Artificial intelligence

Phase 2 : Innovation

Build an NLP model to differentiate real news from fake news

Introduction:

In this ambitious project, we aim to create an advanced Natural Language Processing (NLP) model

using scikit-learn. Our goal is to develop a web application that seamlessly integrates with popular

news sources via their APIs. This web application will offer real-time news predictions, empowering

users to evaluate the credibility of news articles as they emerge, thus promoting a more discerning

readership.

Project Scope and Objectives:

This project has a multi-faceted scope, including the development of a highly accurate news credibility

differentiation model and the creation of a user-friendly web application to make this technology

accessible to the public.

Objectives:

1. Model Development for Credibility Differentiation:

- Algorithm Exploration: We will experiment with various machine learning algorithms, such as

Random Forest and Naive Bayes, to build a model that excels in distinguishing real news from fake

news.

- High Accuracy: Our primary objective is to identify the algorithm or combination of algorithms that

provides the highest accuracy in news credibility assessment.

- Model Optimization: We will fine-tune the selected algorithm(s) to ensure optimal performance.

2. Web Application Development:

- User-Friendly Interface: Our aim is to design and develop a user-friendly web application that

simplifies technical complexities, ensuring an intuitive user experience.

- Python with Django Rest Framework: We will implement the application using Python's Django

Rest Framework for efficient integration with our trained model.

- React.js Front-End: The front-end of the application will be constructed with React.js to deliver a

seamless and responsive user experience.

3. Real-Time News Integration:

- Integration with The Guardian API: We intend to leverage The Guardian news platform's free API to

access real-time news articles, ensuring a continuous flow of data for our system.

- Dynamic Predictions: The system will dynamically predict the credibility of news articles as they

are retrieved in real-time.

4. Database Management:

- Prediction Storage: To prevent redundant predictions, we will implement a database system to store

the outcomes of previous news credibility assessments.

- Efficient Data Handling: The database will facilitate efficient data management, enabling the

avoidance of repetitive analysis.

5. User-Facing Display:

- Presentation of Predictions: The web application will present news articles along with the

predictions made by our model, clearly indicating whether the news is real or fake.

- User Accessibility: Our goal is to ensure that users can easily access and interpret the credibility assessment without requiring an understanding of the underlying technical aspects.

Detailed Project Plan:

1. Model Training in Google Colab

- Preprocess and clean the news dataset.

- Implement and test various machine learning algorithms (e.g., Random Forest, Naive Bayes).

- Analyze the performance metrics of each algorithm (e.g., accuracy, precision, recall).

- Choose the algorithm that demonstrates the highest accuracy for news credibility assessment.

- Save the trained model in Google Colab.

- Export the selected model from Google Colab to your local machine.

2. Web Application Development

- Set up a Django project for the web application.

- Develop the RESTful API to handle incoming news data and predictions.

- Begin developing the user interface using React.js.

- Establish communication with the Django REST API to fetch predictions.

- Integrate the web application with The Guardian's news API to retrieve real-time news articles.

- Implement logic for running background threads every 10 seconds to fetch new news.

- Develop functionality to check whether the news is already in the database.

- Set up a database system to store news articles and their predicted results.

- Establish database connectivity within the Django project.

3. Testing, Deployment, and Monitoring

- Conduct extensive testing, including unit testing and integration testing, to ensure the system

functions correctly.

- Verify the accuracy of news predictions.

- Address and resolve any issues or bugs that arise during testing.

- Plan for regular maintenance, updates, and scalability considerations.

4. Project Evaluation

- Continuously monitor the accuracy of news predictions.

- Stay updated with developments in machine learning and NLP to enhance the model's accuracy.

Conclusion:

Our project represents a fusion of advanced NLP model training with user-centric web application

development, signaling an innovative approach to combat misinformation. With a carefully selected

algorithm, real-time news predictions, and a commitment to user-friendliness, we are positioned to

empower individuals with reliable news assessments. As we move forward, our dedication to accuracy

and continuous improvement remains steadfast, marking the beginning of a new phase in our mission

to promote informed decision-making and accurate information dissemination.

THANK YOU

Submitted by:

D.MUNI BABU

bmuni064@gmail.com

au723921104012