Project Title

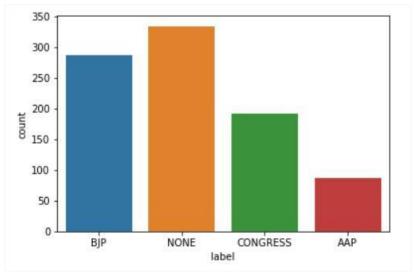
spotBias-Identifying Political Bias in News Article.

Motivation

The major reason we chose this issue was to determine the political bias in a news story. Certain media sources tend to produce articles that support a political party's point of view or propaganda. As a result, they approach news items with a predetermined agenda. Because the general public is frequently unaware of a Media Outlet's preferences, they tend to believe them and the ideas/incidents mentioned in their publications. We believe that the general public must know the truth (before forming an opinion) without succumbing to the bias of media outlets or bloggers' published articles.

Dataset

The dataset was built using BeautifulSoup on News Catcher API and Google News.It currently consists of roughly 900 articles and respective annotations. The annotations represent the bias in the news articles. Link to the <u>dataset</u>



Preprocessing steps:

Preprocessing is needed to clean the dataset so that we get better results. We have used the NLTK library for the preprocessing of our dataset. We have performed preprocessing on the dataset as well as the input query.

Various steps followed are:

- Converted text to lowercase.
- Removed all punctuations from the dataset.
- Removed all the blank spaces.
- After that, we have tokenized the text using word tokenizer from the NLTK library.
- After that, remove the stopwords using the NLTK library.

METHODOLOGY

- First implemented the above-mentioned preprocessing steps.
- Then we followed two approaches: With feature selection and Without feature selection.
- In with feature selection, we did the following steps:
 - Created table dictionary to keep class count for each term.
 - Calculated mutual information for each term in each class.
 - Created a word map and assigned each word to a unique id.
 - o Then we created a method to select features for each class.
- Then we applied Count vectorizer and TF-IDF Vectorizer and tried different classification models such Multinomial Naive Bayes, SVM, KNN, Logistic Regression, and Random Forest with hyperparameter tuning.
- We develop the application using these tech/framework NewsCatcherAPI, HTML,CSS, Bootstrap, FLASK, SQLite3.

Build Status

We have deployed the <u>spotBias-Web Application</u> in our local server- Ubuntu.

Framework

Front End: HTML, CSS, Bootstrap

Back End: FLASK News Catcher API

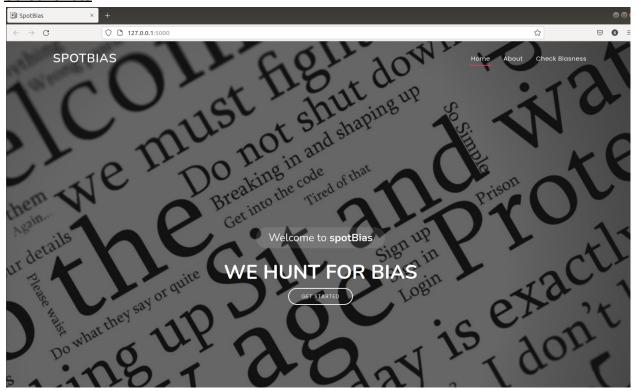
Database Connectivity SQLite3

Google News Republic News NDTV News

Features

- 1. Fetch News: 4 different News Articles will be fetched and their headings will be displayed. Users can select from any of the 4 headings and the selected news article will be displayed.
- 2. Predict Biasness: Our web application will identify if the selected news article is biased towards any political party (BJP, Congress, and AAP) or not.
- 3. Database Connectivity: In our database, we'll save the selected news article, predicted biasness, and what users think about which political party it's biased.

Screenshots

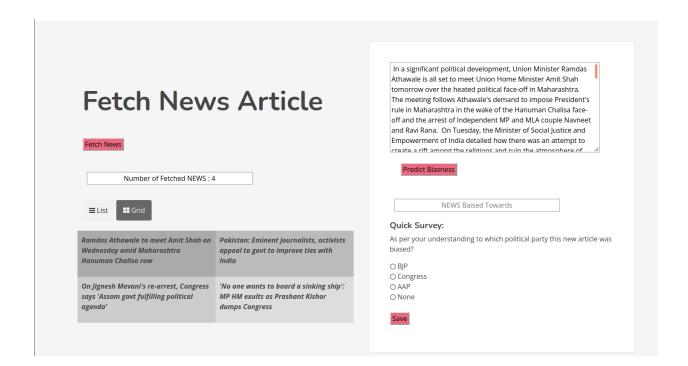


ABOUT

Developing a NLP/ML based web application that in real time reports whether the reporting done in news articles is politically biased or not.

Our main motivation for selecting this topic includes identifying the political bias present in a news article. Certain media outlets tend to publish articles which strengthen a political party's view point or propaganda. Thus, they are approaching news items with a biased agenda. Since the general public most often then not, are not aware of the preferences of a Media Outlet, tend to believe them and the ideas/incidents reported in their articles.

We believe it's highly important for the general public to know the truth (before forming an opinion), without falling prey to bias of the Media outlets or the bloggers'



Database:

