nguage Processing -	IMDB Movie Review		<u> </u>				
	Description	Hyperparameters	Number of Epochs	Training Loss	Training Accuracy	Test Accuracy	Comments
	Given model - Word Embedding	ADAM optimizer with LR=0.001,		3	j	,	This one has more hidden units
	Layer + Mean Pooling + Fully	BatchSize=200,					than custom2. Therefore, its
-	Connected Layer + Relu + Output						training accuracy grew faster, so
Part 1a	Layer	HiddenUnits=500	6	0.1404	97.46%	87.03%	it influences testing accuracy.
							Since I set more hidden layers
		HiddenUnits=500, Reomve					and removed droupout, the resu
	Custom 1	Dropout	10	0.021	99.42%	85.50%	behaved like overfitting.
	- Custom 1	Второш	10	0.021	30.4270	00.0070	The overfitting is not less serious
							than custom1, but the test
		HiddenUnits=200 , Dropout = 0.5,					accuracy is as good as the
	Custom 2	weight decay as 10^-6	10	0.0709	97.38	85.76%	original 1a.
	Custom 3						
	Guotom G						
		ADAM optimizer with LR=0.001,					
		BatchSize=200,					
	Given Model - Fully Connected	VocabularySize=8000,					Given model does not show
Part 1b	Layer + Relu + Output Layer	HiddenUnits=500	6	0.3006	87.23%	84.24%	overfitting problem
	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		-				Since I set more hidden unites
		HiddenUnits=1000 , Reomve					amd removed dropout. Overfitting
			4.4	0.4000	00.550/	0.4.7.4	
	Custom 1	Dropout	11	0.1698	93.55%	84.74	happened.
							Although I lessoned hidden laye
							and added weight decay, this
		HiddenUnits=200, Dropout = 0.5,					model's overfitting problem still
	Custom 2	weight decay as 10^-6	11	0.2848	88.20%	85.44%	more serious than given model
		weight decay as 10°-6		0.2046	00.2076	65.44%	more serious triair giveri moders
	Custom 3						
		ADAM optimizer with LR=0.001,					
	Given Model - Word Embedding	BatchSize=200,					
	Layer + LSTM Layer + Max	VocabularySize=8000,					
							This made used DNN but the t
	Pooling + Fuly Connected Layer	HiddenUnits=500,					This mode used RNN, but the t
Part 2a	+ Output Layer	sequence_length = 100,	20	0.0808	97.23%	87.33%	was not better than 1a.
							I increased the HiddenUnits, but
							overfitting did not happen. The
		HiddenUnits= 1000, sequence					possible reason was that I lower
	Custom 1		20	0.0813	97.38%	88.48%	
	Custom 1	lenth = 50, Dropout = 0.5	ZU	0.0013	31.30%	00.40%	the sequence length.
							I lessoned the HiddenUnits and
		HiddenUnits = 100 sequence =					the result was similar to the give
	Custom 2	200 batch number = 100	20	0.0808	97.20%	87.45%	model's.
	Custom 3						
	0.00.00				†		
	+	ADAM antimizer with LD 0 004			+		
		ADAM optimizer with LR=0.001,					
		BatchSize=200,					
	Given Model - LSTM Layer + Max	VocabularySize=10000,					
	Pooling + Fuly Connected Layer	HiddenUnits=500,					This model showed a little bit
Part 2b	+ Output Layer	sequence_length = 100,	20	0.2129	91.36%	90.28%	unerfitting problem.
i ait Zb	· Output Layor	ocquerice_ierigiri = 100,	20	0.2123	31.3070	30.2070	
							This model showed serous
							underfitting problem. The possib
		HiddenUnits= 1000, sequence					reason was that I used shorter
	Custom 1	lenth = 50, Dropout = 0.5	20	0.367	83.34%	90.82%	sequence length.
	Oustoill 1	161141 = 30, D10pout = 0.3	20	0.001	00.0470	30.0270	This model did not showed
							overfitting or underfitting. But tes
		HiddenUnits = 100 sequence =					accuracy was not better than 1b
	Custom 2	200 batch number = 100	20	0.3942	82.22%	88.66%	too much.
	Custom 3		-				
	Oddioin 0			1	1		