Project 4

CS 525: Deep Learning

Jerry Duncan
jdunca51@vols.utk.edu

April 20, 2020

1 Introduction

For this project, we're testing different RNNs on their ability to predict the next character based on a small input sequence. We'll be looking at different types of RNN layers, such as SimpleRNN and LSTM layers, as well as trying different input sequence lengths and strides. We'll then use these networks to generate some text given an input sequence to see how well they are able to produce text that, at a minimum, consists of real words, and hopefully makes logical sense.

2 Networks

For this project, I'll be using two different networks. The first uses a SimpleRNN layer with a fully connected layer, with a variable number of hidden units. Similarly, the second uses a LSTM layer with a fully connected layer, with a variable number of hidden units. I decided to tweak the number of hidden units instead of experimenting with a variable number of layers due to training time and thoroughness of experiments.

```
# SimpleRNN with a single layer attached to a FCL for
                               classification
def srnn(self, nb_hidden):
return Sequential([
    SimpleRNN(nb_hidden, input_dim=self.vocab_size),
    Dense(self.vocab_size, activation="softmax"),
  ])
```

```
# LSTM with a single layer attached to a FCL for
                               classification
def lstm(self, nb_hidden):
return Sequential([
    LSTM(nb_hidden, input_dim=self.vocab_size),
    Dense(self.vocab_size, activation="softmax"),
  ])
```

3 Results

For my experiments, I tried two configurations for the hidden unit size, the window length, and the stride. I did these for both networks and generated graphs to see the loss of each network over 100 epochs.

In Figure 1, we can see the losses of the SimpleRNN experiments. The first thing to note is that none of them managed to bring the loss below 1.6. Another thing is that the networks with 100 hidden units performed worse in all cases. While it's typically known that there is an upper limit on hidden units to get better performance, 100 hidden units is not very many compared to other modern deep neural networks.

In Figure 2, we can see the losses of the LSTM experiments. This time one configuration managed to bring the loss below 0.5, and all of them were under 1.25. These results seem to make more sense from the perspective of other modern deep neural networks because more hidden units allows the LSTM networks to perform better. It also seems that a larger stride affects loss more than window size regardless of the number of hidden units.

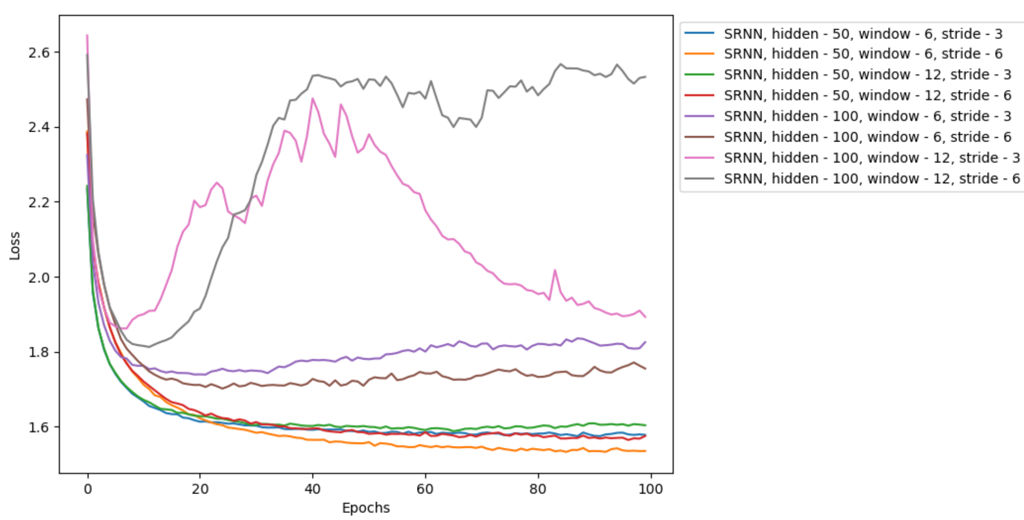

Figure 1: Loss of SimpleRNN networks with different hyperparameter configurations.

