Report

Aadya Ranjan

February 2024

1 Hyperparameters used to train the models

- **Input Size**: The input size refers to the dimensionality of the input features fed into the neural network model. Each word in the input sentence is usually represented as a one-hot encoded vector, and the input size determines the length of these vectors.It is crucial for defining the dimensionality of the input feature space, ensuring that the model can effectively process input data and learn meaningful representations.
- **Hidden Size**: The hidden size specifies the number of neurons or units in the hidden layer(s) of the neural network.It determines the capacity or expressive power of the model to learn complex patterns and representations from the input data.
- **Output Size**: The output size indicates the dimensionality of the output layer of the neural network.In POS tagging, the output size typically corresponds to the number of possible POS tags in the dataset. It determines the number of classes that the model can classify. In POS tagging, a larger output size accommodates a broader range of POS tags, enabling the model to produce more nuanced and accurate predictions.

2 Corresponding Graphs and Evaluation metrics

FFNN

For the **Feedforward neural network**, I have implemented the following measures and parameters:

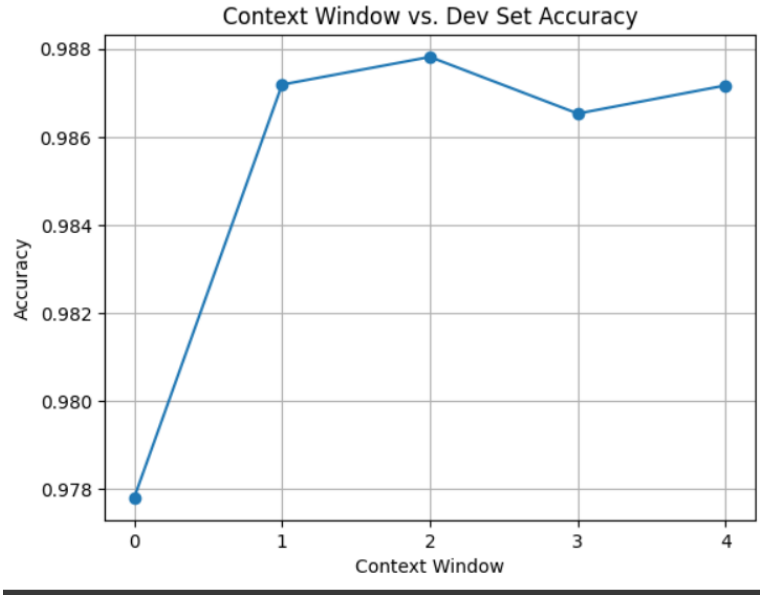
Validation and test set accuracy were assessed with the following parameter settings:

