HYBRID EVIDENCE – GROUNDED AI FOR INDIAN NEWS VERIFICATION

PROJECT REPORT

Submitted by

DEEPAK KUMAR V

in partial fulfilment for the award of the degree

of

MASTER OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

SRI VENKATESWARA COLLEGE OF ENGINEERING AND TECHNOLOGY

THIRUVALLUR-631203

ANNA UNIVERSITY: CHENNAI 600025

MARCH 2025

ANNA UNIVERSITY: CHENNAI 600 025

BONAFIDE CERTIFICATE

Certified that this project report **"HYBRID EVIDENCE – GROUNDED AI FOR INDIAN NEWS VERIFICATION"** is the bonafide work of **"DEEPAK KUMAR V"** who carried out the project work under my supervision.

SIGNATURE

Mr. Dr. B. Gnayanasundara Jayaraja M.E.Ph.D,MBA
HEAD OF THE DEPARTMENT,
ASSITANT PROFESSOR,
SVCET, THIRUPACHUR,
THIRUVALLUR 631203.

SIGNATURE

Ms. M. K. Anbarasi M.E
SUPREVISOR,
ASSITANT PROFESSOR,
SVCET, THIRUPACHUR,
THIRUVALLUR 631203.

Submitted for the project work and viva-voce examination held on ……

INTERNAL EXAMINER

EXTERNAL EXAMINER

ACKNOWLEDGEMENT

We are personally indebted to a number of people who gave us their useful insights to aid in our overall progress for this project. A complete acknowledgement would therefore be encyclopedic. First of all, we would like to give our deepest gratitude to our parents for permitting us to take up this course.

We extend our sincere gratitude to our respected **Chairman Dr. S.K. PURUSHOTHAMAN, Ph.D., Sri Venkateswara College of Engineering and Technology**, for providing facilities in the college premises for carrying out this project work.

We record our sincere thanks **Dr. B. Gnanasundara Jayaraja, M.E.,Ph.D, MBA, Principal of Sri Venkateswara College of Engineering and Technology**, for his encouragement to do this project.

We express our sincere thanks to **Mr. Dr. B. Gnanasundara Jayaraja, M.E.,Ph.D, MBA, Head of the Department, Computer Science and Engineering**, for his immense support to do this project.

We express our thanks to our **Internal Guide Ms. M.K. Anbarasi, M.E.,** Assistant Professor, Computer Science and Engineering Department, for her valuable guidance to complete this project successfully.

We express our sincere thanks to our entire department faculty members for their guidance and support to do this project.

ABSTRACT

In today's digital ecosystem, fake news spreads faster than factual reporting. This project proposes a **Hybrid Evidence-Grounded AI System** that verifies the authenticity of Indian news by combining **AI reasoning** and **real-time evidence gathering**.

The system employs the **Google Gemini API** for deep textual analysis and cross-checks each claim with articles obtained from **trusted Indian news portals** such as *Times of India*, *NDTV*, and *The Hindu* via **NewsAPI**, **GDELT**, and **Google News RSS**.

It supports **multiple Indian languages**, translating regional inputs into English for consistent evaluation.

Results are expressed through **confidence scores** and **factual explanations**. By integrating natural language understanding, evidence retrieval, and data storage, this model provides a scalable foundation to combat misinformation across India's multilingual digital landscape.

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

1.1 BACKGROUND

In today's digital era, information spreads across the internet within seconds through news portals, blogs, and social media. While this has democratized access to knowledge, it has also given rise to a critical problem — **the widespread dissemination of fake news**. Fake news can distort public opinion, damage reputations, create panic, and even influence elections. Detecting misinformation in real time is essential to ensure that people receive accurate and verified information.

Conventional methods of fact-checking are manual and time-consuming. They rely on human experts who verify facts using multiple sources. This approach cannot cope with the **sheer volume and velocity** of online information in today's interconnected society. Hence, there is a strong need for **automated fake news detection systems** that can analyze information quickly and reliably.

