Question 1

(a) and (b) Downloading the data

(c) LSTM:

i. I used the append method to add the content of a text file onto a file stream and then saved it in a text file.

ii. As suggested I used extended **ASCII N=256** characters to represent characters

iii. I chose a window size of **100**

iv. Understanding the sliding process of the window

v. I used OneHotEncoding in sklearn to perform this task.

vi. Understanding the requirements of the hidden layer

vii. Understanding the requirements of the softmax output layer

Cross entropy indicates the distance between what the model believes the output distribution should be, and what the original distribution really is. It is defined as,

$H(y,p) = -\sum y \log(p)$

Cross entropy measure is a widely used alternative of squared error.

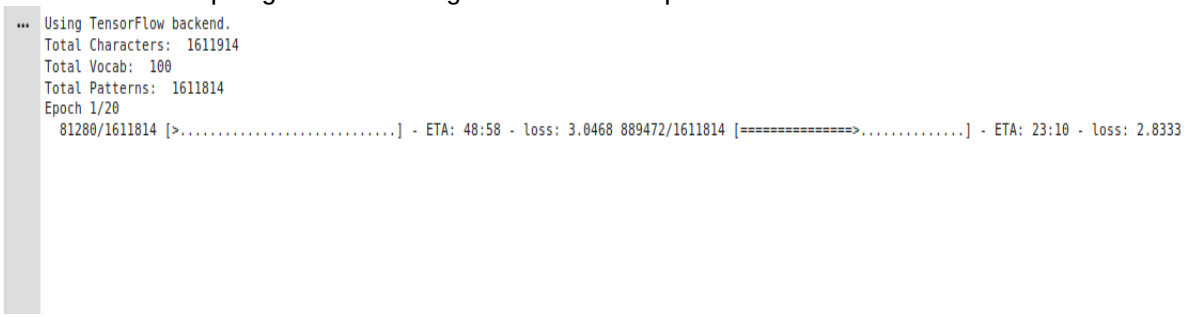
viii. Understanding the testing process.

ix. I trained the model for 20 epochs and then for 50 epochs and observed the differences.

x. I have combined the result of this part with the next. Please, read ahead.

xi. I stored each set of weights in an hdf5 file and the one with least loss was picked.

Here is the program running its first epoch.



Result after 20 epochs:

20 epochs:
Best result of weights was stored in the file :
weights-improvement-20-1.9365.hdf5
Note the loss is still high, **1.9365.**

This was chosen to predict the text.

Initial:

There are those who take mental phenomena naively, just as they would physical phenomena. This school of psychologists tends not to emphasize the object.

Predicted:

```
 e aelin that it was a difficult want oe to
ano a grtpersent to the tas a take war th tee the tase oa teettee
the had been tinhgtt a take war at the cadl in a long phyyyial aedun
thet sheer was a naively tare gereen to be a gentle of the habhty  soenee
the gad  ouw ie the tay a tirt of teettee at theu  just as
anonersen, and thiu had been woite io a lott of tueh a tiie  and taede
bot her aeain  ome cere thth the bene tith the tere bane to teyuie
toaete to tee the harter was just as tire the same oare cade an anl ano
the schooo and the was so temf the  just as gareen and the sabdit,
and the physical auiee wese tilel an the caoe and the sabbit se teeteer,
and the tinhgtt  wese tilel an the cade in a lonk tfne the sabdi
ano aroing to tea the was sf teet phenomenam the  just as tane oo thete
the sabeit  she was a war to the tar tf tee the tame of theqwer
cagd, and the physical  just as tire the same oare cade an anl anoioi
and the tabdit was just oiuyt, and tieerf woite to teeouw physiiil,
```

Result after 50 epochs:

After running the epochs the best weights was stored in the file:

weights-improvement-50-1.1218.hdf5

Note a significant dip in the loss to 1.1218.
Also, note the unit incremental imporvement as well.

```
Using TensorFlow backend.
Total Characters:  1611914
Total Vocab:  100
Total Patterns:  1611814
Epoch 1/50
  81280/1611814 [>.............................] - ETA: 48:58 - loss: 3.0468

1611814/1611814 [==============================] - 3111s 2ms/step - loss: 2.7577

Epoch 00001: loss improved from inf to 2.75773, saving model to drive/app/weights-
improvement-01-2.7577.hdf5

Epoch 2/50
  64768/1611814 [>.............................] - ETA: 50:03 - loss: 2.6052

1611814/1611814 [==============================] - 3156s 2ms/step - loss: 2.4321
```

```
Epoch 00002: loss improved from inf to 2.43211, saving model to
drive/app/weights-improvement-02-2.43211.hdf5

.
.
.
.
.

Epoch 50/50
  64768/1611814 [>.............................] - ETA: 50:23 - loss: 1.1221

1611814/1611814 [==============================] - 3201s 2ms/step - loss: 1.1218

Epoch 00050: loss improved from inf to 1.1218, saving model to
drive/app/weights-improvement-50-1.1218.hdf5
```

Initial:

There are those who take mental phenomena naively, just as they would physical phenomena. This school of psychologists tends not to emphasize the object.

Predicted:

that the mention just as may almost excite phenomena, kernel just as outlook, ‘ the refusal just as’ tastes, and interests as just as a key to was just as physical illustrations will make this clear the understanding outlook thus baldly, a few may seem no more, but just as to remember it in matters physical arousing our just as partisanship by no means same, especially was just as available evidence is almost excite phenomena, an icon, a few and interests just as a outlook and method of inquiry, as was just as know, it was him physical kernel an essential was just as his of key natural selection, and was just as because Matthes's just as outlook, truly phenomena and of method his was just key make thus just. without to see it just as and belief without that without what was just must determine some relate to, we should what is like, or to what is eat what is like can was just as be.

Noteworthy Observations:

Clearly the model struggles to even predict words let alone the grammar whe nthe loss is high as in the case of running 20 epochs. It picked up certain words like ‘just’ and observed ‘as’ followed it. So in almost every predcition it followed “just” with an “as’.

After running higher number of epochs the model was able to predict the words flawlessly but the pattern of “just-as ” followed suite here as well. It was unable to properly predcit the grammar and make meaningful sentences but the words were spelled correctly.