Introduction to NLP (CS7.401)

Spring 2022, IIIT Hyderabad
Assignment 3

Abhinav S Menon

1 Language Model

1.1 Model Details

The model contains one embedding layer, an LSTM layer (unidirectional, single layer), and an output layer.

The model was trained (on each corpus) until the difference in the loss between successive epochs is less than 0.01. The embedding dimension is 100 and the hidden size is 150. The learning rate is 0.001.

The trained models are saved as `{eng,fr}_lm.pth`.

1.2 Results

The perplexities of the train and test splits of the Europarl corpus are given in `2020114001_LM_{train,test}.txt`. The averages are approximately 1.3 for both these splits.

1.3 Analysis

The perplexity scores are very good compared to those obtained by statistical methods like n-grams. This is most likely due to the availability of an arbitrarily long context; it allows for much better prediction.

2 MT Model

2.1 Model Details

The encoder and decoder are both LSTMs with the same hyperparameters as in the LMs. The early stopping condition and the learning rate are also identical.

2.2 Results

The BLEU scores of the model trained from scratch are given in `2020114001_MT1_{train,test}.txt`. The corpus scores are approximately 0.27 on the train split and 0.23 on the test split.