1.2 MOTIVATION

The motivation behind this project stems from the growing influence of **misinformation in the Indian media ecosystem**. India is a multilingual country with diverse news sources operating in several regional languages. Fake news in vernacular languages spreads faster due to limited verification mechanisms.

Existing AI systems for news verification are often trained only in English and do not consider India-specific media sources.

To address these issues, this project proposes a **Hybrid Evidence-Grounded AI System** that verifies news claims using both **AI-based reasoning** and **live evidence** from **trusted Indian news portals**. This ensures factual accuracy, linguistic inclusivity, and real-time validation.

1.3 PROBLEM DEFINITON

The primary problem tackled by this project is the **automatic verification of Indian news articles and claims** by combining artificial intelligence and evidence retrieval.

Current fake news detection systems mainly focus on static text classification, which fails to validate information against real-world data.

The challenge is to design a system that:

- Understands and analyzes **multilingual text** (English and Indian languages).
- Fetches **live supporting or contradicting evidence** from verified news APIs.
- Produces a **confidence-based, explainable verdict**.

By bridging the gap between AI understanding and factual cross-verification, this project builds a **hybrid, evidence-grounded framework** for trustworthy news validation.

1.4 OBJECTIVES

The main objectives of the system are:

1. To design a **hybrid model** that integrates AI reasoning and live data evidence.
2. To utilize **Google Gemini API** for advanced natural language understanding.
3. To retrieve real-time information from **trusted Indian news sources** using APIs like NewsAPI, Google RSS, and GDELT.
4. To enable **multilingual support**, translating regional language input into English.
5. To provide **confidence scores and factual explanations** for every verification result.
6. To create a **user-friendly interface** using Streamlit for easy access.
7. To store all verification data securely for historical tracking and auditing.

1.5 SCOPE OF THE PROJECT

The system is designed as a **Phase 1 prototype** focusing on textual news verification and multilingual analysis. It currently supports:

- Text or URL input for news claims.
- AI reasoning via Gemini API.
- Evidence collection through multiple live APIs.
- Display of confidence-based verification reports.

In the future, the system can be extended to include:

- Image verification using CNN-based models.
- Advanced regional language handling.
- Real-time misinformation alerts through social media integration.

1.6 SIGNIFICANCE OF THE PROJECT

The **Hybrid Evidence-Grounded AI for Indian News Verification** project holds substantial social and technological importance. It contributes to:

- Combating misinformation and promoting truth in digital communication.
- Empowering users to verify news independently.
- Enhancing public trust in media through transparency and explainability.
- Supporting government and fact-checking organizations with automated verification tools.

1.7 CHAPTER SUMMARY

This chapter discussed the growing issue of fake news, the motivation behind building an AI-based verification system, and the project's objectives and scope.

The next chapter presents a **Literature Review**, which analyzes existing systems, related research papers, and technologies relevant to fake news detection.

CHAPTER 2
LITERATURE REVIEW

2.1 OVERVIEW OF EXISTING WORK

The problem of fake news detection has been actively researched over the past decade, particularly after the explosion of social media platforms such as Facebook, Twitter (now X), and WhatsApp.

Researchers have applied a variety of **machine learning**, **deep learning**, and **natural language processing (NLP)** techniques to automatically detect misinformation.

Early studies relied on **linguistic features** such as word frequency, sentiment polarity, and stylistic cues.

Machine learning models like **Naïve Bayes**, **Support Vector Machines (SVM)**, and **Random Forests** were trained on labeled datasets to classify text as “real” or “fake.”

While these models showed moderate accuracy, they lacked the ability to understand deeper context and often failed when encountering cleverly disguised misinformation.

In recent years, the focus has shifted toward **neural network–based models** and **transformer architectures**, such as **BERT**, **RoBERTa**, and **XLNet**, which better capture semantic relationships in text.

These models analyze not just word-level features but also the contextual meaning behind statements. However, they are limited by training data — models trained on English datasets rarely generalize to multilingual or region-specific fake news.

India, with its diverse linguistic landscape, presents an even greater challenge. Most fake news detection systems focus on English or Hindi and do not effectively handle content from regional sources such as Tamil, Telugu, Malayalam, or Kannada.

