations because if the language model has accurate data to generate responses from, the chances of absurd outputs decrease. Additionally, the vastness of the internet contributes to the spread of misinformation. While fact-checking websites such as PolitiFact and Snopes provide valuable services, their manual approach limits their capacity to keep up with the overwhelming volume of content. Several existing models use both text and images to substantiate their claims, but they often struggle with accuracy. These models generally involve retrieving evidence, assessing claims, and generating explanations, revealing significant limitations in their effectiveness.

We believe that a novel approach is essential to achieve substantial progress in this field. Therefore, we propose automating the fact-checking process through the development of a specialized dataset tailored for specific domains like history, science. This dataset would facilitate the customization of language models, enhancing the scope and efficiency of fact-checking by leveraging advanced open-source models to systematically evaluate and verify claims. Through rigorous experimentation with various models, we aim to develop a highly accurate fact-checker.

**1.2 Motivation**

The rise of misinformation has become a global issue, particularly as it infiltrates specialized fields such as science, and history. Unlike general misinformation, which can often be easily identified, misinformation in these domains tends to be more subtle and complex. Inaccuracies can mislead people, distorting public knowledge and influencing opinions based on false or misleading claims.

In science, the issue becomes more critical, as misinformation can spread harmful inaccuracies about discoveries, health guidelines, or technology. Unverified claims regarding vaccines, climate change, or medical treatments can lead to public distrust or poor decision-making. The rise of pseudoscience and manipulated studies have made it harder for the public to discern what is accurate. In history, misinformation often takes the form of revisionism, where facts are twisted or events reinterpreted to serve specific agendas. Misrepresentation of key historical events or figures can distort future generations' understanding, affecting collective memory.

To counter these challenges, a robust fact-checking mechanism is crucial. General-purpose fact-checking systems lack the domain-specific expertise required to handle nuanced claims in these specialized areas. Therefore, domain-specific fact-checking requires a system trained on authoritative data. The challenge lies in creating a system that reliably verifies information across these domains using trusted sources like government websites, encyclopedias, and peer-reviewed journals. The model must be fine-tuned and along with RAG implementation, the nuances of science, and history, can be accurately distinguished between fact, opinion, and misinformation.

.

**1.3 Problem Statement & Objectives**

In today’s information-driven world, the proliferation of misinformation is a significant concern. While general-purpose fact-checking tools exist, they lack the capability to verify facts accurately within specific areas. The challenge arises from the complexity of domain-specific knowledge, requiring a nuanced understanding to discern facts from misinformation. In science, verifying claims about discoveries or health necessitates knowledge of rigorous scientific principles and credible references. Similarly, historical claims need verification against established records, which general-purpose fact-checkers may struggle to interpret accurately.

The absence of domain-specific fact-checking systems allows misinformation to spread unchecked, leading to confusion and mistrust. Thus, there is an urgent need to develop a system capable of addressing these gaps, utilizing advanced tools to verify facts based on curated, credible datasets..

#### Objectives:

1. Domain-Specific Verification: Build a fact-checking system that verifies information in science, and history by fine-tuning LLMs like Llama on datasets from trusted sources such as government websites, encyclopedias, and scientific journals.
2. Enhancing Accuracy and Reliability: Train the model on curated, high-quality data, cross-referencing multiple credible sources to improve accuracy and provide reliable fact-checks.
3. Filling Research Gaps: Focus on in-depth, contextual verification within each domain to address the limitations of general-purpose fact-checkers.
4. Improving Response Time and Efficiency: Optimize the system to deliver fact-checks quickly and efficiently by streamlining data retrieval and verification.
5. Creating a User-Friendly System: Design an intuitive interface that simplifies fact-checking queries for users, making it accessible through text, voice, and search options.
6. Ensuring Transparency: Provide source references for all verified facts, fostering trust and discouraging misinformation by ensuring transparency in the verification process.

**1.4 Organization of the Report**

Introduction :

This chapter introduces the core idea behind the Reality Check project, aimed at building a domain-specific fact-checking system covering sports, science, and history. It discusses the motivation driving the development of such a system, highlights the problem statement, and outlines the key objectives that shape the direction of the project.

Review of Related Work :

This section provides an in-depth review of existing fact-checking systems and relevant literature. It examines previous research, studies, and tools related to fact-checking, with an emphasis on the limitations of current systems in verifying domain-specific information. Additionally, it identifies gaps and challenges within the fields of sports, science, and history, setting the stage for the system's unique contribution.

Proposed System :

The third chapter describes the Reality Check system in detail. It covers the architectural framework, the algorithmic process design, and the hardware and software utilized in the system’s development. This chapter also includes a presentation of experimental results and analysis, summarizing the performance and validation of the system. It concludes with a discussion of potential future enhancements, giving a comprehensive overview of the system's development and future direction.

