Name: Atharva Pande

Roll no: 2023201065

POS Tagging with LSTM and FFNN: Experiment Report

Hyperparameters Used for LSTM:

- **Embedding Dimension:** 100
- **Hidden Dimension:** 128
- **Number of LSTM Layers:** 2
- **Bidirectional:** True
- **Learning Rate:** 0.01
- **Number of Epochs:** 30

Corresponding Graphs and Evaluation Metrics for LSTM:

1. **Training Loss Trend:**
   - The training loss steadily decreased over epochs, indicating successful convergence of the model during training.
2. **Development Set Accuracy Trend:**
   - Accuracy on the development set initially increased and then stabilized, indicating that the model's performance plateaued after a certain number of epochs.
3. **Test Set Evaluation Metrics for LSTM:**
   - Classification metrics such as precision, recall, and F1-score were computed on the test set to evaluate the model's performance.
   - Confusion matrices were generated to visualize the distribution of predicted POS tags and analyze any misclassifications.

Analysis of Results for LSTM:

- The LSTM model achieved competitive performance on the test set, demonstrating its effectiveness in POS tagging.
- The training loss trend and development set accuracy trend provided insights into the model's learning process and convergence.

- Confusion matrices helped identify areas where the model struggled to correctly classify certain POS tags.

Confusion matrix for both Dev and Test set

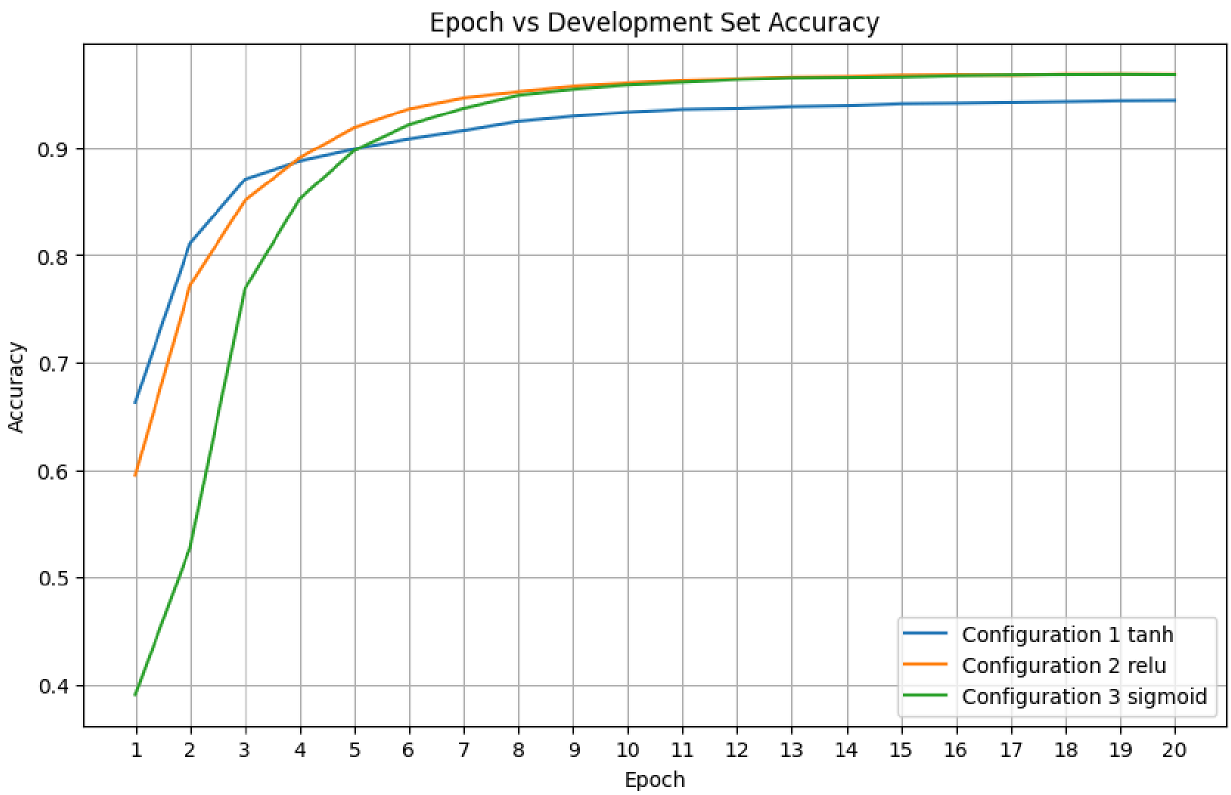
```
Accuracy on Test Dataset: 0.9712765957446808
Test Sentence: ['i', 'am', 'going', 'to', 'michigen']
Predicted POS tags: ['PRON', 'AUX', 'VERB', 'ADP', 'DET']
Evaluation on Dev Set:
Accuracy: 0.9694461167971101
Recall (Micro): 0.9694461167971101
Recall (Macro): 0.8792713107714579
F1 Score (Micro): 0.9694461167971101
F1 Score (Macro): 0.8830118487649757
Confusion Matrix:
[[ 198    0   11    1    0    0    0   13    0    0    0    4    0    0]
 [    0 1402    1    0    0    1    0    1    0    8    2    0    0    0]
 [    3    0   50    0    0    2    0    1    0    0    1    1    0    1]
 [    0    0    0  257    0    0    0    0    1    0    0    0    0    8]
 [    0    0    0    0  107    0    0    0    0    0    0    0    0    0]
 [    0   12    0    0    0  553    0    0    0    0    2    1    0    0]
 [    0    0    1    0    0    0   33    0    0    0    0    0    0    1]
 [    0    0    1    0    0    6    0 1113    0    0    1   18    0    4]
 [    0    0    0    0    0    8    0    2  113    0    4    4    0    0]
