

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN



LAB02

20127103 - Nguyễn Đức An

Môn học: Cơ sở AI

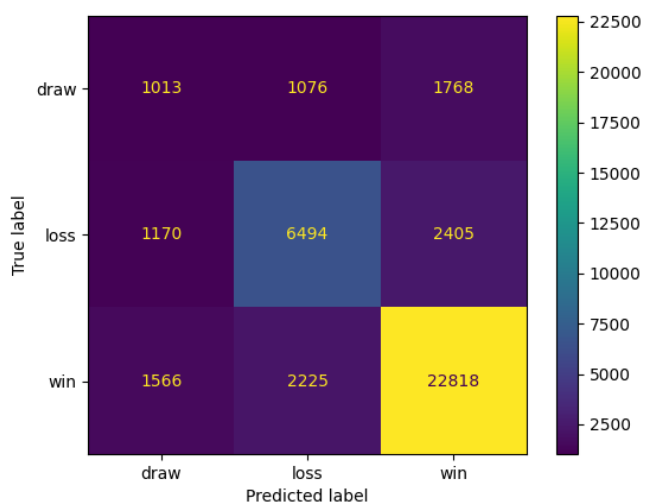
I) Bảng đánh giá

No.	Criteria	Scores	Completion percent
1	Preparing the datasets	20%	20%
2	Building the decision tree classifiers	20%	15%
3	Evaluating the decision tree classifiers		
	Classification report and confusion matrix	20%	20%
	Comments	10%	10%
6	The depth and accuracy of a decision tree		
	Trees, tables, and charts	20%	20%
	Comments	10%	10%
Total	100%	100%	100%

Evaluating the decision tree classifiers

Train/Test is 40/60 with max_depth = 35 (auto)

Confusion Matrix:



Classification Report:

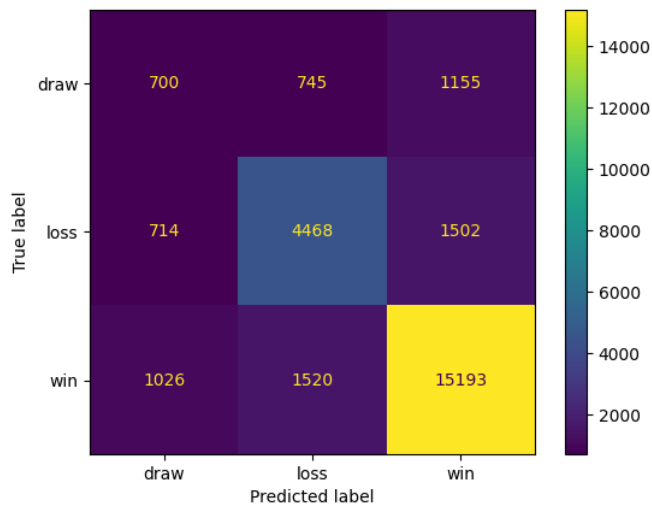
```
max_depth 35
Accuracy: 0.7481189095843098
Start plotting confusion matrix phase...
      precision    recall  f1-score   support

 draw      0.27      0.26      0.27      3857
 loss      0.66      0.64      0.65     10069
 win      0.85      0.86      0.85     26609

 accuracy              0.75     40535
 macro avg      0.59      0.59      0.59     40535
 weighted avg   0.75      0.75      0.75     40535
```

Train/Test is 60/40 with max_depth = 32 (auto)

Confusion Matrix:



Classification Report:

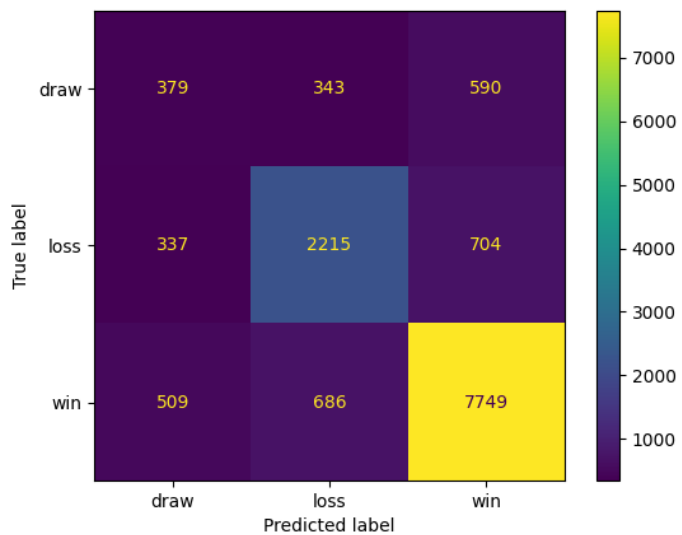
```
max_depth 32
Accuracy: 0.7534692669207712
Start plotting confusion matrix phase...
      precision    recall  f1-score   support

 draw      0.29      0.27      0.28      2600
 loss      0.66      0.67      0.67      6684
 win      0.85      0.86      0.85     17739

 accuracy              0.75     27023
 macro avg      0.60      0.60      0.60     27023
 weighted avg   0.75      0.75      0.75     27023
```