**Literature Survey**

**2.1 Survey of Existing System(LLM MODELS)**

| **Model name** | **Training Data** | **Data Size** | **Accuracy** | **No. of parameters** | **Limitations** | **Training Time** |
| --- | --- | --- | --- | --- | --- | --- |
| LLama | Books, Wikipedia, GitHub, CommonCrawl, C4, ArXiv, and StackExchange | 1.4T tokens | Llama-2-70b : 81.7%. | 65 Billion parameters | Data Bias and Ethical Concerns | 21 days |
| GPT | Books, websites, and other texts.  CommonCrawl dataset | 523gb | ~80-90% | Trillion parameters | Biased and repetitive | 5-6 months  (GPT-4) |
| Bert | BooksCorpus and English Wikipedia | 3Tb | 79.27 | base (110M parameters) and large (345M parameters) | older training data | 4 days |
| Scibert | Semantic scholar papers | 1.14M papers, 3.1B tokens. | 80% | 110 million parameters. | Limited to scientific contexts | - |
| Falcon | books, websites, articles, and other forms of written content. | 5,000 billion tokens, | 76.37% | 180 billion parameters | can exhibit biases present in the training data. | 2 months(Falcon 40B) |
| t5 | C4 dataset | 750 Gb  32,128 subword tokens | 92.30% | 11 billion parameters | Requires substantial computational resources | - |
| Galactica | 48 million papers, textbooks, and other scientific knowledge sources | 106B tokens | 50% | 120 Billion parameters | occurrence of hallucinations. | - |
| Skywork | data filtered from Chinese web pages | 3.2T tokens | 90% | 13B | Biasness and scalability | -- |
| Bloom | ROOTS corpus | 366B tokens | - | 176B | Outdated or incorrect information for current events. | 105 days |
| StarCoder | The Stack with 384 Programming Languages and Github repositories | 1 trillion tokens sourced | 86.6 | 15.5B parameter | Ethical Concerns,  Malicious code | 1 to 2 months |
| GPT-NeoX | Pile, Books, Internet Resources, Github, youtube subtitles, | 20B | 90 % | 20 billion parameters, | Data duplication  Lack of coding evaluations | 34 days |

Table no.1 Comparisons of LLM models

**2.2 Limitation Existing system or Research gap**

| **Paper Title** | **Inference** |
| --- | --- |
| Generating Fact Checking Explanations | 1. DistilBERT Implementation 2. First to Generate Explanations |
| Fake News Detection Using Deep Learning and  Natural Language Processing | 1. Used Word2Vec and LSTM Models 2. Factors Affecting System Accuracy: Training Iterations, Data Diversity, Vector Size 3. Achieved 90% Accuracy |
| End-to-End Multimodal Fact-Checking and Explanation Generation: A Challenging Dataset and Models | 1. Performs a comparative study for the existing datasets used to train fact-checking models. 2. Multimodal |
| Comparative Study of Supervised Learning Algorithms for Fake News Classification | 1. Comparative Study: Logistic Regression, Random Forest, SVM, Gradient Boosting (Best: Random Forest) 2. 99.7% Accuracy with Gradient Boosting Classifier |
| A Novel Text Resemblance Index Method for Reference-based Fact-checking | 1. Performs a comparative study for the existing datasets used to train fact-checking models. 2. Use of Veracity Scanning Model and Text Resemblance Score 3. Achieves 82.31% accuracy |
| Token-Level Fact Correction in Abstractive Summarization | 1. Token-Level Fact Correction for Abstractive Summarization 2. Improved Consistency & Summarization Performance 3. Accuracy : 81.04(BERTScore) |
| A Hybrid Framework Integrating LLM and ANFIS for Explainable Fact-Checking | 1. LLM & ANFIS Model Integration 2. 0.9 F1 Score on FEVER Dataset |
| Automated Fact Checking Using A Knowledge Graph-based Model | 1. ConVe Model is trained on 2 KGs made with Liar datasets 2. 88% precision |

Table no.2 Literature Survey

**2.3 Mini Project Contribution**

Sustainable Development Goal: Peace, Justice, and Strong Institutions

Our objective is to create a system that combats misinformation by enabling rapid and accurate fact-checking across various domains, with a particular emphasis on scientific and historical data.

1. The proposed system combines advanced methods for real-time fact verification using Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), and Knowledge Graphs (KGs) to ensure precise and timely validation of information.
2. Moreover, converting or using the dataset into Knowledge Graphs (KGs) may result in higher accuracy. Knowledge Graphs provide a structured, contextual understanding of interconnected facts, enhancing the analysis of complex relationships for accurate verification.
3. This approach mitigates misinformation by offering reliable, up-to-date facts, reducing the social and economic impacts of false information. Additionally, it raises public awareness about the importance of fact-checking, promoting digital literacy and fostering a more informed society.

