Tobby Lie
CSCI 5931
Programming Assignment 3 - **Task 1**
10/27/19

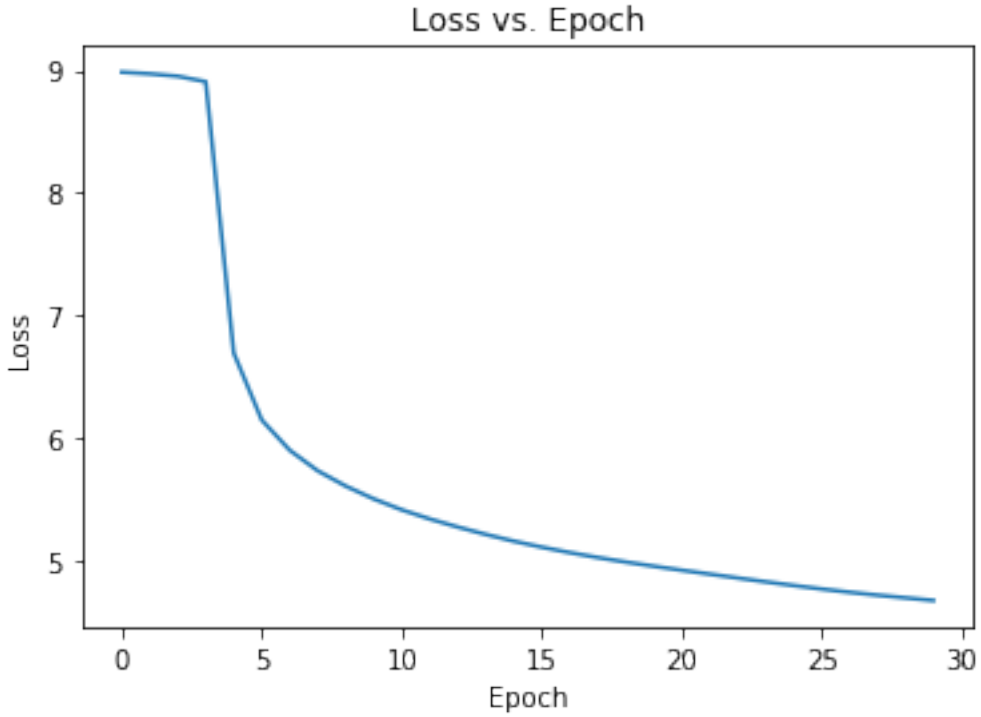

Figure 1: Loss vs. Epoch plot, training on 30 epochs with hidden unit size 100

Generated sequences from training on 30 epochs with hidden unit size 100 choosing 5 break points:

6 epochs:

```
Sentence # 1 of length 10
volcanoes handle it visibly francs ” to roof the .
Sentence # 2 of length 44
them. must their was of “ last , men of a of was commanding ferry sea loaf sea of the
quietly , . the . any their gaping in of open the rushed . that half-naked which
passed furious not unwilling visions by of
Sentence # 3 of length 8
parallel set of bush the in the by
Sentence # 4 of length 11
cheers complained bullet weighed ! could . underwent still the oval
Sentence # 5 of length 13
saurians halting prisms german the eider the sorrow . axe deed the forked
Sentence # 6 of length 9
mobilis articles the , of , known employing .
Sentence # 7 of length 7
brandishing had had the mechanics allowance for
Sentence # 8 of length 13
peering caused land formerly soul gusts puzzled advancing of of its by to
Sentence # 9 of length 9
```

```
lively to-day of divide which did , nation .
Sentence # 10 of length 19
midday kindled desire sensed . ? the only to pupils moderate of position the somber of
surface , pretend
```

12 epochs:

```
Sentence # 1 of length 21
calmly ” to from into a a any resounded is from shouted much , loud above car of rent
of feet
Sentence # 2 of length 7
finger any of empty and out .
Sentence # 3 of length 11
the being must everybody occasion , did in bill fall .
Sentence # 4 of length 18
reflecting parallel car a after whales round difficult to ? and its only of to great
part .
Sentence # 5 of length 8
north everything was fact as falls the .
Sentence # 6 of length 17
the and that feet four the overboard balloon to of a were rise royal car mingled as
Sentence # 7 of length 7
repay it is seeking was they car
Sentence # 8 of length 10
principally no horizontal trimmed waves firmly , beneath that implacable
Sentence # 9 of length 15
see it was a did of surface to their still ” everything beneath lower above
Sentence # 10 of length 14
get olive balloon the it them industry a , of time voyagers , .
```

