Natural Language Processing - IMDB Movie Review							
	Description	Hyperparameters	Number of Epochs	Training Loss	Training Accuracy	Test Accuracy	Comments
Part 1a - BOW without GloVe Features	Given model - Word Embedding Layer + Mean Pooling + Fully Connected Layer + Relu + Output Layer	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500	6	0.1403	94.64%	86.13%	This is the standard model given in the homework.
	Custom 1 - Word Embedding Layer + Mean Pooling + Fully Connected Layer + Relu + Output Layer	ADAM optimizer with LR=0.01, BatchSize=200, VocabularySize=8000, HiddenUnits=4000	6	0.0862	96.22%	85.02%	I increased the number of hidden units for a huge amount. We can see the training accuray increase largely but test accuracy decreased due to overfit.
	Custom 2 - Word Embedding Layer + Mean Pooling + Fully Connected Layer + Relu + Output Layer	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=150	6	0.2031	93.01%	73.47%	I decreased the number of hidden units by a large amount.The model suffer from underfit but it was not as bad as I thought it would be.
Part 1b - BOW with GloVe Features	Given model - GloVe Features + Fully Connected Layer + Relu + Dropout + Output Layer	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=100000, HiddenUnits=500	6	0.3254	85.03%	86.26%	This is the standard model given in the homework.
	Custom 1 - GloVe Features + Fully Connected Layer + Relu + Dropout + Output Layer	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=100000, HiddenUnits=500	60	0.1335	95.12%	98.31%	I increased the number of training epochs by a large amount. Both the training accuracy and the testing accuracy increased dramatically.
	Custom 2 - GloVe Features + Fully Connected Layer + Relu + Dropout + Output Layer	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=200000, HiddenUnits=500	6	0.2832	85.81%	87.46%	I tried to increase the size of the vocabulary, but it does not significantly change the accuracy in this case. I believe it is because we include many vocabularies that do not appear many times.
Part 2a - RNN without Glove Features	Given Model - Embedding layer + LSTM units + BactchNorm1d + DropOut + Fully Connected	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500, SequenceLength = 100	20	0.1173	95.70%	86.86%	This is the standard model given in the homework.
	Custom 1 - Embedding layer + LSTM units + BactchNorm1d + DropOut + Fully Connected	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=1000, HiddenUnits=500, SequenceLength = 100	20	0.4356	90.38%	80.12%	I decrease the vocabulary size and both the training accuracy and the testing accuracy decrease.
	Custom 2 - Embedding layer + LSTM units + BactchNorm1d + DropOut + Fully Connected	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=500, HiddenUnits=500, SequenceLength = 100	20	0.3698	83.11%	74.27%	I decrease the vocabulary size and both the training accuracy and the testing accuracy decrease.
Part 2b - RNN with Glove Features	Given Model - GloVe Features + LSTM units + BactchNorm1d + DropOut + Fully Connected	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500, SequenceLength = 200	20	0.3586	93.05%	90.85%	This is the standard model given in the homework.
	Custom 1 - GloVe Features + LSTM units + BactchNorm1d + DropOut + Fully Connected	ADAM optimizer with LR=0.01, BatchSize=200, VocabularySize=8000, HiddenUnits=500, SequenceLength = 25	20	0.3032	87.49%	86.80%	I decreased the SequenceLength of the model and I notice a decrease in training and test accuracy.
	Custom 2 - GloVe Features + LSTM units + BactchNorm1d + DropOut + Fully Connected	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=50000, HiddenUnits=500, SequenceLength = 250	20	0.1345	95.08%	87.49%	The increase in SequenceLength of the model cause an increase in training accuracy and an decrease in test accuracy, which indicate an overfitting occure.