## 

## 

## 

## 

## 

## Proposed System

## **3.1 Introduction**

## 

## The proposed system aims to build an advanced fact-checking framework by leveraging cutting-edge technologies such as large language models (LLMs), Retrieval-Augmented Generation (RAG), and knowledge graphs (KGs). In an age where misinformation spreads rapidly, real-time verification of information has become crucial. This system addresses this challenge by ensuring the accuracy and reliability of facts across various domains, particularly in historical and scientific contexts.

## Large language models (LLMs) are powerful AI models trained on vast amounts of text data. They excel at natural language understanding and generation, making them ideal for tasks like fact-checking, where precision in interpreting and responding to queries is critical. However, while LLMs are highly capable, they rely on static, pre-existing knowledge, which can lead to challenges in keeping up with rapidly evolving information.

## To overcome this limitation, Retrieval-Augmented Generation (RAG) is introduced. RAG enhances the system by integrating LLMs with a dynamic retrieval mechanism. This is essential for fact-checking, as the relevance and accuracy of information often depend on its timeliness.

## Additionally, to deepen the system’s ability to understand and analyze complex relationships between entities, knowledge graphs (KGs) will be incorporated. Knowledge graphs represent information through entities (nodes) and their relationships (edges), providing a structured, interconnected way of reasoning about facts. By integrating KGs, the system can ensure not only accuracy but also a more profound understanding of the context surrounding the facts being verified.

## This hybrid approach, combining LLMs for processing, RAG for real-time data retrieval, and KGs for structured reasoning, forms a highly adaptable and modular system. Each component works together to enhance fact-checking accuracy, while also allowing for future extensions and improvements as the information landscape evolves.

