Homework 3- Report

Deep Learning – CPSC 8430

Name: Vinay ponugoti

CU ID: C13375304

Git hub link: https://github.com/Vinay-ponugoti/DeepLearning/tree/main/HW3

Introduction:

The primary goal of a Question-Answering Model is to allow a computer to accurately respond to questions, regardless of whether context is given. These models can choose answers from multiple-choice options, pull out essential information from text, generate paraphrased responses, and more. A significant advancement in natural language processing (NLP) has been the introduction of self-attention mechanisms, particularly the Transformer architecture. In this discussion, I will concentrate on BERT, a variant of the Transformer model, and provide a brief overview of its structure and method for answering questions.

Requirements:

1. Pytorch
2. Python
3. GPU
4. Tokenizer
5. Hugging face Transformers

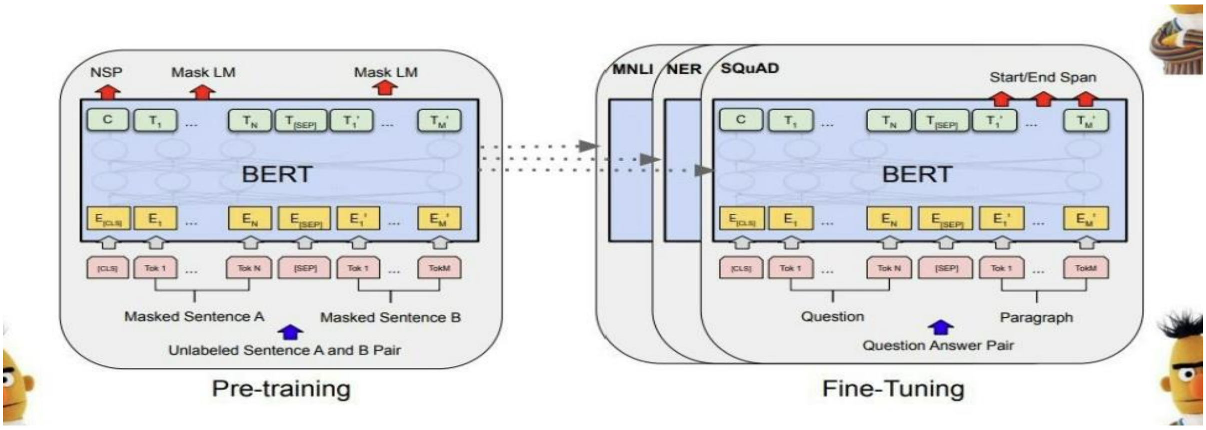
Data Set:

Used a modified version of the original SQuAD dataset, the Spoken-SQuAD dataset, which translates spoken language into text across several themes. Our goal is to evaluate how effectively we can train a BERT model—a type of AI—to answer questions. We will explore different training techniques and settings to determine the most effective method.

Used two types of background noise, referred to as 'white noise,' into the audio files used for testing. The purpose was to evaluate the computer's ability to comprehend speech in unclear sound conditions. As a result, we observed varying degrees of errors in the words recognized by the computer, which we measure as word error rates.

Model That Used:

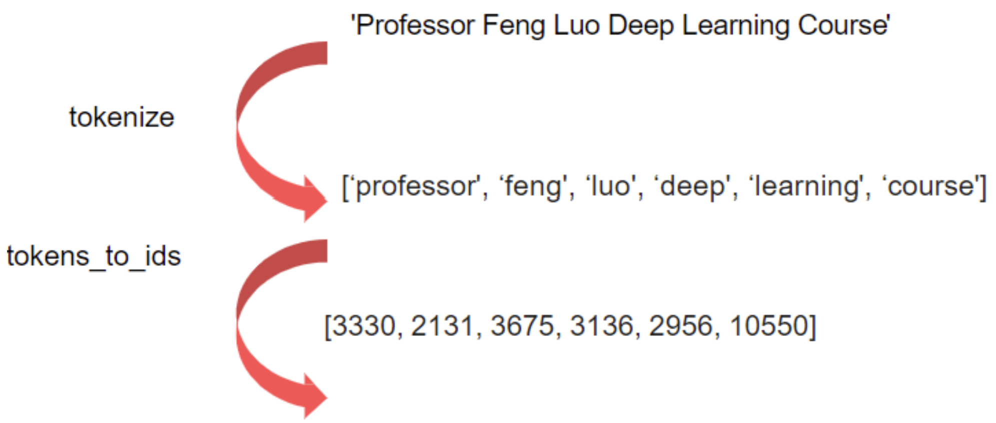
BERT is a pre-trained language model that can take into account both the left and right context at all its layers. This special quality enables BERT to be highly versatile for different natural language processing tasks, like question answering and language inference, with only minor adjustments to its architecture. Usually, adding just one output layer is enough to harness BERT's capabilities for these tasks.



The figure above is pre-training and fine-tuning phases of BERT. The only distinction between the two stages lies in the output layers, while the overall structure remains the same. For various downstream tasks, the same pre-trained model parameters serve as the starting point. During fine-tuning, all parameters of the model are adjusted. Each input example begins with a special token, and when needed, a separator token is included to distinguish between questions and answers.

Tokenization:

Tokenization refers to the method of splitting text into smaller components known as tokens, which typically consist of single words or segments of words. This process is vital for making the analysis and handling of extensive text datasets more straightforward by breaking them into smaller, more manageable parts. Tokenization plays a crucial role in various natural language processing tasks, such as sentiment analysis, language translation, and question answering.



