Abstract

Title: RUMOUR CLASSIFICATION IN FOOTBALL TRANSFER WINDOW TWITTER DATA
Name: Mark Murtagh
Supervisor: Seamus Lawless
Degree Title: Master in Computer Science

Fake news is a topic which has been circulating through mainstream media for a number of years now. Fundamental flaws in social media ranking algorithms are being exploited. Football transfer speculation news is an area in which this problem has been present for a long time. Sports media outlets and individuals have predicted player transfers claiming to be "in the know". This research is aimed to determine to what extent supervised machine learning approaches could be used in predicting the accuracy of a tweet or Twitter account in relation to a football transfer rumour. The research project involved three parts: data gathering; data labelling; classification experiments.

The research details the steps involved in data collection, labelling the data and performing the classification experiments. Two distinct approaches were taken during the classification experiments. One classification approach using a simple multi-layer perceptron model showed promising evaluation metrics when run on unseen data. Another approach using a Separable Convolution Neural Network showed no capability of learning the features of the training data. The problems and causes of overfitting with each approach are also discussed.

Potential issues with the training set collected were considered. Mainly the concerns with potential biasing in the methods used for data collection. Furthermore, the steps taken to accurately labelling the training data are detailed.

The findings of this research add to the extensive body of research in the area of fake news and football transfer markets. Possible sites for future work which builds on the findings are proposed.