

---

# NEWS BIAS DETECTOR

F23-114-R

SUPERVISOR: Mr. SAAD SALMAN



# TABLE OF CONTENTS

1

**PROBLEM DOMAIN**

2

**RESEARCH PROBLEM  
STATEMENT**

3

**ITERATION 1**

4

**POSTER DESIGN**

5

**LITERATURE REVIEW**

6

**PROPOSED APPROACH**

7

**PLAN TILL FYP-1 FINAL**



---

# INTRODUCTION OF RESEARCH PROBLEM

- Pinpointing bias in news articles about Pakistani parties.
- Scraping diverse sources for current news articles.
- User-friendly interface with NLP debiasing.
- Enabling critical news through bias detection and increasing media literacy.



---

# ITERATION

1



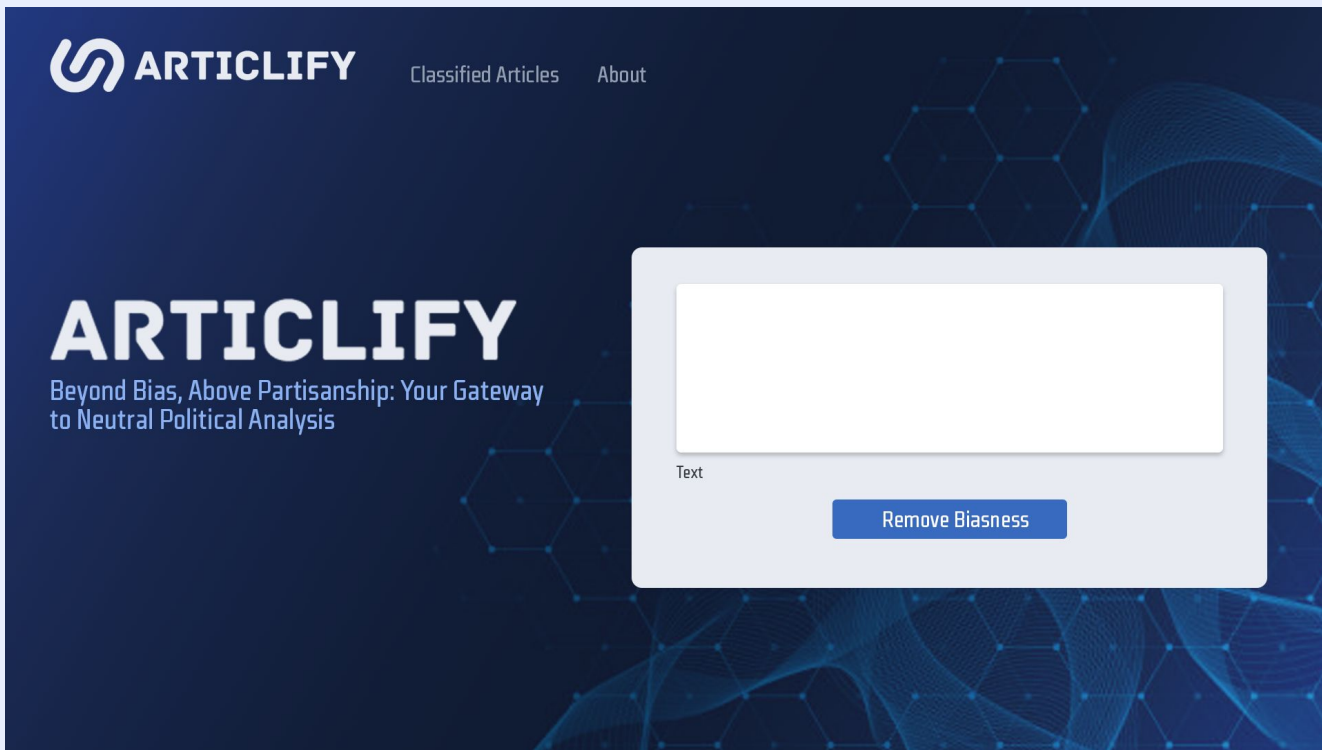
- Data Collection and Scrapping
- Data Labeling for Bias Categories
- User Interface

