Laidlaw Summer Report

Billy Ly

September 2023

It is with great pride and gratitude that I reflect fondly on my Summer 1 experience with the Laidlaw Programme; having had the unforgettable privilege of conducting a self-directed research project under a distinguished academic. The science behind information, data, and decision-making had always piqued my interest ever since I first learned about cognitive biases, and it was the only topic I knew I wanted to do any sort of research project on.
I had no clear ambition of going into academia in the future, and thus had no smattering of knowledge of what it was like to work in an academic setting, other than through the various anecdotes of the previous cohort. From their wildly varying descriptions, I had formed the expectation of long, hard, gruelling hours spent grinding in a lab or library, with the majority of time being spent alone, staring at a computer screen.

To start off, I would like to say that my research experience, in retrospect, did not look like this at all (for the most part).
Firstly, through my supervisor's backing, I had the immense fortune of working in the ADAPT research centre on campus; with my own desk and surrounded by other academics. ADAPT is a world-leading SFI research centre for AI-driven digital content technology, which means this closely aligned with my project's themes and meant I was surrounded by researchers working in very similar fields to my project's. This was absolutely one of the key highlights of

my entire research experience, as this granted me the opportunity to meet, discuss with, and learn from various talented scientists and PhD students. In addition to this, I even had the privilege of attending the annual ADAPT conference and annual Project VIGILANT dinner; both of which organisations my supervisor had major leadership roles in. VIGILANT is a major European collaboration project between researchers and police to combat disinformation, and thus was right in my area of interest. Both events were extremely interesting and enriching experiences, since I not only got an inside view to a lot of the exciting developments happening in the field, but also encountered many different types of leaders too. It was a genuinely fantastic opportunity to network with distinguished researchers in the field, and it truly made me consider what qualities I could incorporate into my own personal style of leadership.

Secondly, whereas most Scholars may have had one or two supervisors, it often felt like I had much more than that because of how collaborative and helpful the ADAPT environment was. There were several ADAPT researchers I had encountered during my tenure, who had kindly spared time to offer advice, explain confusing concepts, and even review my code in times of frustration. It was here that I recognised just how prominent the role of team-working and collaboration was in any research, and just how quickly problems could get ironed out through open dialogue and discussion.

I truly cannot thank my supervisor, Prof. Owen Conlan, and his colleagues enough for making my first experience with academia such an enjoyable one; as they have been nothing but kind, patient and helpful throughout the entirety of my six weeks.

In brief, my research project, titled 'An Investigation into the Features of Disinformation that may lead to it being Spread Online', aimed to identify some of the key aspects of disinformation which lead to its circulation on the web. I used a pre-annotated multimodal dataset, containing 2,593 COVID-19 related news articles and 24,184 related tweets collected between February 2020-May 2021, and conducted various text analyses to identify the key features such as sentiment, readability and topicality. Every tweet in the dataset referenced one of the news articles, and each was manually labelled as either 'True', 'False' or 'Inconclusive' depending on the reliability of the referenced article and the stance of the tweet. For example, if a tweet referenced an unreliable news article and supported its content, that tweet would be labelled as 'False'. If the tweet was instead refuting the unreliable article, it would then be labelled as 'True'.

Tweet Stance
(through Stance Detection Process)

News Reliability
(Media Bias/Fact Check)
+
(Media Bias Chart Checking)

	Support	Refute	Not Enough Information
Reliable	T	F	I
Unreliable	F	T	I

Tweet Labelling Process

T = True
F = False
I = Inconclusive

Figure 1: Tweet Labelling Process

Once I had gathered all the analyses results and stored them in one combined dataset, I split the data into four variations:

1: Tweet Tokens (Control)
2: Tweet Tokens + Tweet Sentiment
3: Tweet Tokens + Article Tokens + Article Reliability Label
4: Tweet Tokens + Tweet Sentiment + Article Tokens + Article Reliability Label

After some deliberation and discussions with my supervisor, it was decided to not use a lot of the analysis results, such as readability and topical features, to avoid overfitting the data and obtaining inaccurate results.

This brings me to the machine learning section of the project. The four machine learning algorithms I used for the project were Naïve Bayes (NB), Support Vector Machines (SVM), K-Nearest Neighbour (KNN) and Extreme Gradient Boosting (XGB). The importance isn't on how they exactly differ, but on the fact that they all perform relatively well at text classification tasks (which this project can be classified as).

In very simple terms, the goal was to train machine learning algorithms on these four different variations of data, and make them predict based on the data given whether a tweet was true or false. We would then compare the results to see how they improve/deteriorate under different variations and algorithms. For example, if the KNN algorithm took in the second data variation, it means it will try predict whether a tweet is true or false based only on the content of the tweet (its tokens) and its overall sentiment.

Each variation of data also underwent two types of vectorisation, which simply means it was prepared for the machine learning algorithm using two different methods; TF-IDF and Word2Vec. The preparation of data also involved other steps, such as stemming and lemmatising all of the text, but in layman's terms we were simply cleaning up and formatting the data so as to make it readable to a machine learning model.

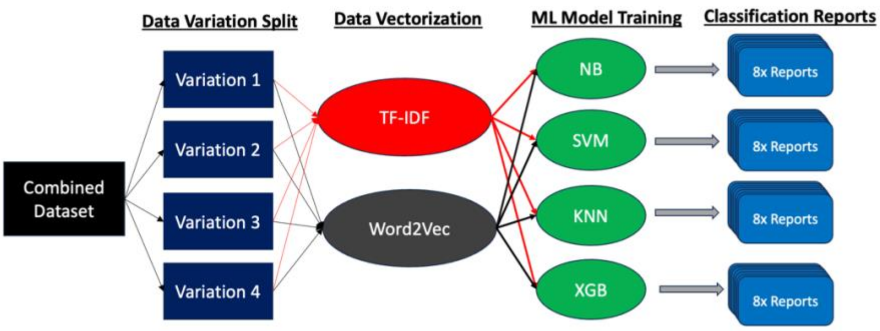

Figure 2: Methodology Visualization

The final findings of the project can be found here:

Machine Learning Model

Dataset Variation	NB	SVM	KNN	XGB
1	F1: 0.829, 0.897 Acc: 0.871	F1: 0.905, 0.933 Acc: 0.922	F1: 0.827, 0.885 Acc: 0.862	F1: 0.857, 0.895 Acc: 0.879
2	F1: 0.819, 0.893 Acc: 0.865	F1: 0.903, 0.933 Acc: 0.921	F1: 0.827, 0.885 Acc: 0.862	F1: 0.858, 0.895 Acc: 0.879
3	F1: 0.813, 0.891 Acc: 0.862	F1: 0.931, 0.951 Acc: 0.942	F1: 0.892, 0.925 Acc: 0.911	F1: 0.949, 0.964 Acc: 0.958
4	F1: 0.898, 0.929 Acc: 0.916	F1: 0.933, 0.952 Acc: 0.944	F1: 0.899, 0.929 Acc: 0.917	F1: 0.948, 0.964 Acc: 0.957

Legend
(SVM,2) = SVM Model trained only on tweet tokens and tweet sentiment score

F1: 0.903, 0.933
Average F1 score of ***unreliable*** tweet label = 0.903
Average F1 score of ***reliable*** tweet label = 0.933

Acc: 0.921
Average accuracy of model across folds = 0.921

Variations
1: Tweet Tokens
2: Tweet Tokens + Tweet Sentiment
3: Tweet Tokens + Article Tokens + Article Reliability Label
4: Tweet Tokens + Tweet Sentiment + Article Tokens + Article Reliability Label

Figure 3: TF-IDF Vectorisation Results

Machine Learning Model

Dataset Variation	NB	SVM	KNN	XGB
1	-	F1: 0.154, 0.751 Acc: 0.615	F1: 0.758, 0.848 Acc: 0.814	F1: 0.831, 0.887 Acc: 0.865
2	-	F1: 0.253, 0.746 Acc: 0.622	F1: 0.752, 0.845 Acc: 0.809	F1: 0.833, 0.889 Acc: 0.867
3	-	F1: 0.741, 0.770 Acc: 0.757	F1: 0.910, 0.936 Acc: 0.925	F1: 0.941, 0.959 Acc: 0.951
4	-	F1: 0.741, 0.770 Acc: 0.757	F1: 0.883, 0.918 Acc: 0.903	F1: 0.945, 0.961 Acc: 0.955

Legend
(XGB,2) = XGB Model trained only on tweet tokens and tweet sentiment score

F1: 0.833, 0.889
Average F1 score of ***unreliable*** tweet label = 0.833
Average F1 score of ***reliable*** tweet label = 0.889

Acc: 0.867
Average accuracy of model across folds = 0.921

Variations
1: Tweet Tokens
2: Tweet Tokens + Tweet Sentiment
3: Tweet Tokens + Article Tokens + Article Reliability Label
4: Tweet Tokens + Tweet Sentiment + Article Tokens + Article Reliability Label

Figure 4: Word2Vec Vectorisation Results

Across almost all models and variations, the addition of extra information, such as article context and text sentiment, seemed to have contributed to an improvement in the algorithm's accuracy in distinguishing between reliable and unreliable information. This certainly warrants further investigation, where future research could use different datasets, models and even features, such as readability, source information and topical category.

This research project in its entirety, has taught me three core lessons about research and leadership.

The first, is the importance of specificity. I genuinely cannot over-emphasise, just how different the research I outlined in my initial Laidlaw proposal is, to the research that I ended up doing. Aside from the first two weeks of literature review and dataset searching, which pretty much went as planned, the rest of the works looks borderline unrecognisable from what I said I would do in the proposal. To give an idea, the term 'machine learning', which essentially made up the majority of my research, is nowhere to be found in my initial proposal. And in hindsight, this was to be expected as everything I had outlined in the beginning was much too vague to be used as an actual research plan. Had I hunkered down and formed a much more detailed plan from the start, a lot of mistakes and wasted time I spent doing pointless work could have been easily avoided. Specificity, and being precise with your method, saves time, effort and brainpower.

This brings me to my second insight though, which is that it is perfectly fine to pivot and change directions; and I believe this applies to leadership settings too. Both my supervisor and I agree that in the end, the research journey I embarked on was far superior to the journey I said I would embark on in the proposal. I learned far more than I expected, especially in the emerging fields of AI and machine learning which had always piqued my interest. I developed my skills not only as a researcher, but also as an analyst, a programmer and a leader.

It was extremely tough making decisions to abandon previous plans, even when entire days had already been spent on executing them. For example, when conducting the various text analyses on the dataset's tweets, I had actually collected a lot more data than I actually used. Data which took hours to extract, code and format. Thus, one could imagine my frustration when being told the risk of overfitting was too great to include and use it all. It was the correct decision, but a tough and difficult one to make. One could attribute the reason for this to the sunk cost fallacy, which I admit I had definitely fallen into over the course of the six weeks.

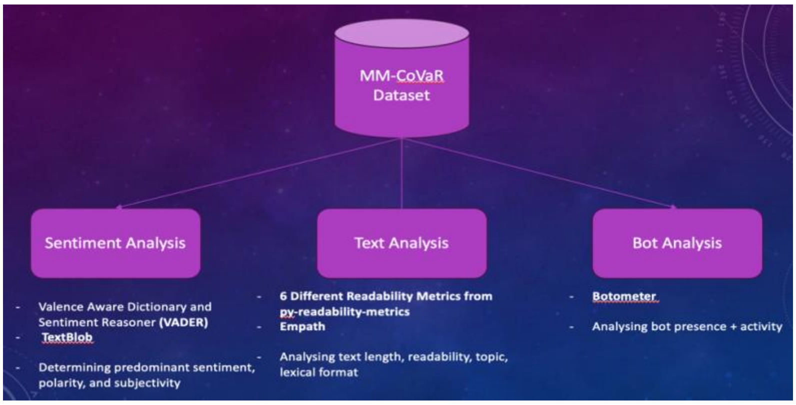

Figure 5: A slide from the initial research plan, detailing what analyses I planned to execute

With everything said and done though, I can now look back on everything and truly appreciate the courage it took to recognise when a pivot was necessary, and acting upon it. There is no doubt in my mind that if I hadn't changed the direction of my research so many times during the six weeks, then I would have come out with far more underwhelming findings and experiences. This lesson will stick with me not only as a lesson gleaned from research, but also as an insight gained from a leadership role.

This brings me now to my final and greatest lesson from this research project; the importance of absolute resilience. For every point of success that happened during my research journey, I can recount twice as many challenges and frustrations. From extremely unlucky problems with Twitter's API that only occurred because of a certain new CEO taking charge, to my models producing rubbish after literal hours of running, I had become a regular at the pits of despair over the course of the six weeks. This isn't even mentioning all the code, (which at the start only God and I knew how worked, and now only God knows) that took days on end to write and fix.

All of this, takes resilience to conquer. In my initial Laidlaw proposal, I mentioned that 'it is only during the toughest and most challenging environments, that a person can truly grow'. I have learnt as a researcher and a leader, how important it is to be resilient in the face of challenge. Not only for the goal of reaching the final destination, but also just for one's own development as a person.

Additionally, while it is good to bask in the achievements of the project, it is just as important to reflect on the failures and challenges too, and appreciate the resilience and determination it took to overcome those challenges. All of what I have learnt and completed during this project, has kept closely in line with most of my PDP goals too. For example, one of the initial goals in my PDP was to develop emotional maturity and resilience. My research experience has not only helped with this, but has also further reinforced it as a core leadership value.

In summary, my Summer 1 of the Laidlaw programme has helped me achieve many things. From completing my first ever research project in academia, to meeting with many exceptional individuals. But most importantly, it has taught me the importance of specificity, resilience, and flexibility with direction. I am extremely thankful for the programme for helping me develop my leadership style and character thus far; and look forward to using as well as building upon these new skills in my future endeavours.

The annual ADAPT scientific conference

My final day in the ADAPT centre