CSC413 Final report

The model we intend to build is a type of text generation model with the Shakespeare style. To achieve that we want to build an RNN model and train it with the famous Shakespeare work--- Hamlet.

Model:

Our model has three parts:

An embedding layer that takes a character index (sparse representation of a one-hot vector representing the character), and returns an embedding x(t) for the character.

A recurrent neural network layer. We will use an nn.GRU unit uses the following function. At each time step, this layer will take the previously hidden state h(t−1) and the embedding of a new generated token x(t) and compute the new hidden state h(t) representing tokens not yet generated.文本

描述已自动生成

The projection MLP takes the hidden state and computes the distribution of the next character to generate.

For example, generate a sentence “He closes with you in this consequence:” is like the following picture.![图示

描述已自动生成]()

Data:

Since all Shakespeare’s work doesn’t have any copyright problems, we just download the Hamlet novel in txt form. We process the data so that it turns into 4167 sentences. We intend to use those 4167 sentences to train our model.

We clear the front space and “\n” in each sentence because we don’t want those to show up too frequently.

Training:

We first train based on one sentence. The training curve is like this:图表

描述已自动生成

Then we try to train the model with batches and change the hyper-parameters like learning rate and batch size to get the best effect on the model.

We eventually use the batch size of 32 and the learning rate to be 0.001 as the final choice.

Results:

In the end, we try to use the trained model to generate articles. Even though some sentences do not follow the grammar, the words our model generates are almost right to read.

However, there are still repeat a sentence that shows up frequently and sentences that do not have a lot of connection. We guess that it might be our training data is still too small. And our model generates character by character, it may have the ability to generate the word and sentences correctly. But it‘s not powerful for the article.

Ethical Consideration:

We don’t face any ethical problems.

Authors:

Zuo Ziqiang