Decision Tree:

Split Criterion

Y_{new} = predict (Mdl, (new)) → out-of-sample predictors.

MATLAB Ex:

check off var 1.

430x8(training samples) ; ∧ All trees. no validation

ROC curve

good (we want it jumps to 1 quickly).

Tree2_u_u   ↓ optimize   Resubmission :
unvalidated   unoptimize.   resubPredict
(validated
cross validated

with validation, accuracy is lower
                    Tree opt . M
see picture   % Tree optimization w/o cross validation.
            Predictors = Tree2_u_u. Classification Tree.X ;
            Response = Tree2_u_u. Classification Tree. Y ;
            Tree2_u_o = fitctree (Predictors, Response, 'optimize Hyper parameters',
            view (Tree2_u_o, 'mode', 'Graph');   'auto');

run : Treeopt :

Output : a tree (optimized).

Hyperparameter for decision tree.

Pruning a tree in the graph.

eg.

Gini & Deviance criterion.

Split Industry

Use the insample I generated
(430 observations).