Josh Mandzak, Doug Lowe
4/24/2022

Project 4

Introduction:
This project aims to generate Beatles lyrics through the use of both simple recurrent neural networks and long short term memory models. The model is trained on a .txt file containing Beatles lyrics all concatenated together. The data is read in and parsed into readable x and y input for the models, then the model is trained and evaluated both with loss and accuracy as well as by the quality of generated lyrics.

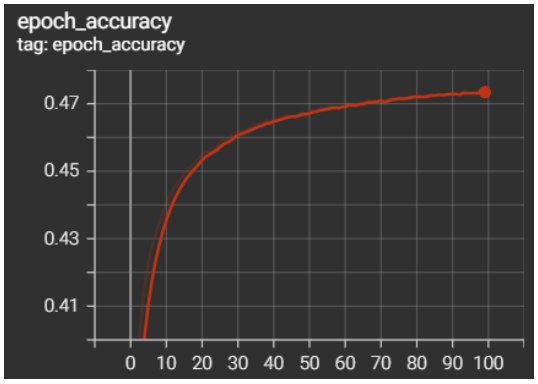
Description:
Eight different networks are trained and evaluated to perform the task of generating Beatles lyrics. All models use a sampling temperature of 0.2 (this number was found to be the best mix of generating new but possibly non-english phrases vs just repeating the same valid english phrases repeatedly. All models also have a dense layer after the LSTM or simple RNN layer of size vocab_size (which in our case is 48) with softmax activation. All of the models also take input of size (number_of_sequences, window_size, vocab_size). All models are compiled with the adam optimizer, use categorical cross entropy as their loss in line with the softmax, and are evaluated quantitatively by loss and accuracy.

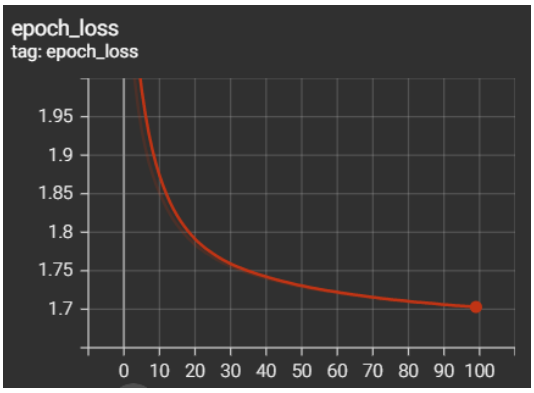
The following list details all of the 8 networks with the parameters that are unique to that network.

- Simple RNN, hidden state size = 100, window size = 5, stride = 3
- Simple RNN, hidden state size = 200, window size = 5, stride = 3
- Simple RNN, hidden state size = 100, window size = 10, stride = 7
- Simple RNN, hidden state size = 200, window size = 10, stride = 7
- LSTM, hidden state size = 100, window size = 5, stride = 3
- LSTM, hidden state size = 200, window size = 5, stride = 3
- LSTM, hidden state size = 100, window size = 10, stride = 7
- LSTM, hidden state size = 200, window size = 10, stride = 7

Results:
All loss + accuracy plots are detailed below for each network, as well as the generated text after running 100 epochs of training.

<u>Simple RNN, hidden state size = 100, window size = 5, stride = 3</u>

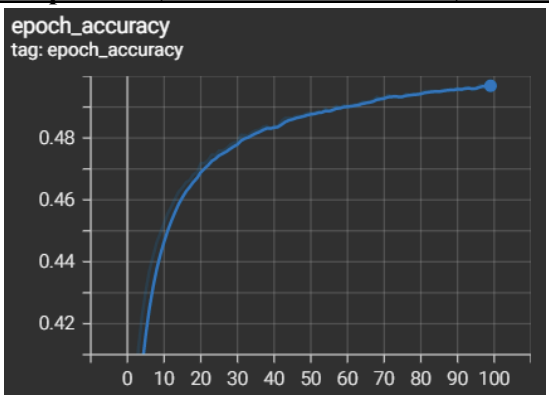


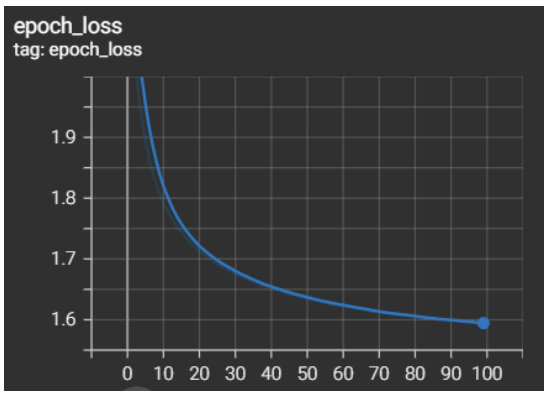


Generated text:

it all the shen sie love you kele.
i le gon the say the work it in you know you go tell you know you go tell you know yo leter 9not the bay you know you go tell you know yo de ntare to shevwell you go they sunshines you know yo laay the say tarr
i want to see it in here to the say i want to yep you know

