Informatics Institute of Technology

In Collaboration With

University of Westminster, UK

Rumor and False News Detection System for Social Media

A dissertation by

Mohamed Nizar Arfaq Hamool

Supervised by

Mr. Saman Hettiarachchi

Submitted in partial fulfillment of the requirements for the

BEng (Hons) Software Engineering Degree

Department of Computing

May 2020

Abstract

The ideology behind the rumor is a story or a statement which holds both importance and ambiguity in general without any evidence or confirmation to support them. The rumors and false information on social media is corrupting our society and those rumors create extremely dreadful situations and lead to harmful consequences. In order to avoid these types of misinformation in the social media in this research project we are going to build a rumor detection system as above mentioned which will have the capabilities to detect and classify social posts as rumors and non-rumors by reading the social media post texts. Generally, the most rumors get spread or get created when a crisis event occurs so for this research will be focusing on specifically on the rumors spreading during and related to the crisis events and then will be detecting and classifying rumors and non-rumors.

In the area of rumor detection the most commonly used technologies are Natural Language Processing and Machine Learning. Other than these technologies there are few state-of-the-art systems that use Deep Learning and provide exceptional performance. Even though there are many systems available for rumors detection yet none of them provide the expected accuracy or the performance and all these systems have various limitations and drawbacks which causes the imperfect accuracy and poor performance.

To overcome this accuracy and performance issues in the rumor detection system in this research the proposed solution is to use the Natural Language Processing with Transfer Learning. This Transfer Learning technology is relatively very new to Natural Language Processing and Machine Learning. The main purpose of using Transfer Learning in this solution is to use pre-trained language models which are available, and these pre-trained models can provide better understanding and performance in the natural language processing tasks. This rumor detection system will be inputted with the rumor and non-rumor data sets and the system will use this to train and fine-tune the transformer language model and the system will use the trained transformer language model to predict the veracity of the social media posts and will return the predicted results. This system will provide better accuracy and performance than the other available rumor detection systems.

Keywords: Natural Language Processing, Transfer Learning, Pre-trained language models, Classification task head, Transformer language model