IS 784 Deep Learning for Text Analysis

Assignment 3 (Deadline 13 June 2021)

In this assignment, you are going to implement a classification task using a recurrent neural network. Then, you need to experiment with numerous parameters in your model and compare their effects on performance using Weights and Biases.

1. LSTM(50pts)

For this part of the assignment, you will use the IMDB dataset we used in the labs. Preprocess the dataset and create your data loaders. You can use torchtext or any other method for preprocessing.

Then create a network model with the following parameters: 1 layer LSTM network with 100 sized hidden dimensions, a dropout layer with 0.5 probability, and a final fully connected layer.

Train the network for 1 epoch with 64 batch size with adam optimizer and save your results.

You are expected to show and explain your steps clearly. Save your work as a single notebook file.

Tip: Before starting on this work check the second part of the assignment, Creating a modular network will help you on the second part of your assignment.

2. 'Weights and Biases'(50pts)

If you are not familiar with Weights and Biases[1] , it is an online platform for tracking your hyperparameters and metrics in machine learning tasks. You can check its tutorial page[2] to learn how to use this platform, its PyTorch tutorial[3] and its accompanying notebooks[4]. You need to create a free account on this platform to access its functions.

In this part of the assignment, you will use the network you created in the first part and try to find optimal parameters with help of *Weights and Biases*. Some parameters you can explore are as follows;

- Depth (number of layers) of LSTM
- Hidden size of the LSTM
- Dropout probability
- Number of epochs,
- Batch size

You are expected to create a report that will show how these parameters affect the accuracy and the loss of the network using graphs you obtained using *Weights and Biases*. You are also expected to give your notebook file that runs these simulations.

---

[1] https://wandb.ai/site
[2] https://docs.wandb.ai
[3] https://docs.wandb.ai/guides/integrations/pytorch
[4] Simple PyTorch Integration and Hyperparameter Sweep tutorials

Note: You don't need to check for every possible combination of parameters. But you need to show your approach clearly. You are expected to show the effects a minimum of 3 different parameters.

3. Bonus: BiLSTM(20 pts)

In this step of the assignment, you are going to compare the efficiency of the BiLSTM architecture with the LSTM network you created in the previous parts. For this purpose, edit your code on the previus part to compare BiLSTM architecture with LSTM. Again, show your comparisons using *Weights and Biases*.

Tip: Check PyTorch's "pack_padded_sequence" and "pad_packed_sequence" functions while creating your network.

4. Deliverables and Late submission

You are expected to submit '*.ipynb' files for each part and one report in PDF or Word. Your file names should contain your student number. Add your name and student number inside your documents as well. Your code blocks should be adequately explained. Your report should explain your approaches and show the results of your network.

Late submission is accepted until 15 June 2021 with a -5 points penalty for each day.