NLP Assignment-3 Report

Name: Saransh Sharma
Roll Number: 20CS30065

LSTM Model

i) Define hyperparameters

- max_seq_length = 50
- batch_size = 30
- num_hidden_units = 10
- n_layers = 3
- dropout = 0.3
- num_output_units = 4
- epoch_count = 100

ii) Data Specification:

- Mean Sentence Length: 26.0216
- Median Sentence Length: 26.0
- Min Sentence Length: 9
- Max Sentence Length: 97

iii) Justification for Hyperparameters:

The choice of hyperparameters was influenced by the characteristics of the data. With a mean/median sentence length around 25 and a maximum length of 97, a max sequence length of 50 was selected to accommodate the majority of sentences without excessively increasing computational complexity.

iv) Model Architecture:

- Input Layer: Embedding layer
- Hidden Layers: 3 layers of RNN with 10 hidden units each.
- Embedding dimension: 100
- Output Layer: Dense layer with 4 output units.

v) Training Details:

- Training Duration: 100 epochs (based on validation set)
- Optimizer: Adam
- Learning rate: 0.01
- Loss Function: Cross Entropy

vi) Performance:

- Accuracy: 0.676
- F1 Score: 0.6726733324181308
- Confusion Matrix:

	Predicted Class 0	**Predicted Class 1**	**Predicted Class 2**	**Predicted Class 3**
True Class 0	**104**	6	7	8
True Class 1	17	**94**	1	13
True Class 2	19	5	**72**	29
True Class 3	17	15	25	**68**

- Classification Report:

	precision	recall	f1-score	support
0	0.66	0.83	0.74	125
1	0.78	0.75	0.77	125
2	0.69	0.58	0.63	125
3	0.58	0.54	0.56	125
accuracy			0.68	500

macro avg	0.68	0.68	0.67	500
weighted avg	0.68	0.68	0.67	500