

Assignment 4

Your name: Bing Yang

Your GTID: byang322@gmail.com (903563092)

Seq2Seq Results

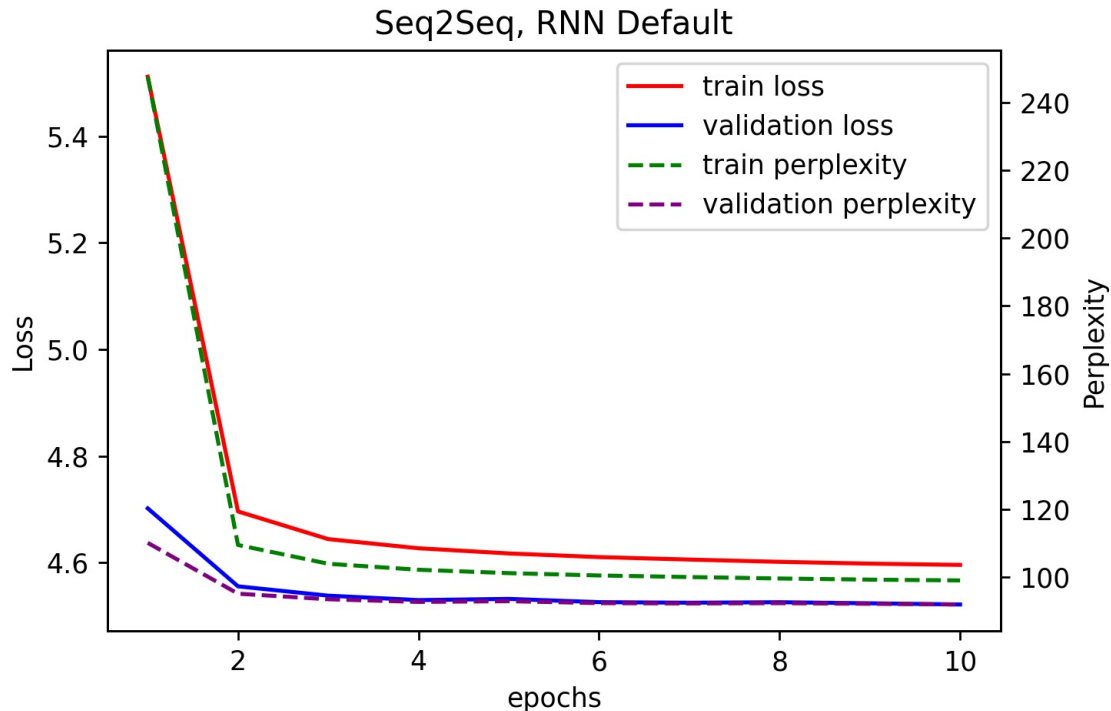
Put your results from training before and after hyperparameter tuning here.

Results for default configuration using RNN		Results for default Configuration Using LSTM	
Training Loss	4.5958	Training Loss	4.3472
Training Perplexity	99.0695	Training Perplexity	77.2650
Validation Loss	4.5215	Validation Loss	4.2230
Validation Perplexity	91.9775	Validation Perplexity	68.2360
Result for your Best Model using RNN after hyperparameter tuning		Resut for your Best Model using LSTM after hyperparameter tuning	
Training Loss	4.1303	Training Loss	3.3429
Training Perplexity	62.1949	Training Perplexity	28.3002
Validation Loss	4.1795	Validation Loss	3.4789
Validation Perplexity	65.3362	Validation Perplexity	32.4235
Your best model configuration for RNN after hyperparameter tuning		Your best model configuration for LSTM after hyperparameter tuning	
EPOCH = 150 B = 5e 4 others same		EPOCH = 150 B = 5e 4 others same	

