Assignment#2

Compare the performance of ResNet-50 (fine-tuned) and CLIP (zero shot learning)

Submission/viva date: 12-12-2024

With reference to Assignment#2 perform the following tasks:

1. **Dataset Selection**

   Choose a dataset of your choice to evaluate the zero-shot learning performance of the CLIP model trained on the Flickr 8K dataset (**as discussed in Lecture 12**). Justify your choice of the dataset and describe its key characteristics.

2. **Fine-Tuning ResNet-50**

   Fine-tune ResNet-50 on the selected dataset. Provide a detailed explanation of the fine-tuning process, including:

   a. Steps involved
   b. Hyperparameter choices (e.g., learning rate, batch size, optimizer)
   c. Challenges faced during training and how you addressed them.

3. **Implementation of Zero-Shot Learning with CLIP**

   Implement zero-shot learning using the CLIP model on the same dataset. Explain the concept of zero-shot learning and describe how CLIP achieves it compared to traditional models like ResNet-50.

4. **Comparison and Insights**

   Compare the performance of ResNet-50 (fine-tuned) and CLIP (zero-shot learning) on your selected dataset. Highlight the strengths and weaknesses of both approaches based on the results.