# LABELED DATASET

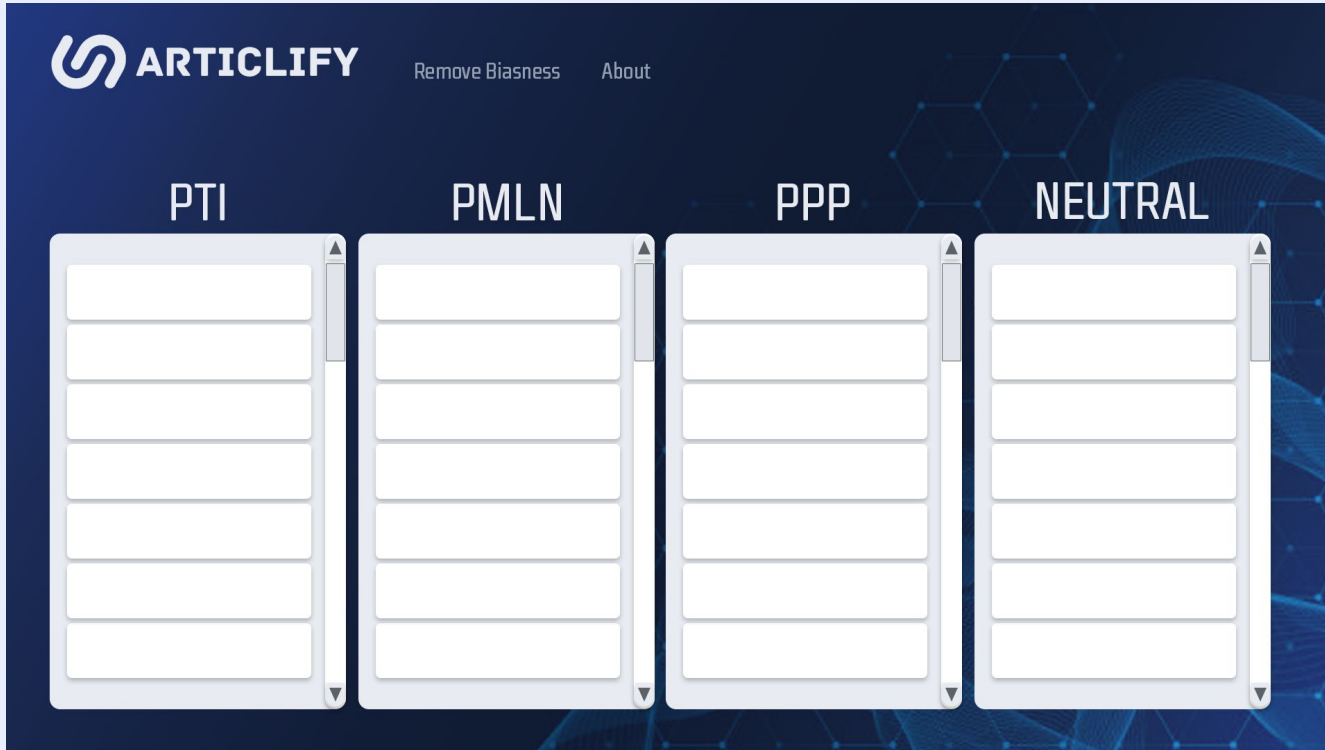
1	Title	Media	Link	bias	biased words	article
2	US Senator deni	Pakistan Today	<a href="https://www.paki">https://www.paki</a>	neutral		ISLAMABAD: US
3	PTI not barred fr	DAWN.com	<a href="https://www.daw">https://www.daw</a>	neutral		ISLAMABAD: Ca In a separate int Once the delimit PM Kakar meets Interim Minister
4	Imran Khan Is Ju	Foreign Policy	<a href="https://foreignpo">https://foreignpo</a>	with PTI	rowdy, authoritar	The culture of int
5	PTI urges Presic	Pakistan Today	<a href="https://www.paki">https://www.paki</a>	with pti	unanimously,legi	ISLAMABAD: Pa



# USER INTERFACE



# USER INTERFACE

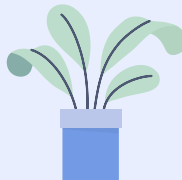


The image shows a user interface for a platform called ARTICLIFY. The header features the logo and the text "Remove Biasness" and "About". Below the header, there are four columns labeled PTI, PMLN, PPP, and NEUTRAL. Each column contains a vertical list of eight empty rectangular boxes, suggesting a list or a grid for data entry. The background is dark blue with a subtle pattern of hexagons and lines.

