

Practical Task

November 20, 2017

1 Task 1

We can train an RNN model to predict the next character given a sequence of characters. Such models are affectionately known as Char-RNNs. Optionally, you can read [this article](#) by Andrej Karpathy to better understand RNNs and the idea of Char-RNNs.

Train a Char-RNN model on the [Tiny Shakespeare](#) dataset. You can use the following repositories as a reference:

- Char-RNN in PyTorch: <https://github.com/spro/char-rnn.pytorch>
- Char-RNN in Tensorflow: <https://github.com/karpathy/char-rnn>

Report the change in perplexity during training. Plot a graph to do so. Smoothen the curve if required. Plots that are not readable or are not self explanatory will not be considered.

2 Task 2

Use the trained Char-Rnn model from Task 1 and report 3 generated sequences of length 100 by priming the model with random character sequences of length 5. E.g., "2 b3n", "bg09Z", etc. While reporting the generated sequences, also report their corresponding priming sequences.

3 Task 3

Like in the previous task, generate character sequences of length 100 for the following priming sequences:

- The
- What is
- Shall I give
- X087hNYB BHN BYFVuhsdb

4 Neural Language Models (OPTIONAL)

You might be thinking, what would one get by training a model to predict the next character? Well, this question could have been asked in the theoretical exercise, but it would not have made any sense there without any context. So we will ask a set of optional questions here for you to understand how these models (models like Char-RNNs) are useful in real world applications. You can choose to not answer the following questions and move on to the next practical task. If you don't feel like writing the answers down, at least give these questions a thought.

4.1 Question 1

What are Language Models? Where and how are Language Models used?

4.2 Question 2

How can you use a trained Char-RNN model as a Character Level Language Model?

4.3 Question 3

How can you train a Word Level Language Model?

4.4 Question 4

Formally describe the inference model that an RNN trained to predict the next word represents?

4.5 Question 5

How will you generate the "probability of existence" of an input sequence of words, given a trained RNN Language Model?

5 Task 4

Extend the Char-RNN model in Task 1 to train a Word-RNN model on the same dataset. Report the perplexity during training. Report 3 generated word sequences for the follow priming sequences:

- The
- .
- which is,
- blah blah blah