Moreover, these systems rely heavily on **static datasets** and **offline training**, which means they cannot validate information in real time.

Hence, there is a need for an adaptive system that not only analyzes text intelligently but also **cross-verifies the claim with live news sources** — ensuring factual consistency and credibility.

2.2 COMPARATIVE STUDY OF TECHNIQUES

Method	Approach	Advantages	Limitations
Rule-Based Models	Handcrafted features and keyword-based matching.	Simple and interpretable.	Poor scalability and low accuracy.
Traditional ML Models (SVM, RF, NB)	Statistical classification based on text features.	Works on small datasets.	Ineffective with contextual
Deep Learning Models (CNN, LSTM)	Learns hierarchical patterns in news text.	Better accuracy with large data.	Requires extensive training data and GPUs.
Transformer-Based Models (BERT, RoBERTa)	Context-aware text embeddings for fake news detection.	High contextual accuracy.	Poor multilingual coverage and slow inference.
Hybrid AI Systems (Proposed)	Combines AI reasoning (Gemini API) with live evidence.	Real-time, explainable, multilingual, and adaptive.	Dependent on external API availability.

The **Hybrid Evidence-Grounded AI** system builds upon the strengths of prior models by integrating **text understanding, real-time evidence retrieval**, and **multilingual translation** in a single framework.

2.3 RESEARCH GAP IDENTIFIED

Through the review of existing studies and systems, several critical gaps have been identified:

1. **Lack of Real-Time Verification:**

   Most systems rely on static datasets rather than live APIs, leading to outdated or incomplete analyses.

2. **Absence of Evidence Correlation:**

   Current systems classify fake news purely based on textual patterns

3. **Low Explainability:**

   AI models like deep neural networks act as black boxes, providing little insight into *why* a decision was made.

4. **Scalability and Speed Issues:**

   Traditional models require retraining for new data and cannot scale efficiently to handle large input streams.

The proposed system directly addresses these challenges by introducing a **hybrid model** that merges **AI-based reasoning** (for contextual understanding) with **evidence-based verification** (for factual accuracy).

2.4 SUMMARY OF RELATED WORKS

- **Vidhya W. S. & Mrs. Kanchana (2023):** Proposed an *AI-Powered Multi-Model Fake News Detection System* combining Gemini API with CNN for visual content detection. Their system demonstrated high accuracy but lacked regional language support and live verification.
- **Zhou et al. (2020):** Explored *Fake News Detection via NLP and Deep Learning*, using linguistic and sentiment features. The model achieved good precision but required large labeled datasets.
- **Sharma et al. (2021):** Implemented a *BERT-based Fake News Classifier* for English news articles. Performance dropped sharply when tested on Indian regional content.
- **Kumar & Singh (2022):** Presented a *Fact-Checking System using Knowledge Graphs* to verify online claims. The method was accurate but computationally expensive.

These studies inspired the development of the proposed system, which merges **Gemini's advanced reasoning capabilities** with **real-time evidence retrieval** for Indian news.

2.5 CHAPTER SUMMARY

This chapter presented a review of existing fake news detection techniques and identified their key strengths and limitations.

It also highlighted the unique challenges faced in multilingual environments such as India.

The findings from this review emphasize the need for a **hybrid, evidence-based AI framework** that can deliver real-time, explainable, and accurate results.

The next chapter discusses the **System Analysis**, focusing on the limitations of the existing system and the feasibility of the proposed approach.

CHAPTER 3
SYSTEM ANALYSIS

4.1 EXISTING SYSTEM

The existing systems for fake news detection primarily rely on **text classification models** that analyze linguistic or statistical features within articles. These systems are trained using supervised learning algorithms such as Naive Bayes, Logistic Regression, or SVM.

However, these methods share several shortcomings:

- They depend heavily on **pre-trained datasets** that may not represent current or region-specific news.
- They are typically limited to **English-language content**.
- Most do not provide **explainable results** — users receive a binary output ("fake" or "real") without understanding the reasoning.

Such limitations reduce the trustworthiness and adaptability of existing fake news detectors, especially in multilingual and rapidly evolving environments like India.