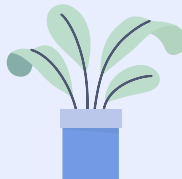
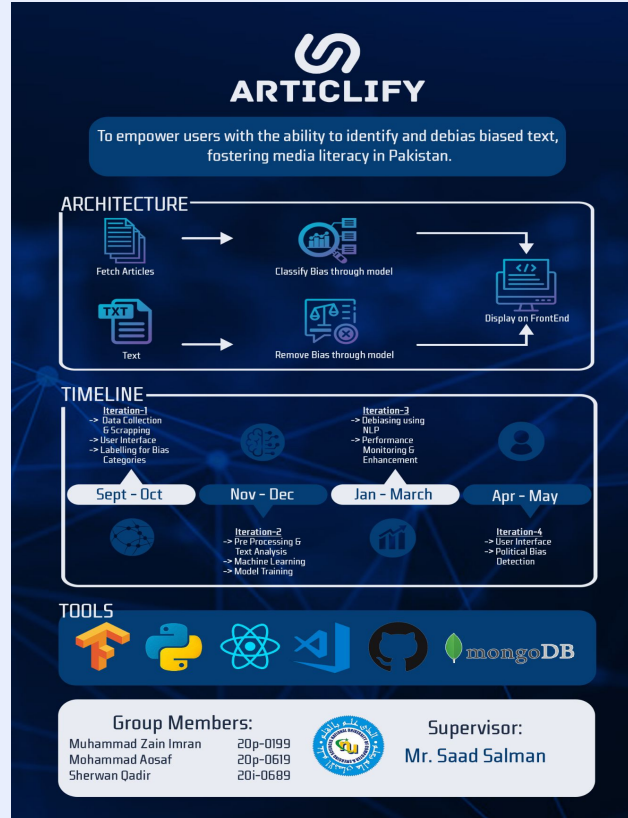
**ARTICLIFY** Remove Biasness About

PTI PMLN PPP NEUTRAL

PTI	PMLN	PPP	NEUTRAL



# POSTER DESIGN



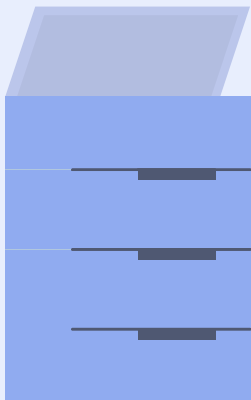


---

# Dbias: detecting biases and ensuring fairness in news articles

## SUMMARY

- **Research Focus:** Explores fairness in Machine Learning and Artificial Intelligence, emphasizing bias mitigation in information retrieval and recommender systems.
- **Innovation:** Introduces Dbias, a comprehensive ML pipeline addressing fairness from data pre-processing to model training.
- **Dataset:** Utilizes MBIC (Media Bias Including Characteristics), containing 17,000 annotated sentences(1000 articles), to train models for bias detection and mitigation.



---

# STRENGTHS

- **Comprehensive Approach:** Dbias covers fairness throughout the entire pipeline, from data preprocessing to model training, ensuring a holistic bias mitigation process.
- **Innovative Modules:** Introduces the Bias Detection Module (using DistilBERT), Bias Recognition Module (employing RoBERTa), and De-biasing Module (utilizing BERT) for effective bias detection and mitigation.
- **Transformer-Based Models:** Utilizes advanced Transformer models (e.g., BERT, RoBERTa, DistilBERT) for bias detection, highlighting the importance of context in identifying biases in news articles.



