KU LEUVEN

FACULTY OF
ENGINEERING SCIENCE

The best master's thesis ever

First Author
Second Author

Thesis submitted for the degree of
Master of Science in
Electrical Engineering, option
Electronics and Chip Design

Supervisor:
Prof. dr. ir. Knows Better

Assessors:
Ir. Kn. Owsmuch
K. Nowsrest

Assistant-supervisors:
Ir. An Assistent
A. Friend

Academic year 2023 – 2024

Preface

I would like to thank everybody who kept me busy the last year, especially my promoter and my assistants. I would also like to thank the jury for reading the text. My sincere gratitude also goes to my wive and the rest of my family. [2]

First Author
Second Author

Contents

Abstract

In today's digital, information-rich society, **media** bias poses a significant challenge to the objectivity and credibility of news reporting. As someone living in our current society, one has inevitably encountered some form of bias in the media, either consciously or unconsciously. Media bias can shape our perceptions, influence our opinions, and affect our understanding on various issues. It is crucial to recognise and address this bias to ensure a well-informed and balanced perspective. By being aware of the inherent biases in media sources, individuals can critically evaluate the information they consume and seek out diverse viewpoints to form a more comprehensive understanding of the world. In this paper...

List of Figures and Tables

List of Figures

List of Tables

List of Abbreviations and Symbols

Abbreviations

LoG	Laplacian-of-Gaussian
MSE	Mean Square error
PSNR	Peak Signal-to-Noise ratio

Symbols

42	"The Answer to the Ultimate Question of Life, the Universe, and Everything" according to [**?**]
c	Speed of light
E	Energy
m	Mass
π	The number pi

Chapter 1

Introduction

1.1 Overview

This thesis project is part of an Advanced Master of Artificial Intelligence programme—Speech and Language Technology, in collaboration with the Media Bias Group [1], who provided the topic and additional guidance along the project.

1.2 Motivation

Media misinformation and manipulation are rampant in today's landscape. Despite being a major societal issue, there has been hardly enough resources and work dedicated to in the realm of media bias and its broader context.

Considerable work has been done on fake news and its detection, media bias operates on a higher level and cannot be described the same simply fake news. Fake news involves intentionally spreading false information, whereas media bias refers to the distortion or manipulation of information by media outlets, which may or may not be intentional, to favor certain perspectives or agendas. Unlike fake news, media bias can influence public perception subtly through selective reporting, framing, or sourcing, making it a complex and challenging issue to address comprehensively.

To combat media bias effectively, it is crucial to develop robust methodologies and technologies for detecting and analyzing biased content across various media platforms. Additionally, raising awareness about media literacy and critical thinking skills can empower individuals to identify and navigate biased information effectively in today's media landscape. Ultimately, addressing media bias requires concerted efforts from researchers, policymakers, media organizations, and the public to promote transparency, accountability, and integrity in news reporting and consumption.

1.3 Goals

To aid in the work of media bias, the goal of this project are as follows:

1. Build a reliable, annotated dataset comprised of articles content and metadata.

2. Propose a method to effectively represent articles in a space vector.
3. Design a system that is able to effectively detect and classify bias on the article level of granularity.
4. Validate the resulting system and dataset.

1.4 About the Media Bias Group

The Media Bias Group [1] was established in mid-2020 by Timo Spinde during his pursuit of a Ph.D. in computer science, having been integrated in the topic since his undergraduate studies, with a vision to aid others perceive news in a more balanced and conscious manner. After a year of planning how a system could uncover bias on a vast scale encompassing millions of articles, he founded the group and forged connections with various partners, particularly those relevant to specific aspects of the project. In just one year, the project has garnered support from multiple other research groups, with around twenty students from seven countries joining to contribute to the system. Since 2021, the group has also begun offering its first Ph.D. positions.

The group is comprised of a collective of scholars across various fields such as Psychology, Linguistics, and Computer Science, with a shared goal to comprehend the factors influencing human perception of news content as biased or one-sided. Currently, the network includes of six main researches and coordinators, twenty one professors and postdocs, as well as eight active students. Numerous publications related to media bias have been published through the network into major conferences such as EMNLP 2021 [23], along with dataset and benchmark creations.

