Laksita Prasanna

HW 1 – deep learning assignment

Question 1)

binary classification

Discussion: As learning rate increases, it takes a lot of epochs for the results to converge. Because the optimizer will take slower and slower steps in the direction of lower gradient. A lower learning rate is always a better choice that taking faster learning rates because the model will take a lot of caution when taking each step. And also, the model will avoid making huge jumps.

Top three models:

Rank 1: number of words=10000, Hidden layers=[(16, 4)], Learning rate = 0.001, Avg. Validation accuracy=0.7666380792856217

Rank 2: number of words=10000, Hidden layers=[(16, 4)], Learning rate = 0.1, Avg. Validation accuracy=0.7660377383232116

Rank 3: number of words=10000, Hidden layers=[(32, 16, 8, 4)], Learning rate = 0.001, Avg. Validation accuracy=0.7648370534181594

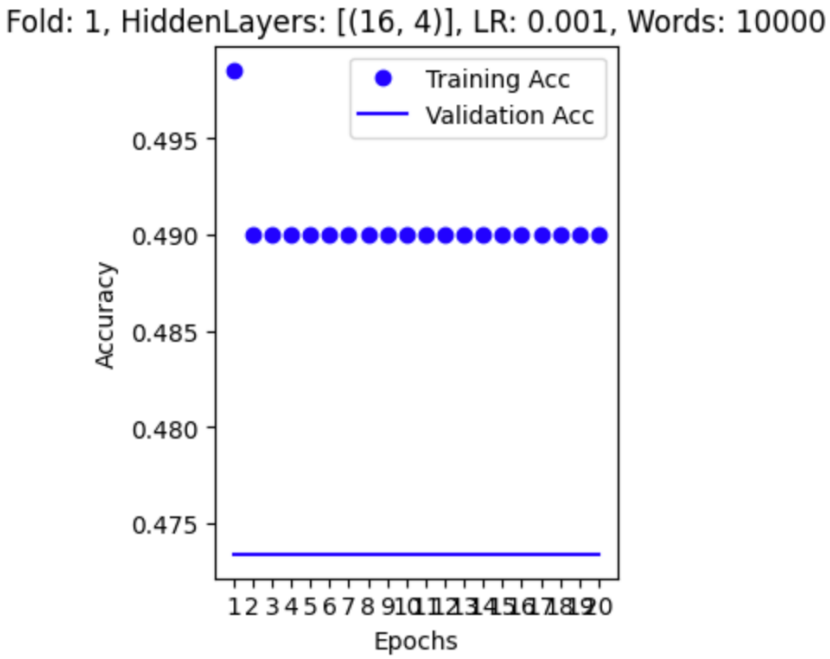


Figure 1

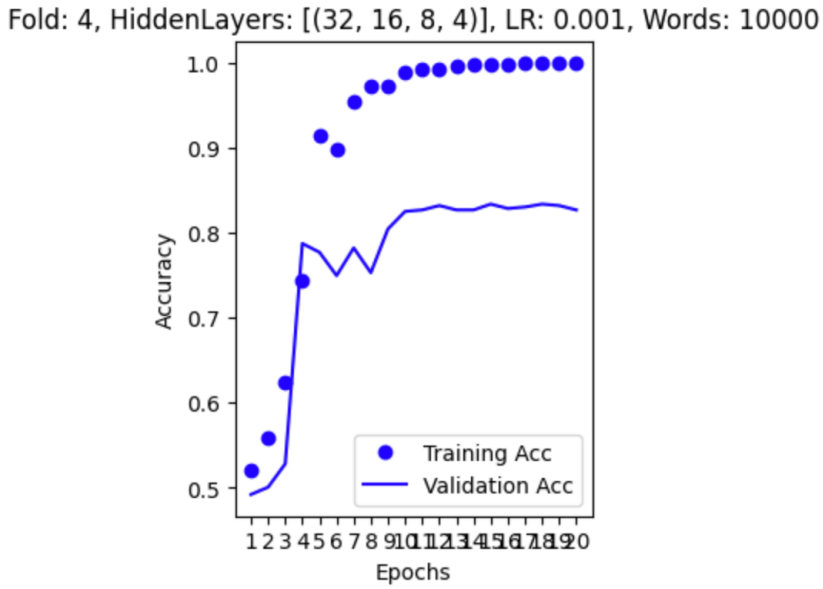


Figure 2

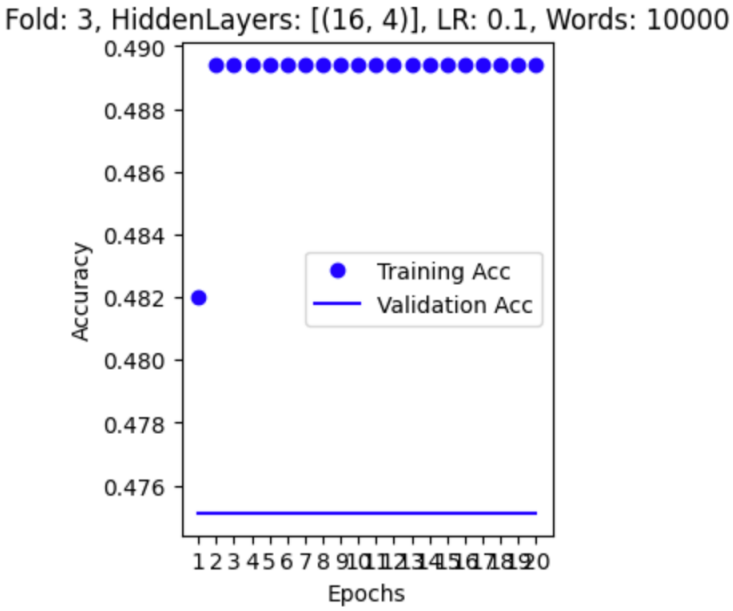

Fold: 3, HiddenLayers: [(16, 4)], LR: 0.1, Words: 10000
0.490
0.488
0.486
0.484
0.482
0.480
0.478
0.476
Accuracy
Training Acc
Validation Acc
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Epochs

Figure 3

Question 2:

Multiclass classification

Discussion: As learning rate increases, it takes a lot of epochs for the results to converge. Because the optimizer will take slower and slower steps in the direction of lower gradient. A lower learning rate is always a better choice that taking faster learning rates because the model will take a lot of caution when taking each step. And also, the model will avoid making huge jumps.

List of hyperparameters:

```
parameters = [
  [250, [(4,)], 0.1],
  [250, [(16,)], 0.1],
  [250, [(32, 16)], 0.1],
  [500, [(16,)], 0.1],
  [500, [(64, 32, 16)], 0.01],
  [1000, [(32,16)], 0.001],
  [1000, [(128, 64, 32, 16)], 0.001],
  [5000, [(64, 32, 16)], 0.0001],
  [5000, [(128, 64, 32, 16)], 0.001],
  [5000, [(128, 64, 32, 16)], 1e-05],
  [5000, [(256, 128, 64, 32, 16)], 0.0001],
  [5000, [(256, 128, 64, 32, 16)], 0.00001],
]
```

Top three models:

3 top choices

Ranking 1:

number of words=500, Hidden layers=[(64, 32, 16)], LearningRate = 0.01, Validation accuracy=0.7194654673337937