- p = 2
- s = 2

- Input size = 4330
- Hidden size = 256
- Output size = 14

1. Accuracy on dev data: 0.9877853881278539
2. Accuracy on test data: 0.9877917414721723

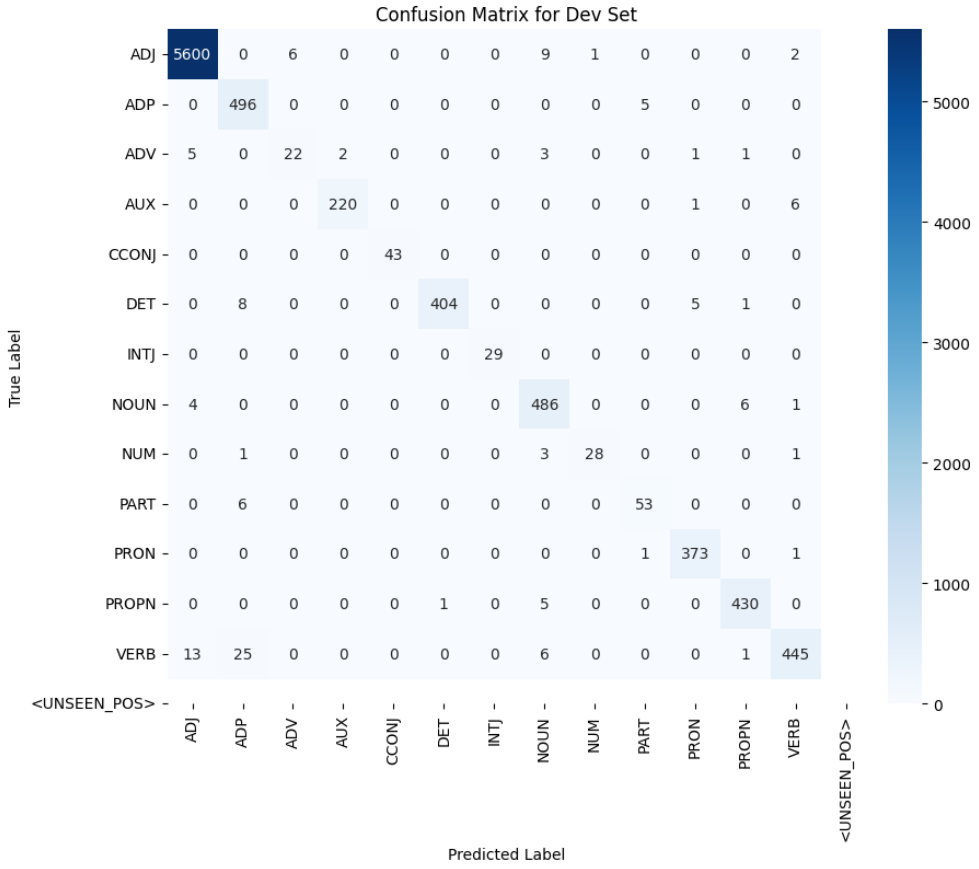
Dev set accuracy vs context window 0-4



3 Accuracy, Recall, F1 and Confusion matrix

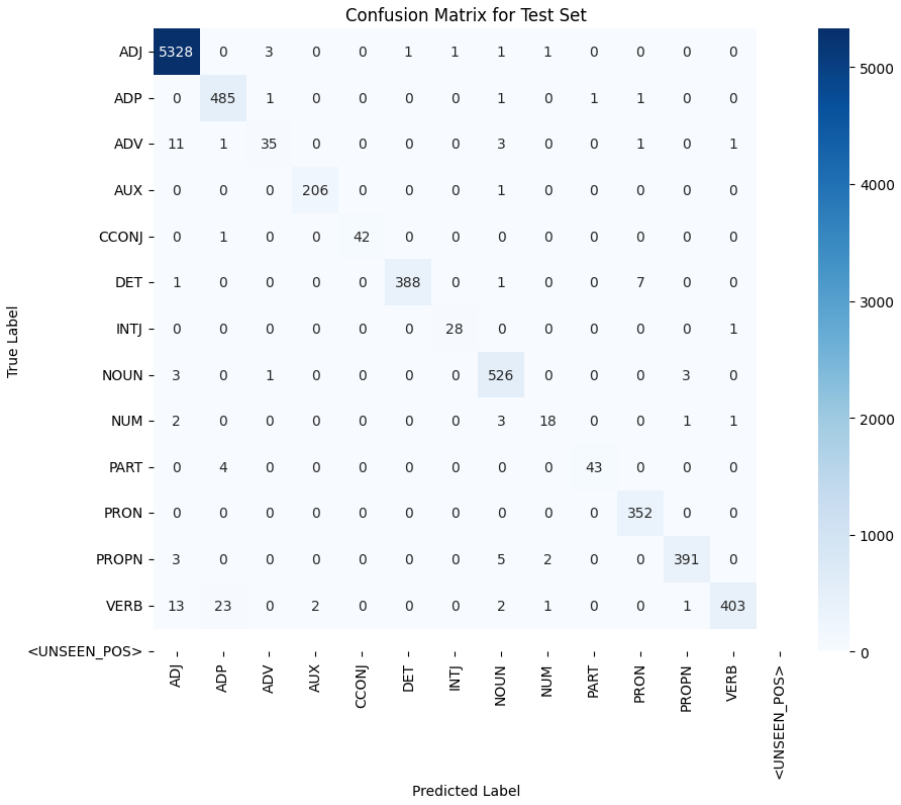
3.1 Validation Data Set

- Accuracy: 0.9850456621004566
- Recall: 0.9371746591989396
- F1 Score: 0.9462349741799996



3.2 Test Data Set

- Accuracy: 0.9868342309994016
- Recall: 0.9292912376000925
- F1 Score: 0.9432626786164058



RNN(LSTM)

The validation and test accuracies for the specified parameters:

- Embedding Dimension: 128
- Hidden Dimension: 128
- Vocabulary Size: 863
- Output Size: 13
- Number of Layers: 1
- Bidirectional: Not enabled
- Activation Function: None