Table 1

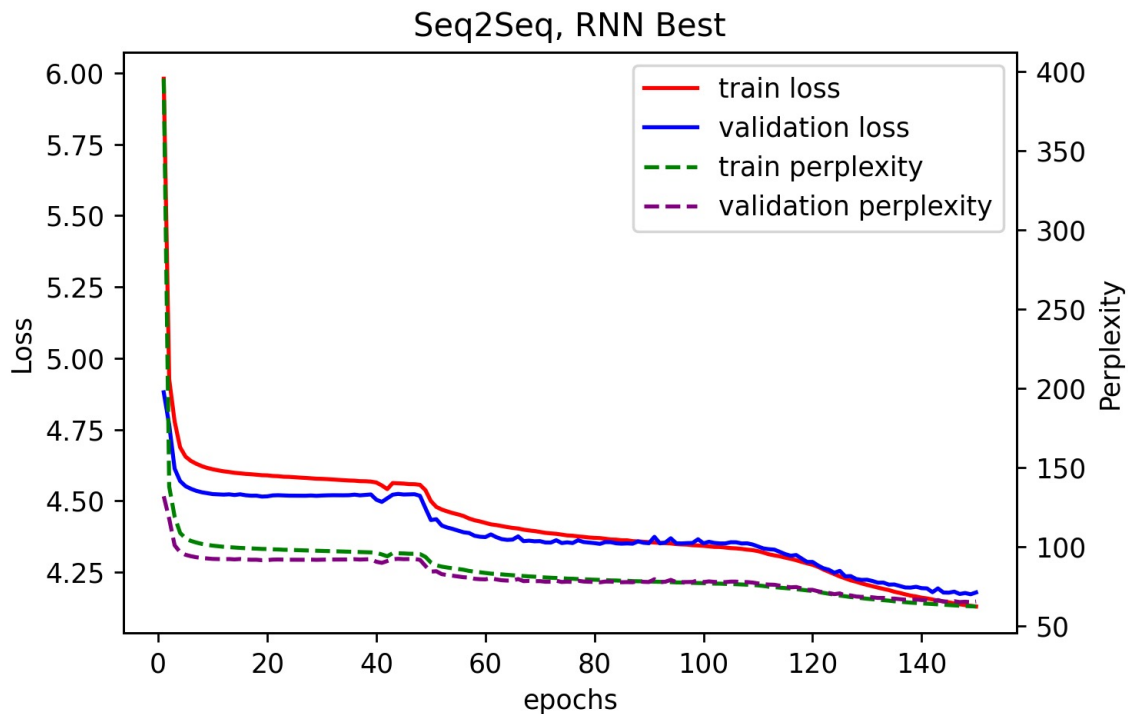
Seq2Seq Curves

Put the plots for loss/perplexity curves (training & validation) for your configuration with default setting and for your best model here.



