Group Project Proposal - Group 3
Andy Christian, Rajkumar Conjeevaram Mohan, Daqian Dang
NLP
DATS 6312
30th March 2023

Text Summarization of News Articles for Daily/Topic News App

Introduction

In many industries, things can change fast. Staying on-top of recent news, trends, and even rumors can be critical for industry decision makers to make well informed choices. News articles and technical reports are critical sources of this information. While the details in these articles and reports are important for implementation, the higher level key points are more relevant for decision makers in understanding the larger picture. An NLP model capable of summarizing the main points of news articles and technical reports, as well as creating a meta-summary of all the relevant information from the article or reports examined from a given day/topic would be a useful tool in helping decision makers stay aware of their industry's situation.

Objectives

1. Develop an NLP model capable of summarizing articles/reports.
2. Find articles/reports on a given topic of interest and produce summaries.
3. Create a meta-summary of the articles/reports from the given topic.

Methodology

Phase 1: Data Source and Preprocessing
● Use a dataset **CNN-DailyMail News Text Summarization** from Kaggle.
● Preprocess the data through tokenization, stemming, lemmatization, word frequency, etc. technique analysis to identify key topics and calculate relevant statistics.

Phase 2: Model Development
● Use LSTM network along with techniques such as Seq2seq and next word prediction to create summary of articles

Phase 3: Future Development
● Create app that lets user input desired articles/reports and produces the summaries and meta summaries, along with highlights of key data

Conclusion

An app that creates summaries of relevant news articles or technical reports of a given topic, as well as as a meta summary of those summaries would be useful in helping industry decision

makers stay well informed across a broad range of developments in high-pace, ever changing operational environments where staying on top of such developments can be critical for making good decisions.