Final Reflective Journal Ezra Bakatubia

Working on the NewsBot Intelligence System felt less like a routine assignment and more like building an actual product. Since this project pulled together every major NLP technique from Modules 1–8, it became a true test of whether I could design and execute a full pipeline on my own. I knew taking on the entire system solo would be demanding, but I also saw it as an opportunity to sharpen my technical ability and create something meaningful for my portfolio.

Throughout the development process, I worked with text preprocessing, TF IDF vectorization, POS tagging, dependency parsing, sentiment analysis, NER, and multi class classification, integrating all of them into a single end to end workflow. The most difficult part was understanding and implementing the syntax related components, especially the semantic role and dependency parsing features that needed to feed correctly into the classification system. Once those relationships improved the model's performance, the effort felt worthwhile. I spent a large portion of time breaking down each part of the codebase just to fully understand what was happening under the hood.

The techniques that proved immediately useful were TF IDF and sentiment analysis. TF IDF helped surface category specific terms and patterns, while sentiment analysis made it easy to compare tone across articles. The real challenge was pulling everything together into one unified pipeline. I tested each module individually, preprocessing, feature extraction, classification, then the analytical layers, and only integrated once each part was stable. Even with the starter notebook providing structure, adding extra features and ensuring they all ran smoothly required careful debugging.

Once the system could process articles end to end, classify them, extract entities, analyze sentiment, and generate insights, it truly felt like a simplified version of an industry media monitoring tool rather than a midterm assignment.

Since I worked independently, I handled the entire scope: data preparation, pipeline construction, feature engineering, model training and evaluation, visualizations, and documentation. Instead of splitting tasks among teammates, I broke my work into stages, stabilizing preprocessing first, then feature extraction, then classification, and finally the advanced analytics like POS, syntax, sentiment, and entity patterning. The benefit of working alone was having full visibility and control over every part of the system, but the trade off was that all troubleshooting fell solely on me.

The completed system has clear real world applications: supporting journalists, media analysts, marketers, researchers, and anyone who needs fast and organized insights from large volumes of

news. For future improvements, I identified a few directions, integrating a transformer based model for better contextual understanding, adding a live news API for real time data, and building a basic user interface so the system is not restricted to a notebook environment.

Professionally, the project gave me a strong portfolio piece that demonstrates end to end system thinking: planning, implementation, analysis, and technical communication. It pushed my patience, but it also showed me that I can independently design a functioning NLP intelligence system using industry relevant methods. More than just learning techniques in isolation, this project helped me understand how to connect them into something cohesive and usable and made me confident that I can apply these concepts beyond the classroom.