COL774 A3 Report

cs1210107

October 2023

1 Decision Trees

1.1 Varying depths without one hot encoding

Depth	Training Accuracy	Test Accuracy
5	0.5510	0.5377
10	0.5592	0.5387
15	0.5592	0.5387
20	0.5592	0.5387
25	0.5592	0.5387

Table 1: Without One-Hot Encoding

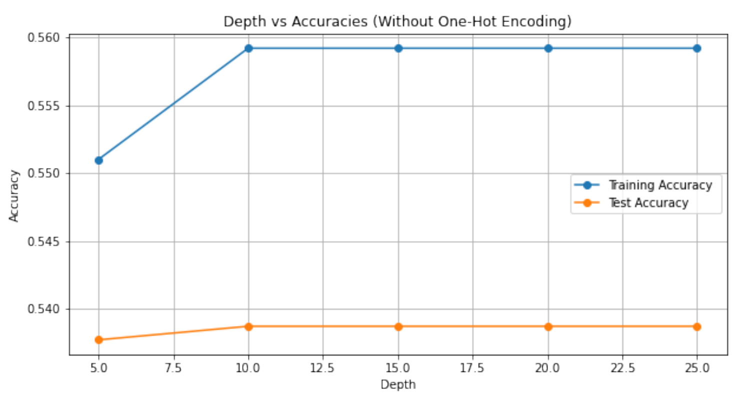


Figure 1: Without one hot encoding

our model is little bit more better than the above two prediction statergies.

Prediction Statergy	Training Accuracy	Test Accuracy
With Only Wins	0.5033	0.4964
With Only Losses	0.4966	0.5036

Table 2: Without One-Hot Encoding - Wins and Losses

1.2 Varying depths with one hot encoding

Depth	Training Accuracy	Test Accuracy
15	0.7053	0.5584
25	0.8483	0.6174
35	0.9245	0.6143
45	0.9900	0.6112

Table 3: With One-Hot Encoding

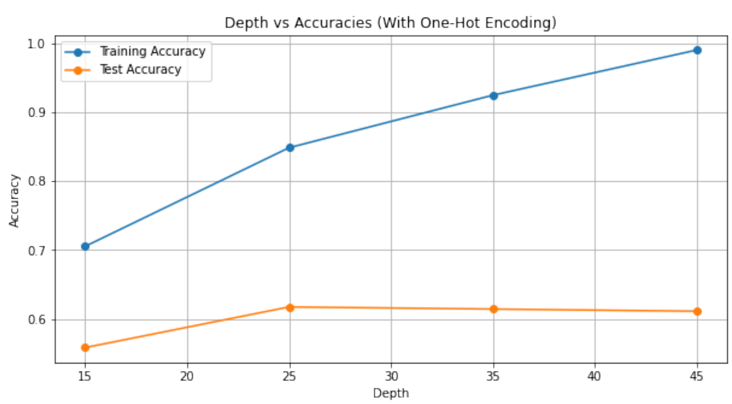


Figure 2: With one hot encoding

Onehot encoding increased the Training accuracy drastically. almost nearly 99% for depth = 45 . also test accuracy also increased a little bit.$\geq 5\%$.
So Decision trees with one hot encoding is a better model.

1.3 Pruning

Depth	Train Accuracy	Test Accuracy	Validation Accuracy
15	0.8243	0.5678	0.6053
25	0.7892	0.6223	0.6126
35	0.7443	0.6254	0.6199
45	0.7089	0.6390	0.6252

