**Initial Plan: “Utilising NLP Models and Visualisations to Identify Bias Within News”**

*Author:* Jose Oscar Dos Santos Nunes

*Supervisor:* Dr. Fernando Alva Manchego

# Project Description

In today’s digital world, news consumers and bombarded with information from a wide variety of mediums and outlets. Each telling their unique view of a story. This project seeks to investigate tools that can be used for the consumer to understand both the bias that the information they are reading holds and how the bias is being represented too them.

Due to the huge growth of online news, it has become ever increasingly difficult to discern objecting and accurate reporting over information being heavily impacted by the authors and stakeholders involved within news. Bias in the news is often demonstrated in subtle forms such as through specific word choice, framing of people and ideas within the articles and through outlets covering/omitting specific aspects of events. Being able to detect and understand the effect that these techniques typically prove to be very challenging, especially without devoting large quantities of time and resources. Understanding the nuances and contexts that contribute to bias is vital for the ongoing development of bias detection methods.

As mentioned before in the digital world of today the news is readily accessible through a range of mediums to an ever-growing audience. With the growth of news accounts on TikTok and other social media it has become a larger issue as many younger people are being exposed to potentially heavily biased information without knowing any better. Uncovering bias ensures a better informed and more united public. Furthermore, existing methods for detecting bias, such as manual analysis are very labour-intensive and thus problematic to scale with the growth in news availability. An automated approach to this problem utilising NLP and visualisations could provide a more efficient and objective way to identify and present bias.

I plan to tackle this project in a series of steps. I will begin with creating a large dataset of news articles both from pre-existing datasets found online and through creating custom datasets using APIs such as “NEWSAPI” if/when it is needed. Once this dataset is completed I plan to develop an initial model of which will analyse the articles stored in my dataset and score it for various types of bias (e.g. political, racial, gender). This model will leverage sentiment analysis, entity recognition and word frequency analysis. In conjunction with this I am going to implement a series of visualisations both as an aid to quickly understand the findings from my models but also as a tool for identifying flaws to remove in further iterations of my models.

I feel that this is an intelligent approach to this project as it follows in line with similar projects of which have dealt with ideas such as predicting the winners of game shows by analysing tweets. Through adaptions to the models and approaches used in these areas to news, I can potentially demonstrate the bias within articles.

Upon completion of the project, I anticipate delivering a robust NLP model trained to identify bias in news reporting. This model will be tested for accuracy and reliability against a benchmark dataset containing a range of news articles with a range of different biases. Additionally, I will produce a detailed report showing my findings at different stages, utilising visualisations to take the abstract idea of bias into a readable and more accessible set of results.

## Problem Tree

One useful exercise that you can do before writing this section is to create a Problem Tree. The goal of this exercise is to brainstorm the **subproblems** that are involved in the problem, their **causes,** and their **consequences**. This can help you determine the context of the problem, and the desired outcomes. It can also help with establishing the scope of your project. For example, it can be the case that due to the time constraints, you will focus on a subproblem of the bigger problem you had initially thought about.

Here are some useful links in case you are interested in performing this exercise:

* <https://mspguide.org/2022/03/18/problem-tree/>
* <https://www.ingenioempresa.com/en/problem-tree/>
* <https://sswm.info/taxonomy/term/2647/problem-tree-analysis>

## Citing Sources

You should justify everything that you are stating in the document. Some of these justifications will come from research papers, articles, surveys, etc. that you have not produced yourself. Therefore, you have to cite them adequately and include their full references in the final section of the document. This is especially important when providing background information, or explaining why you have chosen to use certain resources/tools/methods.

As for the citation style, I personally prefer to use one where the last names of the author(s) appear in text, such as [APA](https://www.mendeley.com/guides/apa-citation-guide/) or [Cardiff Harvard](https://xerte.cardiff.ac.uk/play_4069#harvard). However, you can decide to use your own favourite one. What is important, though, is that you remain consistent with the same style in all your document.

Microsoft Word has a nice [functionality](https://support.microsoft.com/en-us/office/create-a-bibliography-citations-and-references-17686589-4824-4940-9c69-342c289fa2a5) that can help you with adding citations and creating a References section. For example: (Tunstall, Werra, & Wolf, 2022). You do not need to use this functionality if you do not want to, but I thought it could be useful to let you know about it.

# Aims and Objectives

This project aims to develop an efficient and affection model that can be used to detect and illustrate bias within news articles. Leveraging the capabilities of Natural Language Processing and data visualisation, the system will provide an objective measure of bias, enabling readers to get a more accurate and bias free view of an article. The goal is to promote media literacy and contribute to the public knowledge on bias and misinformation/misleading information within the news outlets.

One of the alternatives aims and hopeful outcomes of this project is to provide a stepping on point for others looking to further their understanding of the field. One issue that is often found with the specialised areas of computer science is that they can be very inaccessible to those without a prior knowledge in the field. So, I hope that through this project I can provide insightful findings but also provide my findings and methods used in a way that is both engaging to someone with a prior understanding alongside someone whose just curious about the topic and my work.