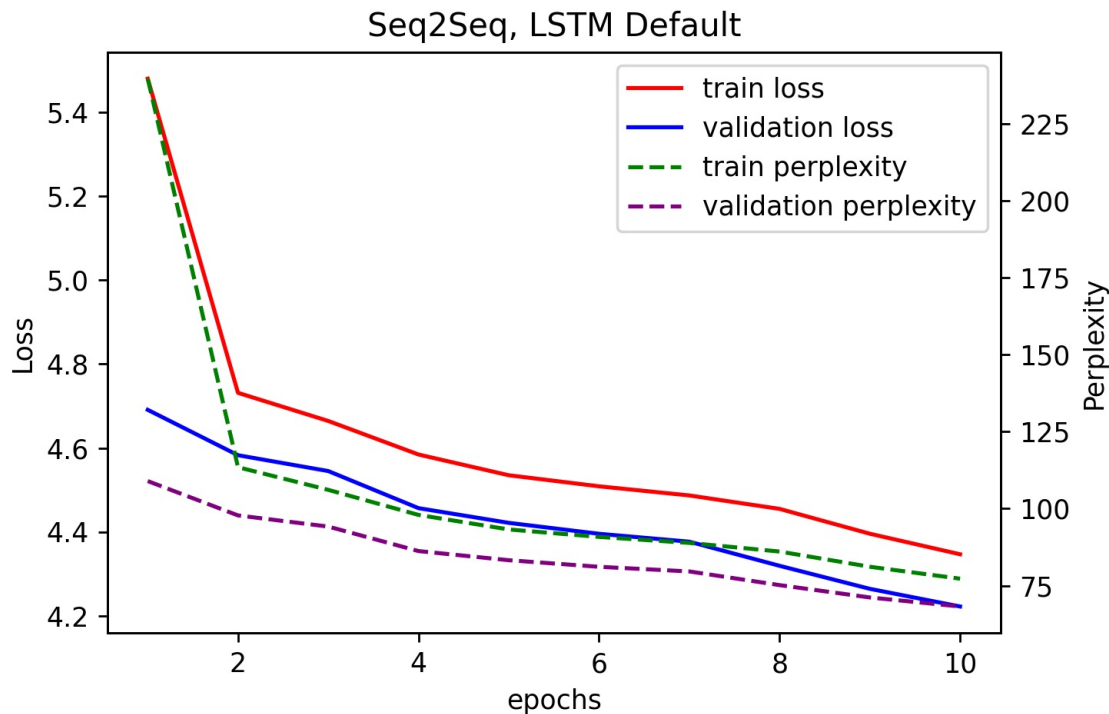
Seq2Seq Curves

Put the plots for loss/perplexity curves (training & validation) for your configuration with default setting and for your best model here.



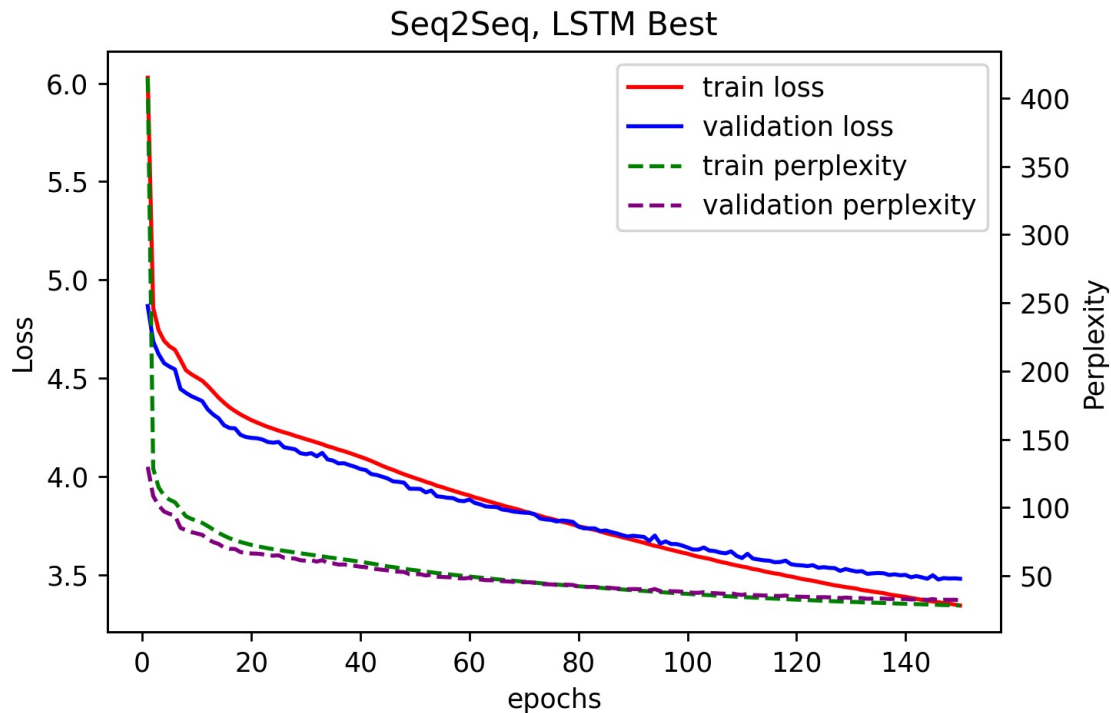
Seq2Seq Curves

Put the plots for loss/perplexity curves (training & validation) for your configuration with default setting and for your best model here.



Seq2Seq Curves

Put the plots for loss/perplexity curves (training & validation) for your configuration with default setting and for your best model here.



