CM3203 Initial Plan

# Using NLP to Identify Propagandist Techniques in Articles

## Project Description

Natural Language Processing (NLP) is a subdomain of computer science concerned with how computers can process large amounts of natural language data and then analyse this data to produce meaningful information. Machine learning is a process through which a program can be constructed to automatically learn principles from training data and then apply these learned principles to new data to make judgements about this new data. This project aims to utilise NLP techniques and machine learning such that journalistic articles can be accurately identified as containing propagandist techniques. In order to achieve this goal, the articles in question will be assessed against the 18 propagandist techniques as defined by the Propaganda Analysis Project[[1]](#endnote-1) (for example; presenting irrelevant data, name calling or labelling, misrepresentation of someone’s position etc.).

The end product of this project will be a program that employees supervised machine learning techniques to perform both binary and multi-class classification. The basis for the classification will be the 18 propagandist techniques mentioned above. The program is expected to read several plain-text files where each file represents a separate article. The program will be expected to only handle data in this format i.e. it will not be expected to perform any pre-processing on the data. All classification tasks must be performed automatically with manual predictions forbidden. The program must produce text files indicating the results of the above two classification tasks.

To complete this project three data sets will be prepared; a training data set that will be used to train the model, a development data set that will be used to fine tune the model’s hyper-parameters, and finally a test data set that will be used to evaluate the performance of the program. The performance of the program will be assessed according to an F1 score[[2]](#footnote-1) computed from the produced results against the expected results for the test data.

This project is based upon *Shared Task on Fine-Grained Propaganda Detection @NLP4IF 2019[[3]](#endnote-2)* conducted by G. Da San Martino, A. Barron-Cedeno, and P. Nakov as part of the Propagandist Analysis Project.

## Project Aims and Objectives

This project has two primary objectives and one potential extension depending on the progress made towards both primary objectives.

The **first objective** of this project is to produce a program that reads several plain-text files and output a new text file that indicates whether the input files made use of any of the 18 propagandist techniques. After reading and analysing the data, the program is expected to perform binary sequence tagging on the data. The program should output a text file that contains the article identifier, and the start and end characters of where a propagandist technique has been detected. The number of lines in the file indicates the number of fragments detected. If two techniques overlap in the same space within the article, then they are merged into one fragment.

The **second objective** of this project is to expand upon the initial implementation described above, such that, given a text fragment has been identified as utilising propagandist techniques, the program then identifies the applied technique used in the fragment. In addition to performing binary sequence tagging, the program must also perform multiclass sequent tagging. The program should output a second text file that contains the article identifier, the technique used, and the start and end characters of where said technique has been used. For this second classification task, overlapping techniques in a text fragment are not merged.

The success of the program in both classification tasks will be determined by its F1 score when ran against the test data. Implicit in both the primary and the secondary objectives is the aim of maximising the respective F1 scores.

Assuming classification of both tasks can be performed with a satisfactorily maximised F1, the extension of the above program will be to produce a *propagandist score* based on a number of factors including weighting of the 18 propagandist techniques, frequency of propagandist fragments etc. The base assumption for this score is that of the 18 propagandist techniques, the techniques are not equal indicators that the author of an article is employing propaganda. To decide upon the weighting of each technique, consenting participants will be asked to complete a survey (*survey 1*) evaluating these techniques. The survey results will be used to decide upon the exact algorithm used to determine the *propagandist score* of an article. Once this algorithm has been implemented as part of the program, participants will then be asked to complete a second survey (*survey 2*) to judge the scores assigned to a select sample of test data articles so that the results can be assessed.

## Work Plan

|  |  |  |
| --- | --- | --- |
| **Period** | **Objectives** | **Milestones** |
| Week 1  *27th Jan – 2nd Feb* | * Begin researching potential approaches. * Supervisor meeting to discuss initial plan. | * **Submission of initial plan.** |
| Week 2  *3rd Feb – 9th Feb* | * Conclude research on potential approaches. * Supervisor meeting to agree approach. | * Document discussing the pros and cons of researched techniques. |
| Week 3  *10th Feb – 16th Feb* | * Collect and prepare training, development, and test data. * Begin implementing the program. * Supervisor meeting to review progress. | * Training, development, and test data collected and prepared. |
| Week 4  *17th Feb – 23rd Feb* | * Implement binary classifier. * Supervisor meeting to assess binary classifier implementation. | * The program is able to successfully perform binary classification. |
| Week 5  *24th Feb – 1st Mar* | * Tune binary classification model to maximise F1 score. * Begin implementing multi-class classifier. * Supervisor meeting to review progress. | * **The program is able to successfully perform binary classification with a sufficiently maximised F1 score.** |
| Week 6  *2nd Mar – 8th Mar* | * Complete implementation of multi-class classifier (*Objective 2*). * Begin writing report discussing the implementation of the binary classifier. * Supervisor meeting to assess multi-class classifier implementation. | * The program is able to successfully perform multi-class classification. |
| Week 7  *9th Mar – 15th Mar* | * Tune multi-class classification model to maximise F1 score. * Continue writing report. * Supervisor meeting to decide on whether to pursue extension goal based on progress to date. * Draft *survey 1.* | * **The program is able to successfully perform multi-class classification with a sufficiently maximised F1 score.** |
| Week 8  *16th Mar – 22nd Mar* | * Continue tuning models if necessary/desired. * Continue writing report moving onto discussing the implementation of the multi-class classifier. * Supervisor meeting to review progress. * *Produce survey 1 and begin engaging participants.* | * Report section on binary classifier implementation drafted. |
| Week 9  *23rd Mar – 29th Mar* | * Continue tuning models if necessary/desired. * Continue writing report. * Supervisor meeting to review progress. * *Conclude engagement with participants for survey 1.* | * Engaged at least 3 participants for survey 1. |
| Week 10  *30th Mar – 5th Apr*  *(Easter recess)* | * Continue tuning models if necessary/desired. * Continue writing report. * *Implement scoring system* | * Report section on multiclass classifier implementation drafted. * Scoring system implemented. |
| Week 11  *6th Apr – 12th Apr*  *(Easter recess)* | * Continue tuning models if necessary/desired. * Continue writing report. * *Produce example article for survey 2* | * Report introduction and approach sections drafted. |
| Week 12  *13th Apr – 19th Apr*  *(Easter recess)* | * Continue tuning models if necessary/desired. * Continue writing report. * *Produce survey 2 and begin engaging participants for survey 2.* | * Any continuing work on the program to stop. |
| Week 13  *20th Apr – 26th Apr* | * Continue writing report. * Supervisor meeting to review and assess program. * *Engage participants for survey 2.* | * Engaged at least 3 participants for survey 2. |
| Week 14  *27th Apr – 3rd May* | * Continue writing report moving onto discussing survey results and future work. * Supervisor meeting to review final report. | * **Report contents finalised.** |
| Week 15  *4th May – 7th May* | * Proofing of final report. * Supervisor meeting to review final submission. | * **Submission of project.** |

Key

* Core Objective/Milestone
* Important supervisor meeting
* *Extension-related optional objective/milestone*
* **Important milestone**

1. <https://propaganda.qcri.org/annotations/definitions.html> [↑](#endnote-ref-1)
2. <https://en.wikipedia.org/wiki/F1_score> [↑](#footnote-ref-1)
3. <https://propaganda.qcri.org/nlp4if-shared-task/> [↑](#endnote-ref-2)