**3.2 Architectural Framework / Conceptual Design**

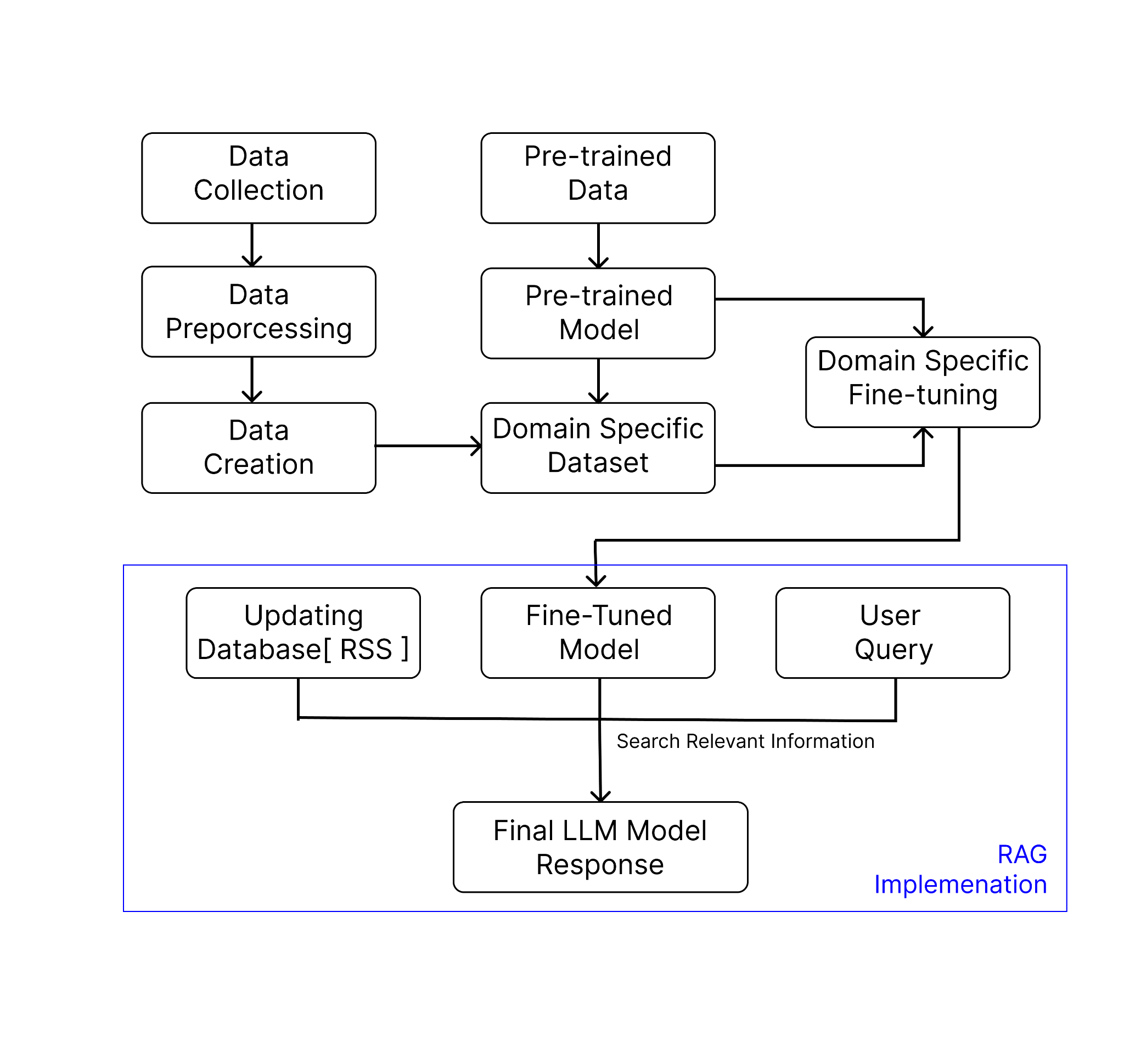


Figure 1: Architectural Framework

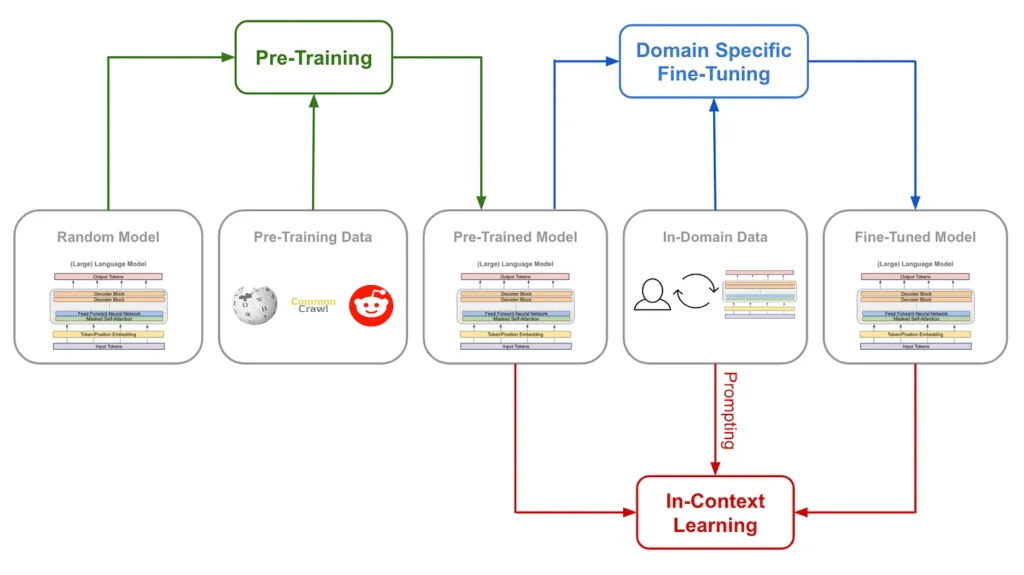


Figure 2 Domain Specific Fine Tuning

**3.3 Algorithm and Process Design**

In this project, the goal is to build a robust fact-checking system by customizing open-source large language models (LLMs), starting with Llama. The process begins with fine-tuning the model using a dataset created through extensive web scraping. To enhance the system's accuracy and adaptability, we plan to implement Retrieval-Augmented Generation (RAG), where a vector database is built using verified data. This allows the system to retrieve up-to-date articles and combine them with the user’s query to generate more informed and accurate responses. Additionally, there is a plan to integrate knowledge graphs into the architecture to further enrich the contextual understanding, although the exact method for this is still being developed. The current priority is dataset creation, which will be tested on various LLM models to evaluate and compare their performance, ultimately providing insight into their respective accuracies. This process is central to the algorithm and system design of the fact-checking framework.