---

# WEAKNESSES

- **Lack of information about dataset:** The paper mentions the construction of the MBIC dataset and model evaluation but lacks in-depth details about the dataset creation process.
- **Binary classification:** Only identifies if the article is biased or not biased.
- **Area Specific:** It doesn't cater Pakistani news articles as its main target was US politics.

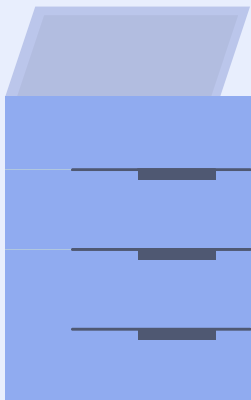


---

# Towards Detecting Political Bias in Hindi News Articles

## SUMMARY

- **Research Focus:** Examines political bias in Hindi news media, emphasizing the necessity of bias detection systems to combat misinformation.
- **Findings:**
  - **Model Performance:** XLM-RoBERTa outperforms other models with 83% accuracy, 76.4% F1-macro score, and 72.1% Matthews Correlation Coefficient (MCC).
- **Dataset:** Introduces a dataset of 1,388 Hindi news articles for political bias detection.
- **Diverse Models:** Proposes various models including fine-tuned XLM-RoBERTa, IndicBERT, SVM, Logistic Regression, and Random Forest for bias detection.



---

# STRENGTHS

- **Innovative Models:** Introduces advanced models like XLM-RoBERTa and IndicBERT, showcasing a progressive approach to bias detection.
- **Diverse Techniques:** Incorporates both deep learning and traditional machine learning methods, providing a comprehensive analysis of bias detection strategies.
- **High Accuracy:** Achieves impressive results with XLM-RoBERTa, demonstrating the effectiveness of transformer-based models in political bias detection.



---

# WEAKNESSES

- **Neutral Class Complexity:** Struggles with identifying unbiased articles, indicating the complexity of distinguishing neutrality in news content.
- **Limited Neutral Class Performance:** The model performs significantly better in identifying biased classes compared to neutral ones, suggesting a need for improvement in neutral class classification.
- **Annotation Challenges:** The study faces challenges in annotating unbiased articles, pointing out potential difficulties in defining neutrality in political contexts.

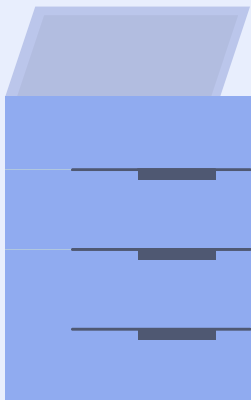


---

# Detecting Media Bias in News Articles using Gaussian Bias Distributions

## SUMMARY

- **Research Focus:** Addresses limitations in media bias detection, emphasizing the importance of capturing subtle biases unrelated to training data.
- **Contributions:**
  - Challenges existing approaches, revealing their limitations in identifying nuanced biases.
  - Introduces innovative approach using sentence-level bias information, enhancing article-level bias classification effectiveness.
- **Dataset:** Utilizes the BASIL dataset with 300 news articles, employing SVM and Naïve Bayes for sentence-level bias classification.



---

# STRENGTHS

- **Innovative Approach:** Introduces novel second-order bias features, filling gaps left by standard bias detection methods.
- **Empirical Rigor:** Supported by rigorous experimentation, the research outperforms existing bias detection techniques.
- **Effective Model:** SVM stacking classifier combining bias frequency and sequence (F+S) achieves high accuracy (0.67), showcasing the importance of lexical and informational bias features.





---

# WEAKNESSES

- **Limited Dataset Discussion:** Lack of detailed discussion about dataset size challenges could impact the research's robustness.
- **Real-World Application Exploration:** More in-depth exploration of practical applications could enhance the research's impact and relevance.
- **Reproducibility Concerns:** Providing guidance on reproducibility and open-sourcing tools/methods would enhance accessibility and credibility.