Seq2Seq Explanation

Explain what you did here and why you did it to improve your model performance. You can use another slide if needed.

I simply increased number of epochs and lower the learning rate. I found out that changing other hyperparameters do not have significant effects on the output.

Training longer can generate better results in general. By experimentation, I found out overfitting is a problem. To solve that, I lowered the learning rate to let the model explore more slowly in the parameter space.

Transformer Results

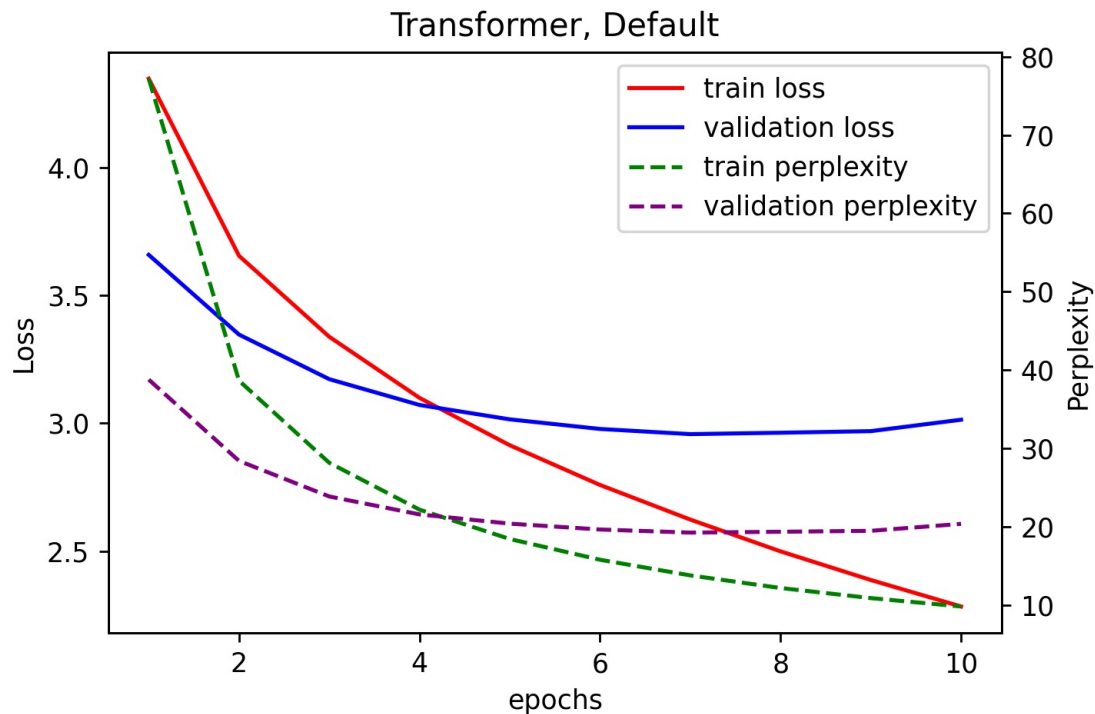
Put your results from training before and after hyperparameter tuning here.

Results for default configuration			
Training Loss	2.2837	Validation Loss	3.0061
Training Perplexity	9.8127	Validation Perplexity	20.2077
Result for your Best Model			
Training Loss	2.5079	Validation Loss	2.9947
Training Perplexity	12.2789	Validation Perplexity	19.9797
Your best model configuration after hyperparameter tuning			
BATCH_SIZE = 64, LR = 5e-4, EPOCHS = 12, others same			