Chapter 2

Media Bias

2.1 Definition

Allsides [11] defines media bias as "The tendency of news media to report in a way that reinforces a viewpoint, worldview, preference, political ideology, corporate or financial interests, moral framework, or policy inclination, instead of reporting in an objective way (simply describing the facts)". This phenomenon has existed and researched since the 1950s (cite), highlighting its enduring presence and impact on public perception. Media bias can manifest in various forms, including the selection of stories, framing of issues, and choice of language. This project would mainly focus on text-level context bias: phrasing bias, spin bias, and statement bias, as described in [22]: statement Bias refers to “members of the media interjecting their own opinions into the text”, phrasing Bias is characterized by inflammatory words, i.e., non-neutral language, spin Bias describes a form of bias introduced either by leaving out necessary information.

While biases may not always be intentional, they might cause significant consequences, possibly leading to inequalities and injustices. Some news outlets tend to use catchy headlines which trick readers into clicking, known as "clickbait", which are often ambiguous or misleading. Biased information can and has been used as a way to shape and influence public opinion [2]. A survey of journalists from United States, Great Britain, Germany, Italy, and Sweden, where they found evidence that journalists’ personal beliefs substantially influence their news decisions, expressed within the stories they choose and the statements they write [20].

2.2 Application

Another particularly dangerous example of media bias is its application in relation to elections. According to a survey done by [3], fake news propagated in social media played a pivotal role in the eventual election of President Trump during the 2016 election. Panagopoulos’ study [18] revealed systemic biases leaning towards Democratic candidates during national and state levels pre-election polls conducted during the 2020 U.S. general election cycle. Rafail et al. [21] examined 201,678 media documents from Tea Party organizations, Fox News, MSNBC, and 785 newspapers, revealing significant differences in how the Tea Party frames itself compared to how other media sources frame the movement, MSNBC portrays it as the worst aspect of the Republican Party, while Fox News sees it as the best, sharply in contrasts with how activists frame the movement as conservative but not strictly Republican, often clashing with Republican Party goals. An analysis by Pew Research Center [12] found that miscalculations observed in the 2020 U.S. election polls would adjust public opinion on issues by an average of just under 1 percentage point. Although errors of such scale would not have produced substantial difference on the American public opinions, this shows the underlying bias within polls specifically and the failure of accurately representing surveys.

The phenomenon of media bias in media has not gone unnoticed, particularly by readers and consumers, lowering public trust on media outlets in general. In the United States particularly, Trust on news media is at an all time low, falling consistently and significantly over the past 20 years [8, 7, 16]. Approximately half of Americans believe that the media is significantly responsible for the political divisions within the United States, with a growing number of Americans losing faith in the media's objectivity and perceive it as actively engaging in ideological wars [7]. Reuters Institute reported only less than half of their respondents (40%) generally trust the majority of news sources, with the US ranked as second-lowest by country (32%) [16].

Readers themselves are not out of the woods, as they are known to prefer to pick, follow, and consume articles that align with their own beliefs and ideology, an issue known as filter bubble [10]. This can be dangerous as it reinforces existing biases and limits exposure to diverse perspectives, creating echo chambers that hinder critical thinking and informed decision-making. The combination of media bias and filter bubbles can distort reality, perpetuate misinformation, and deepen societal divisions. It is essential for readers to seek out a variety of sources and viewpoints to gain a more balanced and comprehensive understanding of the issues at hand.

2.3 Goal

Ideally, unbiased media contents that objectively and fairly represents multiple or a range of perspectives are desirable, news sources should remain neutral and let readers build their own opinions on the subject [17]. However, this is often unachievable due to human capabilities and resources limitations; journalists cannot possibly possess complete knowledge on every topic, be physically present everywhere, or interview every relevant individual on a significant subject. [11]. Truth and journalism objectivity is a complex matter full of choices and dilemma, where ultimately it falls on the journalist own preferences and criteria [4].

Therefore, instead of eliminating media bias, our goal should be to draw attention to its existence, giving readers awareness of such content [22], ultimately building a tool to defend readers from media manipulation, to let readers know on the quality of the article, and if they fall under a victim of political agenda or indoctrination.

Therefore, the primary objective of a media bias classifier should be to develop an automated system capable of classifying media bias at the article level. This system should be able to recognise and categories bias in news articles from diverse sources, ensuring that readers are aware of what they are reading and therefore can make informed choices and opinions. The classification should encompass various dimensions of bias, including political orientation (e.g., left, right, center), sensationalism, and framing techniques.

