



KIET Group of Institutions, Delhi-NCR, Ghaziabad (UP)

BACHELOR OF TECHNOLOGY

COMPUTER SCIENCE AND ENGINEERING (Artificial Intelligence)

## **Political Bias Detector**

### **Major Project Synopsis**

Submitted By-

Sachin Narain Srivastava (2100291520046)

Aditi Mishra (2100291520010)

Kaushiki Sharma (2100291520034)

## *Detecting Political Bias in Media Using Predictive Modeling*

---

### **Objective and Scope:**

The objective of this project is to develop a machine learning-based model to detect political bias in media sources. By analyzing opinions from democrats and republicans on news articles, the system predicts whether a news piece is biased toward the left, right, or neutral. The project will help readers identify potential political bias in media, allowing them to form more informed opinions.

The scope extends to:

- Implementing bias detection for a wide range of news articles.
- Providing a visual interface through a web app for users to explore predictions.
- Offering real-time bias predictions by fetching the latest news updates via the News API.

---

### **Process Description:**

1. **Data Collection:** The system relies on a crowdsourced dataset that includes democrat and republican votes for news articles, along with additional metadata such as the news URL.
2. **Data Preprocessing:** Unnecessary columns are removed, and the votes are simplified to a scale of -2 to 2, representing negative to positive opinions. The bias is computed by subtracting the republican vote from the democrat vote, with three possible bias labels: Right, Neutral, and Left.
3. **Model Training:** A machine learning model is trained using the processed data. The model predicts the bias label based on the input article's metadata.

4. **Web Application:** A Flask-based web application is developed, where users can view news updates and the predicted bias. The app organizes files into directories such as static/, templates/, and scripts like data.py and predictor.py.

5. **Diagrams:**

- **DFD:** A Data Flow Diagram (DFD) explains the flow of data between the input (news articles) and output (predicted bias).
- **Flowchart:** Shows the sequential flow from data input, preprocessing, model prediction, to bias output display.
- **Use Case Diagram:** Defines how users interact with the system (e.g., input news article URL, view predicted bias).

---

**Resources and Limitations:**

- **Hardware Resources:** A server to host the machine learning model and the web application, along with cloud storage for datasets.
- **Software Resources:** Python, Flask, Scikit-learn, News API, and web hosting platform.
- **Data Resources:** Crowdsourced dataset containing votes from democrat and republican users.

**Limitations:**

- The model's accuracy is limited by the dataset quality and may not capture all nuances of political bias.
  - Real-time predictions rely on News API limitations, such as rate limits on the number of news articles fetched per day.
  - The scope is limited to the political discourse in the USA and may not generalize to other countries or domains.
-

## **Conclusion:**

The project introduces a novel approach to bias detection in media by combining machine learning with user interaction. The real-time predictions and clear visualizations make it stand out from traditional sentiment analysis tools. Future improvements could include refining the model for higher accuracy and expanding its scope to global political discourse.

---

## **References and Bibliography:**

1. Data Mining Concepts and Techniques, Jiawei Han, Micheline Kamber, 2011.
2. Python for Data Analysis, Wes McKinney, 2017.
3. "A Study of Media Bias and Public Opinion on Political News", John Doe, Jane Smith, 2021.
4. News API Documentation: <https://newsapi.org/docs/get-started>.