

Decision Tree Classifier

WITH SCIKIT LEARN

Data

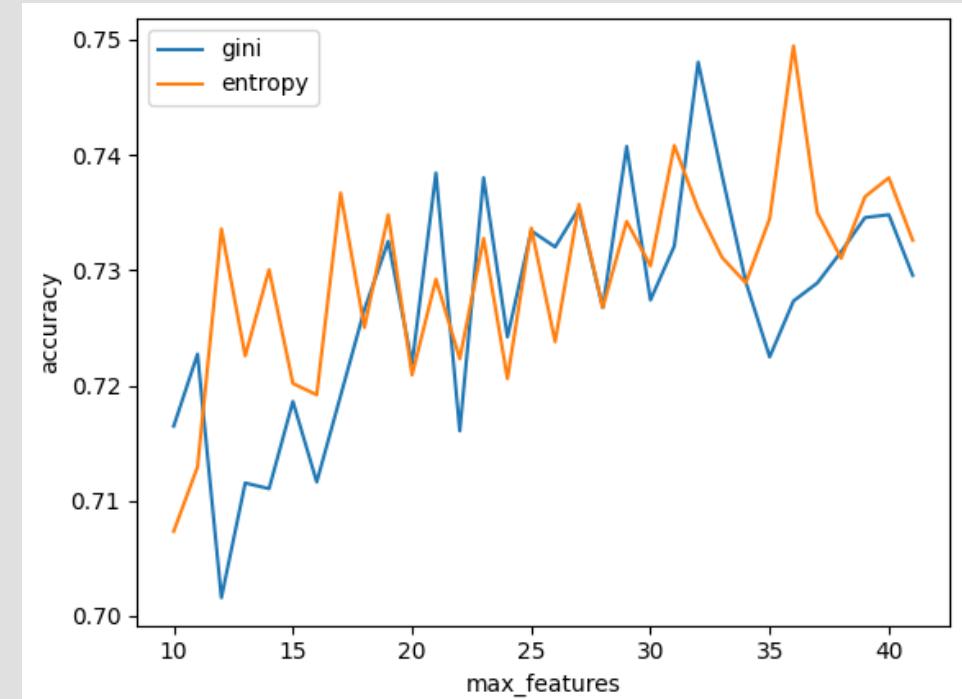
- 42 features → (blank, X or O)
- 3 output classes → (win, loose or draw)
- 67557 samples
- Date is splitted into three parts
- (70% train, 20% validation and 10% test)

Decision tree classifier hyperparameters

- Criterion (gini, entropy , log loss)
- Max features
- Max depth
- Min sample split
- Min sample leaf
- Max leaf nodes
- Cost complexity pruning alpha (ccp_alpha)

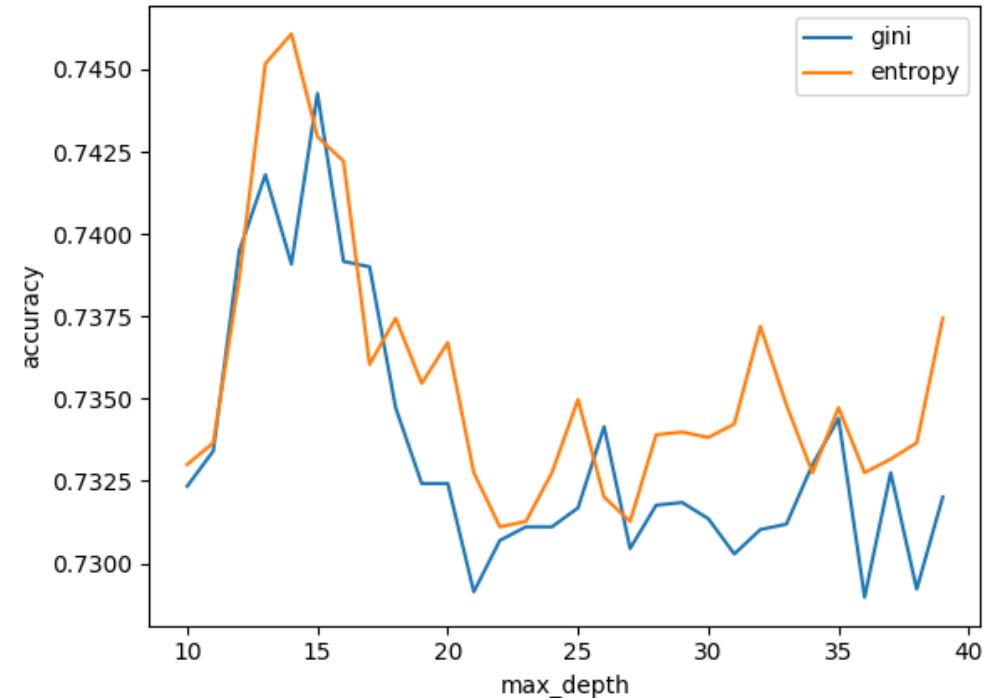
Max features

- Best max feature with entropy criterion = 37
- Best max feature with gini criterion = 33
- Overfit



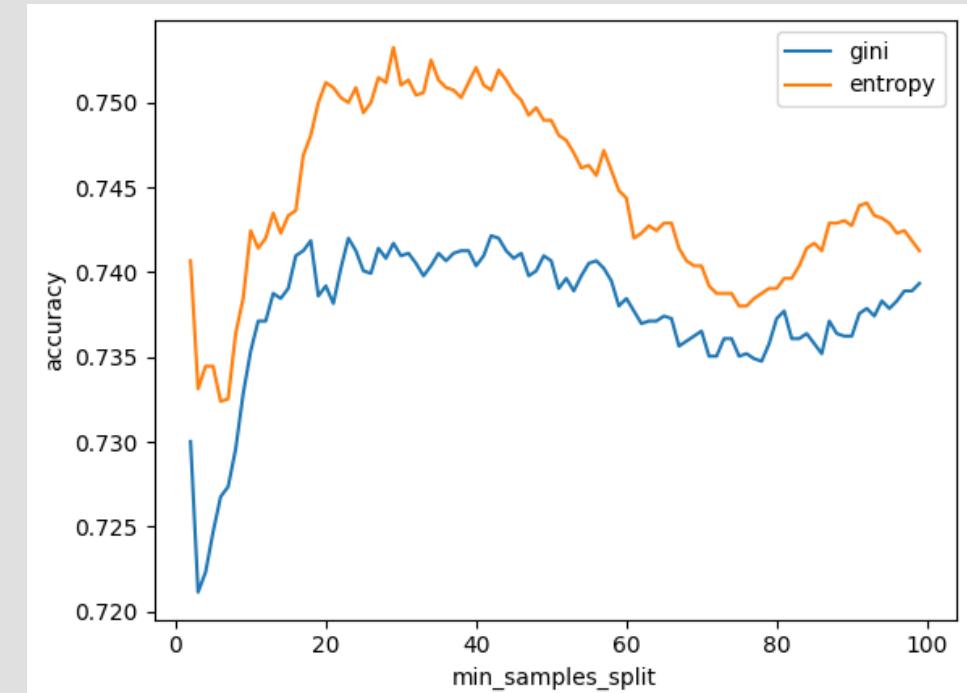
Max depth

- Best max depth with entropy criterion = 15
- Best max depth with gini criterion = 16
- Overfit