Table 4: Varying Depth

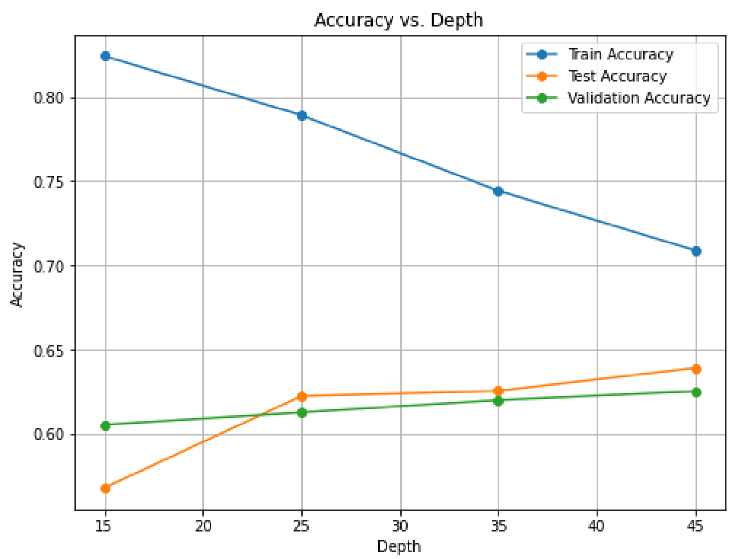


Figure 3: Accuracies after pruning

1.4 using Scikit Decision Tree classifier

1.4.1 varying depth with all other parameters fixed

Using scikit learn also gave us the same resluts as part b. training accuracy nearly 100% for depth=45 . and test accuracies are nearly 60%. by checking validation accuracy for depth = 15 has highest validation accuracy followed by depths 45,35,25.

Depth	Train Accuracy	Test Accuracy	Validation Accuracy
15	0.8243	0.6070	0.6448
25	0.9881	0.6246	0.6126
35	0.9996	0.6184	0.6299
45	1.0000	0.6070	0.6322

Table 5: Varying Depth

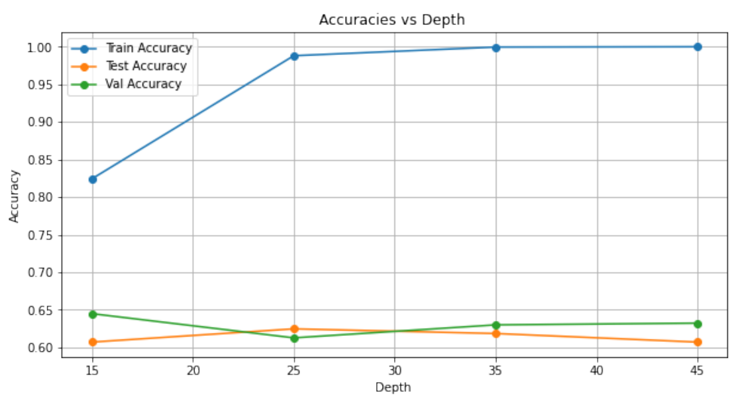


Figure 4: varying depth

1.4.2 varying alpha with all other parameters fixed

for ccp = 0.001 is the best model we got . and as the ccp paramater increased the accuracy gradually decreased.

CCP	Train Accuracy	Test Accuracy	Validation Accuracy
0.001	0.6946	0.6319	0.6402
0.01	0.5344	0.5181	0.5
0.1	0.5034	0.4964	0.4736
0.2	0.5034	0.4964	0.4736

Table 6: Varying CCP Values

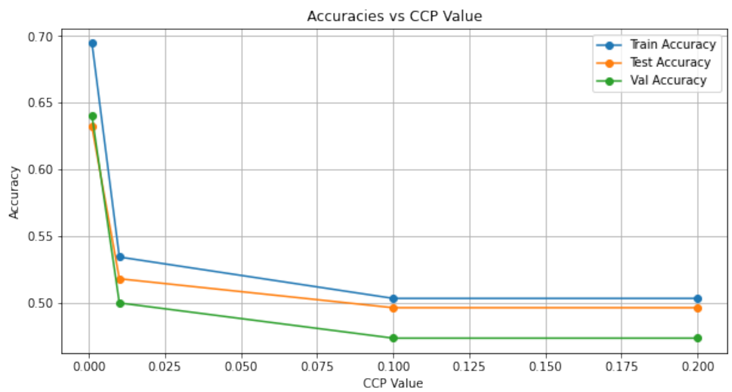


Figure 5: varying pruning parameter

1.5 Random Forests

Best parameters:
max features: 0.5
min samples split: 10
n estimators: 350
Training set accuracy: 0.9593714066692219
Validation set accuracy: 0.7241379310344828
Out-of-bag accuracy: 0.7276095566628338
Test set accuracy: 0.7218200620475698

1.6 Gradient Boosted trees

Best model using XG Boost
Depth: 15
min samples: 4
max feature: 0.7
Validation Accuracy: 0.6942528735632184

2 Neural Networks

2.1 Sigmoid Activation function

Stopping Criteria = 400 epochs

2.1.1 single hidden layers with varying sizes

as the number of units in the hidden layer increases the f1 score increased gradually.

