Detecting Fake News via Machine Learning-
A Literature Review and Comparison of Existing Algorithms

With the extensive influence of the Internet and the amount of oversharing on social media platforms, it is becoming increasingly harder for users to distinguish fake from legitimate news. Users rely heavily on social media as a primary source of news around and beyond their immediate surroundings. With almost 3.4 billion active users worldwide [1], the gradual shift from traditional to social media stems from the convenience and inexpensive cost of obtaining news. News comes in many different forms of media, such as video recordings, photos, and articles, which are all easily doctored nowadays.

Albeit not being a new phenomenon, a great number of people were brought to an awareness of fake news during the period of the 2016 US presidential election amongst the masses of fake news and clickbait about the two candidates. Not only an epidemic in America, it is everywhere within the Internet's grasp. Just last year in India, a viral video of a child being kidnapped, later proven fabricated, was widespread throughout WhatsApp, instigating attacks on strangers, eventually leading to multiple instances of homicide. Online news disseminates quickly and efficiently, this case showing just how dangerous such a combination could be. Perpetrators exploit fake news manipulating people's stances on issues ranging from political election biases to healthcare disinformation to harmful celebrity gossip.

Empirically, utilizing machine learning to solve the fake news detection problem is not something new but we aim to be able to distinguish fake news by running text-based content collected through our program. In this paper, our contributions are two folds. First, we did research and literature reviews on other people's works. Second, we built our own machine learning program to investigate whether machine learning could detect fake news. Our objective was to create a surface-level fake news detection with traditional machine learning implementation. To do this, we focused on the text-based content, training and testing data with specific classifiers and feature extraction models. We classified its fallacy and genuineness through a score returned from our program where the proposed work achieved an accuracy rate of 91.52%. To give us a better understanding on how our program detected fake news and how well our model worked on real-world data, we laid out several hypotheses based on related literature reviews and made a comparison of our experimental results with other people's related works. Drawing from our experiments we found that: i) results are more reliable when the program was fed with data from a certain domain and location; ii)

short input returned insubstantial results; iii) publishing date of an article affected the accuracy rate; and iv) dataset size mattered when calculating credibility.

The problem is complicated and needs more sophisticated approaches to completely eradicate it. To really solve the fake news problem, unsupervised learning and deep learning are the way to go in the long run because unbiased, labeled data is hard to find and requires an enormous amount of time and human resources, which is not efficient nor practical. Unlabeled data does result in lower accuracy and uncertainty in classification, but it is available all over on the Internet and is much easier to acquire than labeled data. Besides that, deep learning allows us to manipulate more text features and dig deeper into text context to provide a more meaningful result, thus, increasing the accuracy rate. The works we have done provide some useful insights into fake news and text classification problem. Fake news detection as an emerging research area has a lot of unfilled potential. We hope to explore more aspects of this area in the coming future.