Informatics Institute of Technology

In Collaboration With

University of Westminster, UK



*University of Westminster, Coat of Arms*

Abstractive Text Summarization for Movie Reviews Using Optimized Transformers

A Project Proposal by

Mr. Nazhim Kalam

w1761265 / 2019281

Supervised by

Mr. -------------------

September 2022

This Project Proposal is submitted in partial fulfilment of the requirements for

the BSc (Hons) Computer Science degree at

the University of Westminster.

**1.1 Introduction**

In this study, the author tries to improve text summarization for movie reviews using customized transforms which is a deep learning approach, which can also be used in any other related domains related to text summarization, such as hotels reviews, ecommerce product reviews, transportation service reviews etc... therefore this on the other hand would seem like a general implementation for all common use cases. The proposed architecture will improve the quality of text summarization to its best than how normal machine learning algorithms perform with.

The problem, the research gap, the research challenge, and the research approach that the author plans to use over the coming several months are all outlined in this document. Additionally, a review of prior research interests and the essential evidence of the issue is done. Finally, in the work plan, the expected schedule of the project’s deliverables is presented.

**1.2 Problem Domain**

**1.2.1 Movie Reviews**

A growing number of websites, like Amazon and the Internet Movie Database (IMBD), a website for movie reviews, allow users to publish reviews for the things they are interested in, in line with the growth of Web 2.0, which places an emphasis on user interaction. *(Khan et al., 2020)*

Online movie reviews are evolving into an important information source for Internet users as the amount of information on the web increases. *(Khan et al., 2020)* Online users, however, publish tens of thousands of movie evaluations every day, making it challenging for them to manually summarize the reviews. The mining and summarizing of movie reviews is one of the difficult issues in natural language processing. *(Khan et al., 2020).*

Text summary aids users or corporate executives in decision-making by gathering and evaluating a significant collection of online reviews *(Alsaqer and Sasi, 2017)*.

These days, most of us check the movie ratings before choosing or watching a movie on any platform, such as Netflix or Amazon Prime, but we also come across mixed movie reviews that can be either positive or negative. This is where the issue arises because we will feel the need to read through all the reviews to determine how and what the movie is specifically about and whether it meets our needs. A summary of the countless reviews a film has gotten can make it easier for viewers (or customers) to quickly peruse the summary and decide about whether or not to watch the film. The summaries of movie reviews, on the other hand, can help streaming services like Netflix quickly discover the viewing habits or preferences of their users. *(Khan et al., 2020)*.

**1.2.2 Text Summarization**

Given that there is a lot of textual information available nowadays, including news articles and reviews, text summarizing allows us to rapidly discover the important points of the entire piece by reducing the amount of text *(**Mahajan, Vast, Mhaske and Barahate, 2020).*

Extractive summarization and abstractive summarization are typically the two methods of text summarization. In the case of extractive summarization, the context or article's most valuable sentences are taken out without changing them in any way. On the other hand, abstractive summarization tries to develop the sentences on its own and produce the summary; this is much better than extractive summarization, since it is more meaningful to build our own phrases inside the context than to use chosen sentences from the context without any alteration *(Etemad, Abidi and Chhabra, 2021)*

**1.2.3 Transformers**

Transformers in NLP is a novel architecture that aims to solve sequence-to-sequence tasks while handling long range dependencies with ease. It has surpassed competing neural models like CNN (Convolutional Neural Nets) and RNN (Recurrent Neural Nets) in terms of performance to appear as the dominant architecture for natural language processing *(Wolf et al., 2020).*

Transformers uses self-attention mechanism to target on selected areas of the input sentence followed by the encoder and decoder architecture *(Etemad, Abidi and Chhabra, 2021).*

* 1. **Problem Definition**

In the domain of movie review summarization, currently there are no in-depth researches done using deep learning approaches to solve this problem, mostly standard machine learning algorithms such as Naïve Bayes have been used, and that’s where the room for the usage of deep learning approaches comes into picture in order to enhance the quality/accuracy of the text summarization process for this domain area.

Although deep learning models take longer to train, they offer more accuracy because they can simultaneously automate feature extraction and classification, whereas machine learning algorithms require feature selection first. Therefore, applying deep learning techniques will help to improve the quality of text summarization and help the user in making better decisions.

* + 1. **Problem Statement**

**References**

Khan, A., Gul, M.A., Zareei, M., Biswal, R.R., Zeb, A., Naeem, M., Saeed, Y. and Salim, N. (2020). Movie Review Summarization Using Supervised Learning and Graph-Based Ranking Algorithm. *Computational Intelligence and Neuroscience*, 2020, pp.1–14. doi:10.1155/2020/7526580.

Alsaqer, A.F. and Sasi, S. (2017). Movie review summarization and sentiment analysis using rapidminer. *2017 International Conference on Networks & Advances in Computational Technologies (NetACT)*. [online] doi:10.1109/NETACT.2017.8076790.

Mahajan, Vast, Mhaske and Barahate (2020) ‌Text Summarization using Deep Learning. *International Journal of Recent Technology and Engineering*, 9(1), pp.2663–2667. doi:10.35940/ijrte.a3056.059120.

‌Etemad, A.G., Abidi, A. and Chhabra, M. (2021). A Review on Abstractive Text Summarization Using Deep Learning. *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*. [online] doi:10.1109/icrito51393.2021.9596500.

‌Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S. and Drame, M. (2020). Transformers: State-of-the-Art Natural Language Processing. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. doi:10.18653/v1/2020.emnlp-demos.6.

‌