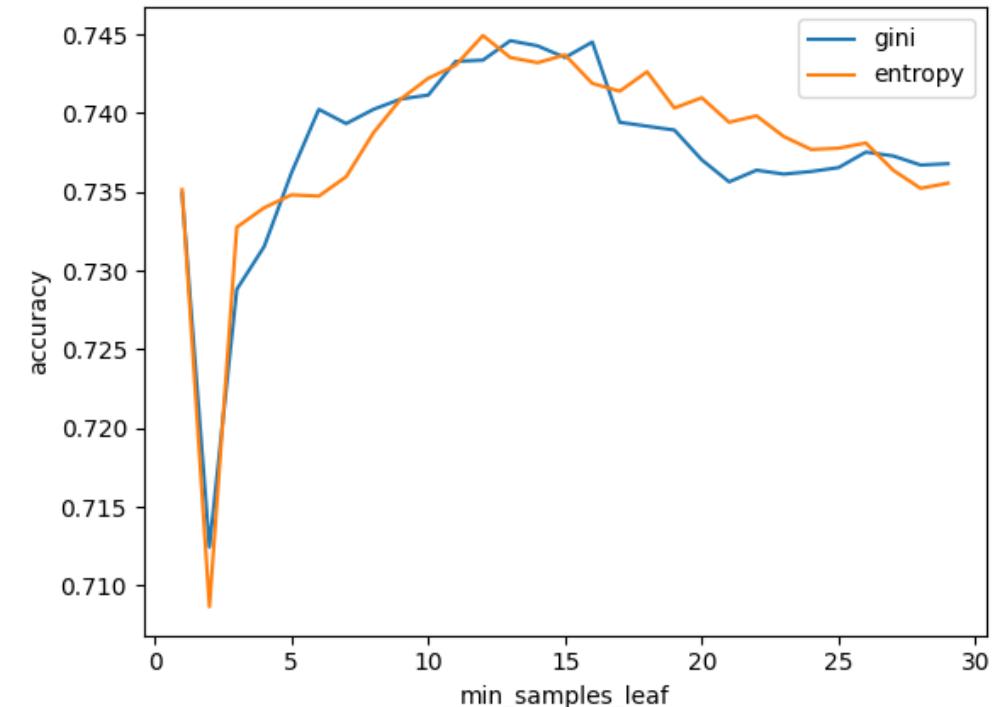
Min sample split

- Best max feature with entropy criterion = 30
- Best max feature with gini criterion = 43
- Overfit



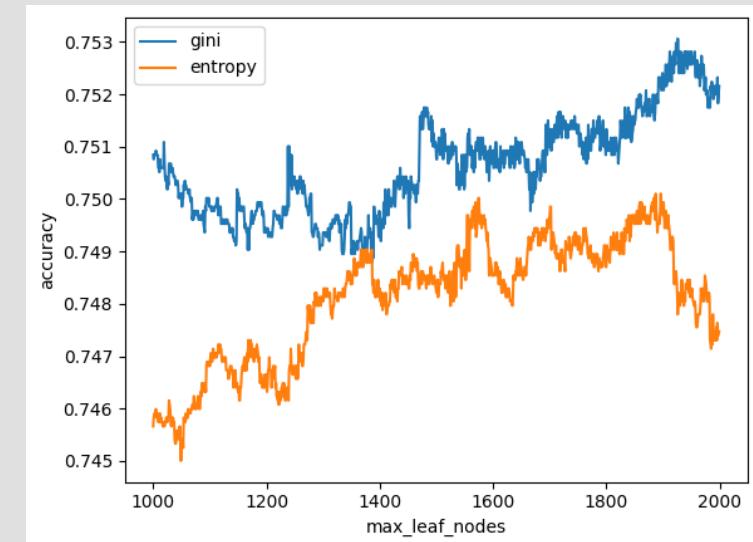
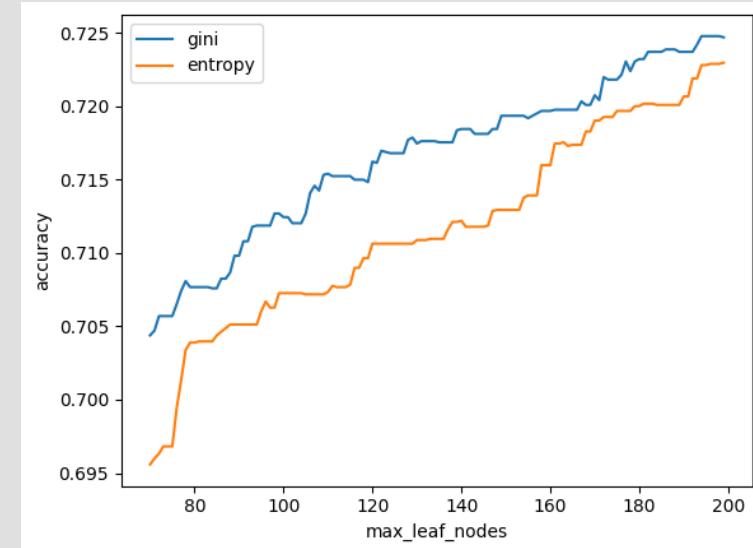
Min sample leaf

