CS 613: NLP

Assignment 3: Fine Tuning & Evaluation

Total marks: **100 Pts.**	**Submission deadline: 23:59:59 Hrs, November 20, 2024**

Assignment Instructions

1. The assignment deadline cannot be extended. A 100% penalty will be incurred.
2. We will follow the zero plagiarism policy, and any act of plagiarism will result in a zero score for the assignment.
3. Please cite and mention others' work and give credit wherever possible.
4. If you seek help and discuss it with the stakeholders or individuals, please ask their permission to mention it in the report/submission.
5. Compute requirement: Use Colab and write the answers in the Colab itself.

Problem Statement (100 Points)

Wherein you gave them a popular pre-trained model (say Llama3/Gemma) and asked them to finetune it for two or three different tasks. Ask them to report performance scores before and after finetuning.

Task 1:

1. Select the Llama3.2-1B/Gemma model.
2. Calculate the **number of parameters** of the selected model from the code. Do your calculated parameters match with the parameters reported in the respective papers of the selected model? **[05 pts]**
3. Fine-tune the pre-trained model on the following two tasks:
    a. Classification: SST-2 **[12.5 pts]**
    b. Question-Answering: SQuAD **[12.5 pts]**

The train-test split should be 80:20, use random/stratify sampling and seed as 1. Fine-tuning should be performed on the Train split.

4. Calculate the scores for the following metrics on the test splits for the pre-trained (zero-shot) and fine-tuned models. Note that metrics depend on the selected task:
    a. **Classification:** Accuracy, Precision, Recall, F1 (Reference) **[20 pts]**
    b. **Question-Answering:** squad_v2, F1, METEOR, BLEU, ROUGE, exact-match [Read this!, and this too!, lastly this!] **[20 pts]**

5. Calculate the number of parameters in the model after fine-tuning. Does it remain the same as the pre-trained model? **[05 pts]**
6. Push the fine-tuned model to 🤗. **[05 pts]**
7. Write appropriate comments and rationale behind:
    a. Lower or higher scores in the metrics. **[10 pts]**
    b. Understanding from the number of parameters between pretraining and fine-tuning of the model. **[05 pts]**
    c. Performance difference for the zero-shot and fine-tuned models. **[05 pts]**

Total = 05+12.5+12.5+20+20+05+05+10+05+05 = 100 Pts.

Submission

1. Submit your code (GitHub) or colab notebook with proper comments to this link.
    a. Ensure the individual contribution is appropriately added (OTHERWISE PENALTY OF 10 MARKS).

Expectations from the team:

1. Properly divide the team into sub-groups and distribute your tasks equally.
2. Negative marks for not creating the report for the assignment (This can be a simple summary report).
3. Write the contributions or tasks completed by each team member. Scores might be different among team members if the tasks are not equally distributed.

TAs to Contact

1. Himanshu Beniwal (himanshubeniwal@iitgn.ac.in)
2. Indrayudh Mandal (24210041@iitgn.ac.in)
3. Mithlesh Singla (24210063@iitgn.ac.in)
4. Alay Patel (alay.patel@iitgn.ac.in)
5. Aamod Thakur (aamod.thakur@iitgn.ac.in)

FAQs

1. *We will add clarifications to doubts here. Please check periodically, as someone might have already asked about the doubt, which will be appended here.*