DD2431 Machine Learning - Lab 1: Decision Trees

Yamada Jun and Philipson Samuel

General Information

MONK-3 has 5% noise in a training data

MONK-1

$(a1 = a2) \vee (a5 = 1)$

MONK-2

($a_i = 1$, for exactly two i ={1..6}

MONK-3

$((a5 = 1) \wedge (a4 = 1)) \vee ((a5 \neq 4) \wedge (a2 \neq 3))$

Assignment0

MONK-3 The comparison of the different attributes in Monk-1 just requires more splitting of the dataset, and isn't a hard thing to do. The random noice presented in Monk-3 makes it harder though to create a good decision tree.

Assignment1

monk-1: 1.0 monk-2: 0.957117428264771 monk-3: 0.9998061328047111

Assignment2

- Uniform Distribution

The entropy becomes maximal value. Suppose that we have $P(X = x_n) = frac1N$ where X takes the value $X = [x_1, x2, ..., x_N]$(N={1.....}). Then, entropy is

$$Entropy(S) = -\Sigma_{n=1}^{N} P(X = x_n) log_2 P(x = x_n) = -\Sigma_{n=1}^{N} \frac{1}{N} log_2 \frac{1}{N} = N \times \frac{1}{N} log_2 N = log_2 N$$

- Non-Uniform Distribution The entropy of Non-Uniform Distribution becomes smaller than the entropy of uniform distribution, because probability of some events in Non-Uniform Distribution are larger than the others. In other words, some events are more predictable than the other states.
- high and low entropy distribution Suppose that we have a normal distribution, the form of entropy is $\frac{1}{2}ln(2\sigma^2\pi e)$. If $\sigma = 5$, then the entropy become high, because the shape of the distribution become flat and it is difficult to predict events. On the other hands, if $\sigma = 0.2$, then the entropy become low, because events around X=0 happen frequently and it is easy to predict it.

Assignment 3