INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

Sp-BERTsum: Speech Summarization Using Prosodic Features Based on BERT

A Project Specification Design and Prototype

by

Mr Sivasamy Dhanush

Supervised by

Mrs Fathima Farhath

Submitted in partial fulfilment of the requirements for the BEng in Software Engineering degree at the University of Westminster.

05/2022

Abstract

Speech summarization is the process of taking in human speech and producing a shorter version of the speech containing the relevant information. Speech Summarization have been extensively researched in various domains like Broadcast News, Meetings and many more. The popular approaches towards summarization are extractive summarization and abstractive summarization. In this research the author will be presenting a novel method of extractive summarization using the newer transformer BERT model and prosodic features of speech.

Transformer neural network models in the recent have shown great promises in the recent in a various natural language processing tasks like semantic analysis, text summarization and more. But research in utilizing these models for speech summarization have been a few. The system that does utilize transformer model only employ the textual features but many researches conducted previously have proven that prosodic features that are present in the speech also aid the summarization process. Hence the aim of this research is to utilize the acoustic features along with the textual features to be able produce summaries for speech data. These results will be evaluated using rouge scores in comparison to other existing speech summarization systems.

Keywords: Speech Summarization; Extractive Summarization; Prosodic Features; Bi-directional Encoder for Transformers; Binary Classification.