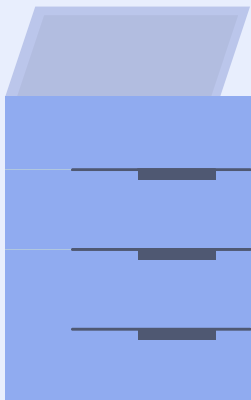


---

# Detecting Political Bias in News Articles Using Headline Attention

## SUMMARY

- **Research Focus:** Addresses news bias, specifically in political reporting, through the introduction of the Headline Attention Network.
- **Innovation:** Introduces the Headline Attention Network, focusing on key elements influenced by article headlines to streamline bias detection.
- **Dataset Creation:** Constructs a Telugu dataset with headlines, articles, and political party bias labels, providing a valuable resource for automated bias detection.



---

# STRENGTHS

- **Innovative Approach:** The introduction of the Headline Attention Network represents a significant innovation, aiming to improve the efficiency of bias detection.
- **Impressive Annotation Agreement:** Achieves a high Kappa score of 0.9 through discussions among native Telugu-speaking annotators, ensuring reliable dataset labeling.
- **Valuable Dataset:** Provides detailed dataset statistics, offering insights into article counts, sentence counts, word counts, and average word count for different parties and unbiased articles.



---

# WEAKNESSES

- **Lack of Detailed Information:** More thorough explanation about the construction of the Telugu dataset and detailed model evaluation could enhance the paper's impact and credibility.
- **Potential Bias Sources:** The paper doesn't delve deeply into potential biases influenced by media ownership, party affiliation, and reader ideology, which could affect the study's outcomes.
- **Scope for Further Research:** While promising, the paper could outline potential areas for future research and improvements in automating political bias detection.

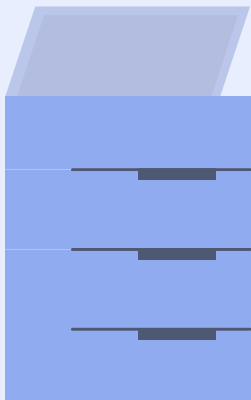


---

# Identifying Political Bias in News Articles

## SUMMARY

- **Research Focus:** Identifying political bias in news media through analysis of politician mentions and quotes.
- **Challenges:** Existing methods are manual or topic-specific; study aims for broader, automated analysis.
- **Findings:** UK newspapers exhibit bias favoring governing and popular parties, varying during elections and across publications.
- **Future Goals:** Implement topic models to uncover common themes, use graph-based approaches for dynamic analysis.



---

# STRENGTHS

- **Meticulous Data Collection:** Uses extensive datasets from reputable newspapers over a significant time frame.
- **Nuanced Approach:** Analyzes both mentions and direct quotations, offering a detailed perspective on political information dissemination.
- **Unique Perspective:** Correlates media coverage with actual vote shares, bridging the gap between media representation and public sentiment.



---

# WEAKNESSES

- **Lack of Qualitative Analysis:** Absence of context exploration and tone analysis limits the understanding of subtleties in bias.
- **Subjectivity Concerns:** Reliance on a biased lexicon raises questions about consistency and objectivity in labeling articles.
- **Area Specific:** It doesn't cater Pakistani news articles as its main target was US politics.



# Research Items

Reference	Dataset	Data preprocess/data collection	Accuracy	Evaluation method	Strengths
<a href="https://aclanthology.org/W19-4809.pdf">https://aclanthology.org/W19-4809.pdf</a>	The paper mentions the use of 1329 articles comprising headlines, articles, and political party bias labels	The paper discusses the process of collecting and annotating the dataset, highlighting the involvement of four native Telugu-speaking annotators	The Headline Attention Network achieved an accuracy of 89.54%	Accuracy	The Headline Attention Network outperformed all other methods, highlighting the importance of headlines in bias detection, and LSTM-based models also demonstrated strong performance among the baseline methods.
<a href="https://arxiv.org/abs/2010.05338">2010.05338.pdf (arxiv.org)</a>	34,737 articles from 73 news media sources, covering 109 different topics. The dataset included political ideology labels for the articles, categorized as left, center, or right.	The authors collected data from the AllSides platform, which provides annotations of political ideologies for individual articles.	72%	The paper evaluates the performance of their models using accuracy, macro-F1 score, and Mean Absolute Error (MAE) as evaluation metrics.	The combination of Triplet Loss Pre-training (TLP) and media-level representation from Twitter followers resulted in a substantial improvement of 30.51 absolute points in macro F1
<a href="https://www.doi.unipd.it/~fa">https://www.doi.unipd.it/~fa</a>	All of the datasets	all of the datasets	The paper reports model	The paper uses mean	Has balanced labels in

