Prediction Challenge 4– Deep NLP

Introduction:

The task of this project was to generate three 25-word poems in the style of William Blake using the 'blake-poems.txt' corpus from the Gutenberg Dataset. This report aims to critically appraise the choice of model parameters and hyperparameters for this task.

Model Explanation:

For this prediction, an LSTM model was used with one LSTM Layer, with 800 units. To prevent overfitting, a dropout layer with a rate of 0.5 is added after the LSTM layer. Dropout randomly sets a fraction of the input units to 0 during each update, which helps to prevent the model from relying too heavily on any one feature.

During training, the model is fed batches of 64 samples at a time, for a total of 100 epochs (iterations over the entire training dataset). The training process aims to minimize the categorical cross entropy loss function, while maximizing the accuracy of the model on the training data.

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_4 (LSTM)               (None, 800)               2566400

 dropout (Dropout)           (None, 800)               0

 dense_1 (Dense)             (None, 1506)              1206306

=================================================================
Total params: 3,772,706
Trainable params: 3,772,706
Non-trainable params: 0
_________________________________________________________________
Epoch 1/100
103/103 [==============================] - 55s 514ms/step - loss: 6.4925 - accuracy: 0.0599
Epoch 2/100
```

Results:

Three 25-word poems were generated using the trained model. The BLEU score was calculated for each poem, and the results were as follows:

```
Poem 1
Glee On Cloud Saw Child And
He Laughing Said To Me
Pipe Song About Lamb So
Piped With Merry Cheer Piper
Pipe That Song Again

BLEU Score for predicted words: 1.0
```

```
Poem 2
Song About Lamb So Piped With
Merry Cheer Piper Pipe That
Song Again So Piped He
Wept To Hear Drop Thy
Pipe Thy Happy Pipe

BLEU Score for predicted words: 1.0
```

```
Poem 3
And He Laughing Said To Me
Pipe Song About Lamb So
Piped With Merry Cheer Piper
Pipe That Song Again So
Piped He Wept To

BLEU Score for predicted words: 1.0
```

Critical Appraisal of the model:

After doing multiple experiment with the parameters and with different models, I found out that the LSTM model works better with the one layer and The addition of a dropout layer with a rate of 0.5 helped to prevent overfitting. By using different memory sizes, I observed that with the less memory size the model performance is not up to mark, with the memory size 800, model performing well and get the blue score ‘1 ‘and by adding many layers the models do not improve well.

Conclusion:

This report shows the use of the LSTM model and generate the poems in the style of William Bake. Finally, this project highlights the effectiveness of LSTM models for natural language generation tasks and the importance of optimizing model parameters for optimal performance.

Word Count: 311

Code:https://colab.research.google.com/drive/1TRLMpoe2b3DG4ayY-55S1rdZ7uOzof9n?usp=sharing