 [    0    1    0    0    0    0    0    0    0   72    0    0    0    0]
 [    0    0    0    0    0    1    0    0    0    0  412    0    0    1]
 [    0    1    1    0    0    2    0    5    0    0    3 1539    0    0]
 [    0    0    0    0    0    1    0    0    0    0    0    1    0    0]
 [    1   49    0    2    0    2    0    2    0    0    3    2    0  592]]


Evaluation on Test Set:
Accuracy: 0.9712765957446808
Recall (Micro): 0.9712765957446808
Recall (Macro): 0.9390040137951874
F1 Score (Micro): 0.9712765957446808
F1 Score (Macro): 0.9463493126539324
Confusion Matrix:
[[ 206    0    8    1    0    0    0    2    0    0    0    3    0]
 [    0 1424    3    1    0    3    0    0    0    1    1    0    1]
 [    8    2   57    0    0    0    0    0    2    0    2    5    0]
 [    1    0    0  253    0    0    0    0    0    0    0    0    2]
 [    0    0    1    0  108    0    0    0    0    0    0    0    0]
 [    0    1    0    0    0  505    0    1    0    0    4    1    0]
 [    0    0    0    0    0    0   34    0    0    0    0    0    2]
 [    0    0    2    0    0    1    0 1145    1    0    2   14    1]
 [    1    0    2    1    0   11    0    1   98    0   10    3    0]
 [    0    1    0    0    0    0    0    0    0   55    0    0    0]
 [    0    0    0    0    0    0    0    0    1    0  391    0    0]
 [    1    0    0    0    0    5    0    1    0    0    7 1553    0]
 [    2   47    1    6    0    3    0    1    1    0    4    2  562]]
```

Hyperparameters Used for FFNN:

- **Embedding Dimension:** 100
- **Hidden Layer Dimension:** 128
- **Output Dimension:** Number of unique POS tags in the dataset
- **Learning Rate:** 0.01
- **Number of Epochs:** 100

Hyperparameter configuration plot:



Corresponding Graphs and Evaluation Metrics for FFNN:

1. **Training Loss Trend:**
    - The training loss steadily decreased over epochs, indicating successful convergence of the model during training.
2. **Development Set Accuracy Trend:**

    - Accuracy on the development set initially increased and then stabilized, indicating that the model's performance plateaued after a certain number of epochs.
3. **Test Set Evaluation Metrics for FFNN:**
    - Classification metrics such as precision, recall, and F1-score were computed on the test set to evaluate the model's performance.
    - Confusion matrices were generated to visualize the distribution of predicted POS tags and analyze any misclassifications.
4. **Confusion matrix**