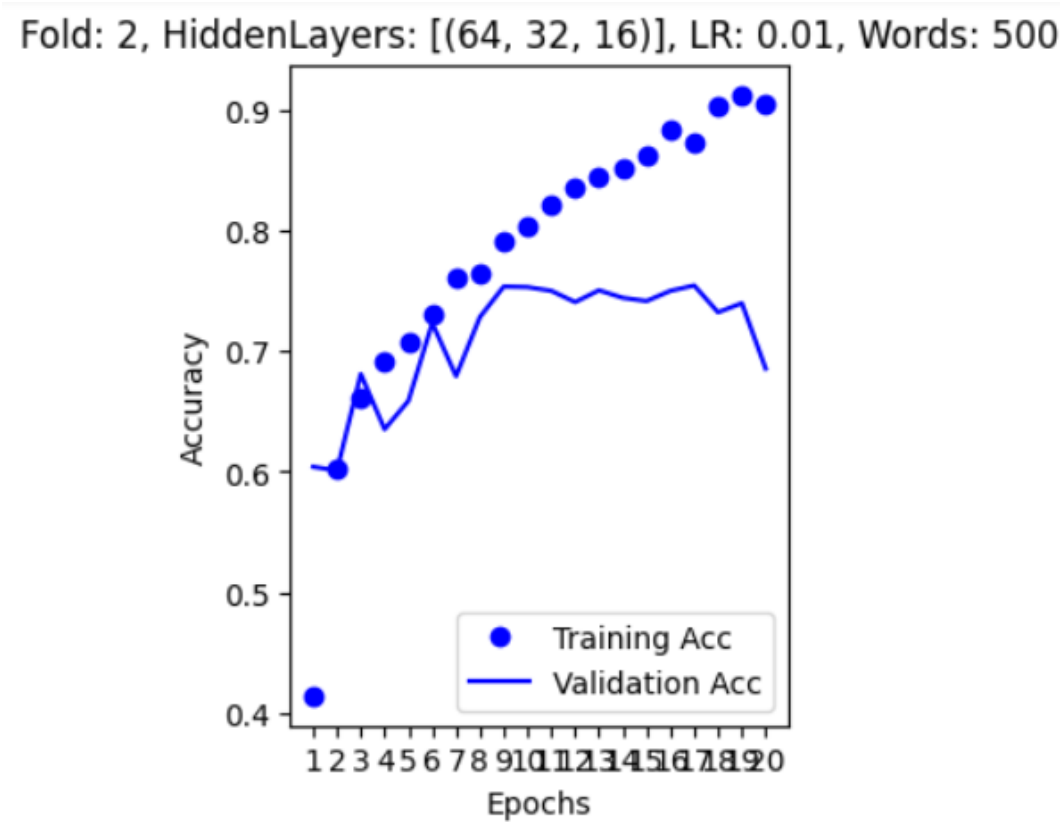


Figure 2a

Ranking 2:

number of words=500, Hidden layers=[(64, 32, 16)], LearningRate = 0.01, Validation accuracy=0.7105075716972351

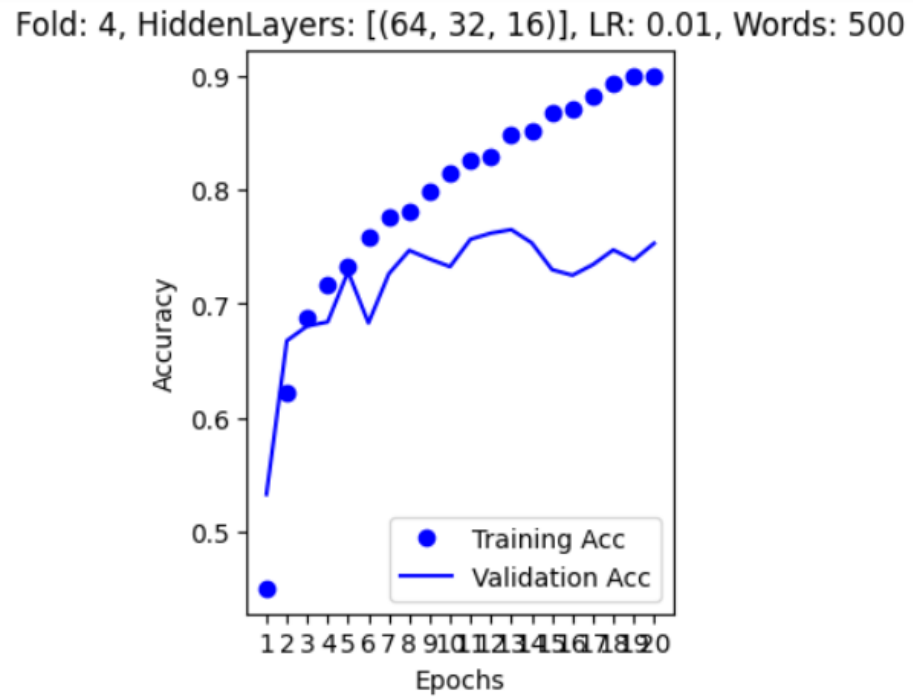
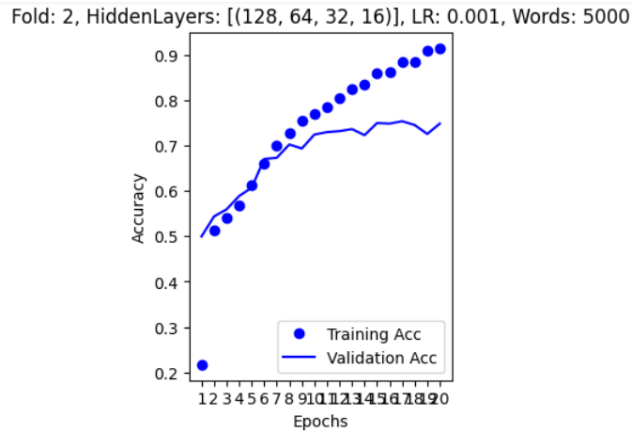


Figure 2b

Ranking 3:

number of words=5000, Hidden layers=[(128, 64, 32, 16)], LearningRate = 0.001, Validation accuracy=0.7088824540376664



Prediction

Model ranking 1: test accuracy = 0.5508, test loss = 1.7085

Model ranking 2: test accuracy = 0.5650, test loss = 1.7538

Model ranking 3: test accuracy = 0.5614, test loss = 2.1615