The next thing we want to look at to see if the losses actually matter, is how well each network can generate text. In Figure 3, we show predicted text given the initial character sequence after every 20 epochs. The first thing to point out is that no network was able to produce the training text sequence that the initial characters were pulled from. The second is that the SimpleRNN networks produces more sequences containing characters that weren't real words. The third is that often the best generated sentences

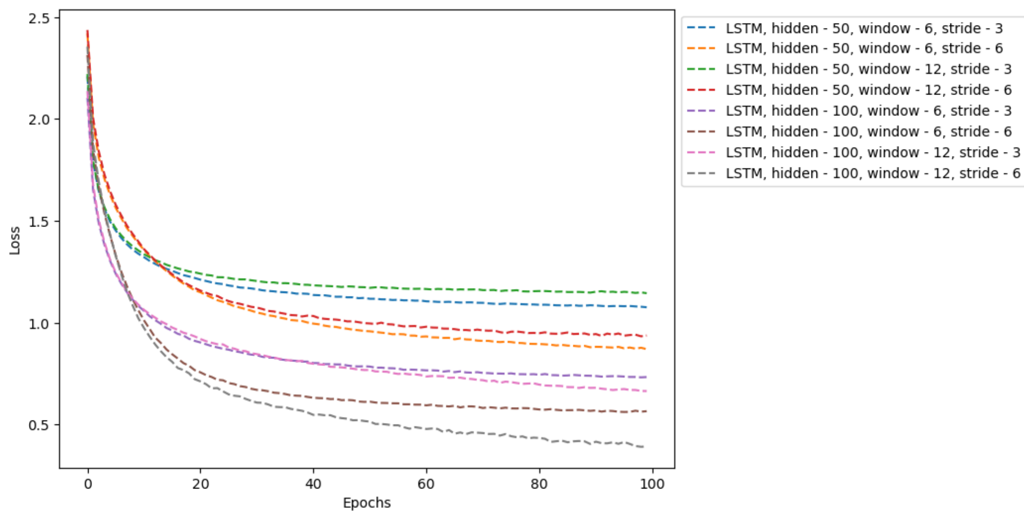

Figure 2: Loss of LSTM networks with different hyperparameter configurations.

came after 20 and 80 epochs and that many seemed to have some serious problems after 100 epochs. And lastly, the two SimpleRNN networks that had losses that spiked above 2.5 essentially generated gibberish.

4 Conclusion

In conclusion, LSTM networks perform better than SimpleRNN ones. LSTM networks also seem to better generalize to more hidden units and can make good use of them as they often have a lower loss the more they have. In all cases, a larger window size worked better than a smaller one, regardless of stride or number of hidden units. Overall, the LSTM network with 100 hidden units, a window size of 12, and a stride of 6 worked the best of all networks tested.

5 Running My Code

My code is able to be ran like asked for in the project requirements.

```
python rnn.py <filename> <model_name> <nb_hidden> \
<window_size> <stride>
```

Figure 3: Text generated by different hyperparameter combinations. Columns are the two types of RNNs and rows are different hyperparameters.

