HW 7

Adam Karl

March 26, 2021

1 Decision Trees

a. The default tree results in an error of 0.2751. When restricting to 25 nodes, the error was 0.2576. When restricting the tree to 20 nodes, the error was 0.2445. With 15 splits the error was 0.2795. 10 splits: 0.2838. 5 splits: 0.2314. And with 1 split the error was 0.2707. Of the settings I tested, the one that minimized the error was restricting the tree to 5 splits with a total error of 0.2314. Since we were able to achieve a better error by decreasing to size of the tree, it would indicate that some backpruning of the tree would be ideal for minimizing error.

b.

- fitctree(x,y,'MaxNumSplits',25, 'MinParentSize',10,'MinLeafSize',2,'splitcriterion','gdi')
  - error = 0.2576
- fitctree(x,y,'MaxNumSplits',5, 'MinParentSize',10,'MinLeafSize',2,'splitcriterion','gdi')
  - error = 0.2314
- fitctree(x,y,'MaxNumSplits',25, 'MinParentSize',25,'MinLeafSize',25,'splitcriterion','gdi')
  - error = 0.2358
- fitctree(x,y,'MaxNumSplits',5, 'MinParentSize',5,'MinLeafSize',5,'splitcriterion','gdi')
  - error = 0.2314
- fitctree(x,y,'MaxNumSplits',25, 'MinParentSize',10,'MinLeafSize',2,'splitcriterion','twoing')
  - error = 0.2576
- fitctree(x,y,'MaxNumSplits',25, 'MinParentSize',10,'MinLeafSize',2,'splitcriterion','deviance')
  - error = 0.2576
- fitctree(x,y,'splitcriterion','gdi', 'PruneCriterion', 'impurity')
  - error = 0.2838
- fitctree(x,y,'splitcriterion','gdi', 'OptimizeHyperparameters', 'all')
  - best error = 0.2795

2 Conditional Independence

We start with the equation (which is true by the reflexive axiom):

$$P(A|B,C) = P(A|B,C)$$

Expand the right side using the definition of conditional probability:

$$P(A|B,C) = \frac{P(A,B,C)}{P(B,C)}$$

Rewrite both the numerator and denominator using the product rule:

$$P(A|B,C) = \frac{P(A,B|C) \cdot P(C)}{P(B|C) \cdot P(C)}$$

Cancel like terms:

$$P(A|B,C) = \frac{P(A,B|C)}{P(B|C)}$$

Substitute the numerator using the definition of conditionally independent variables:

$$P(A|B,C) = \frac{P(A|C) \cdot P(B|C)}{P(B|C)}$$

Cancel like terms:

$$P(A|B,C) = P(A|C)$$

Thus, $P(A|B,C) = P(A|C)$.

3 Bayesian Belief Networks

a.

$$P(E = F) \cdot P(B = F) \cdot P(A = F|E = F, B = F) \cdot P(J = T|A = F) \cdot P(M = F|A = F)$$

b.

$$P(E = F) \cdot P(B = F) \cdot P(A = F|E = F, B = F) \cdot P(J = T|A = F) \cdot P(M = F|A = F)$$

$$0.998 \cdot 0.999 \cdot 0.999 \cdot 0.05 \cdot 0.99$$

$$\approx 0.0493$$