Chapter 3

Literature Review

3.1 Dataset

Datasets for media bias ([23, 24, 6, 5]) often have different metrics and formats, covering different types of bias, and therefore pose a different set of challenges:

- The BABE dataset [23] is a 3700 rows of sentence-level, expert annotated dataset containing binary bias label (biased or non-biased) along with an opinion explaining the factuality of the sentence (ranging from *entirely factual, somewhat factual, no agreement, expresses writer's opinion, etc*). While this is a good bias dataset, its main disadvantage is that it operates solely on the sentence level. Articles contains a huge number of sentences and one of the main problem within detecting on sentence-level is that multiple sentences within the same article could have opposite or different biases (I think there is a cite for this, I remember reading).

- The BASIL dataset [6] (300 articles) is the most commonly used bias dataset out there, containing 300 articles with both article-level and phrase-level annotations (contains one or more sentences. The obvious drawback of this dataset is the low amount of articles included as it is not nearly enough data to get a good working detection model, (find out more about this and cite?)

- The BAT [24] dataset (6345 rows, AdFontes labels) is the most suitable for article-level bias detection as it supplies both bias-score and reliability-score for the whole article. However, the dataset itself does not supply textual content of the articles and therefore need to be extended, something that I have been working on as well during my Master's Thesis this and last year.

- NLPCSS [5] (6964 rows), also annotated via AdFontes labels, containing article-level textual content with three bias labels (bias, neutral, or unknown). This one is also a good candidate for work in article-level granularity, however BAT dataset is richer in labels (scores) fluidity and additional metadata as it contains outlets information as well as twitter comments corresponding to each articles.

Another challenge can be correlated with the task of annotating media bias within a text. The traditional way is to hire experts and journalists to manually annotate and determine how biased the content is. As with [23], annotations are generally compiled and majority voted to achieve the final annotation given a particular text. It is important to use multiple annotators to minimise introducing another form of bias towards the dataset. Annotators' personal background moderately influenced their decisions and should be taken into consideration when building datasets, along with other factors such as topics, reading news habits, and honest mistakes [25]. Clearly, this is not a cheap procedure.

Alternatively, websites such as Allsides and AdFontes have their own experts and annotations to which can be crawled and make use of (cite cite). However, they operate solely within the U.S., coverting U.S. media sources. To my knowledge, there is no other organisation that do what they do outside the U.S. or in a more global scale, at the same level. Therefore, building a dataset that is not only accurate, but also global and diverse would have massive benefits.

3.2 Classification

Current State-of-the-Art in media bias detection tend to employ neural or transformer-based approach, by fine-tuning or exploiting Large Language Models (cite cite). Encoder-based LLMs such as BERT and RoBERTA can be finetuned, zero or few shot learning can be applied to decoder-based LLMs for a downstream task of detecting media bias. While these can work pretty well (cite), these models are general purpose models tailored for general NLP tasks and therefore might lack the necessary attributes to precisely detect bias, in addition to the expensive computational cost for inference.

MAGPIE [9] is the first large-scale, RoBERTa-based, multi-task learning (MTL) model dedicated for bias-related task, a promising approach for media bias detection and can be used to enhance the accuracy and efficiency of existing models. Using MAGPIE's context representation instead of BERT for media detection can potentially improve performance. However, currently the model only trained for sentence-level classification and also outputs binary result.

Many past works on automatic detection of media bias (cite cite) also typically operate on a sentence level and only output binary result (either biased or not) [28, 15, 14]. As media content are often delivered in the form of articles containing a number of paragraphs, detection at the article level is far more useful and desirable. More advanced methods could also include global-level spanning over multiple articles by implementing timeline, graphs, or other types of network to capture relationships among articles with same or similar topics.

Furthermore, The traditional approach of assigning only a single frame label to news articles remains overly simplistic, given that a standard news story often incorporates multiple viewpoints, arguments, or facets, each potentially carrying distinct connotations or framing [27]. An integer label would be a slightly better solution to represent bias from a text, although it is still hardly ideal.

Additionally, it would be good to consider explainability of a model when classifying a bias, as we would need to not just get the result, but to understand why contents are considered bias. Existing media bias detection systems typically concentrate solely on predicting the likelihood of a certain text being biased, offering limited insights into the underlying reason behind the decision.

3.3 Working with articles text

Article texts length generally falls between medium to long sequences, as they are not as long as other types of documents (legal documents, clinical studies, etc.) but not short enough to be classified as short. Most research with long sequence generally have thousands of tokens while most news articles stay between 500 - 2000 tokens (find cite?).