Simple RNN, hidden state size = 200, window size = 5, stride = 3

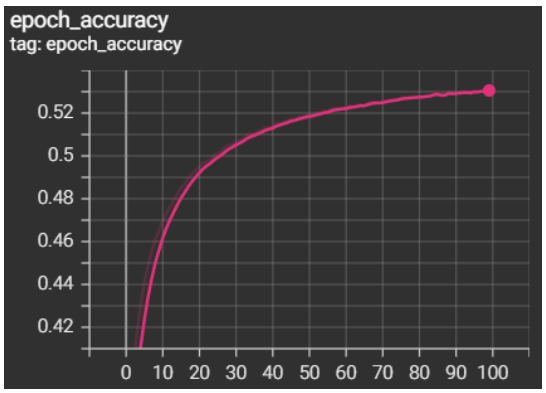


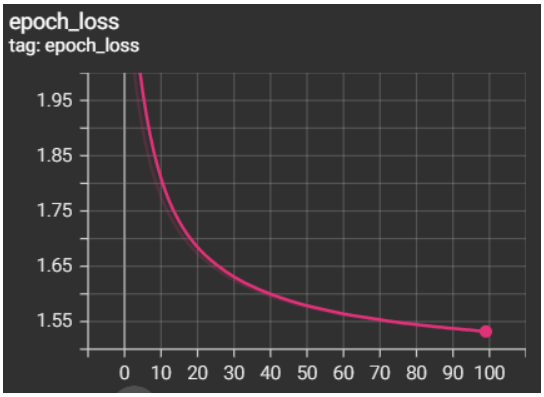


Generated text:

fare going pherecine
but i want to do
daddenings,slong time with you see those why,
i wond, iading me together ming me toom wore thinking long
unint there?
for it in the pelpt a for your mondan withenong or twe, a'd on oug, she's old me, when i'm gonna
betted mryter mind,
i want to do.
when i'm gonna bet

Simple RNN, hidden state size = 100, window size = 10, stride = 7

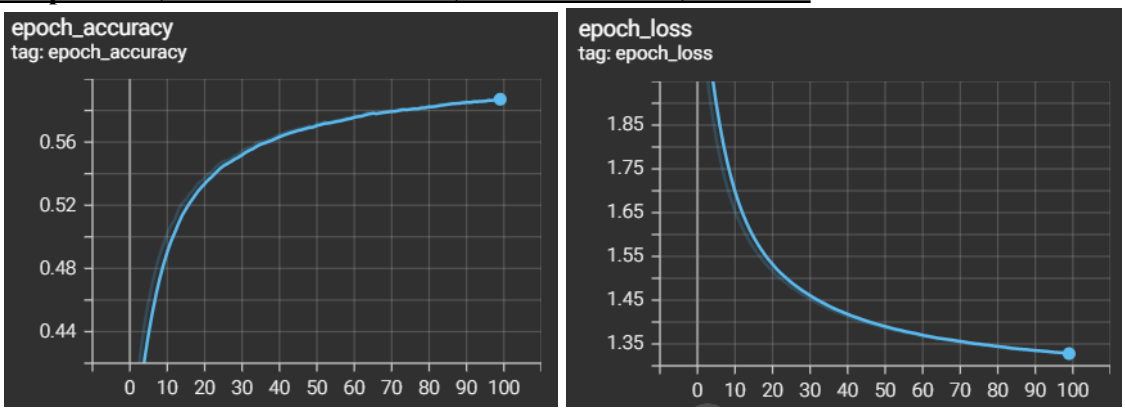




Generated text:

er.
living min the time and i weteystarts to kend bo taxwer, so lone you say hellon the nears co lose
the sand something to ses boy tight to need to take you know that you kill be the sand the waild,
whin i want to save to lest the sand thenne some furean i believe me toun andon't and i want to
have toough to