3.2 LIMITATIONS OF EXISTING SYSTEMS

S. No.	Limitation	Description
1	Language Restriction	Most systems are English-only and cannot process regional Indian languages.
2	Static Dataset	Training data is outdated and not linked to real-time news streams.
3	Lack of Evidence	Systems do not cross-check claims with live or credible sources.
4	Poor Explainability	Users cannot see *why* a statement is labeled as fake.
5	Scalability Issues	Re-training is required for new or evolving datasets.
6	No Image/Multimodal Input	Text-only systems cannot verify multimedia misinformation.

The limitations above highlight the need for a **hybrid approach** that merges AI-based text reasoning with evidence-grounded validation to enhance accuracy and trust.

3.3 PROPOSED SYSTEM

The **proposed Hybrid Evidence-Grounded AI system** overcomes the above limitations by integrating two complementary verification mechanisms:

1. **AI Reasoning Layer (Gemini API):**

   Uses advanced natural language understanding to interpret the semantic meaning of a news claim.

2. **Evidence Verification Layer:**

   Searches live Indian news sources using APIs such as **NewsAPI**, **Google News RSS**, and **GDELT**, comparing retrieved results with the claim.

The integration of these two components results in a **hybrid model** that produces:

- Real-time, evidence-backed verification
- Explainable outputs with confidence scores
- Automated report generation (PDF)

This approach ensures **transparency, adaptability, and inclusivity**, making it suitable for diverse Indian users.

3.4 Feasibility Study

A **feasibility study** assesses the practicality and viability of implementing the proposed system. It evaluates technical, operational, and economic aspects.

Type of Feasibility	Analysis Description
Technical Feasibility	he system uses well-established technologies like Python, Streamlit, and Gemini API. APIs such as NewsAPI and GDELT ensure live data retrieval.
Operational Feasibility	The project workflow is straightforward. Users input a claim or URL, and the system verifies it automatically. It provides explainable results, making it practical for both common users and fact-checking professionals.
Economic Feasibility	The project relies mainly on free or academic-tier APIs. Hardware requirements are minimal — a standard computer with internet connectivity is sufficient. Thus, implementation cost is low.
Schedule Feasibility	Phase I focuses on text verification and AI integration, which can be achieved within one semester.

The feasibility study confirms that the proposed hybrid AI system is **technically sound, operationally practical, and economically viable**.

3.5 ADVANTAGES

- **Real-Time Verification:** Retrieves current evidence rather than relying on outdated data.
- **Multilingual Input:** Detects and translates Indian languages automatically.
- **Transparency:** Provides reasoning behind every classification.
- **User-Friendly Interface:** Simple Streamlit-based web interface.
- **Scalability:** Modular architecture allows easy expansion.
- **Integration Ready:** APIs and AI modules can be updated independently.

3.6 SUMMARY

This chapter presented a comprehensive analysis of the existing fake news detection systems and their shortcomings.

The feasibility study confirmed that the proposed hybrid approach is practical and cost-effective.

The **Hybrid Evidence-Grounded AI** system offers superior scalability, multilingual adaptability, and transparency compared to traditional models.

The next chapter, **System Design**, will describe the overall architecture, methodology, and data flow diagrams illustrating how the system operates internally.

CHAPTER 4

SYSTEM DESIGN

4.1 SYSTEM ARCHITECTURE

The system architecture represents the overall structure of the proposed system, showing how data flows between the modules.

Architecture Components:

- **User Interface (UI):**

  Built using Streamlit, it allows users to enter text or URLs and view the verification results interactively.

- **Language Processor:**

  Detects the input language and translates it into English using translation APIs for standardized analysis.

- **AI Verifier (Gemini API):**

  Analyzes the meaning and factual consistency of the input claim. It provides an interpretive result like *True*, *False*, *Misleading*, or *Unverified* with confidence percentage.

- **Evidence Fetcher:**

  Connects to **NewsAPI**, **Google RSS**, and **GDELT** to fetch the latest news articles related to the input claim.

