

CS 8803 DL Assignment 4

Jingdao Chen

April 6

1 Character Prediction

1.1 Prediction

RNN 10 epochs:

the meaning of life is speak.

PETRUCHIO:

Methink we shall more greck?

HERMIONEN:

As I came dreth, in the gods, all you on propence

Mercution person; shall be kned therew me no have distice us: but he must

Time Pholent give poings in thee.

RNN 50 epochs:

the meaning of life is speak.

POLIXENES:

I shall not rancasion must not great dust that make us, sweet parroos: begitan.

A know some lines, musicifficial-swear of no nury

To pray thee still virtue end on yet the robedience,

That we perforth thou, ho!

KING EDWARD IV:

Now you men is the worth Richard be doubt laul make their figure

By thing master to the vange that contract;

LSTM 10 epochs:

the meaning of life is not

My princime lap in head, like a mind,

Which shall be welcomd.

CLARENCE:

Sir, and know forth my born, in Bepings.

JULIET:

No distage of my moty

So, madam, content us, cold joy, yet the other else I may

LSTM 50 epochs:

the meaning of life is not

Then in your confined and recome;

I'll not do him content the lard to thunder flies.

The luke is love in Bergumer.

MIRANDA:

Call!

ERCEP:

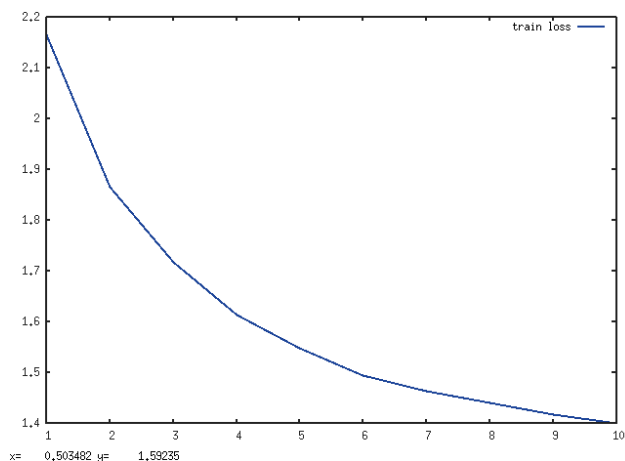
How yet and fellow, what would I find swars;

I pray you! have indeed them.

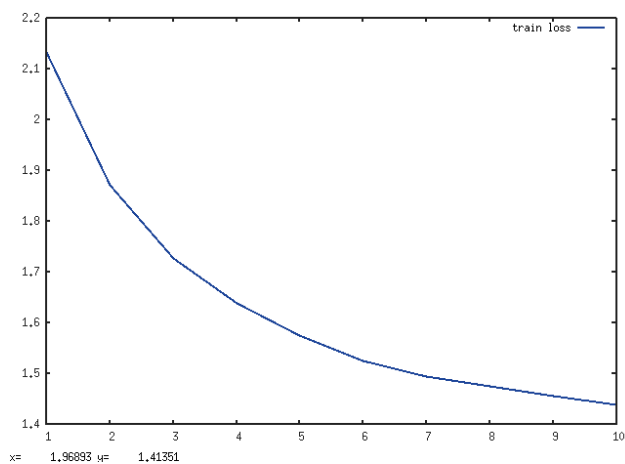
For both RNN and LSTM, the character prediction result improved as the number of epochs increases. LSTM performed better than RNN in terms of word coherence and associations between neighboring words since it is able to model long-term relationships in the input.

1.2 Gate Removal

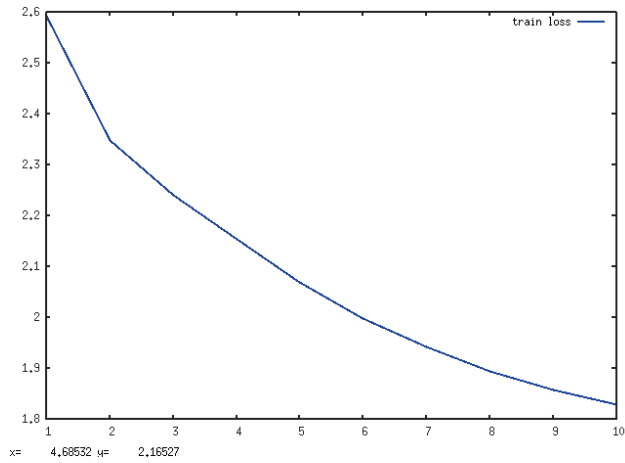
No gates removed:



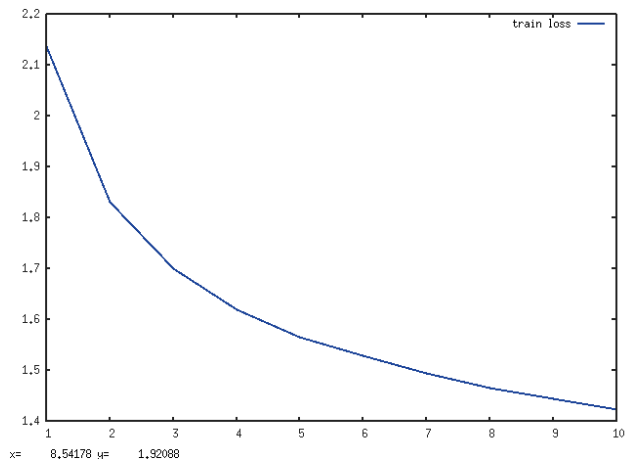
Input gate removed:



Forget gate removed:



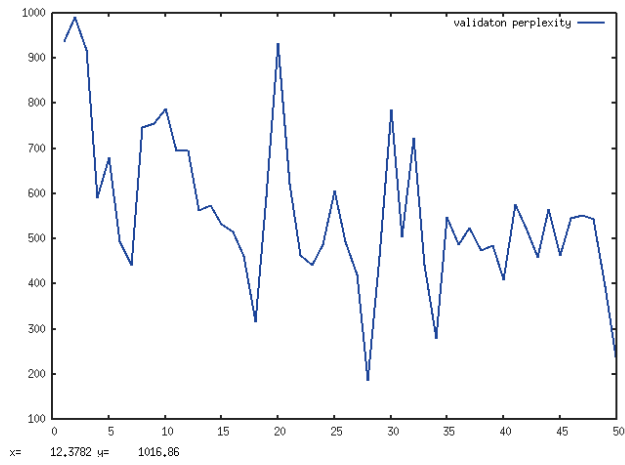
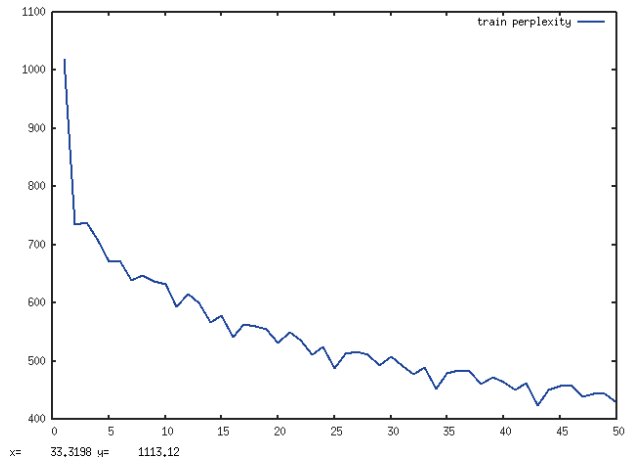
Output gate removed:



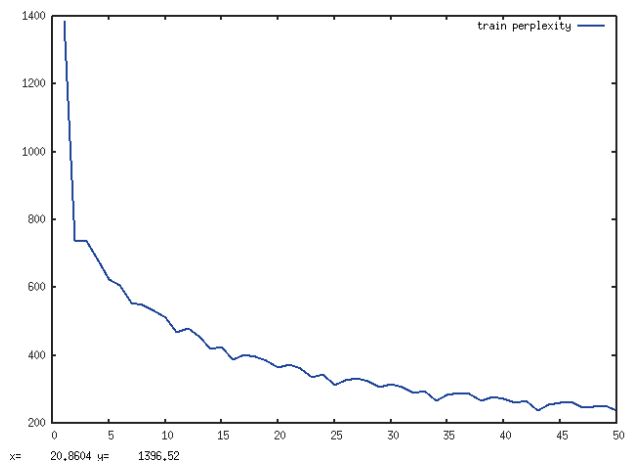
Based on the plots of training error, removing the input or output gate increases the error since information from the input/previous state is not propagated to the neuron. Removing the forget gate also significantly increases the error and slows down the convergence rate.

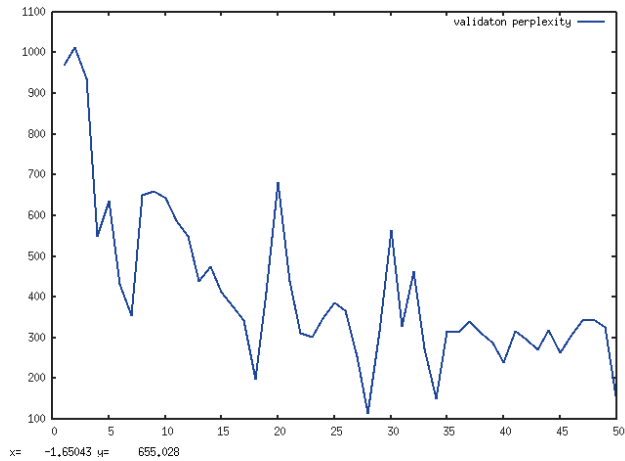
2 Perplexity

The following graphs show the training and validation perplexities respectively for RNN.



The following graphs show the training and validation perplexities respectively for LSTM.





From the graphs of training and validation perplexities, LSTM experiences a smoother decrease in training perplexity compared to RNN. It also has a steadier decrease in validation perplexity whereas the validation perplexity for RNN tends to oscillate wildly. This is because LSTMs are better suited to model long-term dependencies inherent in languages.