1: Decision Tree Learning

(a)

$$I(X,Y) = H(X) - H(X|Y) \tag{1}$$

$$H(X) = \sum_{i=1}^{n} P(X=i) log_2 P(X=i) \tag{2}$$

$$H(X|Y) = \sum_{v \in values(Y)} P(Y=v) H(X|Y=v) \tag{3}$$

$$H(X|Y=v) = -\sum_{i=1}^{n} P(X=i|Y=v) log_2 P(X=i|Y=v) \tag{4}$$

Outlook:

$$\begin{aligned} InformationGain = &-\left(\frac{9}{14} log_2 \frac{9}{14}\right) - \left(\frac{5}{14} log_2 \frac{5}{14}\right) \\ &-\frac{5}{14}\left(-\left(\frac{3}{5} log_2 \frac{3}{5}\right) - \left(\frac{2}{5} log_2 \frac{2}{5}\right)\right) \\ &-\frac{4}{14}\left(-\left(\frac{4}{4} log_2 \frac{4}{4}\right) - \left(\frac{0}{4} log_2 \frac{0}{4}\right)\right) \\ &-\frac{5}{14}\left(-\left(\frac{3}{5} log_2 \frac{3}{5}\right) - \left(\frac{2}{5} log_2 \frac{2}{5}\right)\right) \end{aligned}$$

$$= \mathbf{0.247}$$

Humidity:

$$\begin{aligned} InformationGain = &-\left(\frac{9}{14} log_2 \frac{9}{14}\right) - \left(\frac{5}{14} log_2 \frac{5}{14}\right) \\ &-\frac{10}{14}\left(-\left(\frac{6}{10} log_2 \frac{6}{10}\right) - \left(\frac{4}{10} log_2 \frac{4}{10}\right)\right) \\ &-\frac{4}{14}\left(-\left(\frac{3}{4} log_2 \frac{3}{4}\right) - \left(\frac{1}{4} log_2 \frac{1}{4}\right)\right) \end{aligned}$$

$$= \mathbf{0.015}$$

(b)

Outlook:

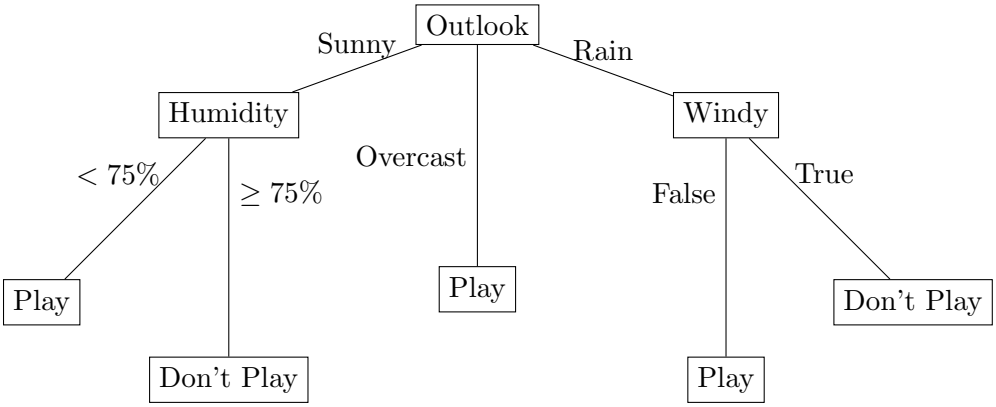
$$GainRatio = -0.247 / \left(\frac{5}{14} log_2 \frac{5}{14} + \frac{4}{14} log_2 \frac{4}{14} + \frac{5}{14} log_2 \frac{5}{14}\right)$$

$$= \mathbf{0.157}$$

Humidity:

$$GainRatio = -0.015/\left(\frac{10}{14}log_2\frac{10}{14} + \frac{4}{14}log_2\frac{4}{14}\right)$$
$$= \mathbf{0.286}$$

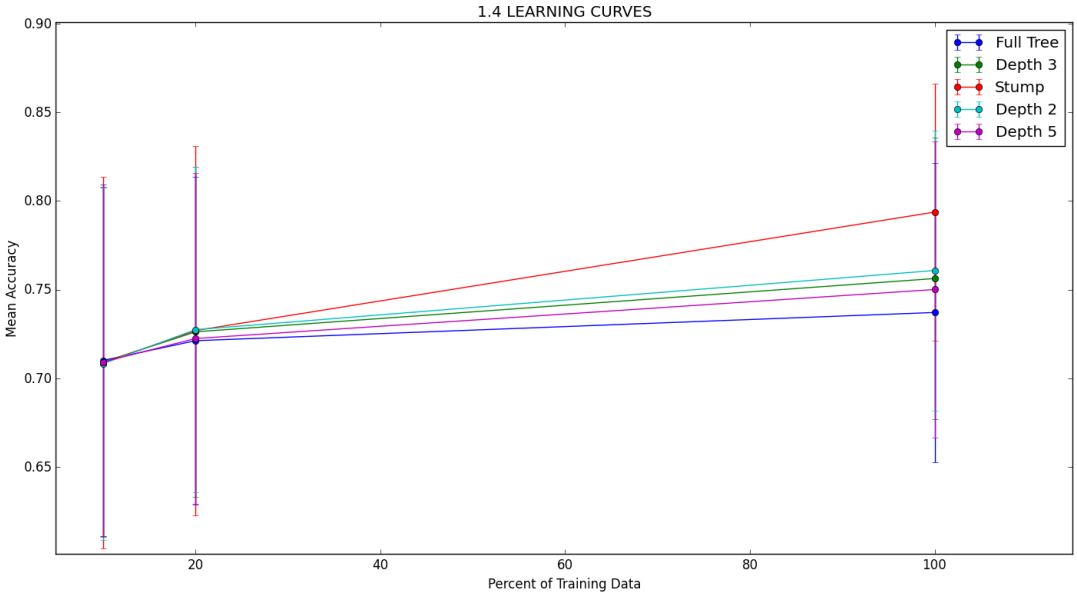
(c)



2: Decision Trees 7 Linear Discriminants

By definition, any split on a single feature of a data set will create an axis parallel split (specifically perpendicular to the feature being split on). In order to create an oblique split, we must split on more than one feature at a time. This could take several different forms, but must involve some relation between the two features. For example if we are looking to create an oblique split on humidity and temperature, we could make a tree that hypothetically says to play ball if $humidity + temperature \geq 150$ then play ball, and if not, do not play ball. We can create relations in may ways. For example we could also create an oblique split with $\frac{humidity}{teperature} \geq 1$. The general premise however is that we have to create a 'new feature' which is some mathematical combination of two other features and compare it to a threshold. What this does is essentially rotate the axis which we are looking at and splitting upon to a new arbitrary axis at a different angle.

Part 2: Graph from 1.4


1.4 LEARNING CURVES
Mean Accuracy
Percent of Training Data
0.90
0.85
0.80
0.75
0.70
0.65
20
40
60
80
100
Full Tree
Depth 3
Stump
Depth 2
Depth 5