

Homework 3

Due: June 4, 2023, 11:59PM PT

Student Name: Andy Franck

Instructor Name: John Lipor

Problem 1 rnn.py (8 pts)

For the first problem of this assignment, you will complete a RNN layer for your MyTorch package. Your task is to complete `rnn.py` as detailed by the documentation in the corresponding file and the lecture notes. In particular, you will implement the forward and backward passes of a standard recurrent layer. For simplicity, we will ignore the bias term in this layer.

Note that the book's derivation of backpropagation does not include the \tanh nonlinearity, which is standard for RNN layers. As a result, you will have to think through the backpropagation equations more carefully. For convenience, your forward equations are as follows

$$\begin{aligned} h_t &= W_{hx}^T x_t + W_{hh}^T h_{t-1} \\ o_t &= \tanh(h_t). \end{aligned}$$

To enable easier backpropagation, you may make use of your `mytorch.nn.activation.Tanh` class.

Turn in the output of the notebook `RNNTester.ipynb`.

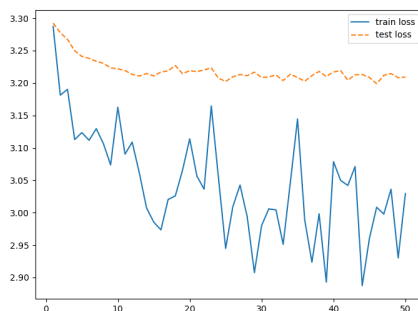
```
Difference in outputs: 1.7939023831631938e-06
Difference in states: 9.217579593529194e-07
```

```
Difference in dLdWxh: 0.09724531
Difference in dLdWhh: 0.0081496372
```

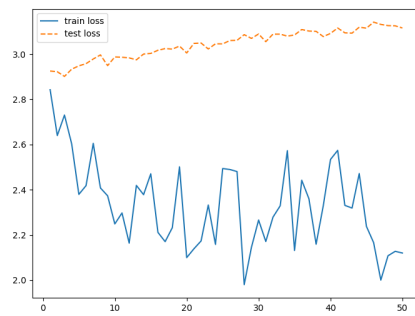
Problem 2 Becoming Shakespeare (10 pts)

For this problem, you will compare a few sequence predictors on Shakespeare's collected works, available from Project Gutenberg. The code to download and preprocess the data is provided in `ShakespearePrediction.ipynb`. Your task is to implement the following networks (use of `torch.nn` is highly encouraged):

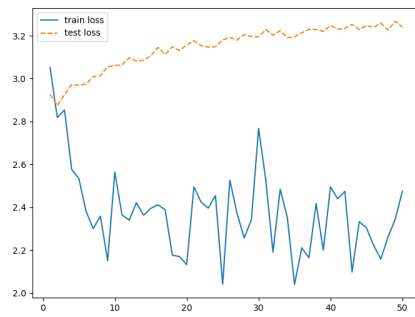
1. A single layer RNN



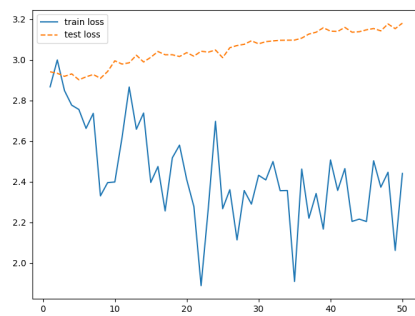
2. A single layer LSTM



3. A single layer GRU



4. One of the above with multiple layers.



Turn in the following:

- (a) A description of your learning rate, batch size, and any other information needed to replicate your results.

Answer:

Learning rate: 0.01

Batch size: 64

Hidden Size: 64

Input Length: 5

Output Size: 27

Note: Dropout was used, but had little to no effect, so it is not necessary to replicate the results.

(b) Plots showing the training and validation loss versus epoch for each model. **Answer:**
Plots displayed above.

(c) A table comparing the prediction error and perplexity for each of the four models. **Answer:**

| Model | Prediction Accuracy | Perplexity |
|------------|---------------------|------------|
| RNN | 29.2% | 19.24 |
| LSTM | 33.2% | 17.59 |
| GRU | 28.6 | 21.66 |
| Multilayer | 33.1 | 17.22 |

(d) A description of which part of this problem was most challenging for you. **Answer:**

Creating the models was fairly easy. The most difficult part was finding optimal sequence output and input lengths, as well as keeping the model from overfitting. It is fairly obvious from my model graphs that I was unable to remove the overfitting in my graphs, however given more time and perhaps a rework of the model, I believe I could have found a better architecture.

Problem 3 Course Evaluation (3 pts + 50 Lipor pts)

You should have received one or more messages requesting that you complete a course evaluation. Please complete this evaluation, and type “I have completed the course evaluation” once you’ve done so. **Note:** To maintain anonymity, I will not review which students answer affirmatively to this question.

I have completed the course evaluation

Problem 4 MP3 Topic (0 pts)

As stated in class, the provided problem for MP3 will involve predicting physiological parameters of interest from ultrasound images of eyes. However, since this project does not involve prediction on sequential data, I am giving you the option to choose a problem of your own to solve for MP3. If you wish to select your own project, it need not be one focused on sequential data, but it must be a problem you have not previously solved. Turn in a description of the problem you wish to solve or a sentence stating that you’ll work on the provided problem for MP3.

I will work on the provided problem.