Instructions:

The goal of this assignment is to build a text classification model using the Hugging Face library to classify a dataset of text into one of multiple categories. The candidate will use a pre-trained model such as BERT or GPT-2 as a starting point and fine-tune it on the classification task and the dataset used is sentiment\_train that is related to the entertainment news articles.

Preprocessing:

Preprocess the text data by cleaning it, removing stopwords, punctuations and other irrelevant characters.

The Architecture of the model used is BertArchitecture.

Fine-tuning:

**\***Froze the pre-trained BERT model's architecture.

**\***Added new layers to the BERT model

**\***Considered the class weights while deifing the loss function, stratified data while splitting into train, test and validation to handle the **\***imbalance in the dataset.

Used BERT tokenizer to process texts before fine-tuning.

Evaluation Metrics:

Train the model on the dataset and evaluate the performance using metrics such as accuracy, precision, recall and F1-score.

Result Obtained:

