Vocabulary Size	8000
Embedding Size	128
Hidden Layer Size	128
Learning Rate	1×10^{-03}
Dropout Rate	0.095
Adam Optimizer (β_1,β_2,α)	$0.09, 0.99, 0.001$