Size	Precision	Recall	F1 score
1	0.61	0.21	0.31
5	0.74	0.71	0.72
10	0.80	0.78	0.79
50	0.84	0.83	0.83
100	0.83	0.83	0.83

Table 7: Training Data

Size	Precision	Recall	F1 score
1	0.59	0.19	0.28
5	0.70	0.68	0.69
10	0.78	0.77	0.77
50	0.79	0.79	0.79
100	0.81	0.80	0.80

Table 8: Test data

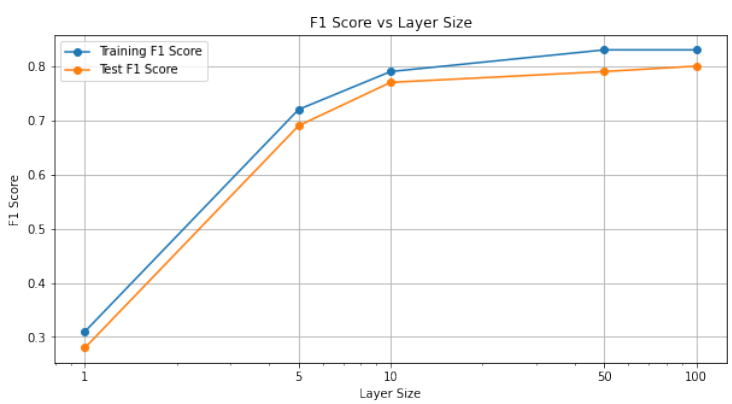


Figure 6: Single hidden layer with varying units

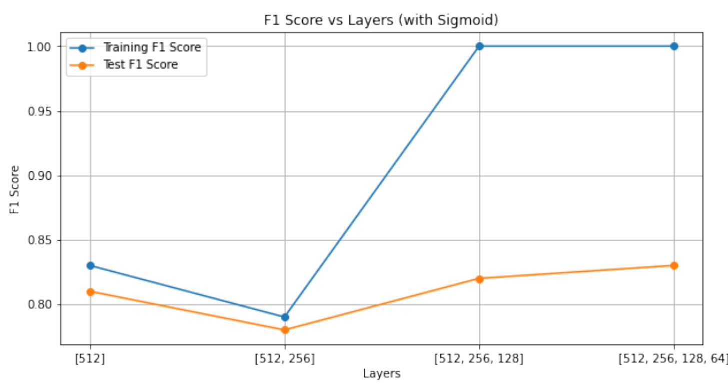
2.1.2 multiple hidden layers with fixed learning rate

Hidden layer Sizes	Precision	Recall	F1 score
[512]	0.83	0.83	0.83
[512,256]	0.80	0.79	0.79
[512,256,128]	1.00	1.00	1.00
[512,256,128,64]	1.00	1.00	1.00

Table 9: Training Data

Hidden layer Sizes	Precision	Recall	F1 score
[512]	0.81	0.81	0.81
[512,256]	0.78	0.78	0.78
[512,256,128]	0.82	0.82	0.82
[512,256,128,64]	0.82	0.82	0.82

Table 10: Test data



2.1.3 multiple hidden layers with adaptive learning

using adaptive learning decreased the f1 score if we still maintain the same stopping criteria 400 epochs it needed 1000 epochs to reach a better f1 score which are nearly the same as the above part. but it took significantly very long time to train around an hour.but it took around 20 minutes without adaptive learning.

