

# IMDb Sentiment Analysis Project Report

## Introduction

I utilized PyTorch and the BERT model for sentiment analysis on IMDb movie reviews for my IMDb Sentiment Analysis project. The major goal was to forecast whether a review would be good or negative based on the substance of the article.

## Methodology

### Model Architecture

#### Encoder-Decoder Architecture:

Bert primarily contains an encoder and a decoder stack where the data is passed through both the stacks but the encoder stack captures the intrinsic relationship among the data passed.

#### Bidirectional Processing:

The model processes data in both directions at the same time whereas previous models only parsed data in a singular direction.

#### Attention Mechanism:

The attention mechanism is the key part of the transformer model as once data is passed in the form of tokens the attention mechanism can determine the impact factor of the token with respect to the passed other tokens.

#### Choice of Data :

I used the IMDb movie review dataset to train and test the model from Kaggle, which contained more than 50 thousand annotated reviews.

#### Challenges faced under implementation :

Due to the large data size, I ran into overfitting scenarios and further optimization of hyperparameters increased model performance and reduced overfitting.

#### Results of Model Evaluation :

The model performed with an f1 score of 93 %.