MD Alamgir 08/17/2017

Udacity Deep Learning Project 4: Language Translation
Hyper parameter Sensitivity Study

Case	Epoch #	Batch size	RNN size	Num layers	Encoding embedding size	Decoding embedding size	Learning Rate	Keep probability	**Training Accuracy**	**Validation Accuracy**	**Loss**	Observations
1	9	256	512	4	128	128	0.01	0.70	0.87	0.88	0.11	
2	9	256	512	4	128	128	0.001	0.7	0.968	0.963	0.017	Case 2 vs. Case 1: Decreasing **Learning Rate** by a factor of ten significantly improves accuracy
3	9	1024	512	4	128	128	0.001	0.70	0.84	0.84	0.17	Case 3 vs. Case 2: Increasing **Batch Size** by a factor of 4 significantly decreases accuracy
4	9	1024	128	4	128	128	0.001	0.70	0.64	0.63	0.63	Case 4 vs. Case 3: Decreasing **RNN Size** by a factor of 4 significantly decreases accuracy
5	14	256	512	4	128	128	0.001	0.70	0.98	0.97	0.009	Case 6 vs. Case 5: When **Embedding Sizes** are doubled, Keep Probability needs to be increased by 0.05 in order to achieve similar accuracy and loss.
6	14	256	512	4	256	256	0.001	0.75	0.98	0.97	0.01	