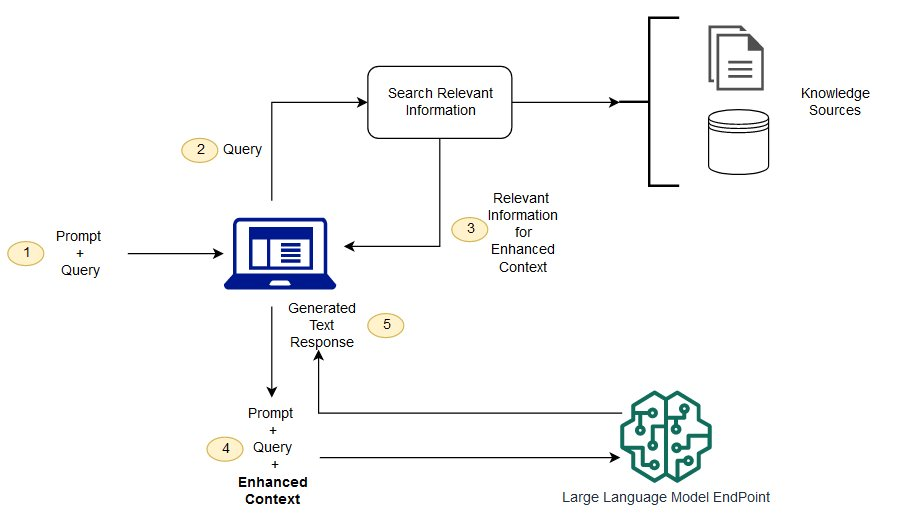


Figure 3 RAG Implementation

**3.4 Methodology Applied**

1. Data Collection

* Use web scraping techniques to extract data from these sources. This includes headlines, article content, publication dates, and author information, etc.
* Compile the extracted data into a structured, personalized dataset in QnA format, ensuring it covers a comprehensive range of topics within the domain .

2. Data Preprocessing

* Removing duplicates, irrelevant content, and incorrect entries.
* Categorize and tag data according to topics, sub-topics, and relevant metadata.

3. Model Training

* Evaluate current models and algorithms for accuracy and suitability.
* Adjust models to improve accuracy and address domain-specific challenges.
* Find areas for improvement and iterate on model development.

4. Claim Verification

* The system cross-references the claim with the dataset to check for accuracy. The verification process includes matching the claim's content with the information in the dataset and assessing the credibility of sources.

5. Explanation and Sources

* After verification, the model provides with the result, indicating whether the claim is

true, false, or uncertain.

* Including a summary of the reasoning behind the verified fact.
* Listing the sources and evidence used in the verification process.

6. UI/UX Design

* Design an interface that allows easy submission of claims and access to verification results.
* Clearly display the verification results, explanations, and sources in an organized manner.

**3.5 Hardware & Software Specifications**

**Tools:**

* **Pre-trained Models:** Using LLMs like LLama, GPT-4, BERT, and specialized models like SciBERT for science.
* **Web Scraping and site testing :** Octoparse, BeautifulSoup, Pabbly, Grist.

**Hardware:**

1. **Server/Cloud Infrastructure:**

* **High-Performance CPU/GPU:** Required for efficient model training and inference (e.g., NVIDIA A100, Intel Xeon processors).
* **High-Speed SSD Storage:** For quick data retrieval and storage.

**Frontend:**

* React.js
* HTML/CSS/JavaScript, Bootstrap

**Backend:**

* Django
* TensorFlow/PyTorch/Cloud–

**3.6 Experiment and Results for Validation and Verification**

Initial dataset:

