Deep Learning – RNN & LSTM – HW #3 :

Introduction:

The model architecture inspired by Elad Hoffer's model:
https://github.com/eladhoffer/recurrent.torch

Model Parameters:

This model allows to decide about many tunable parameters, such that:

"rnnSize" - Size of RNN hidden layer – 200 in this model.

"numLayers" - Number of layers in the LSTM – 2 in this model.

"dropout" -Dropout p value – 0.2 in this model.

"LR" - Learning rate - 0.0025 in this model.

Model Architecture:

The model includes 3 levels of layers:

1. The input layer:
   LookupTable layer - refers to indexes on the inputs vector.

2. The RNN layer:
   Loops until "numLayers" value (2 in this model).
   Each iteration:
   - Adds LSTM layer with the size of hidden layer as "rnnSize" value (200 in this model).
   - Adds Dropout layer with probability = "dropout" value (0.2 in this model).

   This step adds 4 layers.

3. The output layer:
   Linear layer – returns an outputs vector with the probability of each value.

It totals with 2651801 parameters.

The data:

The Penn Treebank (PTB).

The Criterion:

Cross Entropy Criterion.

Training and Evaluate Procedure:

Defines 3 arrays:

- Loss Train
- Loss Test
- Loss Val

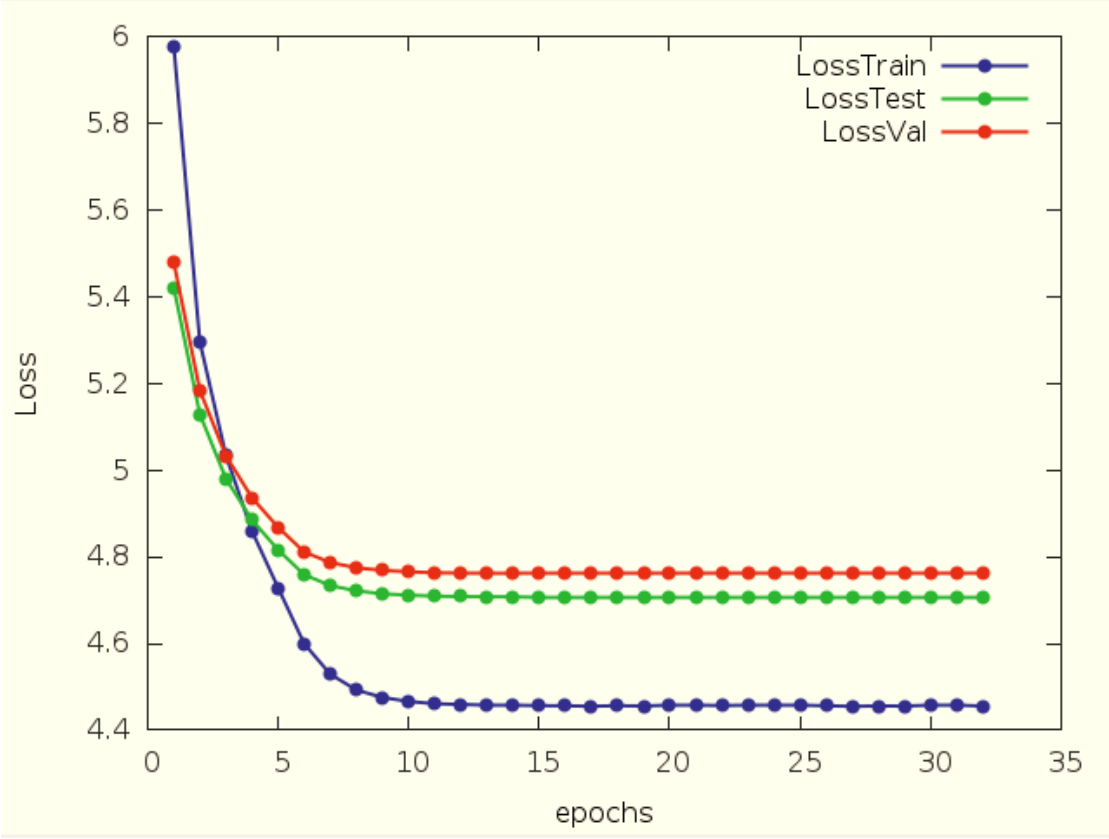
Runs 32 epochs.

On each epoch:

1. Trains the training data and computes the train loss for this epoch.
2. Evaluates validation data and computes the validation loss for this epoch.
3. Evaluates test data and computes the test loss for this epoch.
4. Saves the model as torch format.
5. Generates 5 random sentences continuations to: ”Buy low, sell high is the…”.
6. According to validation loss of this epoch, decides if to decrease the learning rate.

Saves a convergence graph for loss as a function of time (epochs) -
Depicts training, test and validation performance.

The Graph – Loos as a Function of Time:



Perplexity on the Test Set:

Starting from epoch 25, the test perplexity of each epoch was very similar.

The test perplexity - 110.75052.

Generated 5 continuations to: ”Buy low, sell high is the…”:

For example, on epoch 26:

```
Sampled Text:
Buy low, sell high is the... overall price of listed price

Sampled Text:
Buy low, sell high is the... capital
 the partners also

Sampled Text:
Buy low, sell high is the... exchanges and the value of

Sampled Text:
Buy low, sell high is the... capacity of investors such share companies

Sampled Text:
Buy low, sell high is the... maximum value of big N which

Test Perplexity: 110.75052642822
Learning Rate decreased to: 5.9604644775391e-10
```

More examples on epochs 25-32:

```
Sampled Text:
Buy low, sell high is the... industry 's aggressive part in

Sampled Text:
Buy low, sell high is the... signs of economic payments in

 Sampled Text:
 Buy low, sell high is the... stock market to repurchase earnings

Sampled Text:
Buy low, sell high is the... surprise for the next 1980s
```

```
Sampled Text:
Buy low, sell high is the... dollar gains in the best half

Sampled Text:
Buy low, sell high is the... most active lower
 dealers said

Sampled Text:
Buy low, sell high is the... statistics order in the 100-share


Sampled Text:
Buy low, sell high is the... opportunity to signal the capital-gains


Sampled Text:
Buy low, sell high is the... bonds ' portfolio and investment stocks
```

GitHub Link:

https://github.com/hanama/Deep_Learning/tree/master/DL%20-%20hw3