

MSc Project Report

Automatic Sarcasm Detection by leveraging
Conversational Context

Mohomed Shafeek Muhammed Shariq

2018

A report submitted as part of the requirements for the degree of
MSc in Big Data Analytics at Robert Gordon University, Aberdeen,
Scotland

Abstract

The dawn of Web 2.0 has seen a massive growth in user generated content in the web. Almost all web sites and applications today provide some form of interaction with its users, allowing them to share their experience, opinions and experience. Furthermore, the popularity of social media has provided a platform for users to voice their opinions freely. The growth of content shared by the users paved the path to the study of sentiment analysis and opinion mining. While there are many approaches being studied in the field of sentiment analysis, one of the main challenges in identifying sentiment accurately is the presence of sarcasm in social media and online forums. Automatic sarcasm detection is paramount for improving the outcomes of sentiment analysis. The popularity of deep learning in NLP has prompted researchers to investigate the use of deep learning for sarcasm detection. Recent studies have investigated the ability to leverage contextual information in predicting sarcasm.

This research strives to leverage the conversational context to detect sarcasm in a particular domain. Topics such as politics have proved to elicit sarcasm more than some other topic, hence this research primarily focus on predicting sarcasm in discussions on politics. DBPedia Spotlight was used to identify statements related to politics. Many different deep learning architectures were researched on to detect sarcasm by leveraging information from the conversation context. The study shows that it was able to obtain very good results using this approach. Moreover, Transfer Learning methods were researched on to predict sarcasm in conversations related to sports. The study shows the Transfer Learning can be used effectively to train models with comparatively smaller training data set and still receive good results.

Key words:

Deep Learning, Artificial Neural Networks, Automatic Sarcasm Detection, Sentiment Analysis, Transfer Learning