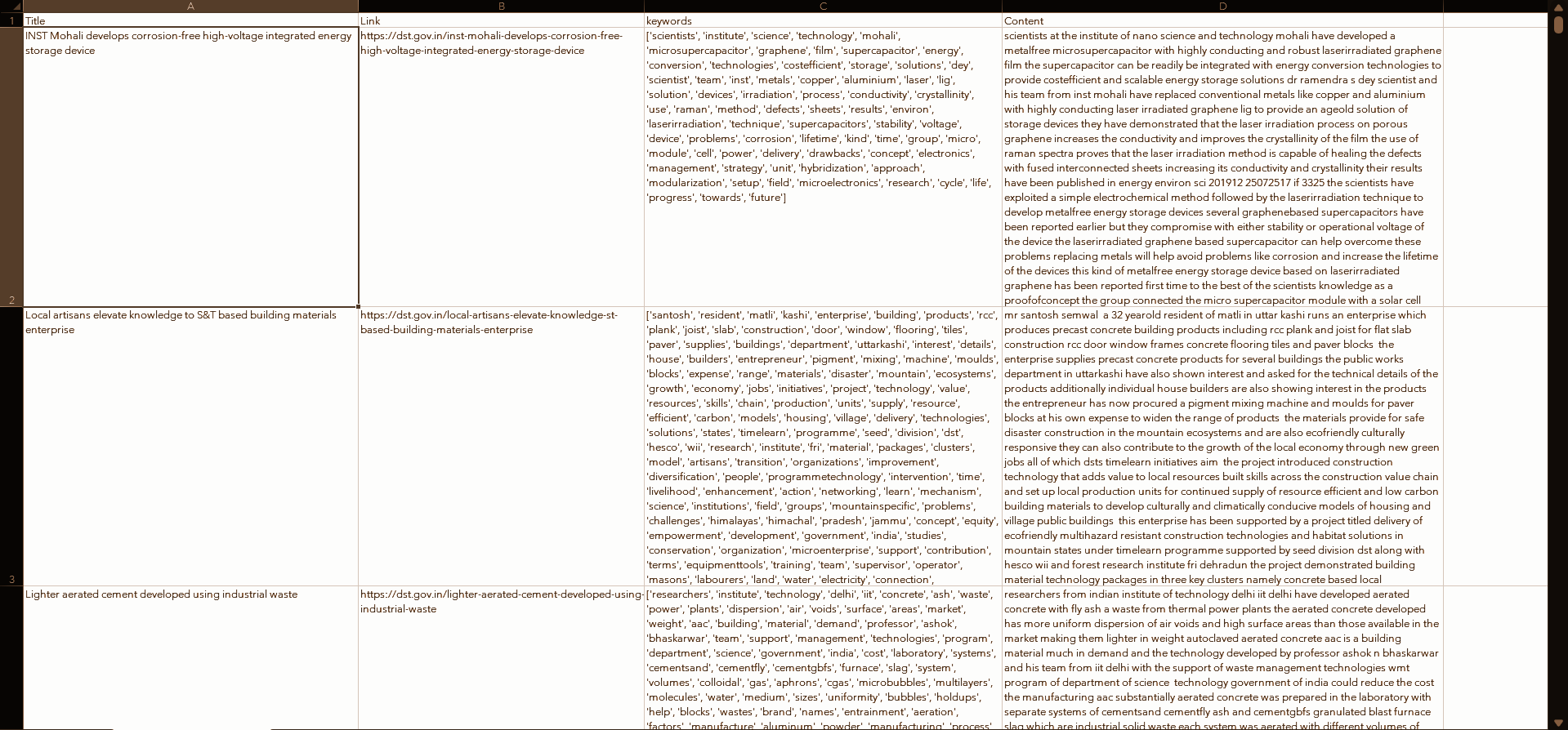
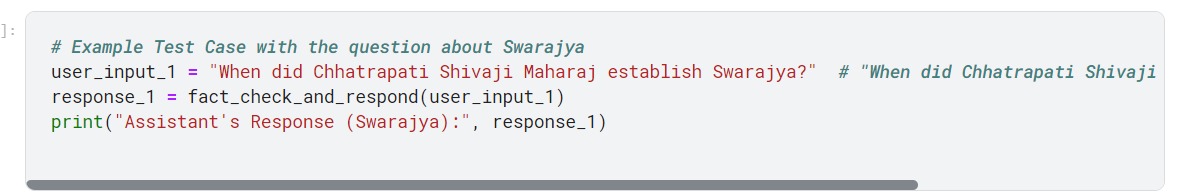


Figure 4. Semi Structured Dataset

Input Query to fine-tuned llama model:

Figure 5. Input Query - 1

Output:

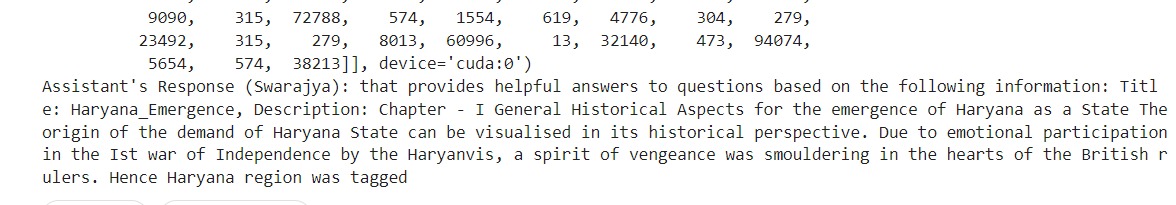


Figure 6. Output for input query - 1

2. Refined as re-formatted dataset

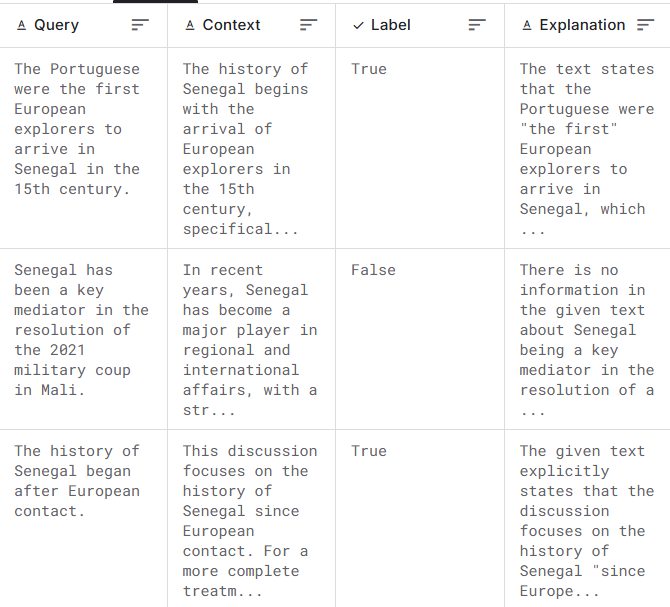


Figure 7. Improvised Dataset

Input feed to fine-tuned llama model:

