MGM's College of Engineering, Nanded
Department of Computer Science & Engineering
"Real-Time News & Social Media Sentiment Analyzer + Summarizer"
Name of Student: 1.Ganesh Kalapad 2. Adnan Khan 3. Aniket Nagare 4. Suhas Dube
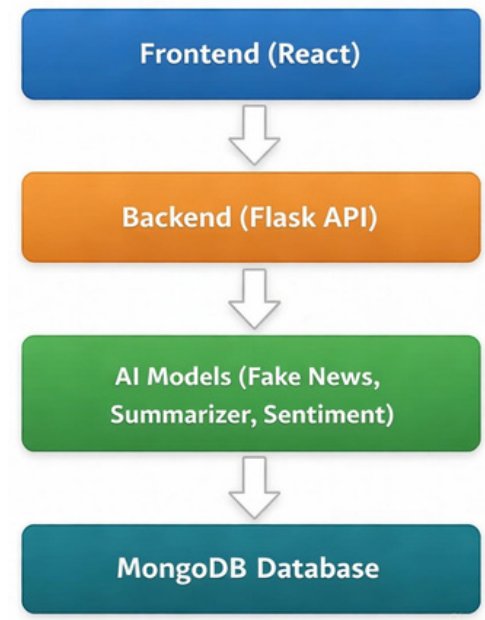Name of Guide: Dr. S. Y. Gaikwad
Academic Year 2025-26

Introduction:

This research proposes an intelligent real-time news analysis system to help users navigate the overwhelming flow of online information. By integrating news fetching, automatic summarization, sentiment analysis, and fake news detection, the system uses advanced NLP models like T5 and BART to condense long articles into concise summaries. Sentiment analysis with the VADER model categorizes content as positive, negative, or neutral, while a fake news module assesses article credibility. This system aims to improve the accuracy, trustworthiness, and accessibility of digital news, reducing cognitive overload and helping users make informed decisions quickly. Future improvements may include multilingual support and optimization for better performance.

System Architecture:

The system architecture of the real-time news sentiment analysis and summarization platform is designed in four layers: User Input, Processing, Analysis, and Output. The User Input Layer allows users to submit text, PDFs, images, or URLs. The Processing Layer extracts content from these inputs using PyPDF2 for PDFs, OCR with pytesseract for images, and web scraping with BeautifulSoup for online articles. The text is then preprocessed through tokenization, stop-word removal, and language detection. The Analysis Layer handles summarization with transformer models like BART and T5, sentiment analysis with VADER, and fake news detection using machine learning based on linguistic features.

Frontend (React) → Backend (Flask API) → AI Models (Fake News, Summarizer, Sentiment) → MongoDB Database

Methodology:

The methodology of the proposed Real-Time News & Social Media Sentiment Analyzer and Summarizer follows a modular and systematic approach to efficiently collect, process, analyze, and present large-scale textual data. The overall workflow is divided into sequential stages to ensure accuracy, scalability, and real-time performance. First, news articles and social media content are gathered through web scraping, followed by text preprocessing and feature extraction. Advanced NLP models then perform summarization, sentiment analysis, and fake news detection, providing users with concise, reliable, and emotionally insightful content.

Conclusion:

This research presents an integrated real-time news analysis system combining summarization, sentiment analysis, keyword extraction, and fake news detection. Leveraging models like BART, T5, and VADER, the system helps users quickly comprehend news content and assess emotional tone and reliability. It reduces information overload and supports informed decision-making. Future work will focus on multilingual support, social media integration, and enhancing fake news detection accuracy.

Maps to	Name of Students	Name of Guide
PO-1, PO-2, PO-5, PO-6, PO-10, PSO-1, PSO-2	Mr. Ganesh Kalapad Mr. Adnan Khan Mr. Aniket Nagare Mr. Suhas Dube	Dr. S. Y. Gaikwad