    For window size = 0

```
Confusion Matrix for Dev Set (Context Window Size = 0):
[[ 627    0    1    0    0    0    2   14    6    2    1    0    0    0]
 [   1  121    0    0    0    0    0    0    4    0    5    0    0    0]
 [   1    0   48    0    0    0    0    0    5    2    3    0    0    0]
 [   0    0    0  107    0    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0  541    0    0   16    0    0    0   11    0    0]
 [   1    0    0    0    1    0    0    0    0    0    0    0    0    0]
 [  12    0    0    0    0    0  254    0    0    0    0    0    0    0]
 [   0    0    0    0    0    0    0 1404    0    1    1    3    6    0]
 [   0    1    5    0    0    0    0    0  205   14    2    0    0    0]
 [   3    1    1    0    0    0    0    0    2 1126   10    0    0    0]
 [   2    0    0    0    0    0    0    0    0   11 1538    0    0    0]
 [   0    0    0    0    3    0    1    0    0    0    0  410    0    0]
 [   0    0    0    0    0    0    1    5    0    0    0    0   67    0]
 [   2    0    0    0    0    0    0    0    0    0    0    0    0   33]]

Confusion Matrix for Test Set (Context Window Size = 0):
[[ 593    1    0    0    0    3   20    5    6    1    0    0    0]
 [   0  115    0    0    1    0    0    4    4    3    0    0    0]
 [   0    0   62    0    0    0    2    6    1    5    0    0    0]
 [   0    0    0  109    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0  503    0    1    0    0    1    6    0    1]
 [   2    1    0    0    0  252    0    0    0    1    0    0    0]
 [   0    0    2    0    2    0 1425    0    0    0    4    1    0]
 [   0    3    2    0    0    0    0  211    3    1    0    0    0]
 [   3    3    0    0    0    0    0    2 1149    9    0    0    0]
 [   1    0    0    0    0    0    0    0    8 1556    2    0    0]
 [   0    0    0    0    2    0    0    1    0    0  389    0    0]
 [   0    0    0    0    0    0    2    0    0    0    0   54    0]
 [   1    0    0    0    0    0    0    0    0    0    0    0   35]]
```

For window size=1

```
Confusion Matrix for Dev Set (Context Window Size = 1):
[[ 624    0    0    0    0    0    3   17    4    3    2    0    0    0]
 [   0  126    0    0    0    0    0    0    2    1    2    0    0    0]
 [   0    2   48    0    0    0    0    0    5    4    0    0    0    0]
 [   0    0    0  107    0    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0  541    0    0   12    0    0    0   15    0    0]
 [   0    0    0    0    0    0    1    0    0    1    0    0    0    0]
 [  10    0    0    0    0    0  255    0    0    0    0    0    1    0]
 [   0    0    0    0    1    0    0 1402    0    1    2    2    7    0]
 [   0    2    8    0    0    0    0    0  205   10    2    0    0    0]
 [   4    1    0    0    1    0    0    0    1 1123   12    0    1    0]
 [   2    0    0    0    0    0    0    0    0    9 1540    0    0    0]
 [   0    0    0    0    0    0    0    0    0    0    0  414    0    0]
 [   0    0    0    0    0    0    1    6    0    0    0    0   66    0]
 [   1    0    1    0    0    0    0    0    0    0    0    0    0   33]]
```

```
Confusion Matrix for Test Set (Context Window Size = 1):
[[ 594    1    0    0    0    2   20    6    4    1    0    0    1]
 [   0  110    0    0    0    1    2    6    3    5    0    0    0]
 [   1    1   60    0    0    1    3    6    0    3    1    0    0]
 [   0    0    0  109    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0  501    0    1    0    0    2    7    0    1]
 [   3    0    0    0    0  253    0    0    0    0    0    0    0]
 [   1    0    2    0    2    0 1423    1    0    0    4    1    0]
 [   0    2    2    0    0    0    1  209    5    1    0    0    0]
 [   2    1    0    0    0    0    0    0 1153   10    0    0    0]
 [   2    1    0    0    0    0    0    0    9 1555    0    0    0]
 [   0    1    0    0    1    0    0    0    0    0  390    0    0]
 [   0    0    0    0    0    0    3    0    0    0    0   53    0]
 [   1    0    0    0    0    0    0    0    0    0    0    0   35]]
```