Query1 :

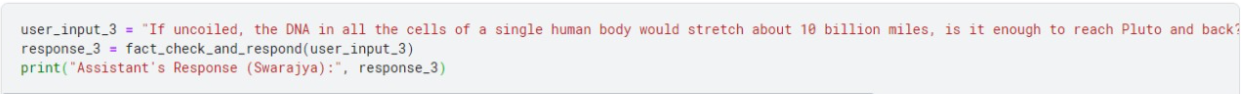


Figure 8. Input Query 2

Output:

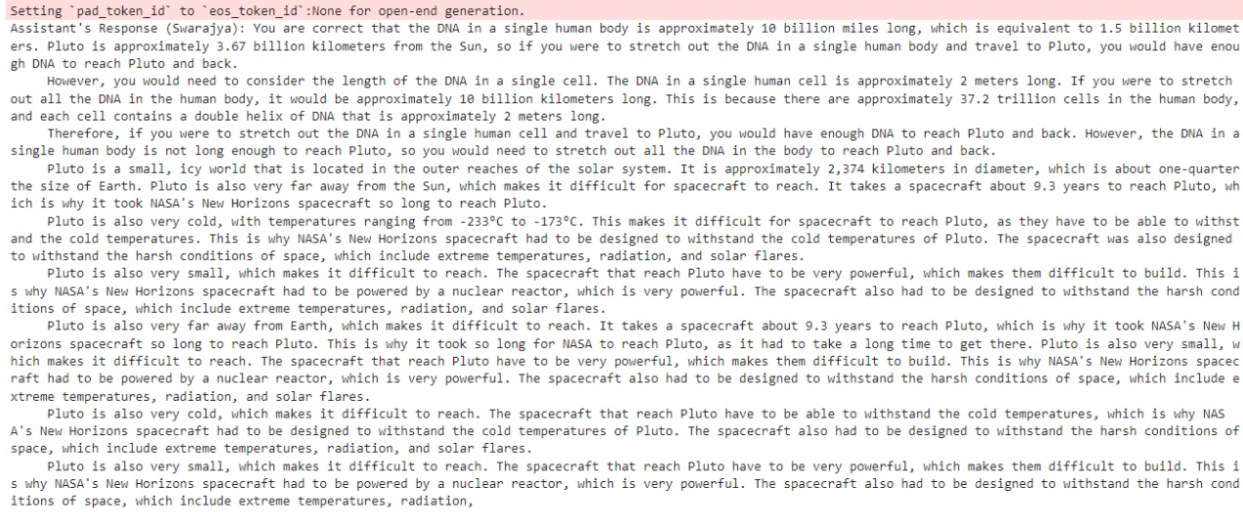


Figure 9. Output for input query - 2

Query2 :

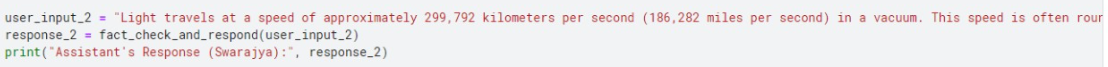


Figure 10. Input Query 3

Output:

