Executive Summary

NewsBot 2.0: AI-Driven News Intelligence System

Student Team: John, Dylan, Milagros, Ola
Course: ITAI 2373 – Final Project
Date: August 2025

Overview

NewsBot 2.0 is an AI-powered system designed to process and analyze news articles, delivering accurate classifications, sentiment evaluations, topic identifications, named entity extractions, summaries, and multilingual translations. The system is accessible through a conversational interface, enabling users to interact naturally and receive precise, relevant insights.

Purpose and Objectives

The main goal of NewsBot 2.0 is to demonstrate the integration of advanced natural language processing (NLP) techniques into a unified tool for news intelligence. By combining multiple modules—classification, sentiment analysis, named entity recognition, topic modeling, summarization, and translation—the system can handle diverse queries while maintaining high accuracy and interpretability.

Key Features

- **High-Accuracy Classification** – Naive Bayes classifier trained on BBC News Dataset achieving 96.64% test accuracy.
- **Sentiment Analysis** – Detects polarity and assigns sentiment labels (positive, negative, neutral).
- **Named Entity Recognition** – Extracts people, organizations, locations, dates, and monetary amounts.
- **Topic Modeling** – Identifies main topics using Latent Dirichlet Allocation (LDA) with keyword lists.
- **Summarization** – Generates both extractive and transformer-based summaries.
- **Multilingual Support** – Detects and translates non-English text into English.

- **Conversational Interface** – Allows natural language interaction with the system for real-time insights.

Implementation Highlights

The system was developed in Python 3.11 using Google Colab for testing and GitHub for version control. It leverages widely adopted libraries such as NLTK, spaCy, scikit-learn, transformers, and FastAPI. A modular architecture ensures that individual components can be improved or replaced without affecting the rest of the system.

Results and Impact

Testing demonstrated that the classifier achieved exceptional accuracy, sentiment analysis matched human expectations, and summaries retained the essential meaning of articles. The conversational interface allows non-technical users to interact with complex NLP pipelines easily, making the tool adaptable for journalists, analysts, and researchers.

Conclusion

NewsBot 2.0 proves that multiple NLP techniques can be integrated effectively into a cohesive, user-friendly news intelligence system. Its strong performance and modular design position it as a foundation for future work, such as real-time API integration, broader language support, and more granular classification categories.