A COVID-19 Fake News Detection System

Muvazima Mansoor, Master of Science in Computer Science

University of Dublin, Trinity College, 2022

Supervisor: Khurshid Ahmad

The pandemic bought along with it an infodemic which is the overabundance of information, truthful or not, that is often spread over social media. The spread of misinformation regarding COVID-19 has dire consequences hampering the efforts of health systems worldwide. Social media has made it easy to share personal opinions which are neither entirely truthful nor fake and can be termed as 'imaginative'. This work presents a fake news detection system which can differentiate between informative, imaginative, and fake news. Each type of news is written concerning a target audience. The differences in the style of writing of three types of news are analysed at multiple levels of linguistic description such as – lexical, syntactic, semantic, and pragmatic. Finally, content-based features are leveraged to construct a fake news detection system utilising a Perceptron model that can classify informative, imaginative, and fake news with an F1 score of 0.74, 0.8 and 0.95, respectively. It performs significantly better than the baseline Naïve Bayes model that gave an F1 score of 0.64, 0.58 and 0.35 for informative, imaginative, and fake news, respectively.