Train/Test is 80/20 with max_depth = 34 (auto)

Confusion Matrix:



Classification Report:

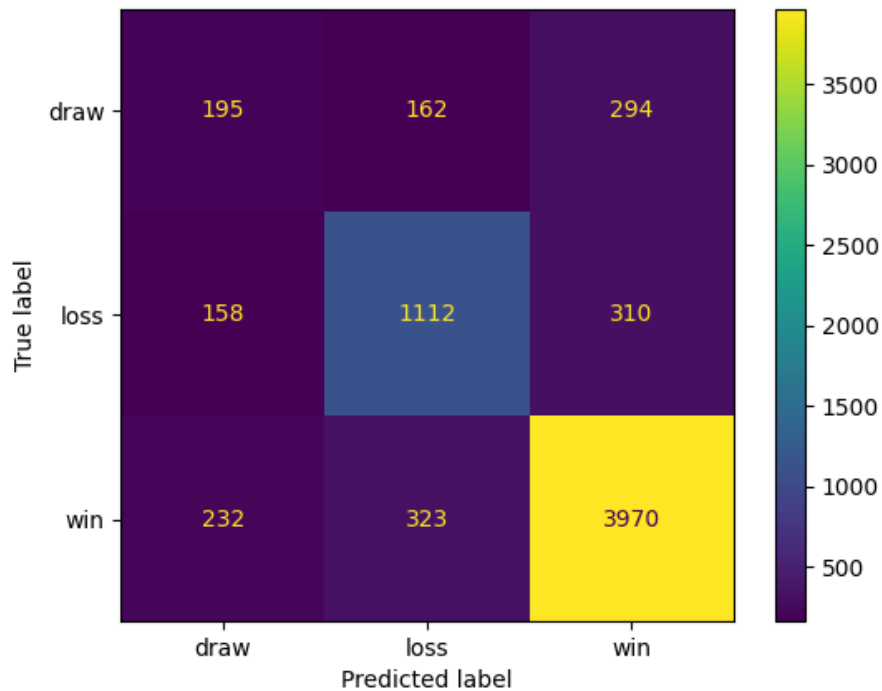
```
max_depth 34
Accuracy: 0.7654677323860273
Start plotting confusion matrix phase...
      precision    recall  f1-score   support

 draw      0.31      0.29      0.30      1312
 loss      0.68      0.68      0.68      3256
 win      0.86      0.87      0.86      8944

 accuracy            0.77      13512
 macro avg      0.62      0.61      0.61      13512
 weighted avg    0.76      0.77      0.76      13512
```

Train/Test is 90/10 with max_depth = 36 (auto)

Confusion Matrix:



Classification Report:

```
max_depth 36
Accuracy: 0.7810834813499112
Start plotting confusion matrix phase...
      precision    recall  f1-score   support

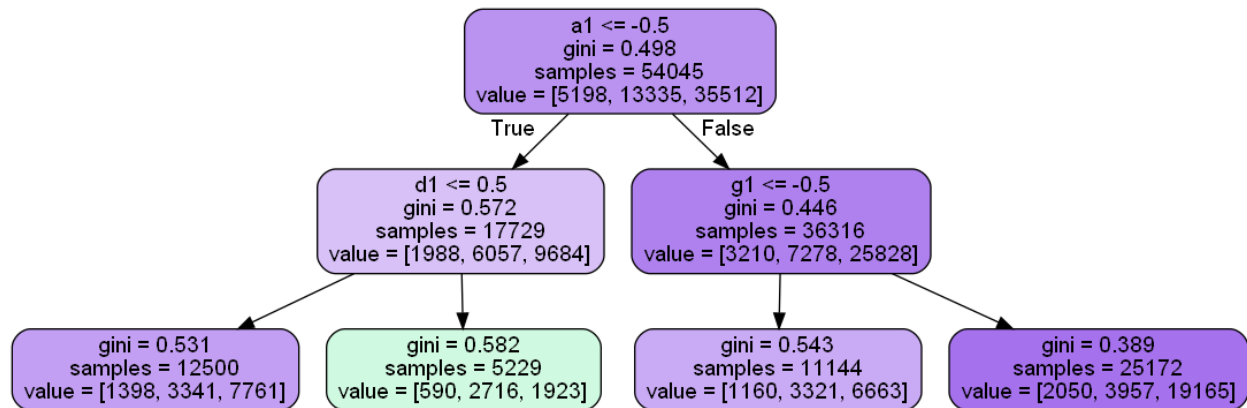
 draw      0.33      0.30      0.32         651
 loss      0.70      0.70      0.70        1580
 win      0.87      0.88      0.87        4525

 accuracy          0.78        6756
 macro avg      0.63      0.63      0.63        6756
 weighted avg   0.78      0.78      0.78        6756
```