1. Dev Accuracy: 0.9525888019265503
2. Test Accuracy: 0.955775075987842

Tuning of Hyperparameters

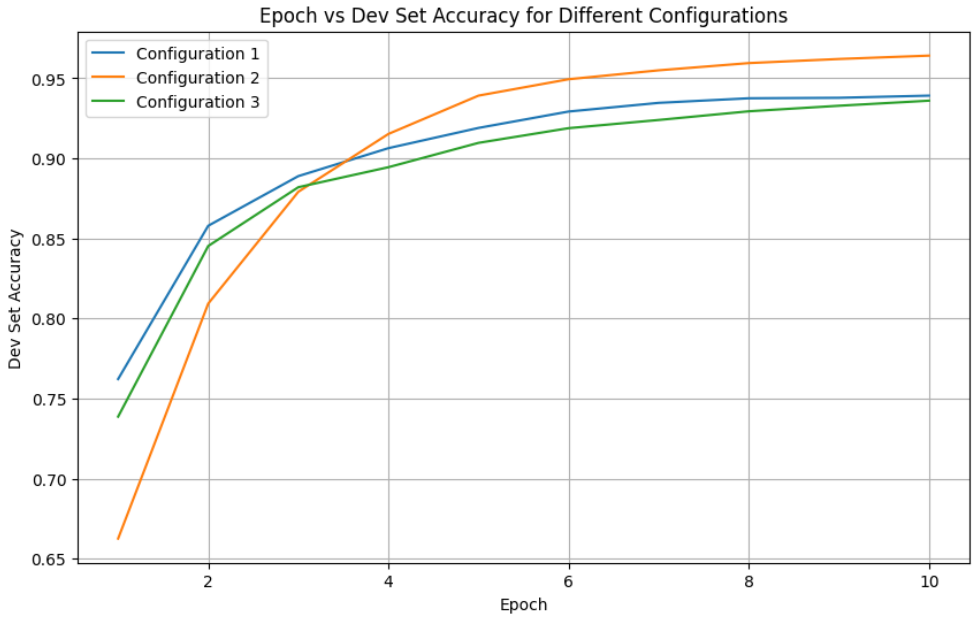
```
Training with configuration:
Number of Layers: 1, Bidirectional: False, Hidden Layer Size: 256, Embedding Size: 128, Activation Function: relu
100%|██████████| 5/5 [01:35<00:00, 19.18s/it]
Dev Accuracy: 0.9230885009030705

Training with configuration:
Number of Layers: 2, Bidirectional: True, Hidden Layer Size: 128, Embedding Size: 256, Activation Function: relu
100%|██████████| 5/5 [03:08<00:00, 37.76s/it]
Dev Accuracy: 0.9388922335942204

Training with configuration:
Number of Layers: 1, Bidirectional: False, Hidden Layer Size: 64, Embedding Size: 64, Activation Function: tanh
100%|██████████| 5/5 [00:35<00:00,  7.19s/it]
Dev Accuracy: 0.9104455147501506

Best performing configuration on dev set:
Number of Layers: 2, Bidirectional: True, Hidden Layer Size: 128, Embedding Size: 256, Activation Function: relu
100%|██████████| 5/5 [02:55<00:00, 35.08s/it]
Test Accuracy for the best configuration: 0.04939209726443769
```

Dev Set Accuracy vs Epoch



Evaluation Metrics on Dev Set

- Accuracy: 0.9360325105358218

- Micro Recall: 0.9360325105358218, Macro Recall: 0.8616604521543635
- Micro Precision: 0.9360325105358218, Macro Precision: 0.9356413513144312
- Micro F1 Score: 0.9360325105358218, Macro F1 Score: 0.8914983182210985

```
Confusion Matrix:
[[ 313    1  100    0    0    0    1    0    1    0    0    0    0]
 [   0  243    0    2    0    2   18    0    1    0    0    0    0]
 [   1    1  549    1   11    4    1    0    0    0    0    0    0]
 [   1    1    6 1101    1   19    7    0    7    0    0    0    0]
 [   1    0    0    4 1397    3    1    2    1    0    0    6    0]
 [   0    0    5   12    0 1531    2    0    1    0    0    0    0]
 [   1    2    4    7   66    4  568    0    1    0    0    0    0]
 [   1    0   11    4    1   10    1   98    5    0    0    0    0]
 [   0    0    1   19    0    4    4    1  193    0    5    0    0]
 [   0    0    0    0    0    1    0    0    0  106    0    0    0]
 [   0    0    2   10    0    5    0    1   11    0   30    0    0]
 [   0    0    0    0   12    0    0    0    0    0    0   61    0]
 [   0    0    0    0    0    0    6    0    0    0    0    0   29]]
```

Evaluation Metrics on Test Set

- Accuracy: 0.935258358662614
- Micro Recall: 0.935258358662614, Macro Recall: 0.8684868689352877
- Micro Precision: 0.935258358662614, Macro Precision: 0.9399451328626903
- Micro F1 Score: 0.935258358662614, Macro F1 Score: 0.8935717920494504

```
Confusion Matrix:
[[ 287    0  104    0    0    0    0    1    0    0    0    0    0]
 [   0  231    2    4    0    1   18    0    0    0    0    0    0]
 [   1    0  500    3    1    3    4    0    0    0    0    0    0]
 [   0    0    2 1131    0   17    8    1    7    0    0    0    0]
 [   0    0    2    3 1418    1    1    4    3    0    0    2    0]
 [   0    0   10    9    1 1542    4    0    1    0    0    0    0]
 [   0   10    7    5   65    5  533    1    3    0    0    0    0]
 [   3    0   17    5    2   12    2   83    3    0    0    0    0]
 [   0    0    0    6    0    8    5    1  200    0    0    0    0]
 [   0    1    0    1    0    1    0    0    1  105    0    0    0]
 [   2    1    0    7    2    4    3    1   18    0   38    0    0]
 [   0    0    0    0    3    0    0    0    0    0    0   53    0]
 [   0    0    0    0    1    0    2    0    0    0    0    0   33]]
```