Name: Sushruth Danivasa Sridhar

Assignment – 4

1. A brief description of the dataset you used for fine-tuning, including its genre of text, classes and size (and if you used only a portion).

   For the fine-tuning of a BERT model, I chose the "Emotion" dataset provided by DAIR-AI, accessible through the Hugging Face Datasets library. This dataset consists of text data specifically designed for emotion classification. The text genre primarily includes short phrases or sentences, each labelled with one of several emotion classes including 'sadness', 'joy', 'love', 'anger', 'fear', and 'surprise', making it an ideal resource for emotion classification tasks. The size of the dataset was 16.1 MB which consist of 16,000 training data and 2,000 data. For processing, I have considered 1,200 data for each.

2. A brief description of the BERT model you chose to use.

   The BERT model I chose to use for fine-tuning is the "distilbert-base-uncased" version of DistilBERT, a lighter and more efficient variant of the original BERT model. DistilBERT is designed to deliver similar performance to BERT but with fewer computational resources and faster processing times. This pre-training involves learning from vast amounts of unlabelled data, which helps the model develop a deep understanding of language context and semantics.

3. A brief description of your network and your training setting.

   I used a DistilBERT base uncased model with neural network and also with additional dense layer and a softmax classification layer, trained with a batch size of 20 for 5 epochs on the classification task.

4. Result of Task 1.

   The result of the Fine-tuned model on the test dataset accuracy is 0.5191666483879089 or 51.91 %.

5. Your comments on the Task 1 results.

   The test accuracy of 0.5191666483879089or 51.91 % achieved in Task 1 suggests that the fine-tuned DistilBERT model has some capability in classifying emotions, but there's significant room for improvement but also indicates that while the model has learned some patterns from the training data, it may not be capturing the complexities of emotional nuances effectively. To enhance performance, further optimization such as hyperparameter tuning, experimenting with different network architectures, or augmenting the training data might be necessary.

6. The 3 observations from Task 2 (include the correct and incorrect examples and their predictions).

Observations:

- **Strong Performance in Identifying Clear Negative Emotions**: The model accurately identifies clear expressions of negative emotions (label 0, likely 'sadness' or similar feelings).

- **Mixed Emotion Detection**: Challenge in interpreting emotions that are not explicitly positive or negative, or that involve a blend of different feelings.

- **Tendency to Misclassify as Joy**: Bias in the model towards predicting joy, or a lack of sensitivity to contextual cues that distinguish other emotions as we have lot of joy label datasets.

Here, the classes sadness, joy, love, anger, fear, and surprise are represented by the numbers 0, 1, 2, 3, 4, and 5, respectively.

Correctly predicted examples:

- ('im updating my blog because i feel shitty', 0, 0)

- ('i never make her separate from me because i don t ever want her to feel like i m ashamed with her', 0, 0)

- ('i left with my bouquet of red and yellow tulips under my arm feeling slightly more optimistic than when i arrived', 1, 1)

- ('i explain why i clung to a relationship with a boy who was in many ways immature and uncommitted despite the excitement i should have been feeling for getting accepted into the masters program at the university of virginia', 1, 1)

- ('i like to have the same breathless feeling as a reader eager to see what will happen next', 1, 1)

- ('i just feel extremely comfortable with the group of people that i dont even need to hide myself', 1, 1)

- ('i feel a little mellow today', 1, 1)

- ('i feel like reds and purples are just so rich and kind of perfect', 1, 1)

- ('i feel like ive gotten to know many of you through comments and emails and for that im appreciative and glad you are a part of this little space', 1, 1)

- ('i survey my own posts over the last few years and only feel pleased with vague snippets of a few of them only feel that little bits of them capture what its like to be me or someone like me in dublin in the st century', 1, 1)

Incorrectly predicted examples:

- ('im feeling rather rotten so im not very ambitious right now', 0, 1)

- ('i was feeling a little vain when i did this one', 0, 1)

- ('i cant walk into a shop anywhere where i do not feel uncomfortable', 4, 1)

- ('i felt anger when at the end of a telephone call', 3, 0)

- ('i jest i feel grumpy tired and pre menstrual which i probably am but then again its only been a week and im about as fit as a walrus on vacation for the summer', 3, 1)

- ('i don t feel particularly agitated', 4, 0)

- ('i feel beautifully emotional knowing that these women of whom i knew just a handful were holding me and my baba on our journey', 0, 1)

- ('i pay attention it deepens into a feeling of being invaded and helpless', 4, 0)

- ('i find myself in the odd position of feeling supportive of', 2, 4)

- ('i was feeling as heartbroken as im sure katniss was', 0, 1)

7. The 5 examples of Task 3 and the results along with your comments.

sentence_pairs = [
    ("The tree shed its leaves.", "The flower bloomed brightly."),
    ("I like pizza.", "She enjoys pasta."),
    ("The sun is in orange color.", "The moon is in white color."),
    ("He read a book.", "She write a letter."),
    ("She cooks dinner.", "He washes the dishes."),
    ("The train whistled loudly.", "The bird sang sweetly."),
]

- cosine_similarity between the following sentences:
'The tree shed its leaves.' , 'The flower bloomed brightly.': 0.7530772165092406

- cosine_similarity between the following sentences:

'I like pizza.' , 'She enjoys pasta.': 0.6976876503222059

- cosine_similarity between the following sentences:
'The sun is in orange color.' , 'The moon is in white color.': 0.7275508269518577

- cosine_similarity between the following sentences:
'He read a book.'   ,   'She write a letter.': 0.6853322826416198

- cosine_similarity between the following sentences:
'She cooks dinner.' ,   'He washes the dishes.': 0.73608623130861

- cosine_similarity between the following sentences:
'The train whistled loudly.' ,   'The bird sang sweetly.': 0.5615370950291125

Comments:

1. "The tree shed its leaves." and "The flower bloomed brightly.": Cosine Similarity = 0.7531
   - This high similarity score indicates that the model perceives a strong contextual relationship between these two sentences. Both sentences describe natural processes involving plants, which likely contributes to the model finding them contextually similar.
2. "I like pizza." and "She enjoys pasta.": Cosine Similarity = 0.6977
   - The moderate cosine similarity score here suggests that the model recognizes a thematic link between these sentences, likely due to their focus on food preferences. It reflects the model's ability to associate different but related activities (liking pizza and enjoying pasta).
3. "The sun is in orange color." and "The moon is in white color.": Cosine Similarity = 0.7276
   - This score indicates a good level of perceived contextual relatedness. Both sentences are descriptive of celestial bodies (the sun and the moon) and focus on their color, which likely leads to this contextual connection being made by the model.
4. "He read a book." and "She write a letter.": Cosine Similarity = 0.6853
   - A similarity score of 0.6853 suggests that the model finds a moderate connection between these activities. Both involve a form of literary engagement, albeit different kinds (reading vs. writing), which the model seems to recognize as somewhat contextually related.
5. "She cooks dinner." and "He washes the dishes.": Cosine Similarity = 0.7361
   - The relatively high similarity score here reflects the model's ability to relate these two sentences through their common context of domestic chores. Cooking dinner and washing dishes are both tasks typically associated with meal preparation and cleanup, explaining the contextual similarity.
6. "The train whistled loudly." and "The bird sang sweetly.": Cosine Similarity = 0.5615
   - The lower similarity score in this pair suggests that the model perceives these sentences as less contextually related compared to others. This could be due to the distinct nature of the subjects – a mechanical train and a natural bird – despite both involving sound.