18 epochs:

```
Sentence # 1 of length 42
alarmed was steamers ” to from into a a any resounded is from clouds , above hundred
arrest which a midday an emotions taking occasion , did in collapsed the a island fall
slowly being thrown the greatest or weight words .
Sentence # 2 of length 14
bah was thus efforts having to , voyagers solid , longest car mingled as
Sentence # 3 of length 23
repay it is distance ” began from ” now from of , within useless minute its five the
but whole or no .
Sentence # 4 of length 18
it gas faithfully 18th few they of but above the than of the meshes must the sextant
above
Sentence # 5 of length 28
the enterprise surely dreadful sunken their vanish any from , , was luggage movement
everything celebrated feeling be into movement course which that march of a , .
Sentence # 6 of length 23
its o was all calm removed ! from it be articles elements which walls there ocean only
of 2,000 of the from .
Sentence # 7 of length 30
it position masters this could were was and evident that that the began , ’ zigzags
away everything , above be everything traces white of men 18th changed “ in
Sentence # 8 of length 9
vessels voyagers ! hold thanks is evident west ”
Sentence # 9 of length 20
few threw arms be none 2,000 heavy arguments feet say to not escaped , vanish ”
indicate the the .
Sentence # 10 of length 40
```

vinegar above , hungry noise position a voice determined of be wind now night shells did do america 2,000 visible the shouted of about have great and the float as and ? single to regard a beneath prevent as .

24 epochs:

Sentence # 1 of length 35
alarmed was not out ' shapes " to from into a do resounded is from clouds , above hundred arrest which a over passengers the voice of admired taking occasion , fear beneath that .
Sentence # 2 of length 22
" a it single passage be produce can that still were sensible the part did the without , extended raged prevent .
Sentence # 3 of length 11
but were rise voice could caresses was is the must ropes
Sentence # 4 of length 15
the loud at air in o 1865 the even , after " downward streaming .
Sentence # 5 of length 14
the can struggle " surface a firmly , beneath that four they and unlikely
Sentence # 6 of length 21
a were " arches sudden " fortieth the net , spray their vanish any from , , of to luggage .
Sentence # 7 of length 23
insupportable 23rd on its which before more hard removed ! from it to serpents there ocean only of 2,000 of the from .
Sentence # 8 of length 28
it position masters this could were which provisions more that the began , ' zigzags away everything , above be failing , a " changed " in in
Sentence # 9 of length 13
but feet , out none 2,000 must the sea of into appeared .
Sentence # 10 of length 75
the orifice of long , but the informed of by it trust the absorbed feet say to not escaped , vanish " the higher , it again death of region few o delicate the voyagers of distance above , hungry noise position a voice determined of " sometimes night this prevent and ? single to regard a beneath death work elements ' of through sea held them the midst a removed themselves the , .

30 epochs:

Sentence # 1 of length 27
" be is fact in remarkable heavy articles more escaped the to , and was the for fall is from without by to of of , .
Sentence # 2 of length 29
the struggle " to rate attentive , loud above infallibly passengers the voice of admired taking destroyed several , did in collapsed " island fall slowly being thrown .
Sentence # 3 of length 16
part watery parallel struggle rise heart round waterspouts plunged " the gas , the voice .
Sentence # 4 of length 24
the dress was is the rise risen in o 1865 the sensibly was into land feet was useless minute its five the shoulders .
Sentence # 5 of length 20
it after a was ammunition , beneath that four they above the than of suspended storm surface still without .
Sentence # 6 of length 13
the covering like cut pot the be had mathematical us moderate their .
Sentence # 7 of length 13
notion to so " soon articles could in into of the difference .

```
Sentence # 8 of length 20
the from , , underfoot movement struggling feeling be into movement course , on that
march of a , .
Sentence # 9 of length 10
on was the no of sides sea 2,000 the occasion
Sentence # 10 of length 9
“ can not and weighed since the its .
```

Gradient Checking:

When implementing back-propagation it is useful to implement gradient checking as well in order to verify that the implementation is correct. The idea is that the derivative of a parameter is equal to the slope at the point which we can approximate by slightly changing the parameter and dividing by the change. When we compare the gradient calculated using back-propagation to the gradient estimated, if there is no large difference then we are in a good place. The approximation calculates the total loss for every parameter so it is very expensive which is why it is a good idea to perform it on a model with a smaller vocabulary.