Table 2

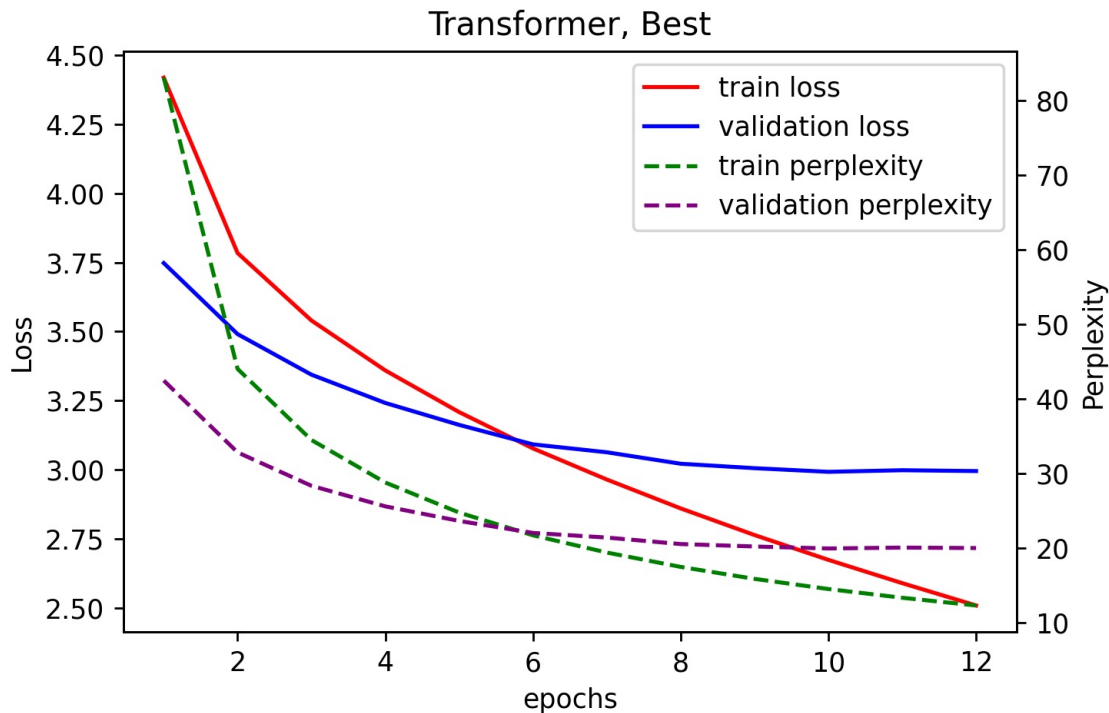
Transformer Curves

Put the plots for loss/perplexity curves (training & validation) for your configuration with default setting and for your best model here.



Transformer Curves

Put the plots for loss/perplexity curves (training & validation) for your configuration with default setting and for your best model here.



Transformer Explanation

Explain what you did here and why you did it to improve your model performance. You can use another slide if needed.

I decreased the batch size from 128 to 64 and adjusted the epochs by experimentation.

The main issue for training transformer is overfitting from my experimentations. Small batch size can usually result in better generalization in practice. Also, I set training time manually by experimentation to avoid overfitting.

Transformer Translation Results

Put translation results for your best model (1st 9 sentences) here

Input sentence	Back translation
'<sos>', 'a', 'young', 'boy', 'jumps', 'into', 'water', '.', '<eos>',	'<sos>', 'a', 'young', 'boy', 'is', 'into', 'into', 'water', '.', '<eos>'
'<sos>', 'a', 'native', 'woman', 'is', 'working', 'on', 'a', 'craft', 'project', '.', '<eos>'	'<sos>', 'a', 'is', 'works', 'on', 'a', 'a', '.', '.', '<eos>'
'<sos>', 'an', 'asian', 'woman', 'sitting', 'outside', 'an', 'outdoor', 'market', 'stall', '.', '<eos>'	'<sos>', 'an', 'asian', 'is', 'sitting', 'in', 'of', 'a', 'a', 'a', '.', '<eos>'
'<sos>', 'woman', 'standing', 'on', 'a', 'brick', 'wall', 'and', 'taking', 'a', 'picture', '<eos>'	'<sos>', 'woman', 'stands', 'standing', 'a', 'a', 'a', 'and', 'photographs', '.', '.', '<eos>'
'<sos>', 'there', 'are', 'construction', 'workers', 'working', 'hard', 'on', 'a', 'project', '.', '<eos>'	'<sos>', 'construction', 'workers', 'workers', 'hard', 'a', 'a', 'a', '.', '.', '.', '<eos>'
'<sos>', 'a', 'man', 'in', 'a', 'cluttered', 'office', 'is', 'using', 'the', 'telephone', '<eos>'	'<sos>', 'a', 'man', 'is', 'on', 'a', 'a', 'office', '.', '.', '<eos>'
'<sos>', 'two', 'chinese', 'people', 'are', 'standing', 'by', 'a', 'chalkboard', '.', '<eos>'	'<sos>', 'two', 'chinese', 'stand', 'at', 'by', 'a', 'a', 'a', '.', '.', '<eos>'
'<sos>', 'children', 'are', 'playing', 'a', 'sport', 'on', 'a', 'field', '.', '<eos>'	'<sos>', 'children', 'are', 'playing', 'playing', 'a', 'a', 'a', '.', '<eos>'
'<sos>', 'a', 'man', 'is', 'working', 'at', 'a', 'construction', 'site', '.', '<eos>'	'<sos>', 'a', 'man', 'is', 'working', 'a', 'construction', 'construction', 'construction', 'site', '<eos>'

