**Sentiment Analysis Using BERT**

**1. Project Overview**

The goal of this project is to classify movie reviews as **positive** or **negative** using state-of-the-art NLP techniques. We used **BERT (Bidirectional Encoder Representations from Transformers)** and **DistilBERT**, which are transformer-based models pre-trained on large text corpora.

**2. Dataset**

* **Source:** IMDB Movie Reviews (datasets library)
* **Split:**
  + Training: 25,000 reviews
  + Test: 25,000 reviews
* **Labels:**
  + 0 → Negative
  + 1 → Positive
* **Preprocessing:**
  + Tokenization using DistilBertTokenizerFast
  + Truncation/padding to a max length of 256 tokens
  + Smaller subsets used for faster training during testing (optional)

**3. Model Architecture**

* **Model:** DistilBertForSequenceClassification (distilled version of BERT)
* **Layers:** Transformer layers + classification head
* **Number of Labels:** 2 (negative, positive)
* **Optimizer:** AdamW (learning rate = 5e-5)
* **Device:** GPU (if available) or CPU

**4. Training**

* **Batch Size:** 16 (train), 32 (test)
* **Epochs:** 2 (for quick experimentation)
* **Loss:** Cross-Entropy Loss (inbuilt in DistilBertForSequenceClassification)

**Training Loop:**

1. Load batches of tokenized input (input\_ids, attention\_mask) and labels
2. Forward pass through the model
3. Compute loss
4. Backpropagate gradients
5. Update weights with optimizer.step()
6. Repeat for all epochs

**5. Evaluation Metrics**

* **Accuracy:** Correct predictions / Total predictions
* **Precision:** TP / (TP + FP)
* **Recall:** TP / (TP + FN)
* **F1-Score:** Harmonic mean of precision and recall

**Sample Output:**

Test Accuracy: 0.878

Classification Report:

precision recall f1-score support

negative 0.89 0.86 0.88 254

positive 0.86 0.89 0.88 246

accuracy 0.88 500

macro avg 0.88 0.88 0.88 500

weighted avg 0.88 0.88 0.88 500

* The model achieves **~88% accuracy** on a subset of the test data.
* Precision, recall, and F1-score are balanced across both classes.

**6. Sample Predictions**

| **Review** | **Predicted Sentiment** |
| --- | --- |
| I loved this movie! | positive |
| It was the worst film ever. | negative |
| The acting was okay, but story was boring. | negative |

* Custom predictions work for any input text after tokenization.

**7. Key Learnings**

* Transformers like BERT/DistilBERT can capture **contextual information** in text for classification.
* Using DistilBERT reduces training time while retaining similar performance.
* Small subsets can be used for **experimentation**, but full datasets improve accuracy.
* Proper preprocessing (tokenization, padding, truncation) is crucial.
* Evaluation requires ensuring that all classes are present in the test subset to avoid errors in classification\_report.

**8. Challenges**

* Running full BERT/DistilBERT models can be slow on CPU. GPU acceleration is recommended.
* Model may predict only one class if test set is small or unbalanced.
* classification\_report fails if the predicted labels don’t cover all classes—must use labels=[0,1].

**9. Conclusion**

* The sentiment analysis model achieves **~88% accuracy** on IMDB reviews using DistilBERT.
* Predictions align well with intuitive sentiment.
* This framework can be extended to **multi-class sentiment analysis**, **other datasets**, or **real-time predictions** in applications like chatbots or review analysis systems.

Submitted by: Dharshini D

Date: 25/08/25