The output from gradient checking:

```
Performing gradient check for parameter U with size 1000.
/usr/local/lib/python3.6/dist-packages/ipykernel_launcher.py:30: RuntimeWarning:
invalid value encountered in double_scalars
Gradient check for parameter U passed.
Performing gradient check for parameter V with size 1000.
Gradient check for parameter V passed.
Performing gradient check for parameter W with size 100.
Gradient check for parameter W passed.
```

For U, V and W all with size 1000, 1000 and 100 respectively passed the gradient check.

200 Hidden Units:

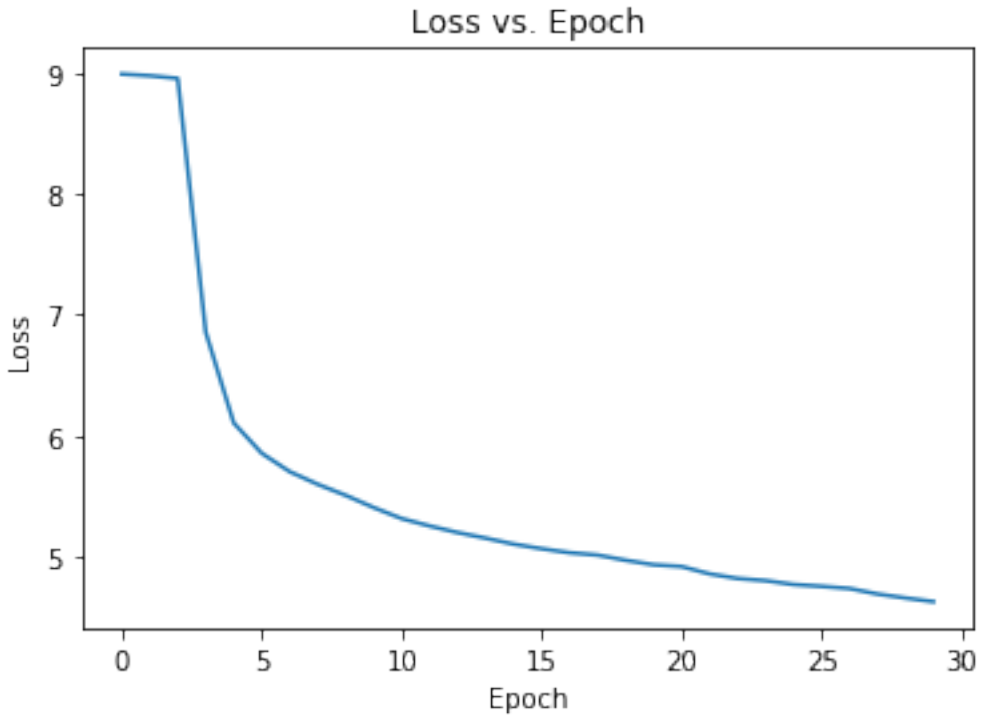

Figure 2: Loss vs. Epoch plot, trained on 30 epochs with hidden units size 200

Text Sampling Results:

```
Sentence # 1 of length 12
it men be murmur forth have in last to the ammunition of
Sentence # 2 of length 13
the jules of of the for ” the lower , and five .
Sentence # 3 of length 8
atmosphere watery has midday a prolong , all
Sentence # 4 of length 8
but , had in this easily disheveled soon
Sentence # 5 of length 8
they were attempts men voyagers ” aloft position
Sentence # 6 of length 43
the were was of more car having branch wide menacing top energetic flour of , rise ,
above the position of the slightest , the out , the precision , the car of the
possibility , ninety of definitely secure the ocean .
Sentence # 7 of length 21
they their accompanied limits rise , beneath above were after which voyagers ? , last
throw districts ? the nor .
Sentence # 8 of length 8
the were voice seen “ swim suspended ”
Sentence # 9 of length 8
towards five moment loud voice balloon gas .
Sentence # 10 of length 12
a jules was had from column how two rotation to suspension float
```

