Table of Contents

Declaration………………………………………………………………………………………………………………………

Summary………………………………………………………………………………………………………………………

Acknowledgments…………………………………………………………………………………………………………

Table of Contents…………………………………………………………………………………………………………………

Table of Figures……………………………………………………………………………………………………………………

Table of Tables……………………………………………………………………………………………………………………

1. Introduction…………………………………………………………………………………………………………………

1.1 Research Question………………………..………………………………………………………………..

1.2 Hypothesis………………………………..……………………………………………………………………

1.3 Overview of dissertation…………………………………………………………………………………

2. Existing Work…………………………………………………………………………………………………………………

3. Methodology……………………………………………………………………………………………………………………

4. Implementation………………………………………………………………………………………………………………

5. Data Analysis……………………………………………………………………………………………………………………

7. Results………………………………………………………………………………………………………………………

8 Conclusions………………………………………………………………………………………………………………………

1. Introduction

Fake news is a term which has been circulating through popular media and culture in recent years. Since the 2016 US presidential election the term has gained traction and has been used to criticize all forms of media, from large television news outlets all the way down to social media accounts. What Initially started as a sound bite for a presidential candidate the term has grown a life of its own and sparked mass debate on the credibility of the news outlets which we had considered to be credible and non-biased for decades before.

The terms traction didn’t stop there however, and even went onto uncover a dark truth about the social media platforms we use today. A truth which exposed a series of vulnerabilities in these platforms which allegedly allowed hacking organizations and alleged Russian Government agencies to manipulate popular opinion using fake social media profiles and accounts. This coupled with news that political consulting firms like Cambridge Analytica were able to develop digital profiles that represented individuals of certain political beliefs provided a scary outlook for the everyday social media user. Not only are there groups creating fake accounts and profiles to manipulate these applications but the data they themselves provided was being used to facilitate political advert targeting.

The fake news frenzy sparked a series of worldwide debates and even forced US Congress and Irish (EU \*?) Government bodies to summon social media companies before them and seriously start thinking about heavier regulation for companies as such. That being said these social media platforms are still being used daily by millions of Users and the same vulnerabilities that were present before arguably still exist today.

Determining the legitimacy of a social media account through language style is a problem which falls out of the imperative programming capabilities. In the last decade\*\* companies have been turning toward Machine Learning solutions to problems. The advancements and research into Machine Learning and AI is leading to trivial human cognitive tasks being replaced. A potential solution to the issue of “fake news” social media may be possible through a machine learning model\*\*.

An area where rumors and fake news is not new is in the Premier League football transfer windows. Twice a year there are transfer window periods where clubs can buy and sell players amongst one another. This leads to a lot transfer rumor speculation amongst the press, specifically on twitter. There are numerous accounts set up with the sole purpose of reporting on this, claiming to be the some of the first individuals “in the know”.

This area gives an ideal sample set of a fake news in the media. By taking a specific transfer period in the past one can look at public tweeted rumors and use a ground truth to label whether this rumor actually became true. Twitter provides a medium to access to hundreds of thousands of rumors and true claims, all of which can be fact checked through official records of confirmed transfers\*\*. This gives a sample set to develop a model capable of determining the veracity of a transfer rumor or a social media account.

* 1. Research Question

The purpose of this research project is to answer the following question:

*“To what extent can an ML model be defined to determine the accuracy of a Twitter rumor or Twitter account in predicting football transfers?”*

From this the following research objectives were created:

* A rumor Twitter API query generation method
* A Natural Language Processing (NLP) technique for Name Entity Recognition (NER).
* Model development using various ML techniques and labelled Twitter data.
  1. Hypothesis

This research is aimed to determine whether the accuracy of a twitter football transfer rumor can be determined by developing a model using supervised machine learning techniques. The research is aimed to find evidence that there is a supervised Machine Learning technique which can be used to train a model capable of predicting rumor accuracy.

*(H0) The research finds no method of developing a supervised machine learning technique capable of producing a model with the aforementioned capabilities.*

*(H1) The research finds a method supervised machine learning method of producing a model with the aforementioned capabilities.*

1.3 Overview of this Dissertation