SEG4300 - Assignment 6 Report - Anthony Nasr - 300170809

Dataset description:

For this assignment, I chose to use a dataset containing posts made to the twitter platform classified by emotion. The dataset is a relatively small text based dataset with 20,000 entries. It is split into training, testing, and validation sets. The dataset can be accessed through this link: https://huggingface.co/datasets/dair-ai/emotion. The labels are encoded into 6 integers ranging from 0 to 6. They are mapped as follows : 0 - Sadness, 1 - Joy, 2 - Love, 3 - Anger, 4 - Fear, 5 - Surprise. For preprocessing, text data was tokenized using BERT-base-uncased and converted into fixed-length sequences (50 tokens per sample) with padding and truncation.

CNN architecture:

The model is a simple 1D CNN that aims to classify the texts into the proper categories. The layers are as follows:

- Embedding Layer: Tokens converted into 100D vectors.
- 3 Convolutional Layers: Each has 128 filters. Followed by ReLU activation.
- Max Pooling Layer: Reduces dimensionality and retains relevant features.
- Fully Connected Layer: Outputs probabilities for each of the 6 possible labels.

Evaluation metrics and key insights from visualizations:

Since the model acts as a classifier, I used accuracy, precision, recall, and F1-Scores to evaluate the model. The final test accuracy is 88%, weighted F1-score is 88%, both indicating that the model performs well. The best performing classes were Sadness, Joy and Fear. The worst ones were Love and Surprise. A simple confusion matrix reveals that the model miss classifies Love as Joy, and Surprise as Fear. This is probably due to similarities in the language used to describe situations involving Joy and Love for example.

Interpretation of the results and their significance:

The CNN effectively captured emotion patterns from the tweets, achieving a pretty high accuracy indicating that the model generalized well enough. During training, the loss decreased over 10 epochs from 1.53 to 0.07. Some of the improvements could be to include more contrasting entries to cement the differences between emotions that can have gray areas. This includes Love, Joy, Surprise and Fear.