- **Fusion Engine:**

  Merges AI reasoning and evidence consistency to finalize the verdict.

- **Database Layer:**

  Stores claims, verification results, evidence links, and timestamps for record-keeping.

- **Report Generator:**

  Uses ReportLab to create downloadable PDF reports summarizing the verification process.

4.2 METHODOOGY

The flowchart represents the operational logic of the system from input to output.

Flowchart Steps:

1. Start
2. User enters text or news URL
3. Language Detection
4. Translation (if non-English)
5. AI Verification (Gemini API)
6. Live Evidence Retrieval (APIs)
7. Fusion and Confidence Scoring
8. Display Result & Generate Report
9. Stop

4.3 USE CASE DIAGRAM

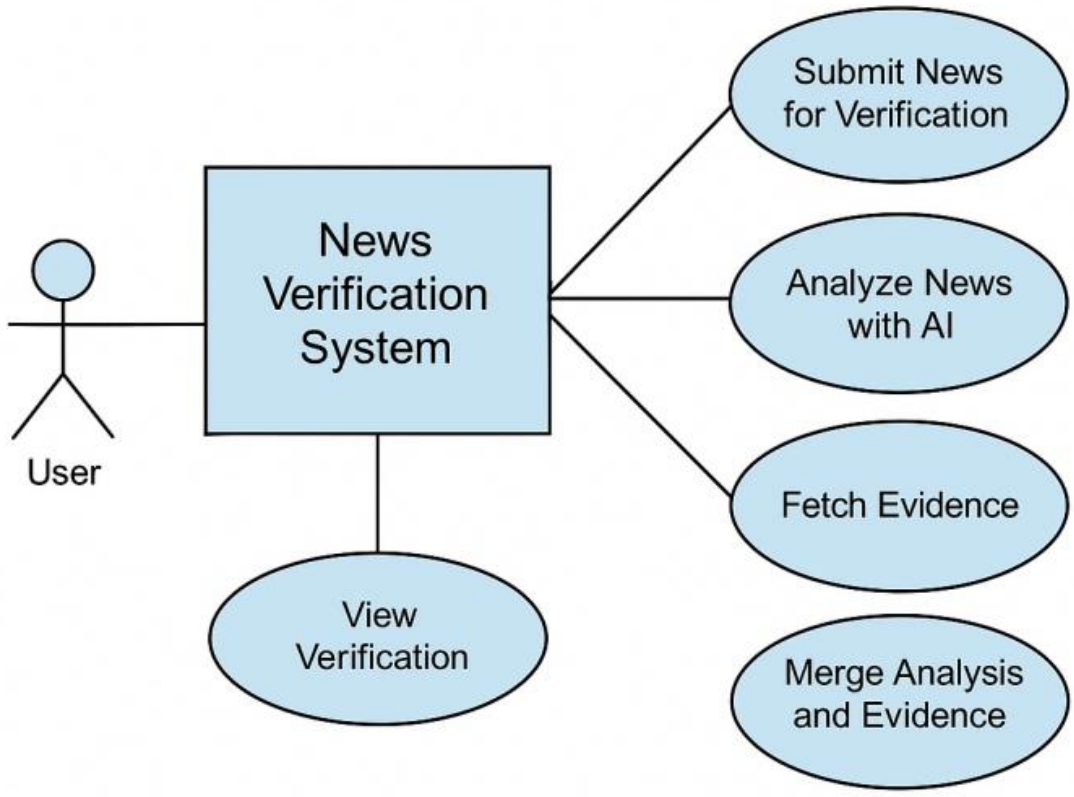
The **Use Case Diagram** describes how external users interact with the system.

Actors:

- **User** – enters the news claim and views results.
- **System** – performs verification and produces reports.

Use Cases:

1. Submit news input
2. Translate and preprocess
3. Verify using Gemini AI
4. Retrieve live evidence
5. Generate confidence score
6. Display result
7. Generate and download report



The diagram shows how a user interacts with the system's main functions.

It ensures that user goals and system operations are clearly defined and traceable to functional requirements.

4.4 ENTITY RELATIONSHIP

The **ER Diagram** models how the system's database stores and links data entities.