The most straightforward approach of a standard fine-tuning of a BERT model is not necessarily effective as the model is only able to process a maximum of 512 tokens. Several techniques have been attempted to address this issue such as the sliding window techniques, CLS techniques [26], where you exploit the 'CLS' token from a BERT embedding, and other more sophisticated models [13].

Additionally, there are other LLMs similar to BERT that are designed to handle longer sequences: Longformer (cite), BigBird (cite). However, most of these models have different architecture than the original BERT and therefore perform differently. Advanced models such as Longformer, ToBERT, and CogLTX do not consistently surpass the baseline models in classification performance, and only performed notably better than the baseline models on only two datasets [19]. Based on my recent work with the BAT dataset, applying these larger models did not provide superior results compared to simply fine-tuning BERT on the initial 512 tokens. It's important to

Nunc sed pede. Praesent vitae lectus. Praesent neque justo, vehicula eget, interdum id, facilisis et, nibh. Phasellus at purus et libero lacinia dictum. Fusce aliquet. Nulla eu ante placerat leo semper dictum. Mauris metus. Curabitur lobortis. Curabitur sollicitudin hendrerit nunc. Donec ultrices lacus id ipsum.

Chapter 4

Dataset Building

4.1 BAT dataset

4.1.1 Characteristics

The BAT dataset [24] is chosen for this project due to its article-level suitability and score labels. It contains 6345 rows of manually labeled news articles from 255 English-speaking news outlets (US-based), originally scraped from Ad Fontes Media's website along with their respective **political bias** and **reliability scores**. Articles in the dataset encompassed a wide arrange of topics such as COVID-19, politics, and lifestyle. The political bias score measures the extent of political influence, ranging from -42 (most extreme left) to +42 (most extreme right). The reliability score reflects the article's truthfulness, with values ranging from 0 (least reliable, containing inaccurate or fabricated information) to 64 (most reliable, original fact reporting).

Both political bias and reliability scores on each article were rated using defined metrics and multiple sub-factors, performed by three randomly selected analysts from Ad Fontes Media's team of over 60 experts. The corresponding three scores were then averaged, producing the final article scores. Moreover, each group consists analysts with different beliefs in the political spectrum i.e., left, center, and right.

4.1.2 Extension

The original BAT dataset only contains news titles and links (along with other metadata) and is missing the body content of articles. To overcome this, a Python script is written and executed, iteratively visiting each of the URL from the dataset and scrape the news content. This was not an easy task as each website has its own unique structures and formats. Furthermore, some of the scraped text contains noises that are almost impossible to remove through the script and requires manual intervention. The current extended dataset contains 5273 rows of articles, due to several unavailable websites and missing articles.

The scraped content are then pre-processed several times, sometimes repeatedly in an attempt to remove as much noise as possible and clean the text. At this step, a lot of scripts and methods are experimented with.

4.1.3 Analysis

Chapter 5

Proposed Methods

5.1 Features and baselines

The main focus of the features will be the textual content of articles. A good article-level bias classifier should be able to generalise solely or mainly from the content of the articles, capturing context of the article will be the key element of a good performance.

As a baseline, traditional methods such as Bag-of-Words and TF-IDF along with standard fine-tuning of BERT are implemented.

5.2 Evaluation

5.3 Proposed Method

5.4 Results

Chapter 6

Conclusion

Appendices

Appendix A

The First Appendix

Appendix B

The Last Appendix

Bibliography

[1] Home - media bias group. https://media-bias-research.org/. Accessed: 2024-05-29.

[2] V. P. Aires, J. Freire, F. G. Nakamura, A. S. da Silva, and E. F. Nakamura. An information theory approach to detect media bias in news websites. In *WISDOM 2020*, 2020.

[3] H. Allcott and M. Gentzkow. Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2):211–36, 2017.

[4] S. Boudana. A definition of journalistic objectivity as a performance. *Media, Culture & Society*, 33(3):385–398, 2011.

[5] W.-F. Chen, K. Al Khatib, H. Wachsmuth, and B. Stein. Analyzing political bias and unfairness in news articles at different levels of granularity. In D. Bamman, D. Hovy, D. Jurgens, B. O'Connor, and S. Volkova, editors, *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science*, pages 149–154, Online, Nov. 2020. Association for Computational Linguistics.

[6] L. Fan, M. White, E. Sharma, R. Su, P. K. Choubey, R. Huang, and L. Wang. In plain sight: Media bias through the lens of factual reporting. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2019.

[7] K. Foundation. American views 2020: Trust, media and democracy. Technical report, Gallup and Knight Foundation, 2020.

