PROJECT 11

INTELLIGENT NEWS CREDIBILITY ANALYSIS & AGENTIC MISINFORMATION MONITORING

From Predictive NLP to Autonomous Fact-Checking

PROJECT OVERVIEW

This project involves the design and implementation of an **AI-driven content analytics system** that evaluates the credibility of news articles and evolves into an agentic AI misinformation monitoring assistant.

In **Milestone 1**, the system applies classical machine learning and NLP techniques to analysis news text, classify articles based on credibility signals, and detect potential misinformation patterns.

In **Milestone 2**, the same system is extended into an agent-based AI application that autonomously reasons about content credibility, retrieves fact-checking sources, and generates structured assessments.

The project emphasizes real-world media integrity challenges, progressive development from predictive NLP to agentic AI workflows, and public deployment.

CONSTRAINTS & REQUIREMENTS

TEAM SIZE	API BUDGET
3–4 Students	**Free Tier Only**
FRAMEWORK	**HOSTING**
LangGraph (Recommended)	**Mandatory**

APPROVED TECHNOLOGY STACK

LLMS (MILESTONE 2)

- Open-source models
- Free-tier APIs

AGENT FRAMEWORK

- LangGraph (Recommended)
- Chroma / FAISS (RAG)

ML & NLP (MILESTONE 1)

- TF-IDF / Embeddings
- Logistic Regression / Trees
- Scikit-Learn (Pipelines)

UI FRAMEWORK

- Streamlit (Recommended)
- Gradio

HOSTING PLATFORM (MANDATORY)

- Hugging Face Spaces
- Streamlit Community Cloud
- Render (Free Tier)

WARNING: Localhost-only demonstrations will not be accepted.

MILESTONE 1: ML-BASED NEWS CREDIBILITY CLASSIFICATION

MID-SEM SUBMISSION

Objective: Design and implement a machine learning-based news credibility classification system that evaluates articles using textual features. Focus on classical NLP and supervised learning *without LLMs*.

Functional Requirements:

- Accept news article text or URLs as input.
- Perform text preprocessing and feature extraction.
- Classify articles based on credibility risk.
- Display credibility scores via user interface.

TECHNICAL REQUIREMENTS (NLP/ML)

- **Preprocessing:** Tokenization, Vectorization.
- **Features:** TF-IDF, Word Embeddings.
- **Models:** Logistic Regression, Decision Trees.
- **Evaluation:** Precision, Recall, F1 Score.

INPUTS & OUTPUTS

- **Input:** Article text / URL.

- **Output:** Credibility Score (High/Low).
- **Metrics:** Pattern detection summary.

MID-SEM DELIVERABLES

- Problem understanding & Media Use-case.
- Input–output specification.
- System architecture diagram.
- Working local application with UI.
- Model performance evaluation report.

MILESTONE 2: AGENTIC AI MISINFORMATION MONITORING

END-SEM SUBMISSION

Objective: Extend the credibility system into an agentic AI misinformation monitoring assistant that autonomously reasons about content, retrieves fact-checking sources, and generates structured reports.

Functional Requirements:

- Accept article analysis and monitoring queries.
- Analyze credibility patterns and risk factors.
- Retrieve relevant fact-checking information.
- Generate structured credibility assessments.

TECHNICAL REQUIREMENTS (AGENTIC)

- **Framework:** LangGraph (Workflow & State).
- **RAG:** Retrieval of fact-checks (Chroma/FAISS).
- **State:** Explicit state management across steps.
- **Prompting:** Strategies to reduce hallucinations.

STRUCTURED OUTPUT

- **Summary:** Article overview & Risk Factors.
- **Analysis:** Cross-source verification results.
- **Verdict:** Credibility assessment (Confidence).
- **Disclaimer:** Ethical/Misinformation notices.

END-SEM DELIVERABLES

- Publicly deployed application.
- Agent workflow documentation.
- Structured credibility report.
- GitHub Repository & Complete Codebase.
- Demo Video (Functionality walkthrough).

Final Artifacts: Hosted Link, GitHub Repo, Demo Video.

EVALUATION CRITERIA

PHASE	WEIGHT	CRITERIA
Mid-Sem (Milestone 1)	25%	• Correct application of NLP & ML techniques • Quality of Preprocessing & Feature Engineering • Appropriateness of evaluation metrics • UI Usability & Code modularity
End-Sem (Milestone 2)	30%	• Quality of Agentic Reasoning & Fact-checking

PHASE	WEIGHT	CRITERIA
		• Correct RAG integration & State management • Utility of Credibility Insights • Responsible AI & Ethical Practices

---