Entities:

- **User_Input** (Input_ID, Claim_Text, Language, Input_Date)
- **AI_Result** (Result_ID, Input_ID, Status, Confidence_Score)
- **Evidence_Data** (Evidence_ID, Result_ID, Source_Name, Article_Link)
- **Report** (Report_ID, Result_ID, File_Path, Timestamp)

Relationships:

- One User_Input → One AI_Result
- One AI_Result → Many Evidence_Data
- One AI_Result → One Report

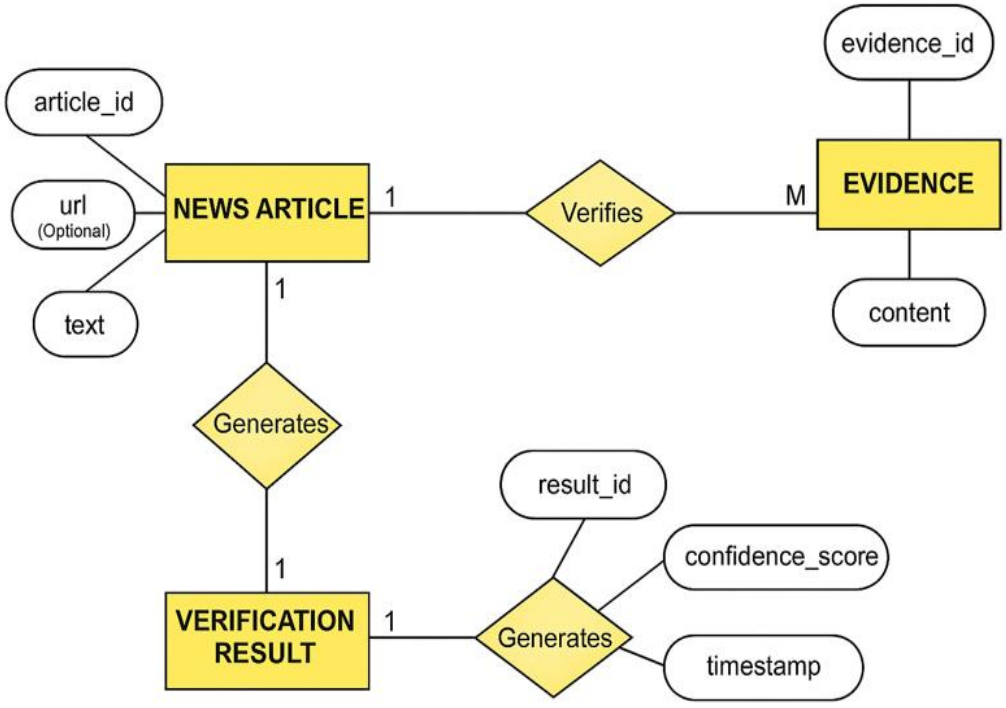


Figure 4.3: Entity Relationship Diagram (ERD)

This diagram ensures efficient data organization, easy traceability, and future scalability for adding new verification types such as image or video analysis.

4.5 SUMMARY

This chapter detailed the architectural design and internal workflow of the proposed hybrid AI system.

It presented all major design diagrams — from high-level architecture to data-level modeling — and explained their roles in system operation.

The design ensures that the system is **modular, scalable, and adaptable**, meeting the technical requirements identified in earlier chapters.

The next chapter, **System Implementation (Phase I)**, describes how the designed modules were developed and integrated to create a working prototype.

CHAPTER 5

SYSTEM IMPLEMENTATION (PHASE I)

5.1 TECHNOLOGIES USED

The following technologies, APIs, and frameworks were used in

implementing Phase I of the system:

Technology	Purpose / Description
Python 3.10+	Core programming language used for logic, API handling, and NLP integration.
Streamlit	Front-end framework used for developing an interactive, web-based UI.
Google Gemini API	Provides advanced text understanding and reasoning to evaluate news claims.
NewsAPI	Fetches recent articles related to the given claim from verified Indian sources.
Google News RSS	Fetches recent articles related to the given claim from verified Indian sources.
GDELT API	Adds global news coverage and multilingual support for contextual evidence.
BeautifulSoup (bs4)	Used for web scraping and extracting text content from URLs when the user inputs a link.