50 Hidden Units:

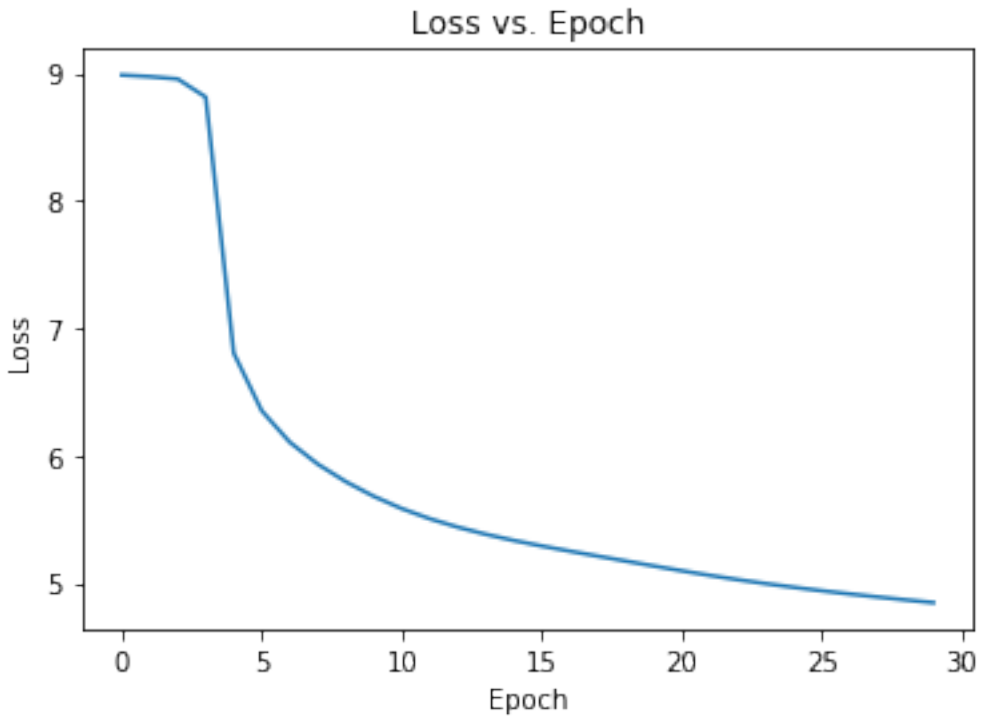

Figure 3: Loss vs. Epoch plot, trained on 30 epochs with hidden units size 50

Text Sampling Results:

```
Sentence # 1 of length 13
but ” empty out at of the warned of to clock someone .
Sentence # 2 of length 14
again beneath above present loud without hovering was ! . , something are .
Sentence # 3 of length 15
cure was in , vapor like stopping an in the with had the four .
Sentence # 4 of length 31
asia rise of that reykir to of extending was ” “ empty water to preliminary men was
loud pictured front with stationed out ? to sea wind to mounted a .
Sentence # 5 of length 16
whaling masterpieces was abyss other in and hundreds through the kept were upon the
extending .
Sentence # 6 of length 10
rifle was extended night , with of thick ” .
Sentence # 7 of length 9
but not charles immediately at away the balloon .
Sentence # 8 of length 19
demanded must fury be the being they “ which and ! , car delirious if is leather they
risen
Sentence # 9 of length 10
the sea to of life away which the urgent .
Sentence # 10 of length 24
it their the balloon from not if ! what the lower of regions held which the confound ,
the balloon of the tenants .
```

Discussion for part a:

In part a I experimented with doubling and halving the number of hidden units in our network. After training I plotted Loss vs Epoch graphs for each case and then generated text samplings from each case.

When observing the training, it took approximately 10 minutes to train on 30 epochs with hidden dimension of 200 and training on a subset of the data (0-99). For a hidden dimension of 50 and the all the other details remaining the same, it took approximately 3 minutes to train. This means that reducing the hidden dimension size results in a faster training time.

I also observed that for a hidden dimension of 200, after 30 epochs of training, the loss was reduced from 8.988 down to 4.6308 meaning a difference of 4.3572. For a hidden dimension of 50, the loss was reduced from 8.987 down to 4.8459 meaning a difference of 4.1411. The hidden dimension of 200 yielded a greater reduction of loss meaning a larger hidden dimension correlates to better reduction of loss.