- Best max feature with entropy criterion = 13
- Best max feature with gini criterion = 14
- Overfit



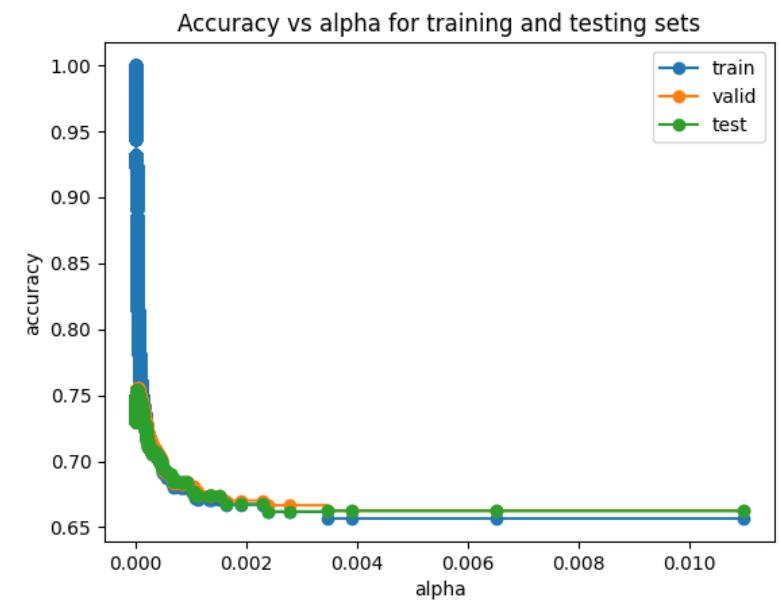
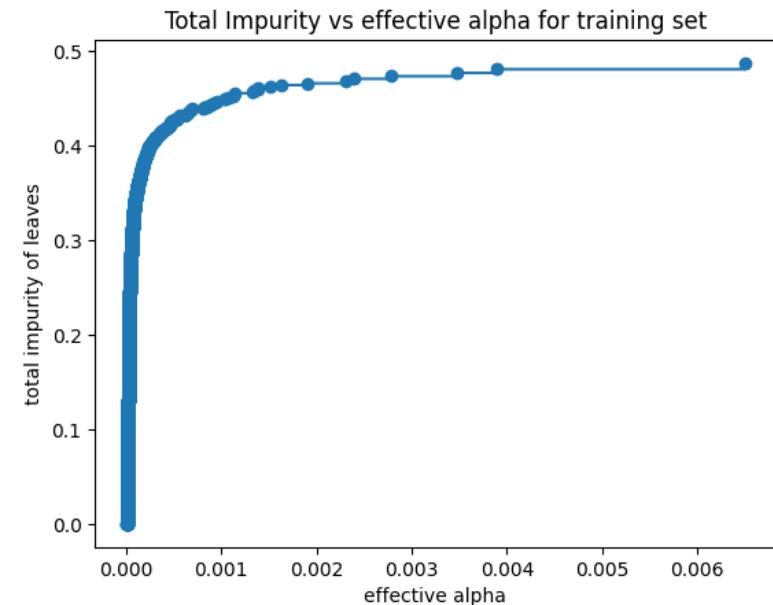
Max leaf nodes

