ADVANCED NLP ASSIGNMENT 2
2022701012 Darshana S
Report

Hyperparameters

	Language Modeling	Classification
Batch Size	256	256
Loss Function	Cross Entropy Loss	Cross Entropy Loss
Optimizer	Adam	Adam
Learning Rate	0.001	0.001
Epochs	1	5

Language Modeling Results

Train Perplexity	Dev Perplexity	Test Perplexity
134.84248416454332	137.47907827155166	138.31136286139875

Classification Results

Train Set

```
              precision    recall  f1-score   support

           0       0.67      0.72      0.69      9983
           1       0.46      0.56      0.50      9988
           2       0.50      0.40      0.45      9978
           3       0.52      0.36      0.42      9984
           4       0.61      0.74      0.67      9987

    accuracy                           0.56     49920
   macro avg       0.55      0.56      0.55     49920
weighted avg       0.55      0.56      0.55     49920
```
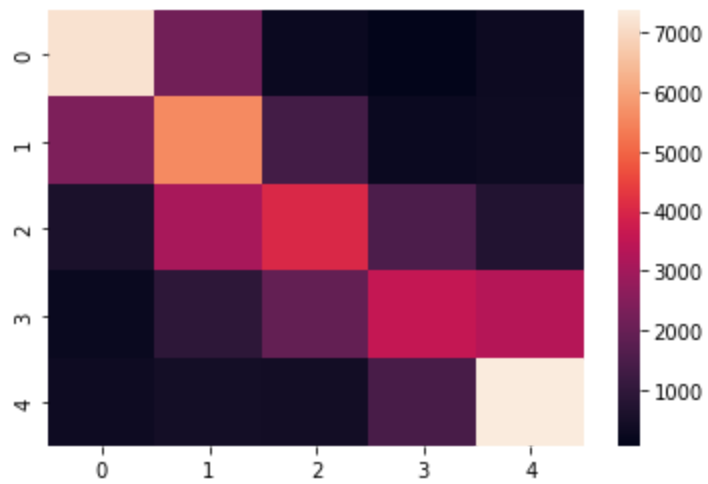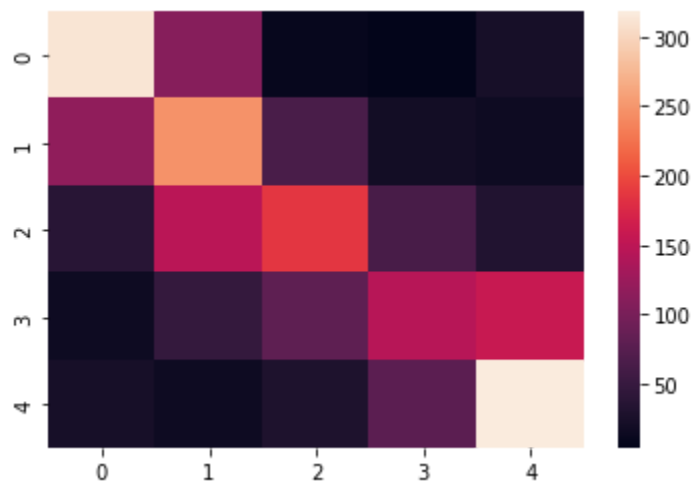
Dev Set

	precision	recall	f1-score	support
0	0.62	0.68	0.65	460
1	0.43	0.53	0.48	465
2	0.50	0.40	0.44	467
3	0.47	0.33	0.38	448
4	0.58	0.69	0.63	464
accuracy			0.53	2304
macro avg	0.52	0.52	0.52	2304
weighted avg	0.52	0.53	0.52	2304

Test Set

```
              precision     recall   f1-score     support

           0       0.62       0.65       0.63         970
           1       0.40       0.51       0.45         970
           2       0.45       0.37       0.41         975
           3       0.46       0.29       0.36         974
           4       0.58       0.70       0.63         975

    accuracy                             0.50        4864
   macro avg       0.50       0.50       0.50        4864
weighted avg       0.50       0.50       0.50        4864
```
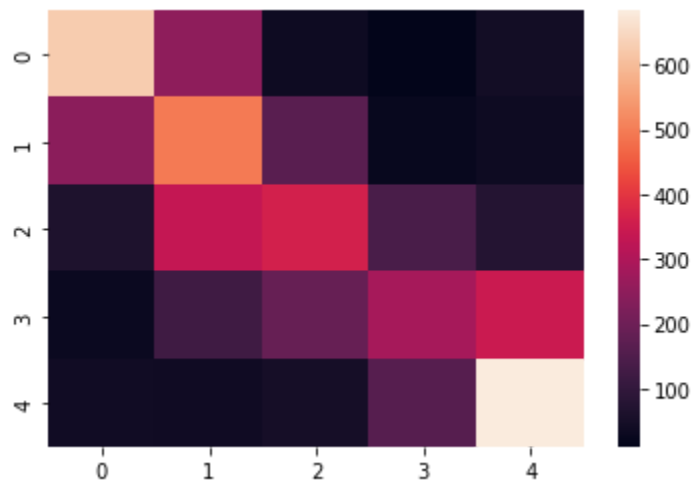

Bonus 1

Accuracy comparison for different weighting schemes:

Input Only	0.29
Layer 1 Only	0.31
Layer 2 Only	0.31
Layer 1 Favoured	0.44
Equal	0.50

1. Input Only. Weights = [1, 0, 0]

	precision	recall	f1-score	support
0	0.36	0.29	0.32	970
1	0.24	0.38	0.29	970
2	0.26	0.33	0.29	975
3	0.29	0.06	0.10	974
4	0.36	0.41	0.38	975
accuracy			0.29	4864
macro avg	0.30	0.29	0.28	4864
weighted avg	0.30	0.29	0.28	4864

2. Layer 2 Only. Weights = [0, 0, 1]

	precision	recall	f1-score	support
0	0.29	0.47	0.36	970
1	0.27	0.20	0.23	970
2	0.31	0.39	0.34	975
3	0.37	0.07	0.12	974
4	0.38	0.43	0.40	975
accuracy			0.31	4864
macro avg	0.32	0.31	0.29	4864
weighted avg	0.32	0.31	0.29	4864

3. Layer 1 Only. Weights = [0, 1, 0]

	precision	recall	f1-score	support
0	0.30	0.49	0.37	970
1	0.22	0.01	0.02	970
2	0.26	0.54	0.35	975
3	0.32	0.26	0.28	974
4	0.55	0.23	0.32	975
accuracy			0.31	4864
macro avg	0.33	0.31	0.27	4864
weighted avg	0.33	0.31	0.27	4864

4. Equal Weighting. Weights = [1/3, 1/3, 1/3]

	precision	recall	f1-score	support
0	0.62	0.65	0.63	970
1	0.40	0.51	0.45	970
2	0.45	0.37	0.41	975
3	0.46	0.29	0.36	974
4	0.58	0.70	0.63	975
accuracy			0.50	4864
macro avg	0.50	0.50	0.50	4864
weighted avg	0.50	0.50	0.50	4864

5. Layer 1 favoured weighting. Weights = [1/6, 2/3, 1/6]

	precision	recall	f1-score	support
0	0.58	0.51	0.54	970
1	0.42	0.21	0.28	970
2	0.34	0.35	0.35	975
3	0.34	0.38	0.36	974
4	0.49	0.73	0.59	975
accuracy			0.44	4864
macro avg	0.43	0.44	0.42	4864
weighted avg	0.43	0.44	0.42	4864