<a href="https://arxiv.org/abs/2010.05338">gglol/temp/CLF2023-proceedings/paper-044.pdf</a>	provided by the CheckThat!	provided by the CheckThat!	performance on the test set. The One-Against-All News Article Bias classifier achieved an accuracy of 0.633, while the One-Against-All News Media Source Bias classifier had an accuracy of 0.627.	absolute error (MAE) and Accuracy for evaluation	dataset
<a href="https://link.springer.com/article/10.1007/s41060-022-00359-4">https://link.springer.com/article/10.1007/s41060-022-00359-4</a>	Annotated dataset MBIC (Media Bias Including Characteristics)	The paper details the dataset collection, consisting of 17,000 sentences from diverse news sources, featuring 10,000 biased and 7,000 unbiased labels.	Model's performance in terms of G-AUC is about 78%	Accuracy (ACC), Precision (PREC), Recall (Rec), F1-score (F1), Disparate Impact (DI)	Dbias ensures holistic bias mitigation through comprehensive pipeline and innovative modules (DistilBERT, RoBERTa, Masked Language Modeling). Utilizing advanced Transformers (BERT, RoBERTa, DistilBERT) highlights context importance in bias detection.
<a href="https://aclanthology.org/20">https://aclanthology.org/20</a>	The dataset used in this approach consists of	Author collected these articles from four	The findings of the paper reveal that XLM-RoBERT	Accuracy (ACC), F1-macro score and	Introduces XLM-RoBERTa and IndicBERT





