

Technical Documentation – Final Project

Student Name: Sharise Griggs

Group Name: TeamSharise

Course: ITAI-2373 – Natural Language Processing

Project Title: NewsBot Intelligence System 2.0

Date: August 3, 2024

The Vision

From the start, I didn't want to just build another text classifier. I wanted this project to feel like a real product, something you could click around and get real value from. That's why NewsBot 2.0 is split into two parts: one that predicts the type of news, and another that goes deeper to show what topics are being discussed and how the content emotionally lands.

Notebook 02: Classification

I used the BBC News dataset and focused on predicting categories like business, tech, sports, politics, and entertainment. After cleaning the text and vectorizing it using TF-IDF, I trained two models: Logistic Regression and Multinomial Naive Bayes. I evaluated both using accuracy and confusion matrices. Logistic Regression was the strongest, so I saved it using pickle and used it in my web app.

Other things I added:

- Category-wise WordClouds for EDA
- Flask frontend with emoji labels + confidence score
- Input validation and live results

Notebook 03: Topic Modeling + Sentiment

This notebook used LDA to break down deeper patterns in the text. I set `n_components=5` and trained a model that revealed topics like:

- Topic 0 – 🎬 Entertainment: film, awards, music, star
- Topic 1 – 📱 Tech & Digital: technology, mobile, digital, phone
- Topic 2 – 🗳️ Politics: government, election, party, minister
- Topic 3 – ⚽ Sports: game, england, win, players
- Topic 4 – 💼 Economy: market, company, economy, firm

I also added sentiment analysis using TextBlob to measure polarity (positive or negative) and subjectivity (fact vs. opinion). I tested this on sample lines and integrated both into a second web app. It's simpler, but it works — and it's real NLP.

Flask Web Apps

1 News Category Classifier (main app)

- Predicts the category using Logistic Regression
- Styled with emoji output + live typing feedback using JS
- Shows confidence score and result instantly
- Built with Flask, custom HTML, and .pkl model files

2 Topic + Sentiment App (bonus app)

- Predicts the dominant topic using LDA
- Shows emotional tone using TextBlob
- Integrated into Flask app for real-time feedback

File Structure (GitHub Ready)

```
newsbot-flask/  
├── app_classifier.py  
├── app_sentiment_topic.py  
├── models/  
│   ├── classifier.pkl  
│   ├── lda_model.pkl  
│   └── vectorizer.pkl  
├── templates/  
│   └── index.html  
├── notebooks/  
│   ├── Notebook_02.ipynb  
│   └── Notebook_03.ipynb  
└── README.md
```

Final Thoughts

This project challenged me but also showed me what I'm capable of. I built this from start to finish, and every piece from training models to writing Flask routes taught me something I can use. NewsBot 2.0 is my proof that I don't just know the theory I can build the real thing.