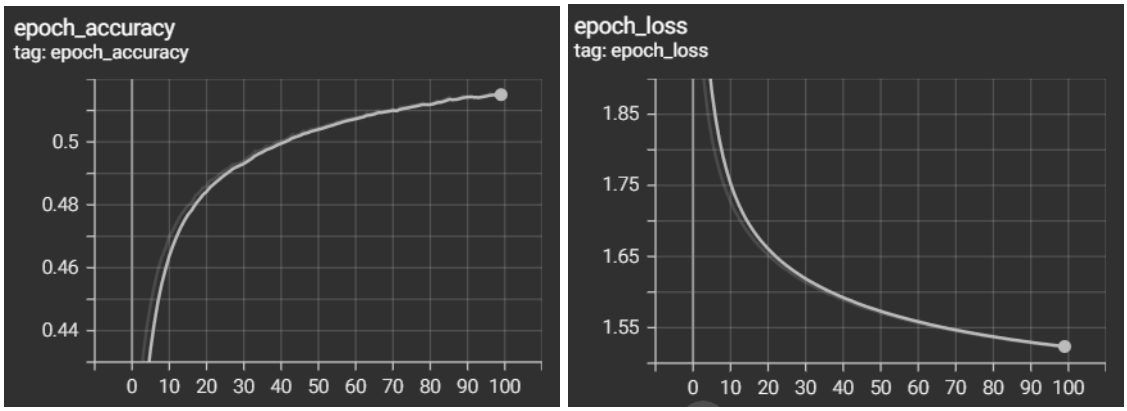
Simple RNN, hidden state size = 200, window size = 10, stride = 7



Generated text:

tend.
i want to know is are glade hade sleep pealy with a girl, you come on an ean you kee to me.
took just to yourselfer toome good don't dance whot i want it a weider the time
we don't care to sleeperby girl
fine her see me why you crytc night ih liok sum in mise i could nothing to huld your hand i don't ca

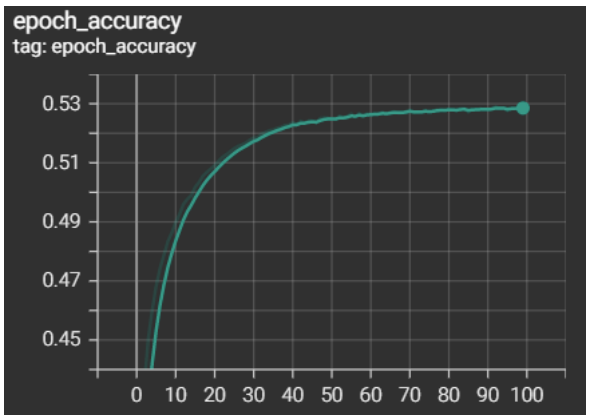
LSTM, hidden state size = 100, window size = 5, stride = 3

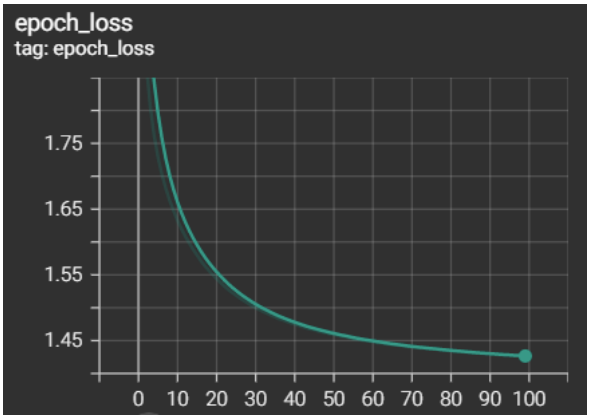


Generated text:

entary i say hello hold man to bes,
but i say too fun tonight nemmonking oun, taring
what you kill the sun snienve
inobaching in the sun is the ded you know itway to be joke, yould know mobores bie each one
come understand the sun shine
be boy the sun shne
helle me aight i can't bulled hart to be aise ov

LSTM, hidden state size = 200, window size = 5, stride = 3

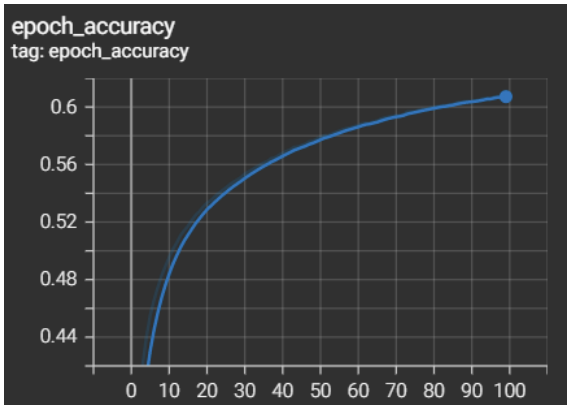


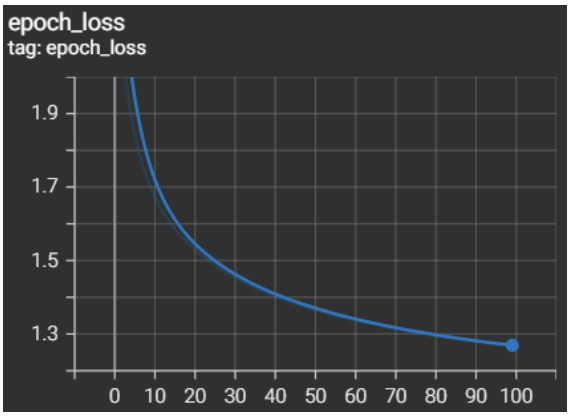


Generated text:

ine
we can see.
ifthere s nothqure in the said tonight bye
(byep bye baby cry.
and i fones born wond'ring me mine.
girl!
inded knees,
but it's gonna be you kidw lopr togeating for your monder what i don't kombad
it's gonna be you kidw lig, bang, maxy bad.
we'le sein.
see thas do you know where witches
sh

