Report

The first dataset I use to test is balloonsthe Attribute Information is: Color: yellow, purple ;size: large, small ;act: stretch, dip ;age: adult, child ; Classes( inflated: T, F).

I separate the attribute set and values set as input, and the out tree look like this :

(: act

{ STRETCH:age

(: age

{ ADULT:T

{ CHILD:F

{ DIP:F

The model can handle this dataset well and get goodresult;

Also, I tried another dataset, The Car Evaluation Database was derived from a simple hierarchical decision model. The data evaluates cars according to car acceptability, overall price , buying buying price, price of the maintenance, technical characteristics, comfort, number of doors, persons capacity, the size of luggage boot, and safety of the car.  The data includes three intermediate concepts: PRICE, TECH, COMFORT. The data information as follow：Class Values: unacc, acc, good, vgood；Attributes: buying: vhigh, high, med, low; maint: vhigh, high, med, low;doors: 2, 3, 4, 5more; persons: 2, 4, more; lug\_boot: small, med, big.   
safety: low, med, high.

Input: attributes set, values set

Output: a decision tree

The result shows in the code “hw1\_part2”, I got a really complicated tree, the result is not optimal, the tree is huge and it is hard to get the useful information, the ID3 model I create is hard to handle the bigger size real world data.