Text political bias classifier

MAIS Winter 2025
Roman Sejnoha

We will be building an application that allows users to determine the political lean of a given piece of text, with special emphasis placed on news media. As of right now, we only have this writeup and our code as our teammate dedicated to the webapp has yet to update us on it.

1. Choice of Dataset

For this project, we will use a dataset containing over 3000 political news articles with bias classifications. The dataset includes:

- Title (The headline of the article)
- Link (The URL to the full article)
- Text (The full content of the article)
- Source (The name of the news outlet)
- Bias (labeled as *left, lean left, center, lean right, right*)

We chose this dataset because political bias detection is a pressing issue in media literacy. The dataset is well-structured and provides a diverse set of articles from various sources, making it feasible for supervised learning.

2. Methodology

This project will involve natural language processing (NLP) and machine learning (ML) classification techniques to predict an article's political bias.

a. Data Preprocessing

To prepare the dataset for training, we will:

- Clean text data (investigating the need for the removal of punctuation, stopwords, special characters, and any other artifacts of the transfer from an article's webpage to the dataset)
- Tokenize and lemmatize words for efficient processing
- Feature extraction

b. Machine Learning Model

Our goal is to classify an article's political bias based on its text. We will experiment with the following models:

1. Baseline model: Logistic Regression
2. Traditional ML models: e.g. Random Forest
3. Deep Learning models

c. Evaluation Metrics

Split the data into training and testing sets, and since this is a multi-class classification problem, we will use:

- Confusion matrix (to analyze misclassifications)
- Accuracy, Precision, Recall, and F1-score (to evaluate performance)
- Word cloud analysis (to visualize the most influential words per bias category)

We will compare model performance against a baseline accuracy (e.g., predicting the most common class) and aim for at least 80% accuracy.

3. Application

We will integrate our model into a web-based app where users can:

1. Input: Paste a news article or provide a URL.
2. Processing: The model will analyze the article text and classify its bias.
3. Output: Display the predicted bias (*left, lean left, center, lean right, right*), explain key contributing words, and provide a bias heatmap.

4. Conclusion

This project aims to develop an AI-powered bias detection tool that helps users critically evaluate news articles. By leveraging NLP, machine learning, and interpretability methods, our model will provide transparency into political bias in news media.