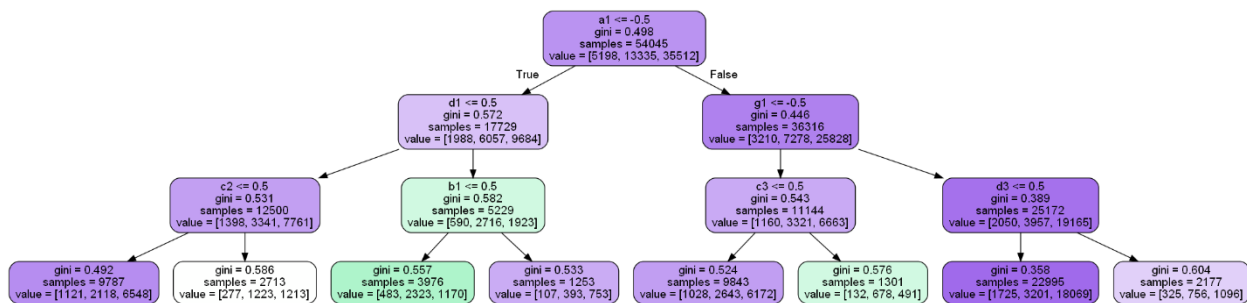
Comment: if max_depth of decision tree is set to None, nodes will expand to all leaf nodes, and here is max_depth reason to make up to always the 30

Max depth = None (auto max_depth = 35)

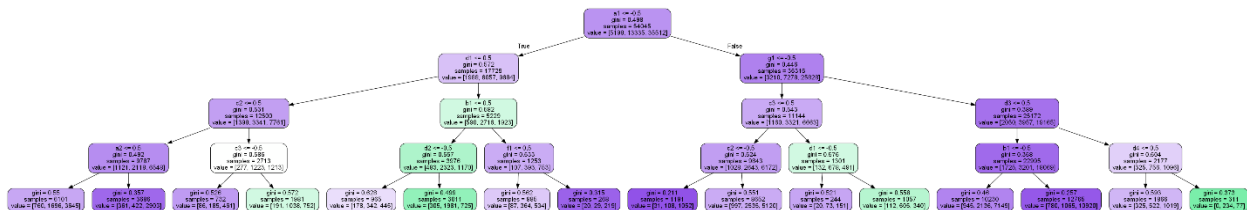
Max depth = 2



Max depth = 3



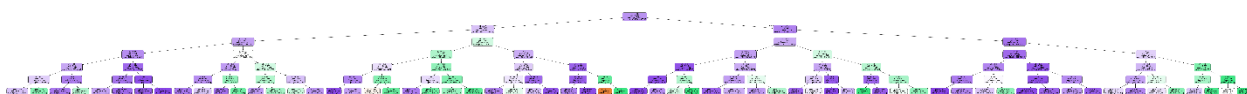
Max depth = 4



Max depth = 5



Max depth = 6



Max depth = 7



```
max_depth 2
Accuracy:  0.6759177027827117
Start plotting tree phase...
```

```
max_depth 3
Accuracy:  0.6898312611012434
Start plotting tree phase...
```

```
max_depth 4
Accuracy:  0.7005624629958556
Start plotting tree phase...
```

```
max_depth 5
Accuracy:  0.7087773830669035
Start plotting tree phase...
```

```
max_depth 6
Accuracy:  0.7212847838957963
Start plotting tree phase...
```

```
max_depth 7
Accuracy:  0.7252812314979278
Start plotting tree phase...
```

Comment: Accuracy increases proportionally to max_depth. The higher the max_depth is, the higher the accuracy is.

References

[1] Scikit-learn decision trees: <https://scikit-learn.org/stable/modules/tree.html>

[2] Tutorial: <https://github.com/lamnguyen5464/Connect-4-decision-tree>

[3] How to use sklearn.tree:

<https://scikitlearn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html#sklearn.tree.DecisionTreeClassifier>