Hidden layer Sizes	Precision	Recall	F1 score
[512]	0.70	0.69	0.69
[512,256]	0.86	0.85	0.85
[512,256,128]	0.87	0.87	0.87
[512,256,128,64]	0.89	0.89	0.89

Table 11: Training Data

Hidden layer Sizes	Precision	Recall	F1 score
[512]	0.69	0.69	0.69
[512,256]	0.80	0.80	0.80
[512,256,128]	0.83	0.83	0.83
[512,256,128,64]	0.82	0.82	0.82

Table 12: Test data

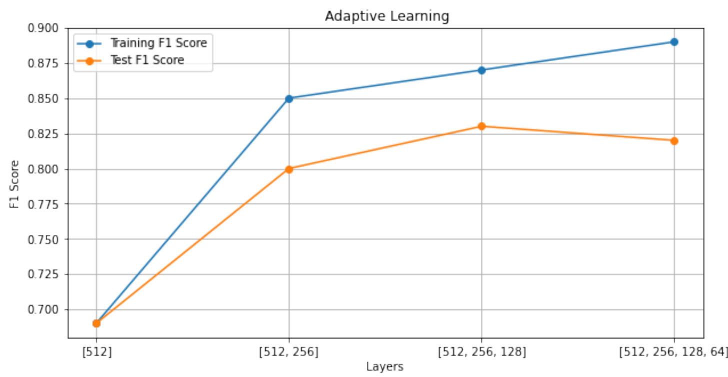


Figure 7: Adaptive Learning

2.2 ReLu activation function

Using relu increased the accuracy significantly and the training time also decreased a little around 45 minutes but significantly higher than the part c. overall it increased the accuracy. compared to part d and did so in lesser time.

Hidden layer Sizes	Precision	Recall	F1 score
[512]	0.80	0.79	0.80
[512,256]	0.86	0.85	0.85
[512,256,128]	0.81	0.81	0.81
[512,256,128,64]	0.80	0.80	0.83

Table 13: Training Data

Hidden layer Sizes	Precision	Recall	F1 score
[512]	0.78	0.78	0.78
[512,256]	0.83	0.83	0.82
[512,256,128]	0.81	0.81	0.80
[512,256,128,64]	0.83	0.83	0.82

Table 14: Test data

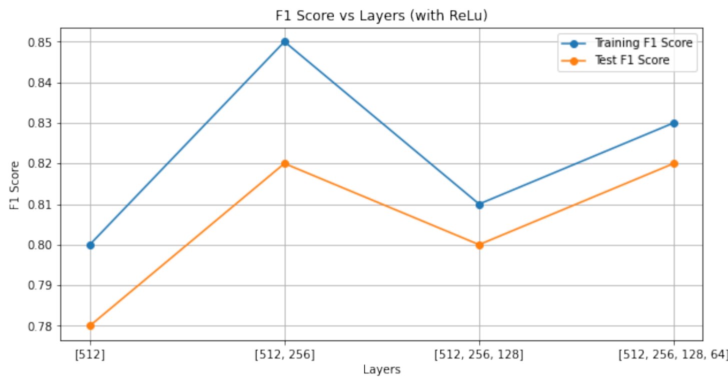


Figure 8: Caption

2.3 MLP Classifier

stopping criteria = 1000 epochs
compared to our model, MLP gave lesser F1 score.

Hidden layer Sizes	Precision	Recall	F1 score
[512]	0.56	0.58	0.56
[512,256]	0.59	0.61	0.60
[512,256,128]	0.61	0.62	0.62
[512,256,128,64]	0.62	0.63	0.63

Table 15: Training Data

Hidden layer Sizes	Precision	Recall	F1 score
[512]	0.53	0.55	0.54
[512,256]	0.59	0.60	0.59
[512,256,128]	0.61	0.61	0.61
[512,256,128,64]	0.62	0.62	0.62

Table 16: Test data

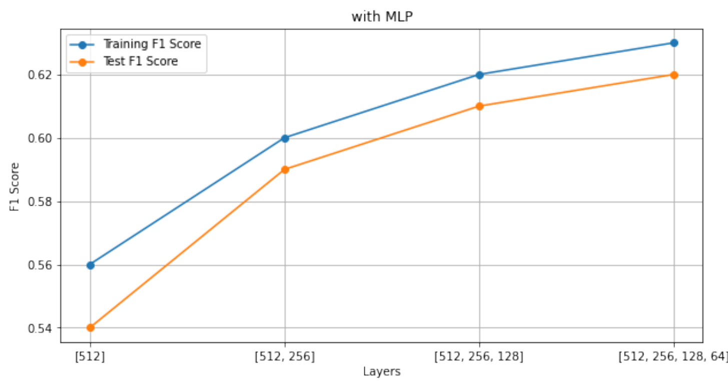


Figure 9: using MLP