[8] J. Gottfried and J. Liedke. Partisan divides in media trust widen, driven by a decline among republicans. Technical report, Pew Research Center, 2021. [Online; accessed 29-April-2024].

[9] T. Horych, M. Wessel, J. P. Wahle, T. Ruas, J. Waßmuth, A. Greiner-Petter, A. Aizawa, B. Gipp, and T. Spinde. Magpie: Multi-task media-bias analysis generalization for pre-trained identification of expressions, 2024.

[10] S. jeong Lim, A. Jatowt, and M. Yoshikawa. Understanding characteristics of biased sentences in news articles. In *CIKM Workshops*, 2018.

[11] Julie Mastrine. What is media bias? https://www.allsides.com/blog/what-media-bias, 2022. [Online; accessed 28-April-2024].

[12] S. Keeter, N. Hatley, A. Lau, and C. Kennedy. What 2020's election poll errors tell us about the accuracy of issue polling. Technical report, Pew Research Center, 2021. [Online; accessed 30-April-2024].

[13] V. Kulkarni, J. Ye, S. Skiena, and W. Y. Wang. Multi-view models for political ideology detection of news articles. In E. Riloff, D. Chiang, J. Hockenmaier, and J. Tsujii, editors, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3518–3527, Brussels, Belgium, Oct.-Nov. 2018. Association for Computational Linguistics.

[14] Y. Lei and R. Huang. Sentence-level media bias analysis with event relation graph, 2024.

[15] I. Maab, E. Marrese-Taylor, and Y. Matsuo. Target-aware contextual political bias detection in news, 2023.

[16] N. Newman, R. Fletcher, K. Eddy, C. T. Robertson, and R. K. Nielsen. Reuters institute digital news report 2023. Technical report, Reuters Institute for the Study of Journalism, 2023. [Online; accessed 30-April-2024].

[17] N. Newman, R. Fletcher, A. Schulz, S. Andı, C. T. Robertson, and R. K. Nielsen. Reuters institute digital news report 2023. Technical report, Reuters Institute for the Study of Journalism, 2021. [Online; accessed 29-April-2024].

[18] C. Panagopoulos. Polls and elections accuracy and bias in the 2020 u.s. general election polls. *Presidential Studies Quarterly*, 51(1):214–227, 2021.

[19] H. H. Park, Y. Vyas, and K. Shah. Efficient classification of long documents using transformers, 2022.

[20] T. E. Patterson and W. Donsbach. News decisions: Journalists as partisan actors. *Political Communication*, 13(4):455–468, 1996.

[21] P. Rafail and J. D. McCarthy. Making the tea party republican: Media bias and framing in newspapers and cable news. *Social Currents*, 5(5):421–437, 2018.

[22] T. Spinde, S. Hinterreiter, F. Haak, T. Ruas, H. Giese, N. Meuschke, and B. Gipp. The media bias taxonomy: A systematic literature review on the forms and automated detection of media bias, 2024.

[23] T. Spinde, M. Plank, J.-D. Krieger, T. Ruas, B. Gipp, and A. Aizawa. Neural media bias detection using distant supervision with BABE - bias annotations by experts. In M.-F. Moens, X. Huang, L. Specia, and S. W.-t. Yih, editors, *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 1166–1177, Punta Cana, Dominican Republic, Nov. 2021. Association for Computational Linguistics.

[24] T. Spinde, E. Richter, M. Wessel, J. Kulshrestha, and K. Donnay. What do twitter comments tell about news article bias? assessing the impact of news article bias on its perception on twitter. *Online Social Networks and Media*, 37-38:100264, 2023.

[25] T. Spinde, L. Rudnitckaia, J. Mitrović, F. Hamborg, M. Granitzer, B. Gipp, and K. Donnay. Automated identification of bias inducing words in news articles using linguistic and context-oriented features. *Information Processing & Management*, 58(3):102505, 2021.

[26] X. Su, T. Miller, X. Ding, M. Afshar, and D. Dligach. Classifying long clinical documents with pre-trained transformers, 2021.

[27] G. Vallejo, T. Baldwin, and L. Frermann. Connecting the dots in news analysis: A cross-disciplinary survey of media bias and framing, 2023.

[28] E. van den Berg and K. Markert. Context in informational bias detection. In D. Scott, N. Bel, and C. Zong, editors, *Proceedings of the 28th International Conference on Computational Linguistics*, pages 6315–6326, Barcelona, Spain (Online), Dec. 2020. International Committee on Computational Linguistics.