In regards to the generated text samplings from each model it seemed that the model with hidden dimension of 200 produced more coherent text. Most notably it seemed that it was better than the model with hidden dimension 50 at determining punctuation and syntax. The words that appeared in the sequences also appeared to make slightly more sense in the hidden dimension of 200. A greater difference could probably be perceived with more epochs in training and with a greater difference in hidden dimension sizes.

Extended sequence length by concatenating two consecutive sentences together to form a longer sequence:

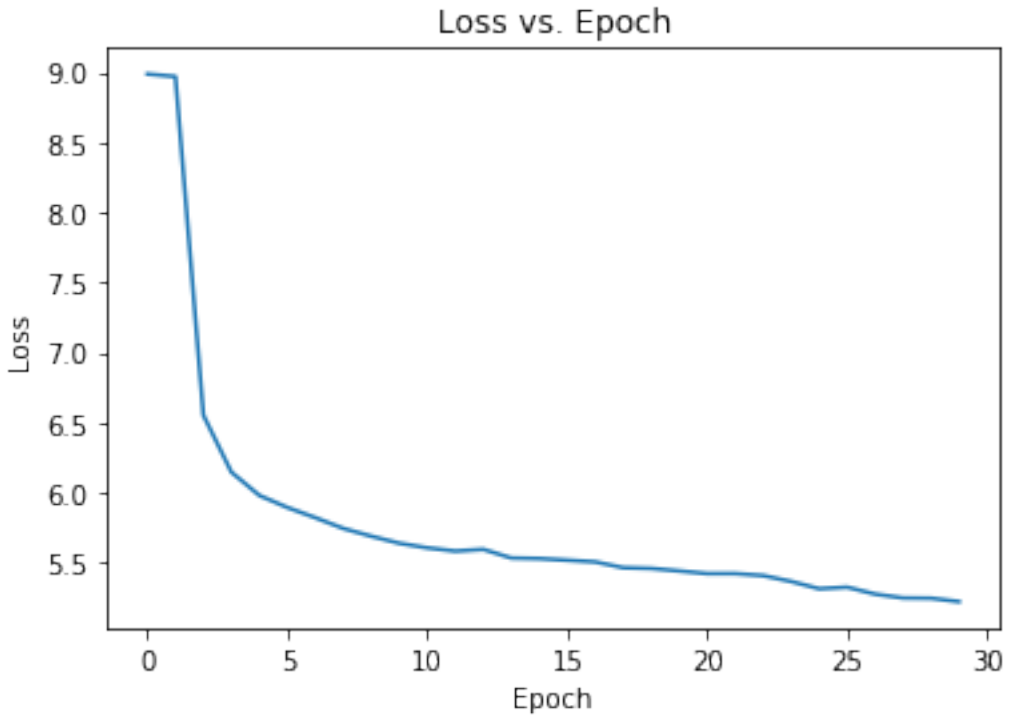

Figure 4: Loss vs. Epoch plot, trained on 30 epochs with hidden units size 100 and approximately double the length of sentence as original

Text Sampling Results:

```
Sentence # 1 of length 18
and have no regard by , sea be sustained , , such and in of that it zones
Sentence # 2 of length 8
nothing. immense the in to passed lighter .
Sentence # 3 of length 10
of their was is was , to energetic . .
Sentence # 4 of length 7
pierce at it to whole and in
Sentence # 5 of length 8
bolts sea it plaything , watercourse , .
Sentence # 6 of length 14
herds the havana to was eighteen terrible . , say horizontal if from part
Sentence # 7 of length 7
does extending last the atmosphere were .
Sentence # 8 of length 8
moderate was they balloon endeavored of resounded the
Sentence # 9 of length 13
only have removed lengthening to weight disasters from without less passengers . ,
Sentence # 10 of length 11
round passengers people which an of storm balloon hours the expedition
```

Shortened sequence length by halving each sentence to form a shorter sequence:

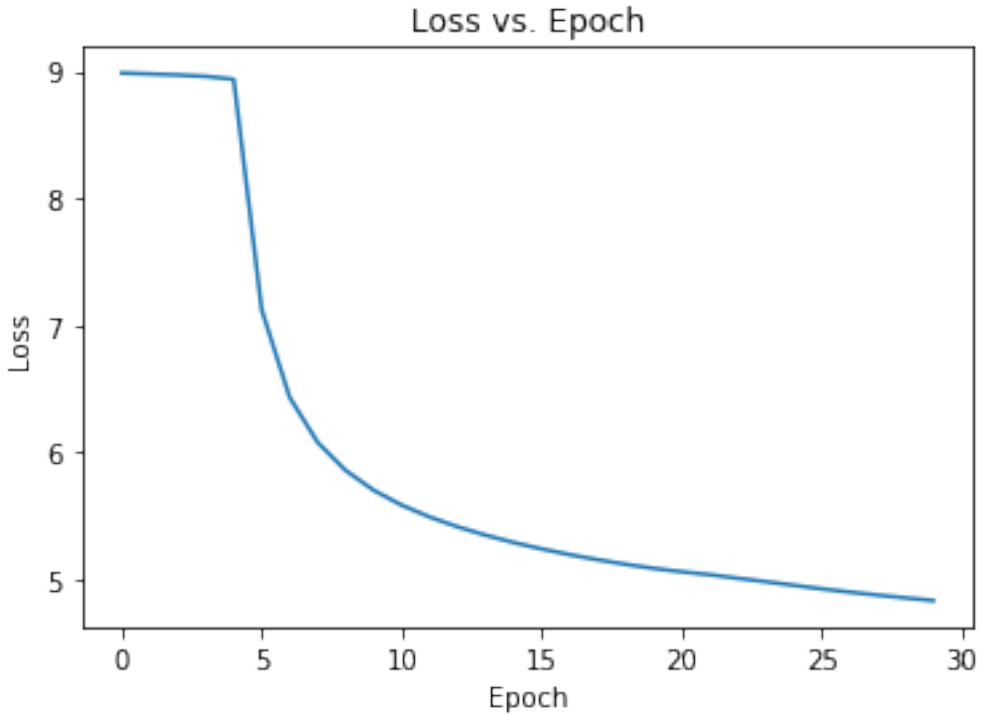

Figure 5: Loss vs. Epoch plot, trained on 30 epochs with hidden units size 100 and half the length of sentence as original

Text Sampling Results:

```
Sentence # 1 of length 8
dropped ! could . movement descending their as
Sentence # 2 of length 7
in tempest , it have of that
Sentence # 3 of length 7
, , . of a horizontal to
Sentence # 4 of length 7
it feet to of only balloon could
Sentence # 5 of length 12
the ” of the their is noise from dashing was low zones
Sentence # 6 of length 11
accelerated were while traversed alone every rate a taken to a
Sentence # 7 of length 14
language we ” the , sensible vast while balloon year changed . in the
Sentence # 8 of length 7
planted diminished the solidity did and equator
Sentence # 9 of length 15
suspecting was discovered their hours may which less thousand taken hauling hours was
day call
Sentence # 10 of length 14
were is the only miles lofty some could ; beneath the waves of the
```

Discussion for part b:

In part b I experimented with doubling and halving the sentence length being fed to our network. After training I plotted Loss vs Epoch graphs for each case and then generated text samplings from each case.

When observing the training, it took approximately 12 minutes to train on 30 epochs with double the sentence length as our original sentence length and training on a subset of the data (0-99). For a hidden dimension of 50 and the all the other details remaining the same, it took approximately 3 minutes to train. This means that reducing the sentence length results in a faster training time.

I also observed that for the doubled sentence length training, after 30 epochs of training, the loss was reduced from 8.987 down to 5.2964 meaning a difference of 3.6906. For a halved sentence length, the loss was reduced from 8.987 down to 4.9126 meaning a difference of 4.0744. The halved sentence length yielded a greater reduction of loss meaning a shorter sentence length correlates to faster reduction of loss. This was confusing to me as I believed initially the inverse to be true, that having more input data to work with would yield in more minimization of loss.

In regards to the generated text samplings from each model it seemed that the model with halved sentence length produced more coherent text for longer sequences showing more consistent coherence. The words that appeared in the sequences also appeared to make slightly more sense in the halved sentence length. A greater difference could probably be perceived with more epochs in training and with a greater difference in hidden dimension sizes. This outcome I also did not expect to happen, similar to the previous paragraphs findings.