

這段程式碼展示了如何使用`DistilBERT`模型來處理IMDb電影評論的情感分析任務。主要內容包括：

1. **資料下載與處理**：

- 下載IMDb電影評論數據集，並將其從gzip壓縮檔解壓縮為CSV格式。
- 將數據集拆分為訓練集、驗證集和測試集。

2. **數據預處理**：

- 使用`DistilBERT`的tokenizer對文本數據進行編碼，將文本轉換為模型可以處理的格式。

3. **模型構建與訓練**：

- 定義了一個自定義的PyTorch數據集類`IMDbDataset`來處理編碼後的數據。
- 下載並加載預訓練的`DistilBERT`模型，並使用Adam優化器進行訓練。
- 實現了一個手動的訓練循環，計算訓練和驗證的準確率，並記錄訓練時間。

4. **使用`Trainer` API 進行訓練**：

- 使用`transformers`庫中的`Trainer` API來簡化模型訓練過程，這比手動訓練循環更為方便。
- 使用`Trainer` API進行模型訓練和評估，並使用`datasets`庫中的`load\_metric`來計算準確率。

5. **模型評估**：

- 訓練完成後，計算測試集的準確率以評估模型性能。

這段程式碼展示了從數據處理、模型訓練到評估的完整過程，並比較了手動訓練和使用`Trainer` API 的效果。