anguage Processing - IMDB Movie Review							
	Description	Hyperparameters	Number of Epochs	Training Loss	Training Accuracy	Test Accuracy	Comments
Part 1a	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	0.1404	97.46%	87.03%	This one has more hidden units than custom2. Therefore, its training accuracy grew faster, so it influences testing accuracy.
	Custom 1	HiddenUnits=500 , Reomve Dropout	10	0.021	99.42%	85.50%	Since I set more hidden layers and removed droupout, the result behaved like overfitting.
	Custom 2	HiddenUnits=200 , Dropout = 0.5, weight decay as 10^-6	10	0.0709	97.38	85.76%	The overfitting is not less serious than custom1, but the test accuracy is as good as the original 1a.
	Custom 3						
Part 1b	Given Model - Fully Connected Layer + Relu + Output Layer	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500	6	0.3006	87.23%	84.24%	Given model does not show overfitting problem
	Custom 1	HiddenUnits=1000 , Reomve Dropout	11	0.1698	93.55%	84.74	Since I set more hidden unites amd removed dropout. Overfitting happened.
	Custom 2	HiddenUnits=200 , Dropout = 0.5, weight decay as 10^-6	11	0.2848	88.20%	85.44%	Although I lessoned hidden layers and added weight decay, this model's overfitting problem still more serious than given model's.
	Custom 3						
Part 2a	Given Model - Word Embedding Layer + LSTM Layer + Max Pooling + Fuly Connected Layer + Output Layer	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500, sequence_length = 100,	20	0.0808	97.23%	87.33%	This mode used RNN, but the test was not better than 1a.
	Custom 1	HiddenUnits= 1000, sequence lenth = 50, Dropout = 0.5	20	0.0813	97.38%	88.48%	I increased the HiddenUnits, but overfitting did not happen. The possible reason was that I lower the sequence length.
	Custom 2	HiddenUnits = 100 sequence = 200 batch number = 100	20	0.0808	97.20%	87.45%	I lessoned the HiddenUnits and the result was similar to the given model's.
	Custom 3						
Part 2b	Given Model - LSTM Layer + Max Pooling + Fuly Connected Layer + Output Layer	ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=10000, HiddenUnits=500, sequence_length = 100,	20	0.2129	91.36%	90.28%	This model showed a little bit unerfitting problem.
	Custom 1	HiddenUnits= 1000, sequence lenth = 50, Dropout = 0.5	20	0.367	83.34%	90.82%	This model showed serous underfitting problem. The possible reason was that I used shorter sequence length.
	Custom 2	HiddenUnits = 100 sequence = 200 batch number = 100	20	0.3942	82.22%	88.66%	This model did not showed overfitting or underfitting. But test accuracy was not better than 1b too much.
	Custom 3						