- Overfit



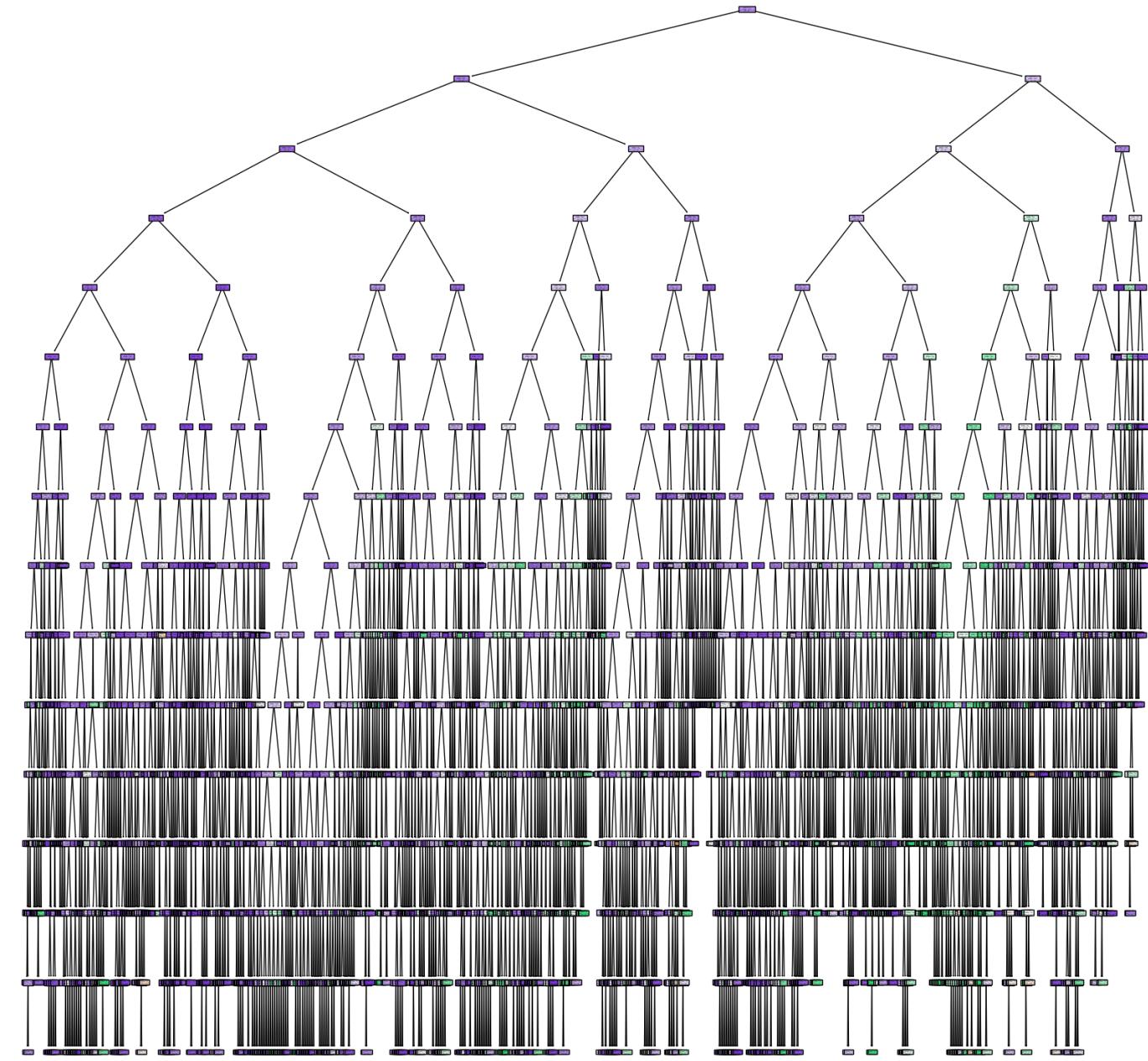
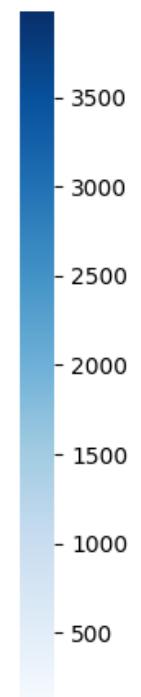
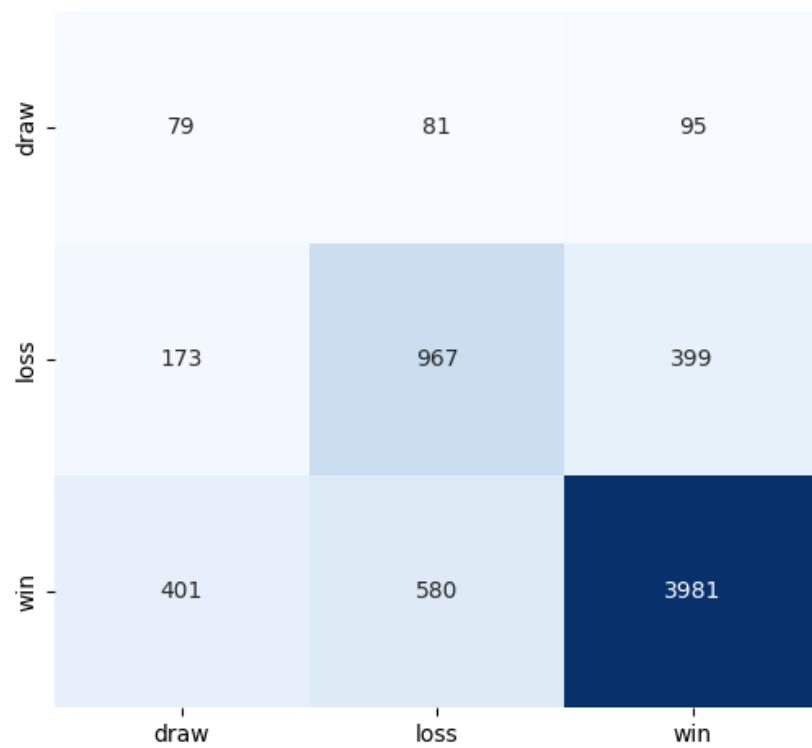
Cost complexity pruning alpha