## Objectives

1. Data collection and preparing:
   1. Utilising either a premade dataset or API. I will extract data to be used within my project in its various use cases. (Testing, training, etc..)
   2. Once data sources are finalised I will clean the dataset of an unnecessary details and perform any needed updates and additions
   3. Finally, the cleaned and prepared data will be stored within a database for later use.
2. Initial model
   1. Plan and outline how my model will work.
   2. Train the model on my training dataset
   3. Calculate my baseline of performance metrics that will be used to test further iterations.
   4. Within this objective I will experiment with visualisation methods to determine the best fitting options for my project
3. Improved model
   1. Using the data collected from my initial models’ visualisations and performance metrics I will adapt and make changed to the model to further increase its capability in speed, accuracy, and scope.
4. Final evaluations
   1. Taking my improved model, I will revaluate its performance using the same training data alongside potentially testing with differently structured data.
   2. Compare the accuracy of my model against the views of real people when looking at articles.
5. Analysis
   1. Once all visualisations and models are finalised an extensive analysis will be performed on how they two models are related and how the differences have led to more accurate results.
   2. I will attempt to pinpoint the specific changes of which have made key differences in the performance as well as looking at specifically which changes have made the biggest differences to my metrics.

# Feasibility

While this project does have some potential ethical and legal issues I have taken several steps to mitigate/remove any issues.

Data collection is an area in which there is typically many hurdles to deal with . For example, for this project I could potentially utilise web scraping to collect my articles and their surrounding data. While there are several websites in which web scraping is allowed for research purposes or in general it can be very difficult to differentiate cases where it is and isn’t okay. To deal with this, I have decided to utilise premade data sets from websites such as “[Kaggle](https://www.kaggle.com/datasets)” or other data websites. Within these websites the permissions are outlines by the creators on each of the dataset’s pages. The other option is to use an API such as “NEWSAPI”, these APIs are set up for a wide range of uses and within the documentation all the licences are available along with an outline of what the APIs can be used for.

Secondly as previously mentioned I have a plan to use a survey or similar method as part of my evaluation of my final model. As is standard with many projects any data/feedback coming straight from people typically requires ethical approval. In the case of my project the interaction with human participants will be minimal and is a potential plan so at this time there will be no steps taken to deal with any ethical issues however should the time arise approval will be sought from the university before perusing any further on the matter.

Aside form the typical ethical and legal issues within the project there is also the need to look at the more technical limitations that could pose a threat to the completion of the project.

Deepening on the progress made within the models they can begin to require large computation power and storage that isn’t readily available to myself as a student. While in the case of my project I feel that the issues here are somewhat mimical there are several options to dealing with this such as google colab, utilising the universities super computers or looking externally and reaching out to businesses.

Lastly one issue with feasibility comes from the timeframe and the expertise in the necessary fields. While this is very typical for a final project at the university level it is once concern that could produce a sizable bottleneck in my progress.

# Work Plan

Below is a table that will define the rough structure of how I plan to allocate the time for the project. While I am using a relatively simple structure for this plan I will also be utilising Microsoft to-do and clockify. Microsoft to-do is going to be used to ensure that each week all my necessary tasks are clearly outlined and to ensure that everything that needs to get finishes are completed. Alongside this I will be using clockify as a time tracker, as is mentioned in the previous section one of the more sizable risks to this project is the delegation of how much time is spent on each of my tasks. Clockify solves this by giving a dashboard of details showing time how much time has been spent on individual sections to ensure certain aspects are not being prioritised and thus other key sections are falling behind.

|  |  |  |
| --- | --- | --- |
| Week | Tasks | deliverables |
| Week 1 | * Plan and write the initial report. | * First draft of the initial plan |
| Week 2 | * Meeting to discuss the plan draft. * Plan and begin writing the first 3 chapters of the final report |  |
| Week 3 | * First draft of the first 3 chapters * Flesh out and finalise the dataset structure and content |  |
| Week 4 | * Research and plan initial model. * Begin development of model |  |
| Week 5 | * Finalise initial model. * Evaluate and create visualisations of model’s performance and stats |  |
| Week 6 | * Write up report regarding the initial model | * Give in first draft of introductory chapters and initial model report |
| Week 7 | * Being work on the improved model. * Make any necessary changes to the report in line with the suggestions |  |
| Week 8 | * Continue working on the improved model * Begin evaluation and analysis of new model * (potentially test against humans for a new metric) |  |
| Week 9 | * Complete evaluation and analysis sections of the report | * Draft containing both models and beginning sections handed in |
| Week 10 | * Conclusions and future work to be completed * Reflection on the project planned and started |  |
| Week 11 | * All sections completed and submitted for a final review | * Final draft for last notes and suggestions |
| Week 12 | * Final tweaks and changes in accordance with draft feedback | * Final report submission |

# References

Tunstall, L., Werra, L. v., & Wolf, T. (2022). *Natural Language Processing with Transformers, Revised Edition.* O'Reilly Media, Inc.