LSTM, hidden state size = 100, window size = 10, stride = 7

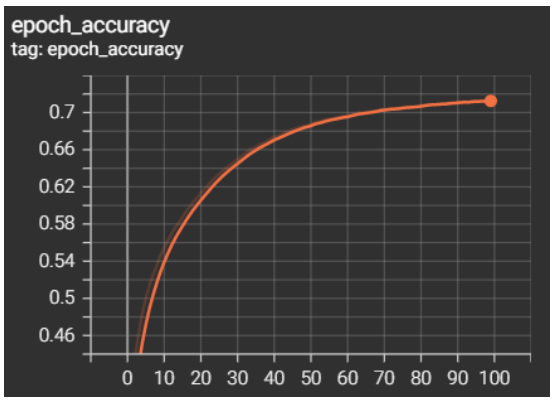


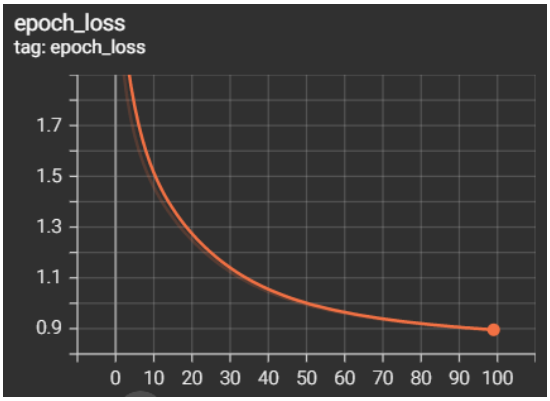


Generated text:

ner
oughta wall she's the one tramed to yello
shens time at answing me
it's been a luck are sones and shout)
yom something is tet a lucky youn hern
as boinsthen won't see me man getta da cul cry onemat a raways bedarstarday mymber
all tan the onder she look up the number
you don't wondy downing here i want yo

LSTM, hidden state size = 200, window size = 10, stride = 7





Generated text:

hey)
hey no rell
call dey lerry back to where you're doing to chance toub youh make a mighen')
hey, byb the worrt
my lovi
girl you know my name.(that my there your beligg and on apart.
you believe out yourselt
girl sare!h ould . hole woud i'ver shan.
the neggea thar eig that isn away
bade?
he'll le lear no he

Conclusion:

Ultimately, a few things were noticed. Firstly, the LSTM seemed to perform better than a simple RNN when all other parameters were the same, showing why LSTM seemed to take over simple RNNs in practice. Secondly, a larger window size provided for statistically significant improvements on loss and accuracy, and can also be seen as qualitatively better (although maybe not by as much) when examining the generated lyrics from the smaller window size vs larger window size.

One of the most interesting things learned through this project was the sampling temperature and its impact. It proved fairly difficult to find the line between generating the exact same lyrics over and over again versus creating new lyrics that were still understandably English. The sampling temperature proved to be fickle and extremely sensitive to change, as values such as 0.15 showed several cases of this undesired repetition while 0.25 at times produced incomprehensible language. For this reason, 0.2 was chosen, although it still produces a fair amount of non-english words.

Another interesting observation was that the models at times actually seemed to learn how to rhyme, creating lines of text that rhymed with the previous line (even with misspellings). However, this model is clearly not good enough to take the place of any of the Beatles songwriting (unless you use an extremely low temperature and cause significant repetition, which quite honestly seems more Beatles-esque than several real Beatles songs).

In future work we could attempt to prune the original lyrics to only include english words in an attempt to increase the performance.

How to run:

This code is run by executing the command

python main.py [input_file] [lstm/simple] [hidden_state_size (int)] [window_size (int)] [stride (int)] [sampling_temp (float)]

For ease of access the 8 commands used to generate the networks described above are as follows:

python main.py beatles.txt simple 100 5 3 0.2

python main.py beatles.txt simple 200 5 3 0.2

python main.py beatles.txt simple 100 10 7 0.2

python main.py beatles.txt simple 100 10 7 0.2

python main.py beatles.txt lstm 100 5 3 0.2

python main.py beatles.txt lstm 200 5 3 0.2

python main.py beatles.txt lstm 100 10 7 0.2

python main.py beatles.txt lstm 100 10 7 0.2