Result:

When working with query and answer models, it's crucial to preprocess the input text to help the model accurately interpret the query and produce a relevant response. The important technique in this process is "document stride," which means dividing the input document (context) into smaller segments and generating answers for each segment separately.

```
Running Epoch : 100%|██████████████████████████████| 2320/2320 [18:03<00:00,  2.14it/s]
Epoch - 0
Accuracy: 0.4134909944673037
Loss: 2.1613702123792007
Running Evaluation: 100%|████████████████████████████| 15875/15875 [03:20<00:00, 79.29it/s]
Running Epoch : 100%|██████████████████████████████| 2320/2320 [18:03<00:00,  2.14it/s]
Epoch - 1
Accuracy: 0.6143184267369838
Loss: 1.1449001652991464
Running Evaluation: 100%|████████████████████████████| 15875/15875 [03:19<00:00, 79.51it/s]
Running Epoch : 100%|██████████████████████████████| 2320/2320 [18:03<00:00,  2.14it/s]
Epoch - 2
Accuracy: 0.7279075585048774
Loss: 0.6961757051822697
Running Evaluation: 100%|████████████████████████████| 15875/15875 [03:19<00:00, 79.65it/s]
Running Epoch : 100%|██████████████████████████████| 2320/2320 [18:03<00:00,  2.14it/s]
Epoch - 3
Accuracy: 0.8162830973108267
Loss: 0.4206507699130553
Running Evaluation: 100%|████████████████████████████| 15875/15875 [03:19<00:00, 79.38it/s]
WER (after adding scheduler) -  [3.835464566929134, 3.5479685039370077, 3.457826771653543, 3.8600314960629922]
```

Preprocessing:

```
Running Epoch : 100%|██████████| 2320/2320 [17:08<00:00,  2.26it/s]
Epoch - 0
Accuracy: 0.4663542949690901
Loss: 1.7670415444492267
Running Evaluation: 100%|██████████| 15875/15875 [11:29<00:00, 23.04it/s]
Running Epoch : 100%|██████████| 2320/2320 [16:50<00:00,  2.30it/s]
Epoch - 1
Accuracy: 0.6726639470535105
Loss: 0.8616831026393278
Running Evaluation: 100%|██████████| 15875/15875 [11:57<00:00, 22.12it/s]
Running Epoch : 100%|██████████| 2320/2320 [16:21<00:00,  2.36it/s]
Epoch - 2
Accuracy: 0.7700623460884752
Loss: 0.5189079064811613
Running Evaluation: 100%|██████████| 15875/15875 [12:26<00:00, 21.28it/s]
Running Epoch : 100%|██████████| 2320/2320 [15:52<00:00,  2.44it/s]
Epoch - 3
Accuracy: 0.8477832512608889
Loss: 0.3119803127110133
Running Evaluation: 100%|██████████| 15875/15875 [12:57<00:00, 20.43it/s]
Running Epoch : 100%|██████████| 2320/2320 [15:21<00:00,  2.52it/s]
Epoch - 4
Accuracy: 0.8945774322953718
Loss: 0.1958243275900243
Running Evaluation: 100%|██████████| 15875/15875 [13:27<00:00, 19.67it/s]
Running Epoch : 100%|██████████| 2320/2320 [14:52<00:00,  2.60it/s]
Epoch - 5
Accuracy: 0.927667025862069
Loss: 0.13154962705648987
Running Evaluation: 100%|██████████| 15875/15875 [12:52<00:00, 20.55it/s]
WER (after adding preprocessing) -  [4.403527559055118, 4.463811023622047, 3.8922834645669293, 4.254740157480315, 3.254236220472441, 3.8236
220472440943]
```

```
Running Epoch 6:  54%|██████
Running Epoch 6:  54%|██████
Running Epoch 6:  54%|██████
Running Epoch 6:  54%|██████
Running Epoch 6:  54%|██████
Running Epoch 6:  54%|██████
Running Epoch 6:  55%|██████
Running Epoch 6:  55%|██████
Running Epoch 6:  55%|██████
Running Epoch 6:  55%|██████
Running Epoch 6:  55%|██████
Running Epoch 6:  55%|██████
Running Epoch 6:  55%|██████
Running Epoch 6:  55%|██████
Running Epoch 6:  55%|██████
Running Epoch 6:  55%|██████
Running Epoch 6: 100%|██████████| 2320/2320 [04:24<00:00,  8.77it/s]
Running Evaluation:  95%|█████████
Running Evaluation:  95%|█████████
Running Evaluation:  95%|█████████
Running Evaluation:  95%|█████████
Running Evaluation:  95%|█████████
Running Evaluation:  95%|█████████
Running Evaluation:  95%|█████████
Running Evaluation:  95%|█████████
Running Evaluation:  96%|█████████
Running Evaluation:  96%|█████████
Running Evaluation:  96%|█████████
Running Evaluation:  96%|█████████
Running Evaluation:  96%|█████████
Running Evaluation: 100%|██████████| 15875/15875 [02:48<00:00, 93.95it/s]
WER (base model) -  [2.980465815176559, 2.784130292528053, 2.3946826300865225, 2.6052931339521583, 2.5828991056930275, 2.897554591502872]
```