NLP Assignment 2 – UFC Comments Sentiment Analysis Report

Submitted by: - Sharad Verma

1. System Architecture

Data Flow & Preprocessing

The sentiment analysis pipeline begins with loading the dataset consisting of labeled text data classified into three categories: *Negative, Positive, and Neutral*. The preprocessing steps include:

1. Text Cleaning: Removal of punctuation, HTML tags, URLs, and extra whitespace.
2. Tokenization: Sentences are tokenized using a pretrained tokenizer from transformers.
3. Encoding: Texts are converted into input IDs and attention masks.
4. Splitting: Data is split into training and testing sets using an 80:20 ratio.

Model Architecture

The model is built on bert-base-uncased, a transformer-based language model from Hugging Face’s Transformers library. It is fine-tuned for multiclass classification with the following structure:

1. BERT Encoder
2. Dropout Layer
3. Fully Connected Linear Layer (output = 3 classes)
4. SoftMax Activation

This architecture supports multi-class classification and is optimized using the AdamW optimizer.

2. Model Implementation

Model Selection

The bert-base-uncased model was selected for its strong performance on downstream NLP tasks like sentiment classification.

Training Configuration

Epochs: 5

Batch Size: 16

Loss Function: CrossEntropyLoss

Optimizer: AdamW

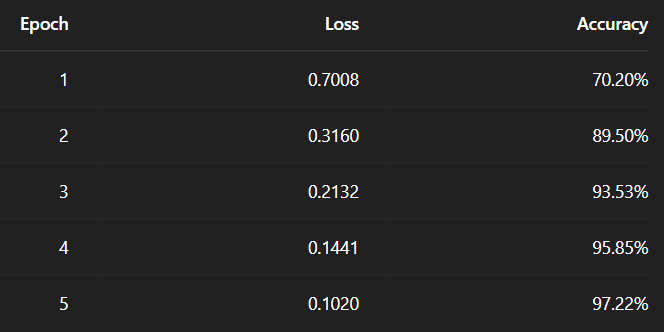
Learning Rate Scheduler: Applied to improve convergence

Mixed Precision: Implemented using PyTorch’s autocast () to accelerate training and reduce memory usage.

Hyperparameter Tuning

The model was trained with gradually improving learning accuracy without overfitting, due to proper dropout and batch size. No grid/random search was explicitly mentioned, but the chosen parameters led to efficient training.

3. Experimental Result



Training Progress

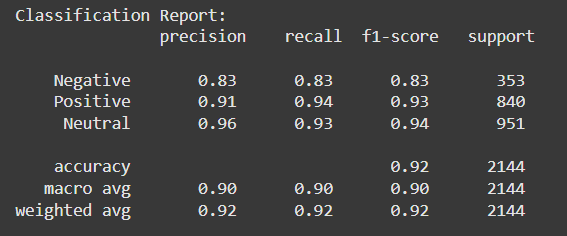
The training accuracy increased steadily while the loss decreased, reflecting effective learning and model convergence. The model shows consistent and rapid improvement across epochs, with accuracy rising from 70.2% to 97.22%. Loss consistently decreases, indicating that the model is learning well without overfitting during training.

Evaluation Metrics

Test Accuracy: 91.79%

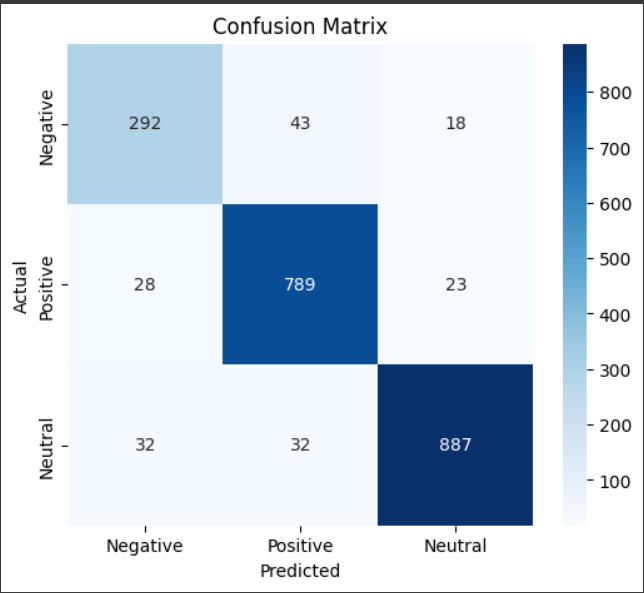
R² Score: 0.7133

The R² score further validates the model's ability to generalize well to unseen data.

 Classification Report

The model performs best in classifying Neutral sentiments, followed by Positive, with slightly lower performance for Negative sentiments.

Confusion Matrix Analysis



Most misclassifications happen between *Negative* and *Positive*, and between *Positive* and *Neutral*. The model confuses fewer Neutral cases, suggesting strong representation learning for Neutral sentiment.

F1-score=92%

Conclusion

The BERT-based model demonstrates excellent performance on sentiment classification, especially considering the limited number of epochs. With a final test accuracy of 91.79% and an F1-score of 0.92, the model is well-suited for practical applications in sentiment analysis based on UFC(ultimate fight championship) comments.