Software Requirements

- Operating System: Windows 10 / Ubuntu 22.04
- IDE: Visual Studio Code / PyCharm
- Libraries: requests, streamlit, google.generativeai, pandas, altair, reportlab, bs4, dotenv
- Browser: Any modern browser supporting Streamlit UI

Hardware Requirements

- Processor: Intel i5 or equivalent
- RAM: Minimum 8 GB
- Internet: Required for API connections

5.2 SYSTEM FLOWWORK

The **workflow** of the proposed system follows a hybrid, evidence-grounded approach combining AI reasoning and real-time data collection.

Workflow Steps

1. **Input Stage:**

   User enters news text or a URL through the Streamlit interface.

2. **Preprocessing & Translation:**

   If the input is in a regional Indian language, it is automatically detected and translated into English using language APIs.

3. **AI Reasoning (Gemini API):**

   The translated or raw text is sent to the Gemini API, which performs context-aware analysis to predict if the claim is *True*, *False*, *Misleading*, or *Unverified*.

   It also generates a **confidence score** based on internal reasoning.

4. **Evidence Retrieval:**

   Simultaneously, the system queries **NewsAPI**, **Google RSS**, and **GDELT** for related news articles.

Retrieved articles are filtered by source credibility and relevance to the claim.

5. **Fusion Module:**

   Both outputs — the AI verdict and live evidence — are merged.

   The system checks how many sources support or contradict the claim and adjusts the confidence score accordingly.

6. **Result Presentation:**

   The final output is displayed in the UI as:

   - Verification Status (True / False / Misleading / Unverified)
   - Confidence Score (%)
   - Number of Supporting Articles
   - AI Explanation and Evidence Summary

7. **Report Generation:**

   A **PDF report** is created summarizing the verification result, including AI analysis, evidence URLs, and timestamps.

5.3 MODULE DESCRIPTION

1. Input Module

- **Purpose:** To receive news claims or article URLs from the user.
- **Function:** Accepts input via a Streamlit text box or URL field.
- **Key Feature:** URL extraction using BeautifulSoup and language detection for regional content.

2. Language Processor Module

- **Purpose:** To ensure linguistic compatibility for analysis.
- **Function:** Automatically detects Indian languages and translates them to English using language translation APIs.
- **Importance:** Enables verification of multilingual content, which is essential for Indian digital media.

3. AI Verifier (Gemini API Module)

- **Purpose:** Performs semantic analysis and predicts truthfulness.
- **Process:**
    - Sends the claim to Gemini API.

- Receives analysis text, confidence score, and status.
- Categorizes results into one of five classes: *True*, *False*, *Partially True*, *Misleading*, *Unverified*.

- **Output Example:**
- VERIFICATION_STATUS: MISLEADING
- CONFIDENCE_SCORE: 72%
- **Advantage:** Context-aware reasoning, unlike static classifiers.

4. Evidence Fetcher Module

- **Purpose:** To collect supporting or contradictory evidence from real-time sources.
- **Process:**
  - Queries multiple APIs: NewsAPI, Google RSS, and GDELT.
  - Extracts article headlines, publication dates, and links.
  - Filters duplicates and low-quality results.
- **Output:** A list of 10–15 related articles used for cross-verification.

5. Fusion and Scoring Module

- **Purpose:** To combine AI reasoning with evidence for a hybrid decision.
- **Process:**
  - Compares AI predictions with evidence overlap.
  - Tags each article as *supportive*, *contradictory*, or *irrelevant*.
  - Computes weighted confidence score and final verdict.
- **Result Example:**
  - 6 supportive, 2 contradictory, 2 irrelevant → Final Verdict: *True (Confidence 84%)*.

6. Database Module

- **Purpose:** To store all verification history and metadata.
- **Schema Overview:**
  - User_Input table stores claim and timestamp.
  - AI_Result table stores prediction, reasoning, and confidence.
  - Evidence_Data table stores fetched news URLs and sources.
  - Report table stores report file path and creation date.
- **Technology:** MySQL with Python connector.