Table 3

LSTM Translation Results

Put translation results for your best model (1st 9 sentences) here

Input sentence	Back translation
'<sos>', 'a', 'young', 'boy', 'jumps', 'into', 'water', '.', '<eos>',	'<sos>', 'a', 'little', 'boy', 'is', 'jumping', 'in', 'the', '.', '.', '<eos>'
'<sos>', 'a', 'native', 'woman', 'is', 'working', 'on', 'a', 'craft', 'project', '.', '<eos>'	'<sos>', 'a', 'asian', 'is', 'is', 'a', 'a', 'a', '.', '.', '<eos>'
'<sos>', 'an', 'asian', 'woman', 'sitting', 'outside', 'an', 'outdoor', 'market', 'stall', '.', '<eos>'	'<sos>', 'a', 'asian', 'woman', 'stands', 'standing', 'front', 'front', 'of', 'a', '.', '.', '<eos>'
'<sos>', 'woman', 'standing', 'on', 'a', 'brick', 'wall', 'and', 'taking', 'a', 'picture', '<eos>'	'<sos>', 'woman', 'sitting', 'on', 'a', 'a', 'a', '.', '.', '<eos>'
'<sos>', 'there', 'are', 'construction', 'workers', 'working', 'hard', 'on', 'a', 'project', '.', '<eos>'	'<sos>', 'construction', 'workers', 'are', 'working', 'a', 'a', 'a', '.', '.', '<eos>'
'<sos>', 'a', 'man', 'in', 'a', 'cluttered', 'office', 'is', 'using', 'the', 'telephone', '<eos>'	'<sos>', 'a', 'man', 'is', 'sitting', 'a', 'a', '.', '.', '<eos>'
'<sos>', 'two', 'chinese', 'people', 'are', 'standing', 'by', 'a', 'chalkboard', '.', '<eos>'	'<sos>', 'two', 'women', 'are', 'sitting', 'in', 'a', 'a', '.', '.', '<eos>'
'<sos>', 'children', 'are', 'playing', 'a', 'sport', 'on', 'a', 'field', '.', '<eos>'	'<sos>', 'children', 'are', 'on', 'a', 'a', 'a', 'a', '.', '<eos>'
'<sos>', 'a', 'man', 'is', 'working', 'at', 'a', 'construction', 'site', '.', '<eos>'	'<sos>', 'a', 'man', 'is', 'a', 'a', 'a', 'a', '.', '.', '<eos>'

Table 4

Compare LSTM to Transformer

Compare your LSTM results to your Transformer Results both quantitatively and qualitatively and explain the differences.

Even though transformer has a significant lower loss compared to LSTM seq2seq model (perplexity 32 for LSTM and 19.9 for transformer), neither model gives satisfactory translations.

I would say transformer does a slightly better job because it captures most pieces for most sentences. LSTM model generates many repeated 'a' at the end of sentences.

The transformer I implemented in this assignment is a very simple version (not even the complete one) of the original model. This might explain the poor performance.