



INFORMATICS
INSTITUTE OF
TECHNOLOGY

INFORMATICS INSTITUTE OF TECHNOLOGY

In collaboration with

UNIVERSITY OF WESTMINSTER, LONDON

**NEUTRALYZE:
Automated Media Bias Detection
and Neutralization in News Articles**

A dissertation by

Chandira Gunatilleke

W1789932 | 2019593

Supervised by

Ms. Krishnakripa Jayakumar

May 2023

This document is submitted in partial fulfilment of the requirements for the
BSc (Hons) in Computer Science degree at the University of Westminster.

ABSTRACT

Media bias is a pervasive problem that can significantly impact public discourse and decision-making processes. With the growth of digital media, the potential for bias to spread quickly and widely has become even more pronounced leading to harmful consequences, such as social polarization, misinformation, and distrust in institutions. Manual methods for detecting and neutralizing bias are time-consuming and expensive. Though significant progress has been made in the field of automated media bias detection, the majority of the work has been focused on sentence level classification. Automated bias neutralization is a relatively new domain with limited research contribution. As a result, based on the limitations identified in existing research in related domains, this work proposes a token level classification task that operates at the level of individual words, allowing for more granular and accurate analysis of media bias. The research also contributes to the limited body of work on text neutralization, by utilizing a sequence-to-sequence models for conditional generation of neutralized texts. This report details the approach, design, implementation, and evaluation for the research conducted.

Keywords: Media Bias Detection, Text Neutralization, Token Classification, Natural Language Processing, Seq2Seq, BERT, T5, BART

Subject Descriptors:

- Computing Methodologies → Artificial Intelligence → Natural Language Processing → Natural Language Generation
- Information Systems → Information Retrieval → Retrieval Tasks and Goals → Clustering and Classification
- Computing Methodologies → Machine Learning → Machine Learning approaches → Neural Networks