7. Report Generation Module

- **Purpose:** To produce professional verification reports for download.
- **Library Used:** ReportLab
- **Features:**
    - AI and evidence summary
    - Verification confidence
    - Timestamp and data source list
- **Output:** PDF with title, date, and structured sections.

8. User Interface (Streamlit App)

- **Purpose:** Provides an interactive environment for user interaction.
- **Features:**
    - Text and URL input fields
    - Progress spinners during verification
    - Real-time metrics (confidence, evidence count, etc.)
    - Expandable AI analysis window
    - Download report button
- **Outcome:** Intuitive, minimal, and modern interface for users.

CHAPTER 6

CONCLUSION AND FUTURE WORKS

6.1 CONCULSION

The proposed project, **Hybrid Evidence-Grounded AI for Indian News Verification**, aims to address one of the most pressing challenges of the digital era — the rapid spread of misinformation and fake news.
Traditional fact-checking mechanisms are slow, manual, and often limited to a single language or static dataset.
This system introduces a **hybrid AI-powered approach**, combining the **reasoning strength of Gemini API** with **real-time evidence gathering** from trusted Indian and global news APIs.

The **Phase I implementation** successfully demonstrates how an automated system can analyze a news claim, retrieve factual context, and deliver explainable verification outcomes to users.
Unlike traditional models that classify text based only on linguistic features, this hybrid system correlates AI predictions with **live, verifiable data sources** to ensure factual accuracy and transparency.

Key Achievements of Phase I:

1. **AI Reasoning Integration:**

   Implemented **Gemini API** to analyze claims and generate detailed reasoning with confidence scores.

2. **Live Evidence Retrieval:**

   Integrated **NewsAPI**, **Google RSS**, and **GDELT** for collecting recent, India-focused news data.

3. **Hybrid Fusion Model:**

   Combined AI results with live data correlation for stronger factual assessment.

4. **User-Friendly Interface:**

   Designed an intuitive web interface with real-time metrics, expandable analysis, and PDF report download features.

The outcomes of this phase validate the project's feasibility and demonstrate that **AI-assisted, evidence-grounded verification** is an effective tool for combating digital misinformation in India.

6.2 LIMITATIONS

Despite its success, the current system has a few constraints:

1. **Dependency on APIs:**

   Real-time verification relies on the availability and performance of third-party APIs like NewsAPI and GDELT.

2. **Language Translation Accuracy:**

   While translation enables multilingual processing, slight inaccuracies in meaning can influence results.

3. **Lack of Image Verification:**

   The current phase handles only text-based content, limiting its scope against multimedia misinformation.

4. **Limited Offline Support:**

   The system requires internet connectivity for real-time evidence collection.

These limitations will serve as the foundation for enhancements in the next development phase.

6.3 FUTURE ENHANCEMENT

The next phase of the project (Phase II) will focus on **broadening the system's capability** and **deepening its accuracy**.

Proposed Enhancements:

1. **Integration of Image Verification:**
    - Implement a **Convolutional Neural Network (CNN)**-based image verification module.
    - Detect fake or manipulated images using pixel-based analysis and pretrained visual models.
2. **Expanded Multilingual Support:**
    - Extend translation coverage to more regional Indian languages such as Kannada, Marathi, Bengali, and Malayalam.
    - Use **context-aware language models** to preserve meaning during translation.
3. **Enhanced Evidence Scoring:**
    - Develop a **credibility scoring algorithm** for ranking sources based on their historical accuracy.

6.4 EXPECTED IMPACT

Once fully developed, the **Hybrid Evidence-Grounded AI system** can be deployed as:

- A **public fact-checking platform** for users and journalists.
- An **AI-powered API service** for media organizations to integrate verification features.
- A **support tool** for educational and governmental institutions to combat misinformation.

This project, by combining **AI reasoning**, **multilingual processing**, and **live factual validation**, can significantly enhance media reliability and public trust in the digital ecosystem.