	LSTM	SimpleRNN
	Initial chars: ‘ as fr‘	Initial chars: ‘oked s‘
	Correct text: as from today, well, i’ve seen some	Correct text: oked so fierce, his mother butted in
50 hidden,	Predicted text: as from me and she love to and i sa	Predicted text: oked she want to the way, if you say
6 window,	Predicted text: as from my friend. i say the way th	Predicted text: oked she love you a lift a long the
3 stride	Predicted text: as from my friends i’m the only lon	Predicted text: oked she want to know i want to know
	Predicted text: as from my friend i said i’ll go it	Predicted text: oked so when it and i want somethare
	Predicted text: as fret rebyb ofoyean the only to t	Predicted text: oked she want to to the worll now to
	Initial chars: ‘ee tha‘	Initial chars: ‘i got ‘
	Correct text: ee that he’s just a fool, and he nev	Correct text: i got a whole lot of things to tell
50 hidden,	Predicted text: ee that i want to do, so how you say	Predicted text: i got ond on your right i want to kn
6 window,	Predicted text: ee that i wanta me a now i’m the tim	Predicted text: i got mead the sing the sing the sin
6 stride	Predicted text: ee that she does, she said i’m on th	Predicted text: i got that you alone you make you ma
	Predicted text: ee that’s right i can do so ho to ca	Predicted text: i got to know you know you know you
	Predicted text: ee that you know i wore you and i wa	Predicted text: i got a when they show and that and
	Initial chars: ‘it’s like to‘	Initial chars: ‘t down below‘
	Correct text: it’s like to listen to your fears ch	Correct text: t down below his knee hold you in h
50 hidden,	Predicted text: it’s like to see you want to like yo	Predicted text: t down below sean sad she sad the s
12 window,	Predicted text: it’s like to hold you know that she	Predicted text: t down belown the word wain a little
3 stride	Predicted text: it’s like to you she wanted it straw	Predicted text: t down below she don’t got the way i
	Predicted text: it’s like to let you know you know i	Predicted text: t down below on the cry can in the c
	Predicted text: it’s like to know the world it a lit	Predicted text: t down belowe she she’s need you kno
	Initial chars: ‘eal love...‘	Initial chars: ‘waiting her‘
	Correct text: eal love... revolution you say you w	Correct text: waiting here for you, wond’ring wha
50 hidden,	Predicted text: eal love... but i don’t know why you	Predicted text: waiting here houn hore, herely hor
12 window,	Predicted text: eal love... say her me the cry the m	Predicted text: waiting her saive when a want to k
6 stride	Predicted text: eal love... but no one you say the b	Predicted text: waiting here the that’s when you th
	Predicted text: eal love... word no know what i want	Predicted text: waiting her love to come on out lik
	Predicted text: eal love... doesn with me with a lov	Predicted text: waiting here here i want i know i k
	Initial chars: ‘ of lo‘	Initial chars: ‘an, ’c‘
	Correct text: of love got a hold on me. please be	Correct text: an, ’cause when i get you alone you
100 hidden,	Predicted text: of love you know that she does, she	Predicted text: an, ’crong the like the like the lik
6 window,	Predicted text: of love you know this like you know	Predicted text: an, ’could i when i won’s she love i
3 stride	Predicted text: of love, love you know i know i kno	Predicted text: an, ’can he sun and i don’t the want
	Predicted text: of love, love with a geamong a dris	Predicted text: an, ’could i like it i look it it’s
	Predicted text: of love, love you. i can’t be stay	Predicted text: an, ’caus to my love baby bar the wa
	Initial chars: ‘oom on‘	Initial chars: ‘ a swe‘
	Correct text: oom only to find gideon’s bible gide	Correct text: a sweater by the fireside sunday mo
100 hidden,	Predicted text: oom onlow oh, that she magade i will	Predicted text: a sweend a sall me when i don’t you
6 window,	Predicted text: oom only heart be man, yea-midhor bl	Predicted text: a swees yearle you and are you and
6 stride	Predicted text: oom onle hady triam on the sky is i’	Predicted text: a swerd somes on won’s oh, so sevin
	Predicted text: oom onl maney think tw ywain, and th	Predicted text: a swee how holl he, i do make help
	Predicted text: oom onle wanty you say good day suns	Predicted text: a swes it and i ne do, it it and i
	Initial chars: ‘llow submari‘	Initial chars: ‘car yes i’m‘
	Correct text: llow submarine we all live in our ye	Correct text: car yes i’m gonna be a star baby yo
100 hidden,	Predicted text: llow submaring and way to make you t	Predicted text: car yes i’m men’t the me the seenee
12 window,	Predicted text: llow submaring all the love there i	Predicted text: car yes i’m on tou tou tou bat ond
3 stride	Predicted text: llow submaring what i do i’ve know t	Predicted text: car yes i’me mane took tome tome to
	Predicted text: llow submaring what can i do, and i	Predicted text: car yes i’m she she she she she she
	Predicted text: llow submaring we and knew of the li	Predicted text: car yes i’m that that that that tha
	Initial chars: ‘ out. what g‘	Initial chars: ‘old me you d‘
	Correct text: out. what goes on what goes on in y	Correct text: old me you didn’t need me anymore w
100 hidden,	Predicted text: out. what go. be, i can. you wand	Predicted text: old me you das it it i wist i was i
12 window,	Predicted text: out. what going to know a have her	Predicted text: old me you dove i kane th aive t at
6 stride	Predicted text: out. what good a love to es. i can	Predicted text: old me you d lantetthettitet tutttel
	Predicted text: out. what going home. the be to sel	Predicted text: old me you d gongi w ging honging w
	Predicted text: out. what goob, wind and i should w	Predicted text: old me you de i in im win i i in th

It can also be run with combinations of parameters using main.py instead.