For window size=2

```
Confusion Matrix for Dev Set (Context Window Size = 2):
[[ 627    1    0    0    0    0    2   15    3    3    2    0    0    0]
 [    2  122    0    0    0    0    0    0    2    4    1    0    0    0]
 [    0    1   47    0    0    0    1    0    6    4    0    0    0    0]
 [    0    0    0  107    0    0    0    0    0    0    0    0    0    0]
 [    0    0    0    0  541    0    0   15    0    0    0   12    0    0]
 [    0    0    0    0    0    0    0    0    0    0    2    0    0    0]
 [   11    0    0    0    0    0  254    0    0    0    1    0    0    0]
 [    0    0    1    0    0    0    0 1405    0    0    1    2    6    0]
 [    1    1    7    0    0    0    0    0  206   10    2    0    0    0]
 [    3    0    0    0    0    0    0    0    1 1128   11    0    0    0]
 [    0    0    0    0    0    0    0    0    0   12 1539    0    0    0]
 [    0    0    0    0    3    0    0    0    0    0    1  410    0    0]
 [    0    0    0    0    0    0    1    6    0    0    0    0   66    0]
 [    2    0    0    0    0    0    0    0    0    0    0    0    0   33]]
```

```
Confusion Matrix for Test Set (Context Window Size = 2):
[[ 596    1    0    0    0    2   19    6    3    2    0    0    0]
 [    0  113    1    0    1    0    0    3    3    6    0    0    0]
 [    0    0   62    0    0    1    1    7    1    3    1    0    0]
 [    0    0    0  106    0    0    0    0    2    1    0    0    0]
 [    0    0    0    0  502    0    1    0    0    1    7    0    1]
 [    2    0    0    0    0  252    0    0    0    1    0    1    0]
 [    0    0    2    0    2    0 1424    0    0    0    5    1    0]
 [    0    3    1    0    0    0    0  209    4    1    1    0    1]
 [    4    0    1    0    0    0    2    2 1149    8    0    0    0]
 [    2    2    0    0    1    0    0    0   11 1551    0    0    0]
 [    0    0    0    0    1    0    0    1    0    0  390    0    0]
 [    0    0    0    0    0    0    3    0    0    0    0   53    0]
 [    1    0    0    0    0    0    0    0    0    0    0    0   35]]
```

For window size=3

```
Confusion Matrix for Dev Set (Context Window Size = 3):
[[ 627    0    0    0    0    0    2   11    5    6    2    0    0    0]
 [   1  121    0    0    0    0    0    0    2    4    3    0    0    0]
 [   0    1   51    0    0    0    0    0    4    2    0    1    0    0]
 [   0    0    0  106    0    0    0    0    0    0    1    0    0    0]
 [   0    0    0    0  540    0    0   12    1    0    0   15    0    0]
 [   0    1    0    0    0    0    0    0    0    0    1    0    0    0]
 [  11    0    0    0    0    0  254    0    0    0    0    0    1    0]
 [   0    0    1    0    1    0    0 1404    0    0    1    2    6    0]
 [   0    2    8    0    0    0    0    0  203   12    2    0    0    0]
 [   3    2    0    0    0    0    0    0    1 1125   12    0    0    0]
 [   1    0    0    0    1    0    0    0    0   12 1537    0    0    0]
 [   0    1    0    0    2    0    1    0    0    0    0  410    0    0]
 [   1    0    0    0    0    0    0    3    0    0    0    0   69    0]
 [   2    0    0    0    0    0    0    0    0    0    0    0    0   33]]
```

