Sentiment Analysis project - Positive_or_Negative

Installation

In the folder “Positive_or_Negative” are present a few files:

- ➢ ReadMe.pdf (this file)
- ➢ Positive_or_Negative.ipynb (the actual notebook)
- ➢ [TMNLP 2022-2023] Positive_or_Negative report.pdf (the report for the project)
- ➢ IMDB_dataset.csv (the dataset used, that can also be found at this link)
- ➢ new_reviews.txt (a text file containing 20 new reviews, then analyzed with the best model)

To visualize and use this project you must upload the folder into Google Drive in the My Drive folder and then run the notebook in the preferred workspace.

Project Overview

This project aims to classify IMDB movie reviews into positive or negative sentiments using various deep learning models. By leveraging natural language processing (NLP) techniques, we preprocess textual data and apply models such as LSTM, GRU, Conv1D, SimpleRNN, and Transformers to understand and predict sentiments expressed in movie reviews.

Dataset

The dataset used is the IMDB movie reviews dataset, which consists of 50,000 reviews split evenly into 25,000 training and 25,000 testing sets. Each set has an equal number of positive and negative labels.

Models

- Long Short-Term Memory (LSTM)
- Gated Recurrent Unit (GRU)
- Convolutional Neural Network (Conv1D)
- Simple Recurrent Neural Network (SimpleRNN)
- Transformer

Results

The Conv1D model demonstrated the highest validation accuracy and was thus selected for making predictions on new reviews.