DL/DLOps (2023)
DLOps Assignment 3: Transformer [100 Marks]
Deadline: 12/04/2023, 23:59:59

Programming instructions:

1. Programming language: Python
2. Use of PyTorch is compulsory. Marks shall not be given for TensorFlow implementation.

Reporting instructions:

1. Please submit all your working codes as .py or .ipynb files.
2. A single report (PDF) file should be submitted containing all relevant information including data pre-processing, observations, results, and analysis across the problem. Do not put snapshots of code in the report.
3. The report should be detailed and clearly explain every step you have followed. All the intermediate outputs, their inferences should be present in the report. The PDF file should be properly named with your complete roll number XYZ (ex:"XYZ_DLOps_Assignment-1.pdf"), with your name and roll number mentioned inside the report as well.
4. Record a video of your terminal/ipynb file which shows that your written codes are working. (max. video size - 3 minutes)
5. Mention any resources/articles/GitHub links that you may have used as a reference to solve any question of the assignment in the references section of the report.
6. Make sure all the submission files along with the working codes are included in a single zip file.

General instructions:

1. DO NOT plagiarize from the internet or your peers. The institute's plagiarism policy will be strictly enforced.
2. The assignment will be evaluated out of 50% of the total marks in case a report is not submitted.
3. We highly suggest using Google Colab with GPU runtimes for this assignment.

Question 1 [50 marks]

Perform image classification on selected classes[check below in Note] of CIFAR-10 dataset using transformer model with following variations:

1. Use cosine positional embedding with six encoders and decoder layers with eight heads. Use relu activation in the intermediate layers. Marks [20]
2. Use learnable positional encoding with four encoder and decoder layers with six heads. Use relu activation in the intermediate layers. Marks [20]
3. For parts (a) and (b) change the activation function in the intermediate layer from relu to tanh and compare the performance. Marks [10]

Note: Those who have

Even roll number : select odd classes

Odd roll number: select even classes

Reference Blog:
https://medium.com/mlearning-ai/vision-transformers-from-scratch-pytorch-a-step-by-step-guide-96c3313c2e0c
https://theaisummer.com/positional-embeddings/

Questions 2 [50 marks]

Based on the lecture by Dr. Anush on DLOPs, you have to perform the following experiments :

- Load and preprocessing CIFAR10 dataset using standard augmentation and normalization techniques [10 Marks]
- Train the following models for profiling them using during the training step [5 *2 = 10Marks]
    - Conv -> Conv -> Maxpool (2,2) -> Conv -> Maxpool(2,2) -> Conv -> Maxpool(2,2)
        - You can decide the parameters of convolution layers and activations on your own.
        - Make sure to keep 4 conv-layers and 3 max-pool layers in the order describes above.
    - VGG16
- After the profiling of your model, figure out the minimum change in the architecture that would lead to a gain in performance and decrease training time on CIFAR10 dataset as compared to one achieved before. [30 Marks]

You are free to use any tool/library that was discussed in the lecture to perform the above task. Describe your analysis and all your steps in detail, with relevant screenshots of the model's profiling in your report.