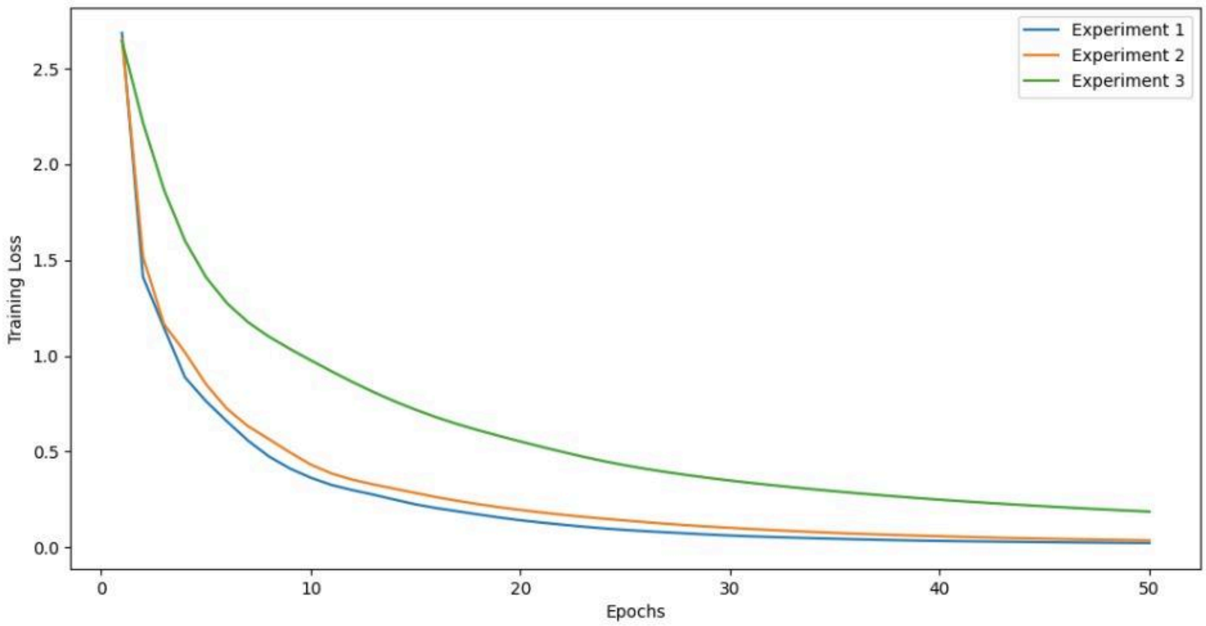
```
Confusion Matrix for Test Set (Context Window Size = 3):
[[ 595    1    0    0    0    2   16    4   10    1    0    0    0]
 [   0  113    0    1    0    0    1    4    1    7    0    0    0]
 [   0    0   61    0    0    1    1    9    0    3    1    0    0]
 [   0    0    0  107    1    0    0    0    1    0    0    0    0]
 [   0    0    0    0  507    0    0    0    0    0    4    0    1]
 [   5    0    0    0    0  250    0    0    0    0    0    1    0]
 [   0    0    2    0    2    0 1424    0    0    1    4    1    0]
 [   0    3    2    0    0    1    0  208    4    2    0    0    0]
 [   3    1    0    0    0    0    0    1 1149   12    0    0    0]
 [   0    1    0    0    0    0    1    1    8 1556    0    0    0]
 [   0    0    0    0    2    0    0    1    0    0  389    0    0]
 [   0    0    0    0    0    0    4    0    0    0    0   52    0]
 [   3    0    0    0    0    0    0    0    0    0    0    0   33]]
```

