Due Date: 17th Nov 2023   **Max.Marks:15**

Implement a network for any one tasks from the following list of tasks using BERT/RoBERTa /ELMo/GPT-2 Pretrained-model:
1. Question-Answering model
2. Machine translation
3. Named Entity Recognition
4. Sentiment analysis
5. Document Classification

You have to perform zero-shot, few-shot and fine-tuning for the chosen task and model.
You should modify the number of layers in the model architecture. For instance, you might choose to use only a subset of layers from the pre-trained model and you can vary the number of hidden units in each layer. you can experiment with the number of attention heads.
Make sure your neural network code is bug free. You should visualize model variants and tune various model parameters.
You should vary the following:
1. Loss function
2. Optimizer
3. Single training item vs mini-batch
4. Freeze certain layers

This assignment will help you to learn how to implement a network and what is the role of each item listed above.

Dataset:
Select a suitable dataset according to the task and get it registered with the course team. Registration will be done on a first come first serve basis. No two groups are allowed to take the same dataset.

Network Analysis:
It is possible to get higher training accuracies by trying out different values of:
1. the regularization constant λ,
2. the learning rate α (good try is 0.001)
4. the hidden layer size H,
5. the number of iterations or epochs

Innovation:
Bringing out a simple and novel mechanism to improve the performance of the network will fetch extra credit up to 20%.

Deliverables:
Running Code with readme file and input sample file
Prepare a report which should contain the following:
1. Short note on data preparation and pre-processing
2. Ground truth
3. Network details
4. Results
    a. Gradient check results
    b. Choose the prominent or interesting variations for showing the performance of the network (as suggested in (a)) with varying parameters. Showing only basic results may not attract credits.
    c. Try to evaluate your code with extrinsic and intrinsic tasks.
    d. Appropriate measures should be taken to evaluate your task.

Group Information: The assignment should be done in a group of TWO students. Group size 1 and more than two won't be considered.