ASSIGNMENT-4 REPORT

Dataset: “dair-ai/emotion”

The dair-ai/emotion dataset, which can be found on Hugging Face, consists of English Twitter messages annotated with six core emotions: anger, fear, joy, love, sadness, and surprise. It's designed specifically for training text classification models and proves to be a valuable tool for those interested in understanding and analyzing emotional sentiments in the context of social media. Despite its compact size, the dataset is ideal for beginners entering the field of emotion detection. It offers a concise yet insightful glimpse into human emotions expressed through brief and informal language on Twitter. By focusing on the nuanced landscape of emotions in a casual online environment, this dataset aids in gaining a deeper understanding of emotional dynamics. It serves as a valuable resource for individuals looking to explore and improve their skills in emotion analysis through natural language processing.

Genre: Dair-ai/emotion dataset on Hugging Face captures casual tweets with spontaneous, opinionated sentiments—perfect for emotion analysis tasks.

Classes: With six emotions, this dataset is great for tasks like multi-class classification, helping categorize texts into various emotional states.

0-sadness, 1-joy, 2-love, 3-anger, 4-fear, 5-surprise

Size: The dataset, falling between 10,000 to 100,000 entries, strikes a good balance, offering a substantial volume with manageable complexity.

Applications: Perfect for training emotion recognition models, this dataset is invaluable for analyzing sentiment on social media, providing insights into how emotions express in written text.

BERT Model: "bert-base-uncased"

BERT, or Bidirectional Encoder Representations from Transformers, is a Google-crafted pre-trained natural language processing (NLP) model. The name "bert-base-uncased" specifies its configuration, where "base" indicates a medium-sized architecture, and "uncased" signifies training on lowercase text without differentiating between upper and lower case. With 110 million parameters, this version excels in capturing contextual nuances and semantic relationships due to its bidirectional training. Trained on extensive datasets, bert-base-uncased performs well in tasks like text classification, named entity recognition, and question answering. Its contextual embeddings enable it to understand intricate nuances and connections within sentences, making it versatile for various natural language understanding tasks across domains. Researchers and practitioners often fine-tune bert-base-uncased for specific applications, leveraging its foundational language understanding capabilities for more specialized tasks in the ever-evolving landscape of natural language processing.

Network and Training setting:

I used the BERT-base-uncased model in my code for three specific natural language processing tasks: emotion classification, sentiment analysis, and cosine similarity computation. For emotion classification, I fine-tuned the model on a selected part of the "dair-ai/emotion" dataset, using the Adam optimizer and sparse categorical cross-entropy loss for efficient training. Afterward, I applied sentiment analysis to the same dataset, presenting the model's predictions and thoroughly evaluating its performance. Additionally, I demonstrated how to calculate cosine similarity between sentence pairs using BERT embeddings. The training setup involved a single epoch with a batch size of 16. In my coding experience, the BERT model, with its robust transformer-based architecture, proves versatile in handling various language understanding tasks. It showcases its effectiveness in transfer learning, emphasizing its ability to capture intricate relationships in textual data with a high level of adaptability.

Task-1 :

Result:

```
PS C:\Users\Vivek Vardhan\OneDrive\Desktop\NLP\assign4> python A4_T1.py
All PyTorch model weights were used when initializing TFBertForSequenceClassification.

Some weights or buffers of the TF 2.0 model TFBertForSequenceClassification were not initialized from the PyTorch model and are newly initialized:
'classifier.weight', 'classifier.bias']
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
188/188 [==============================] - 2062s 11s/step - loss: 1.7172 - accuracy: 0.3123
32/32 [==============================] - 181s 6s/step - loss: 1.7343 - accuracy: 0.3590
Test Accuracy: 0.35899999737739563
```

The Test Accuracy reported is 0.35899999737739563

Comments:

The code uses a pre-trained BERT model for predicting emotions in text classification, but the achieved test accuracy of approximately 36% suggests the model struggles to understand the complexities of emotions in the dataset. Improving performance may involve fine-tuning the model or trying different architectures and hyperparameters. The current result suggests there's room for enhancement, prompting the exploration of more advanced strategies to better capture the subtle nuances of emotion in textual data.

