**Fine-Tuning BERT for Sentiment Classification on Yelp Polarity Dataset**

**(1Problem Statement**

The goal of this project is to build a sentiment classifier capable of determining whether a given review is positive or negative using a Transformer-based model (BERT). The task is a text classification problem based on the Yelp Polarity dataset.

**(2 Dataset and Preprocessing**

* **Dataset:** [Yelp Polarity dataset from Hugging Face Datasets](https://huggingface.co/datasets/fancyzhx/yelp_polarity)
* **Data Source:** Hugging Face Hub (in Parquet format)
* **Preprocessing steps:**
  + Removed HTML tags, numbers, and extra spaces (optional step)
  + Tokenized the text using bert-base-uncased tokenizer
  + Converted to Hugging Face Dataset class
  + Added attention masks and labels for the model
  + Split data into training and testing sets

**(3 Model Selection Rationale**

We selected **DistilBERT (distilbert-base-uncased)** because:

* It’s a distilled version of BERT, pre-trained on large English corpora.
* Provides lighter, faster training and inference while retaining about 95-97% of BERT's performance.
* Well-suited for text classification tasks.
* Compatible with Hugging Face’s Trainer API for efficient fine-tuning.

(4 **Implementation Details**

* Fine-tuned distilbert-base-uncased using Hugging Face Transformers library.
* Evaluation metrics: **Accuracy**, **F1-Score**, **Precision**, **Recall**
* Techniques applied:
  + **Mixed precision training (fp16=True)** for faster training and memory efficiency
* Visualized:
  + Training loss curve
  + Validation accuracy curve
* Deployed final model using **Gradio live demo**

**(5 Results and Analysis**

* **Validation Accuracy:** 96.7%
* **F1-Score:** 96.7%
* **Precision:** 96.7%
* **Recall:** 96.7%
* **Observations:**
  + Grid search improved model performance
  + Gradio demo successfully predicts sentiment in real-time.