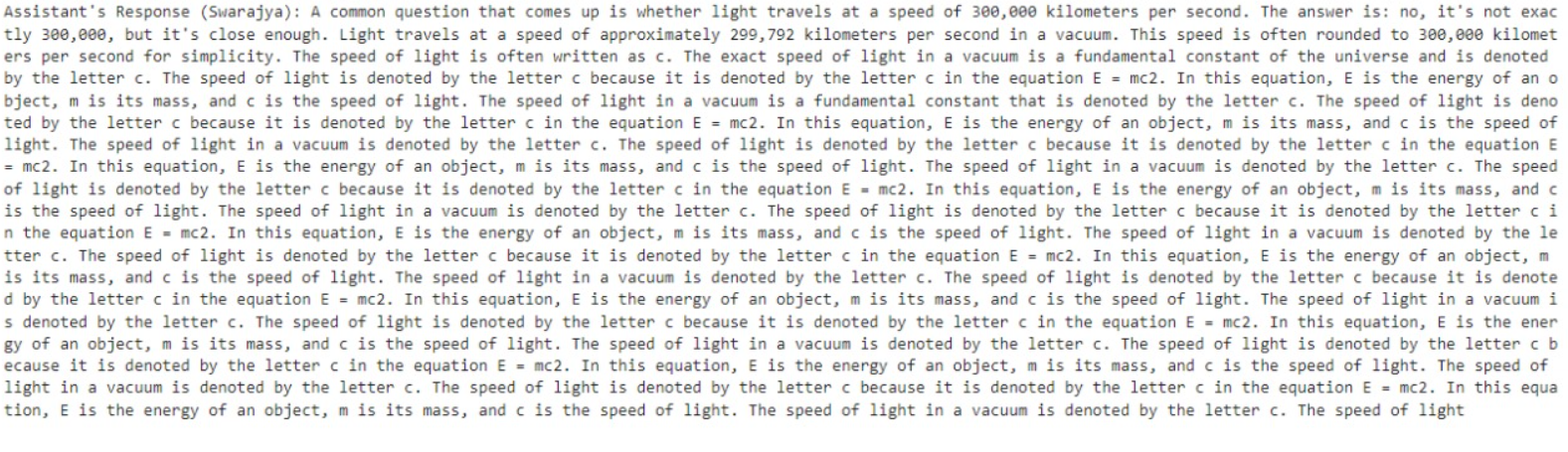
****

Figure 11. Output for input query- 3

**3.7 Result Analysis and Discussion**

The process so far aimed to improve the accuracy of a fact-checking model by fine-tuning the Llama language model. Initially, a semi-structured dataset was created as shown in Fig(5) and (6), as Llama models are typically trained on datasets of similar format. This semi-structured data consisted of text collected from various sources through web scraping and was used to train the model. However, the results from this initial model were unsatisfactory. The output for given queries was often vague and sometimes incorrect, failing to accurately verify the claims.

To improve performance, we reformatted the dataset into a question-answer format, which is more suitable for fine-tuning language models. This approach provided better results(as shown in Fig (4) and (5)) of structured reasoning and fact verification, providing the model with clear input-output pairs to learn from. The newly fine-tuned model demonstrated a better understanding of the queries, producing relevant and factually accurate responses, unlike the previous model which gave only redundant answers irrespective of the questions asked. This confirmed the importance of structured data in improving the learning capabilities of LLMs.

**3.8 Conclusion and Future work.**

Looking ahead, several strategies will be employed to further improve the accuracy and real-time relevance of the model. One key enhancement is the implementation of Retrieval-Augmented Generation (RAG), which will integrate dynamic data retrieval into the system. This involves developing a process to automatically fetch and clean data from RSS feeds, convert it into a structured format, and update the knowledge graph or vector database used by the model. Such a system will enable the model to stay continuously updated with new and relevant information, thereby improving the fact-checking capabilities.

The RAG-based approach will also involve continuous model updates, where the fine-tuned Llama model will regularly incorporate the latest information. This real-time updating will help the model keep pace with rapidly evolving data, addressing limitations observed in existing systems where outdated information can compromise accuracy. Overall, the future work aims to build a self-sustaining fact-checking system that integrates LLMs, RAG, and knowledge graphs to provide an automated, accurate, and continuously evolving solution.

**Reference**

1. Y. Barve, J. R. Saini, “A Novel Text Resemblance Index Method for Reference-based Fact-checking”, International Conference on Computing, Communication, and Intelligent Systems, 2022, pp- 829-836
2. M. L. Bangerter, G. Fenza, D. Furno, M. Gallo "A Hybrid Framework Integrating LLM and ANFIS for Explainable Fact-Checking," IEEE Transactions on Fuzzy Systems, Aug. 2024, doi: 10.1109/TFUZZ.2024.3431710.
3. J. Shin, S. Park, H. Song, "Token-Level Fact Correction in Abstractive Summarization," ‘IEEE Access’, vol. 10, pp. 1234-1245, 2022, doi: 10.1109/ACCESS.2022.3233854.
4. A. Matheven, B. V. D. Kumar, "Fake News Detection Using Deep Learning and Natural Language Processing," in Proc. IEEE Int. Symp. on Computational Intelligence (ISCMI), 2022, pp. 123-130, doi: 10.1109/ISCMI56532.2022.10068440.
5. P. Atanasova, J. G. Simonsen, C. Lioma, I. Augenstein, "Generating Fact Checking Explanations," arXiv, 2020, arXiv:2004.05773v1.
6. T. Ahuja, T. Borkar, "Comparative Study of Supervised Learning Algorithms for Fake News Classification," in Proc. IEEE Int. Conf. on Electronics and Information Engineering (ICOEI), 2022, pp. 45-53, doi: 10.1109/ICOEI53556.2022.9777118.
7. B. Yao, V. Tech, A. Shah, L. Sun, J. Cho, L. Huang, “End-to-End Multimodal Fact-Checking and Explanation Generation: A Challenging Dataset and Models”, *SIGIR ‘23*, July 23–27, pp- 2733-2743