Task-2 :

The correct predictions:

1. **Text:** i mentioned in my last blog that i have started to get the feeling that i have been pressured into studying things i do not like which has also made me into a person i might not fully be
   Predicted: 4, Actual: 4
2. **Text:** im feeling insecure at the moment
   Predicted: 4, Actual: 4
3. **Text:** i feel agitated and annoyed more than worried or fearful but these feelings can easily lead to being short tempered with my family and feelings of disharmony
   Predicted: 4, Actual: 4
4. **Text:** i was feeling frantic
   Predicted: 4, Actual: 4
5. **Text:** i feel myself falling into the pit of buying it from her i think he s for real i m just skeptical of the women
   Predicted: 4, Actual: 4
6. **Text:** i managed to re learn feeling insecure again
   Predicted: 4, Actual: 4
7. **Text:** i think were on a level of understanding though i still feel hes hesitant
   Predicted: 4, Actual: 4
8. **Text:** i do feel insecure sometimes but who doesnt
   Predicted: 4, Actual: 4
9. **Text:** i was feeling particularly vulnerable in a specific area so i began to talking to my friends and interestingly enough there was an incredible understanding of my struggle
   Predicted: 4, Actual: 4
10. **Text:** i feel terrified because my landlord has not changed our locks yet
    Predicted: 4, Actual: 4

The incorrect predictions:

1. **Text:** i was feeling really troubled and down over what my dad said
   Predicted: 4, Actual: 0
2. **Text:** i feel so thrilled to have three such distinguished individuals such as yourselves here
   Predicted: 4, Actual: 1
3. **Text:** i feel is that the most likeable characters aren t important enough to the plot
   Predicted: 4, Actual: 1
4. **Text:** i tune out the rest of the world and focus on the rhythm of the needles and the softness of the yarn and for that time i feel my most peaceful
   Predicted: 4, Actual: 1
5. **Text:** i sit here writing this i feel unhappy inside
   Predicted: 4, Actual: 0

6. **Text:** im feeling and if ive liked being pregnant
   Predicted: 4, Actual: 2
7. **Text:** im very hurt and i feel unimportant
   Predicted: 4, Actual: 0
8. **Text:** i used to be able to hang around talk with the cashier when i was putting away my money now i feel rushed and stressed if i take a second to fumble with the coins and put them in my purse
   Predicted: 4, Actual: 3
9. **Text:** i don t have the feeling of divine vibrations
   Predicted: 4, Actual: 1
10. **Text:** i vented my feelings towards the pathetic excuse of a communicat
    Predicted: 4, Actual: 0

Observations:

The model tends to correctly predict a high level of emotional distress (score 4) in instances where the text explicitly expresses negative emotions, such as feeling troubled, agitated, annoyed, or insecure.

The model struggles when the text contains positive or neutral sentiments, as seen in the incorrect predictions where the actual emotional state is 0 (no distress), 1 (mild distress), or 2 (moderate distress). It tends to overpredict a high level of distress in these cases.

The model may have difficulty discerning between different levels of distress within the high-distress category (score 4). For example, it may assign a score of 4 even when the actual distress level is lower, such as in situations where the text expresses being thrilled or liking a particular experience.

These observations suggest that the model's performance may be influenced by the explicit presence of negative emotional language, while it may struggle with subtleties in emotional expression and context.

Task-3:

The examples that I took for this task are:

1. "Golden rays warmed the chilly morning, painting the world in soft hues."
2. "Rain tapped on the window, a soothing lullaby for a lazy afternoon nap."
3. "Street vendors sizzled up savory treats, tempting passersby with irresistible aromas."
4. "Children giggled, chasing butterflies through the sunlit meadow."
5. "Antique books beckoned with tales of forgotten adventures."
6. "The moonlit waves whispered secrets to the quiet shore."

The cosine similarities are:

1. Between:
   'Golden rays warmed the chilly morning, painting the world in soft hues.'
   and
   'Rain tapped on the window, a soothing lullaby for a lazy afternoon nap.'
   is **"0.9425695538520813"**

2. Between:
   'Street vendors sizzled up savory treats, tempting passersby with irresistible aromas.'
   and
   'Children giggled, chasing butterflies through the sunlit meadow.':
   is **"0.9311324954032898"**

3. Between:
   'Antique books beckoned with tales of forgotten adventures.'
   and
   'The moonlit waves whispered secrets to the quiet shore.':
   is **"0.8604601621627808"**

Comments:

The cosine similarities reveal how closely related pairs of sentences are in terms of meaning. In the first comparison, describing morning warmth and afternoon rain shows a high similarity (0.94), capturing shared atmospheric elements. The second pair, contrasting street vendors and children in a sunlit meadow, also exhibits significant similarity (0.93), likely due to the common theme of sensory allure. However, the third comparison, involving antique books and moonlit waves, has a lower similarity (0.86), indicating the dissimilarity between the contexts of literary allure and natural serenity. These cosine similarities provide insights into the semantic relationships and thematic connections within the given sentences.