

**UNIVERSITY OF
WESTMINSTER**

Informatics Institute of Technology

in collaboration with
The University of Westminster (UK)

6COSC012C Final Year Project 2021/22
Thesis

Sarcastically

Sarcasm detection using a Multi-Modal Approach
via Text and Audio based Sentiment Analysis

Written By:
Ryan Joshua Kuruppu
2018351

Supervised by:
Mr. Banu Athuraliya

Submitted in fulfillment of the requirements for the BSc (Hons) Computer
Science Course, Department of Computing

October 2021

Abstract

Sarcasm is one of the most well-known but easily misunderstood parts of the English language. While sarcasm is understood well to a vast majority, the uniqueness of sarcasm is that no one thinks about it the same way, when it is understood by the majority, it does its job but when it isn't is when questions start to arise in terms of meaning, understanding and so on. Most research on sarcasm has come in the form of a textual understanding of it using context from previous statements or blocks of statements and more recent research has led to the discovery of supplemental forms of associative data such as video and images which help to understand a machine's understanding of sarcasm. Sound plays just as much a part as other social and visual cues which indicate sarcasm and will be the focus of the current research project.

Sarcastically attempts to make use of this area of sarcasm to better classify sarcasm by making use of vocal cues accompanied by basic text sarcasm detection to better evaluate the results of machines understanding sarcasm. It will make use of a multi-model sarcasm algorithm which contains not only a text evaluator but also an audio and hybrid evaluation model that will be cross-compared and extracted to get the highest possible result. For this research, the core content has been based on acted-sarcasm for the purpose of generating generic sarcasm rather than advanced sarcasm which might take considerably longer time to prepare for. As such, Sarcastically's job is to evaluate the benefits or the losses of using audio in accompaniment with text to justify that audio is more or less important to understanding sarcasm as other visual or social cues.

Keywords: Multi-Modality, Sarcasm Detection, Natural Language Processing, Audio Processing, Ensemble Learning, Shared Representation