Name : Tirth Pandit
ID : 2019201017

Assignment-2 Report

Neural Language Model

---

1) Model Architecture

Model Consist of 3 layers

```
Neural_LM(
  (embed): Embedding(41505, 256)
  (lstm): LSTM(256, 1024, batch_first=True)
  (linear): Linear(in_features=1024, out_features=41505, bias=True)
)
```

2) Model Parameters

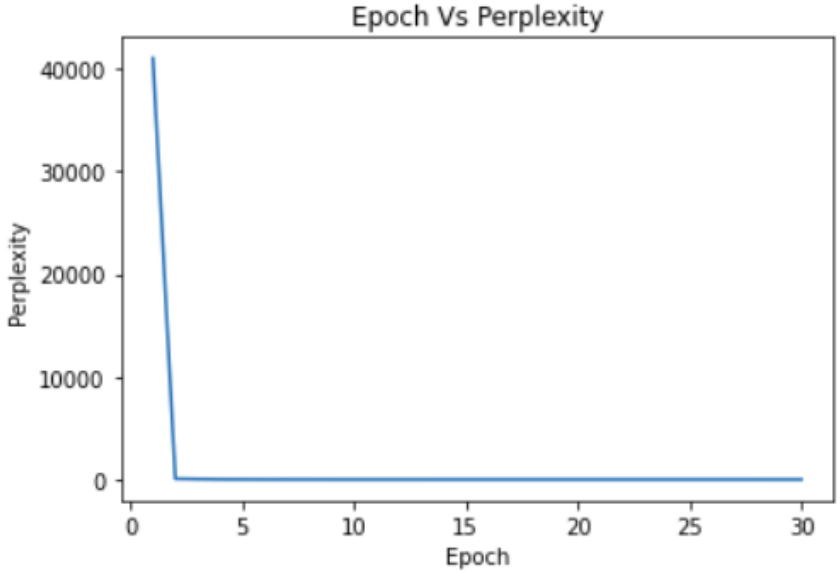
Sequence length = 5
Batch Size = 80
Learning rate = 0.005
Hidden state size = 1024
LSTM layer(Units) = 1
Embeddings = 256
Epoch = 30

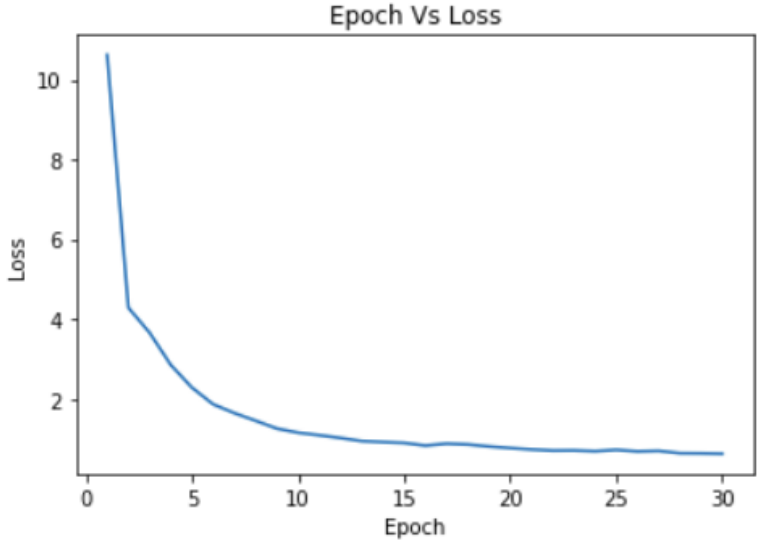
3) Data

Total Sentences = 5460
Train : Validation : Test = 7:2:1

Train Set size = 3822
Validation Data size = 1092
Test Data size = 546
Vocabulary Size = 41505

4) **Model Performance**





Avg Test Perplexity : 19.8468

5) Model Comparison

1) **4-gram + Kneser-ney Smoothing**
   Avg test Perplexity = 1885.4296

2) **4-gram + Witten-bell Smoothing**
   Avg Test Perplexity = 217.6923

3) **Neural LSTM based Model**
   Avg Test Perplexity = 19.8468