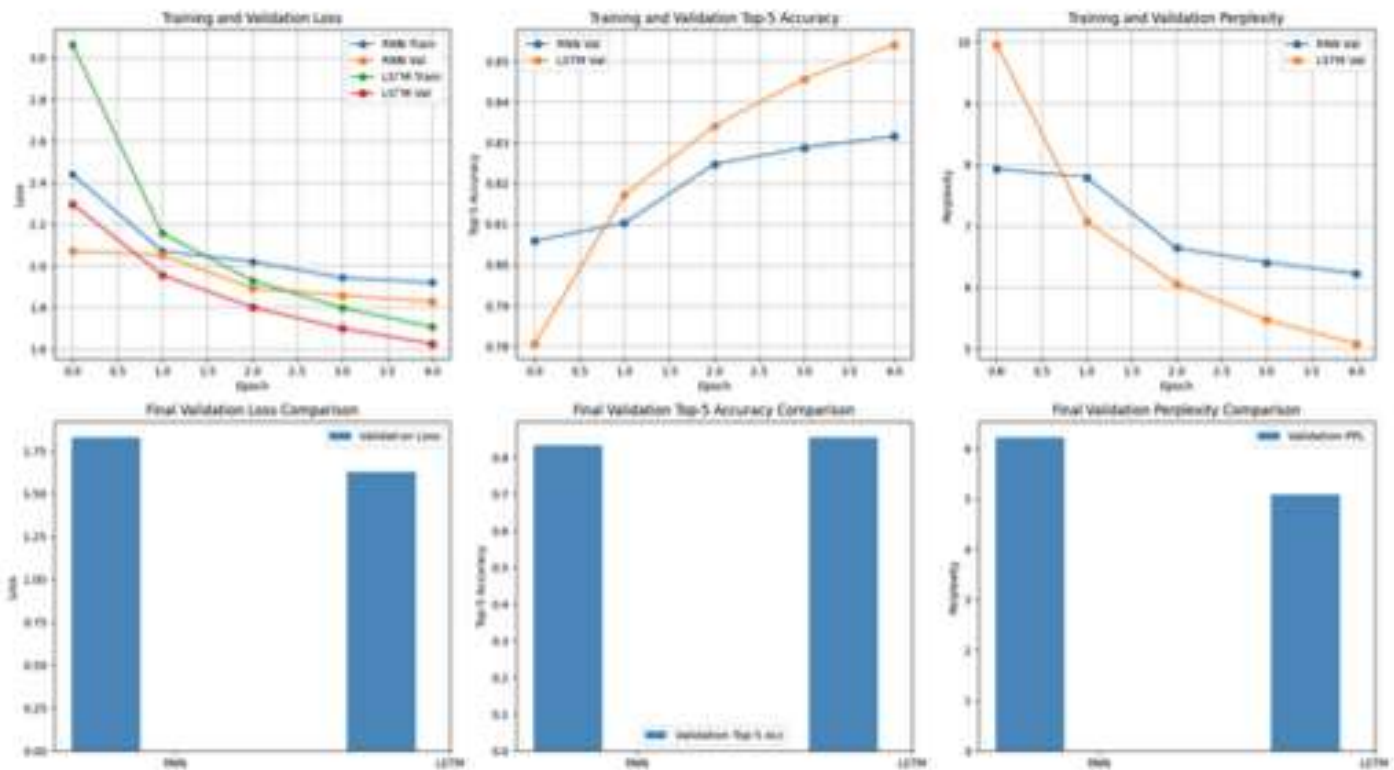


1. Baseline Results on test set (RNN and LSTM)

Model	Loss	Top-5 Accuracy	Perplexity
RNN	1.74	0.84	5.71
LSTM	1.54	0.86	4.70



- Best performing model on test set: LSTM
 Test Top-5 Accuracy: 0.8614
 Test Perplexity: 4.7045
 Total Parameters: 11,371,284
 RNN Parameters: 8,612,628
- **A HIGHER PARAMETER COUNT CORRELATES TO BETTER PERFROMANCE IN BASELINE MODELS**
- I am using Embedding layer which is pre-trained from scratch
- I am using tanh activation function for all models
- I am using dropout layers at specified locations with a rate of 0.3

2. Residual model results on test-set

Best Baseline (LSTM) - Test Top-5 Acc: 0.8614, Test PPL: 4.7045

Residual Model - Test Top-5 Acc: 0.8880, Test PPL: 3.7908

Model	Test Top-5 Accuracy	Test Perplexity	Total Parameters
Baseline LSTM	0.86	4.704540	11371284
Residual Model	0.888023	3.790807	11371284

- **Percentage improvement= 2.32 %**
- **There is no change in number of parameters because we havent added any additional weights nor layers but just a residual connection which only increase the sum operation.**