Homework 4

Due Mar 29 by 11:59pm **Points** 100 **Submitting** a file upload
Available Mar 20 at 12am - Mar 30 at 11:59pm

This assignment was locked Mar 30 at 11:59pm.

Homework 3 (100 Points)

Homework instructions:

- You will submit: (1) a report (pdf file) that explains your answers for each problem. The report MUST contain the link to your Github repository to access the source code you have developed for this homework. Your report should contain your name, student ID, and homework number. (2) a second pdf file that contains both source codes, results, and plots you have developed and plotted for this homework. You can easily generate this pdf in Jupyter Lab or Jupyter Notebook (sample pdf provided on Canvas.). The pdf file name should contain your name, student ID, and homework number.
- In your report, provide separate and clear responses for each problem. Make reasonable assumptions where necessary and clearly state them! Be sure to show all the work involved in deriving your answers! If you just give a final answer without explanation, you may not receive credit for that question.
- You may discuss concepts with your classmates. This fosters group learning and improves the class' progress. However, make sure to submit your own independent and individual solutions.
- **Make sure to use Google Collab (with GPU add-on option), or your personal GPU card for the training.**

Problem 1 (60pts)

In this homework, we focus on the language model we did in the lectures.

1. Use the GRU example, adjust the hyperparameters (fully connected network and the number of hidden states) and analyze their influence on running time, perplexity, training and validation loss, and the output sequence (try a few examples).
1. Use the LSTM example, adjust the hyperparameters (fully connected network and the number of hidden states) and analyze their influence on running time, perplexity, training and validation loss, and the output sequence (try a few examples)
1. Compare runtime for training and inference, computational and mode size complexities, training and validation loss, and the output sequence (try a few examples)for rnn.RNN, rnn.LSTM and rnn.GRU implementations with each other use the same hyperparameters for your comparison.

Problem 2 (40pts)

This homework focuses on the Deep RNN problem we did in the lectures.

1. Build the model by replacing the GRU with an LSTM and compare the training and validation loss, and the output sequence (try a few examples) against GRU.
2. Compare runtime for training and inference, computational and mode size complexities, and the output strings for nn.LSTM and rnn.GRU implementations with each other.
3. Adjust the hyperparameters (fully connected network, number of hidden layers, and the number of hidden states) and compare your results (training and validation loss, computation complexity, model size, training and inference time, and the output sequence). Analyze their influence on accuracy, running time, and computational perplexity.