FRI Natural language processing course 2023
Project report

An Automatic Movie Summary Generator

Luka Pavićević, Andrija Stanišić, Stefanela Stevanović

Abstract

Advisor: Slavko Žitnik

Introduction

In today's world, there is an abundance of films being produced every year. It can be difficult for audiences to keep up with all the latest releases. Movie fans often resort to reading summaries or synopses to get a sense of what a particular film is about before deciding whether or not to watch it. However, it can be a time-consuming and difficult task to manually summarise a long film script.

Here Natural Language Processing comes in. Through the use of Natural Language Processing techniques, we can automate the task of summarising film scripts, saving both time and effort.

In this project, our aim is to create a system that can accurately summarise film scripts into concise and informative synopses through the use of Natural Language Processing. By doing so, we hope to provide film enthusiasts with a quick and efficient way to get a sense of a film's plot, characters and key events. Ultimately, this will help them in their decision-making process when choosing which films to watch.

Related works

One of the related works in text summarization is the work by Aleš Žagar and Marko Robnik-Šikonja [1], who explored the use of various summarization approaches, including neural models, to produce short summaries from larger texts. In their work, they addressed the problem of selecting the most appropriate summarization model for a given text, and proposed a solution that uses a neural metamodel to automate the selection process based on the input document's representation. This work presents an innovative approach to text summarization, particularly in addressing the issue of model selection, which is important in ensuring the quality of generated summaries.

Another related work [2] describes a system for automatically generating movie descriptions using subtitles and other metadata. The authors use a combination of NLP techniques such as named entity recognition, sentiment analysis, and topic modeling to extract relevant information from the subtitles and generate concise descriptions.

In addition, another work [3] present a movie summarization system that uses subtitles and user ratings to generate short descriptions of movies. The authors use a combination of sentence clustering, topic modeling, and sentiment analysis to extract key information from the subtitles.

Furthermore, a recent work [4] presents a neural language model for generating movie summaries from subtitles. The authors use a combination of convolutional and recurrent neural networks with a latent variable model to generate concise and coherent summaries.

Lastly, a recent work [5] proposes a system for generating movie synopses using subtitles and plot summaries. The authors use a combination of named entity recognition, sentiment analysis, and topic modeling to extract relevant information and generate summaries that capture the essence of the movie.

References

[1] A. Žagar and M. Robnik-Šikonja. Slovenian text summarization models. 2022.

[2] Sandeep Subramanian, C. Lawrence Zitnick, and Irfan Essa. Automatic generation of movie descriptions. 2017.

[3] Yen-Yu Huang, Hsin-Hsi Chang, and Liang-Chih Lin. Movie summarization based on subtitles and ratings. 2020.

[4] Sihao Chen, Zhaochun Ren, Shu Yang, and Tao Mei. Generating movie summaries with latent variable conditional neural language models. 2019.

[5] Akash Jhanwar, Qingyu Liu, Siliang Tang, and Tat-Seng Chua. Movie synopsis generation using subtitles and plot summaries. 2020.