- Best max feature with entropy criterion = $4e-05 - 6e-05$
- Overfit



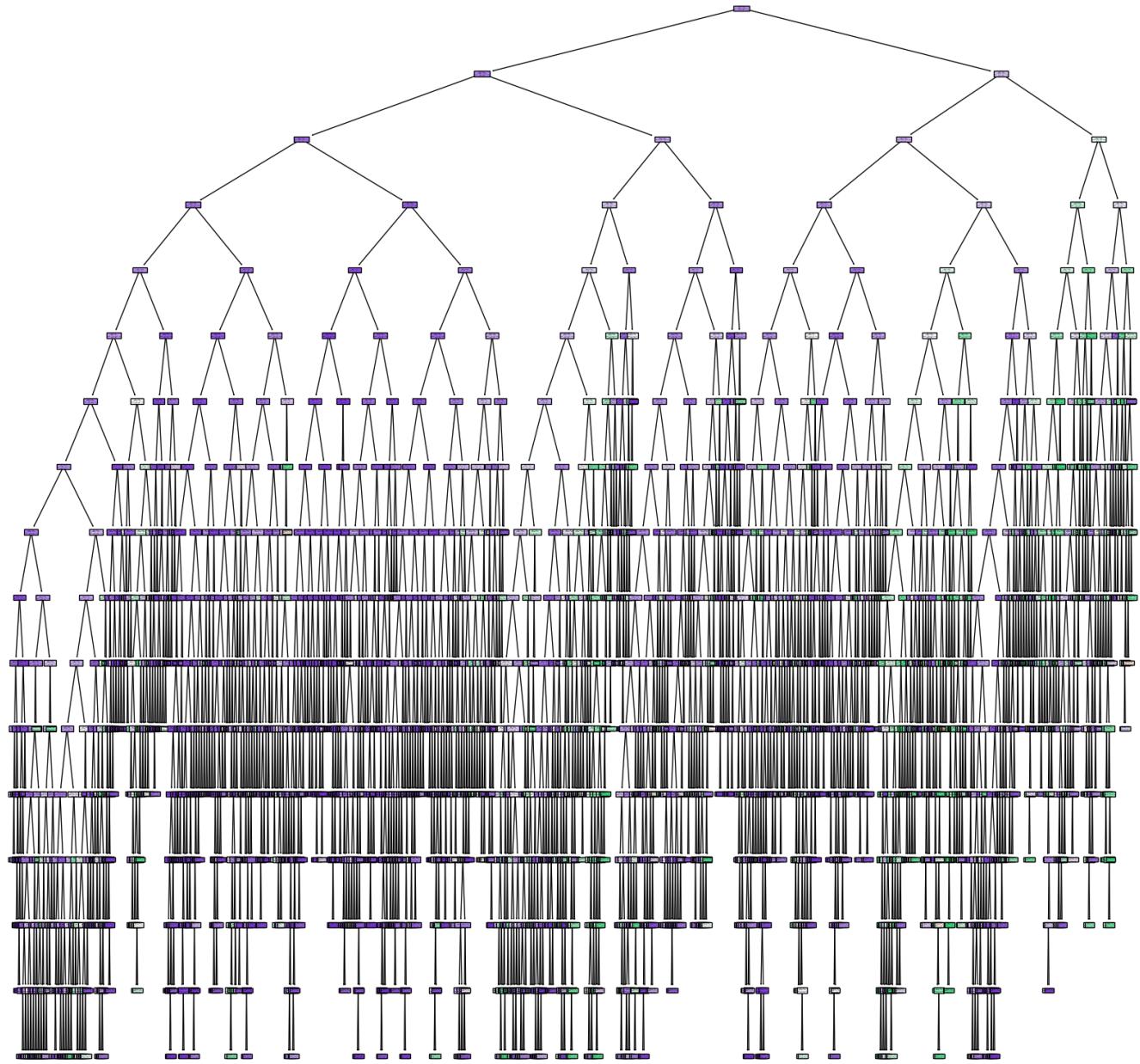
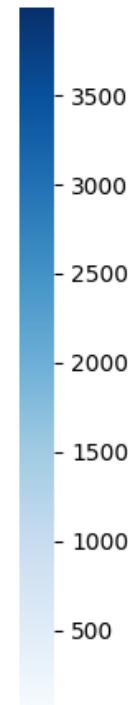
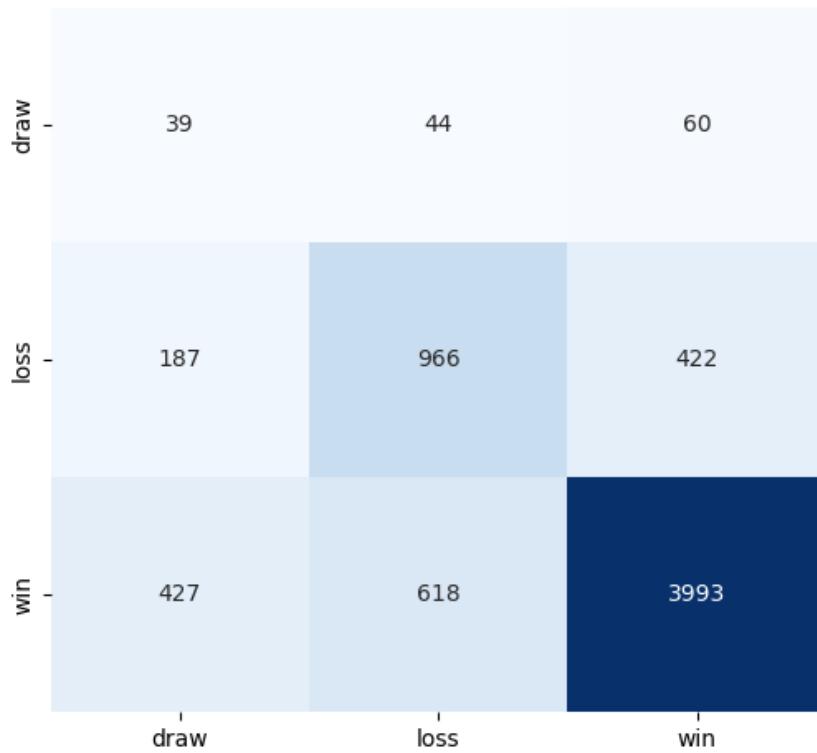
Entropy

- 0.743 on validation set
- 0.739 on test set



Gini index

- 0.743 on validation set
- 0.739 on test set



Log loss

- 0.749 on validation set
- 0.744 on test set