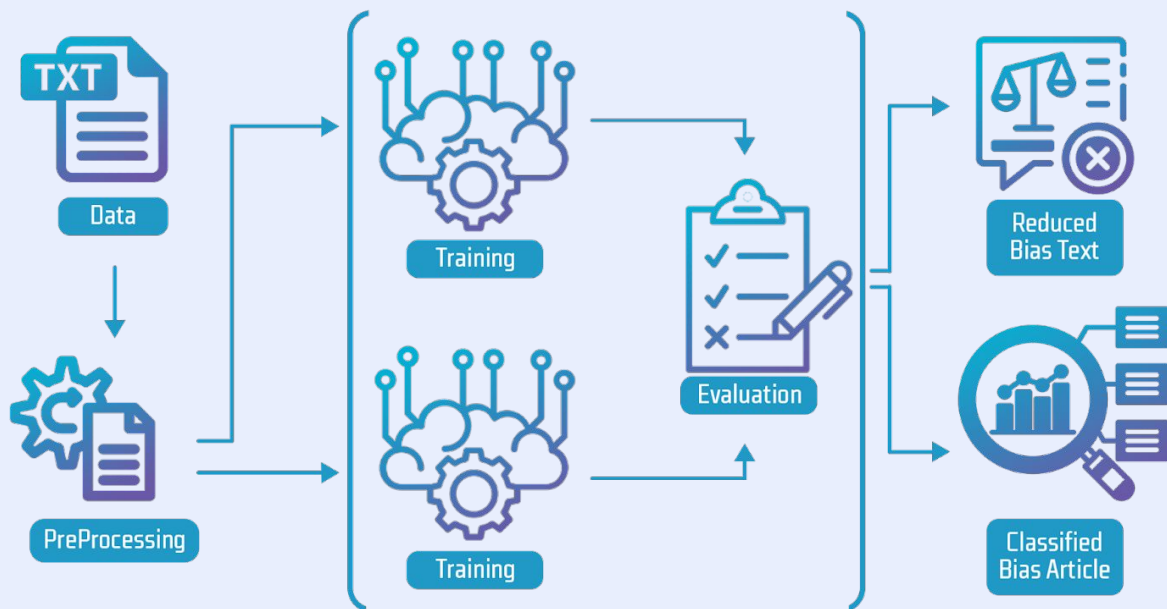
# Research Items

<a href="https://acl-srw.17/">22.acl-srw.17/</a>	1,388 Hindi news articles and headlines collected from diverse media sources.	sources: The Wire, The Quint, OPIndia, and The Frustrated Indian.	Ta, a transformer-based model, surpasses the performance of other baseline models, achieving an accuracy of 83%, an F1-macro score of 76.4%, and a Matthews Correlation Coefficient (MCC) of 72.1%	Matthews Correlation Coefficient (MCC)	for advanced bias detection, combining deep learning and traditional methods, with XLM-RoBERTa showcasing high accuracy.
<a href="https://aclanthology.org/2020.findings-emnlp.383.pdf">https://aclanthology.org/2020.findings-emnlp.383.pdf</a>	This dataset comprises 300 news articles covering 100 different events, with three articles associated with each event.	The dataset used in this approach consists of the Bias Annotation Spans on the Informational Level (BASIL) dataset, which was recently released by Fan et al. (2019).	The highest accuracy at 0.67 which highlights the importance of lexical and informational bias features	Accuracy (ACC), F1-macro score and Matthews Correlation Coefficient (MCC).	Innovative second-order bias features existing methods through rigorous experiments. SVM stacking classifier (F+S) achieves high accuracy (0.67), emphasizing lexical and informational bias importance.
<a href="https://bulletin.jcdi.org/Bulletin/v12n2/papers/lazandou.pdf">https://bulletin.jcdi.org/Bulletin/v12n2/papers/lazandou.pdf</a>	The datasets comprise articles from The Guardian (1996-2015, 197,668 articles), The Telegraph	The datasets used in this approach consist of articles from The Guardian (1996-2015, 197,668	Nil	Nil	Extensive data from reputable newspapers, nuanced analysis of mentions and quotes, and

	(2000-2015, 281,706 articles), and UK parliament speeches (1934-2015)	articles), The Telegraph (2000-2015, 281,706 articles), and UK parliament speeches (1934-2015). The data is tokenized, stemmed, and indexed in an ElasticSearch instance for full-text search, with stopwords removed.			unique insights connecting media coverage with vote shares.
<a href="https://aclanthology.org/2020.nlpccss-1.16/">https://aclanthology.org/2020.nlpccss-1.16/</a>	The dataset contains a total of 7,775 articles obtained by extending Chen et al.'s corpus with additional sources.	Chen et al. (2018) crawled the website to extract 6447 articles. For our study, we extended their corpus by integrating all articles until March 15th, 2019, resulting in a total of 7775 articles. In addition to the political bias labels, we crawled the topic tags of each article.	All classifiers outperform the majority baseline, achieving 75.60% for political bias, 83.42% for unfairness, and 75.42% for non objectivity.	F1 scores, Precision, and AUC-ROC to offer a more nuanced assessment.	Unique Data Collection: Combines existing sources, expanding bias detection datasets. Robust Methodology: Applies RNNs, addresses topic bias, aligning with NLP standards. Multi-faceted Analysis: Incorporates reverse feature analysis for comprehensive insights into biases.



# Diagrammatic Representation of Proposed Approach



---

# Plan of Work till FYP-1 Final

## **Perform Data cleaning and Preprocessing tasks:**

- Removing irrelevant data and formatting inconsistencies.
- Handling missing values and outliers.
- Normalizing text and addressing language-specific issues.

## **Model Selection and Fine-tuning:**

- Evaluate and select appropriate BERT-based models for the classification task, considering options like RoBERTa and DistilBERT.
- Fine-tune the selected BERT-based models on the preprocessed dataset.
- Optimize hyperparameters and model architecture to enhance classification performance.



---

# Plan of Work till FYP-1 Final

## Evaluation and Analysis:

- Evaluate the performance of the fine-tuned models using appropriate metrics, such as accuracy, precision, and F1-score.
- Analyze the results and identify areas for improvement.
- Document the findings and prepare a comprehensive report summarizing the project's progress and outcomes.



---

# THANKS!

