Natural Language Processing - IMDB Movie Review

	Description	Hyperparameters	Number of Epochs	Training Loss	Training Accuracy	Test Accuracy	Comments
Part 3a	Given Model - (Reset_state + 3 Lstm + 3 Batch Norm + 3 Dropout + Output Layer)	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500, Sequence_len=50	75	3.96	25.74	28.58	The accuracy is very low since the task of predicting the next word is far more complicated than predicting sentiment for the entire phrase. One training epoch takes around 550 seconds.
Part 3b	Generated Review	Temperature=1.0	**i love this movie** . a definite collector 's worst of all ! ! ! ! ! ! ! and , now , i 'm addicted to new jersey , but everything else but the script seems better written and above the whole film . there are some tedious twists , slow motion , etc . nothing special . ally is her lovable self - so , the only characters in it stand that great . i hope the movie could have been made better . my negative rating was that me and i watched it . maybe if i am not a mature person , i would like to know if bruce lee was too talented since he has done it in the past scene . he is coming from an alternate place and is quite annoying in his portrayal of a small boy but the woman who she played . overall , a **i hate this movie** . i wore when it came out , and it put me to sleep and i just got the impression that the angry moral guy boy at a dream , said that he has a deeper problem , and comes back to life to get killed . what a robin leigh i was thinking was n't going to be clear . for me he did n't disappoint me and i hate this film . it is normal , touching , true . i laughed at the last half hour and my friend had her too long forward to her and you were stunned one woman just blew her heart or threw the money together . i liked the 10 's that eva starts to live up to her words , and may someone else enjoy her soul was very authentic . i find many lessons against the series , both shocking	Using temperature=1.0 is the same as the regular softmax function. The reviews generated are readable and we are able to understand what the main ideas are. Both reviews are relatively negative and the sentiment is not maintained throughout the generated reviews. It is hard to determine if they are generated automatically or written by human.
	Generated Review	Temperature 0.25	**i love this movie** . it is a very good movie . the acting is great , the story is great , and the acting is great . the story is very well written and the actors are excellent . the story is very well written and the actors are great . the story is very well written and the actors are excellent . the story is very interesting and the characters are well developed . the story is very interesting and the characters are well developed . i think the director has a lot of potential . the movie is a little slow but the story is very good . i recommend this movie to everyone . i would recommend this movie to anyone who wants to see a good movie . i hope that the next day i will see it again . i hope it will be available on dvd . i **i hate this movie** . it 's a shame that the director is trying to make a movie about a man who is a little boy . he is not a bad guy , but he is n't a bad guy . he has a very good role and is a good actor . i do n't know if he was a good actor or a good actor but he was a good actor . i think he was a good actor . i also liked the way he played the lead character . i think that he was a good actor and i liked the movie . i think it is a good movie for kids . i think it is a great movie to watch with friends . i would recommend it to anyone . i hope that they will get more money to make a sequel . i would love to see	As the temperature decreases, the most likely word will approach a probability of 1.0. Using low temperature=0.25 causes looping reviews. The sentiment of the positive review is maintaned but the sentiment of the negative review is not maintained. Hence, it is easy to tell that they are fake reviews.
Part 3c	Given Model - (Reset_state + 3 Lstm + 3 Batch Norm + 3 Dropout + Output Layer) Fine-tune last LSTM layer and output layer.	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500, Sequence_len=100	20	0.1167	95.51	89.06	Fine-tunning only the last LSTM and output layers gives pretty good result with a little overfit. Better results may be achieved using GloVe features.
	Given Model - (Reset_state + 3 Lstm + 3 Batch Norm + 3 Dropout + Output Layer) Fine-tune output layer.	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500, Sequence_len=20	20	0.6207	65.4	70.94	Fine-tunning only output layer and significantly lowerer sequence_length would make model underfit since the model is not able to capture enough features. Both training and test accuracy are much lower.

Given Model - (Reset_state + 3 Lstm + 3 Batch Norm + 3 Dropout + Output Layer) Tune all layers.	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500, Sequence_len=100	20	0.0804	97.15	86.27	Tunning all the layerers would make model overfit(high training accuracy, low test accuracy). More parameters are learned to fit the trainning data and make it harder to generalize on the test data.