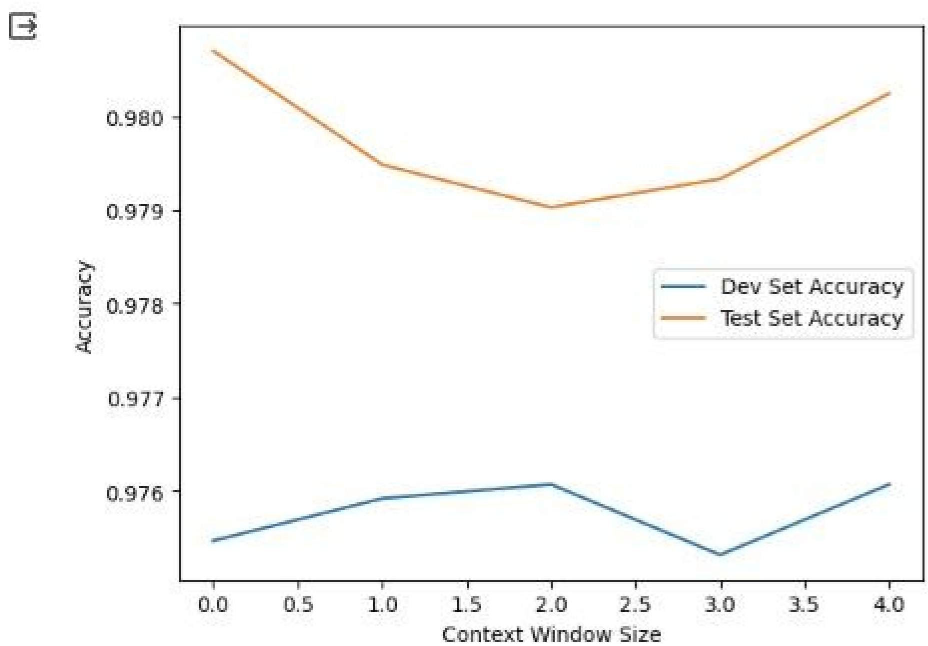
For window size=4

```
Confusion Matrix for Dev Set (Context Window Size = 4):
[[ 626    0    0    0    0    0    2   14    7    4    0    0    0    0]
 [   1  125    0    0    0    0    0    0    2    1    2    0    0    0]
 [   0    1   49    0    1    0    0    0    6    2    0    0    0    0]
 [   0    0    0  105    0    0    0    1    0    0    1    0    0    0]
 [   0    0    0    0  542    0    0   15    0    0    0   11    0    0]
 [   0    0    0    0    0    0    0    1    0    1    0    0    0    0]
 [  12    0    0    0    0    0  254    0    0    0    0    0    0    0]
 [   0    0    1    0    1    0    0 1404    0    0    1    2    6    0]
 [   0    1    6    0    0    0    0    0  206   12    2    0    0    0]
 [   5    0    0    0    0    0    0    0    2 1124   12    0    0    0]
 [   3    0    0    0    0    0    0    1    0   10 1537    0    0    0]
 [   0    0    0    0    3    0    0    0    0    0    0  411    0    0]
 [   0    0    0    0    0    0    2    3    0    0    0    0   68    0]
 [   1    0    0    0    0    0    0    0    0    0    0    0    0   34]]
```

```
Confusion Matrix for Test Set (Context Window Size = 4):
[[ 596    1    0    0    0    1   21    3    6    1    0    0    0]
 [   3  114    0    0    0    0    1    6    2    1    0    0    0]
 [   0    0   62    0    0    0    2    5    1    5    1    0    0]
 [   0    0    0  106    1    0    0    0    1    1    0    0    0]
 [   0    0    0    0  503    0    1    0    0    1    7    0    0]
 [   4    0    0    0    0  252    0    0    0    0    0    0    0]
 [   0    0    1    0    3    0 1427    0    0    0    2    1    0]
 [   0    3    2    0    0    0    0  210    4    1    0    0    0]
 [   4    2    0    0    0    0    1    0 1151    8    0    0    0]
 [   2    1    0    0    0    0    1    0    9 1553    1    0    0]
 [   0    0    0    0    2    0    0    1    0    0  389    0    0]
 [   0    0    0    0    0    0    2    0    0    0    0   54    0]
 [   2    0    0    0    0    0    0    1    0    0    0    0   33]]
```

Plot for Epoch vs Development set Accuracy





Analysis of Results for FFNN:

- The FFNN model achieved competitive performance on the test set, demonstrating its effectiveness in POS tagging.
- The training loss trend and development set accuracy trend provided insights into the model's learning process and convergence.

- Confusion matrices helped identify areas where the model struggled to correctly classify certain POS tags.

Conclusion:

Both the LSTM and FFNN models demonstrated competitive performance in POS tagging, achieving satisfactory accuracy on the test set. By carefully selecting and tuning hyperparameters, it's possible to optimize model performance and generalize well to unseen data. Further experimentation with different architectures and optimization techniques could lead to even better results