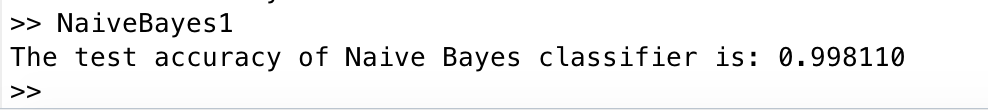
Problem Set 5

I used matlab to build two functions, **NaiveBayes1.m**, **ChowLiuTree2.m**, which training a naïve Bayes classifier and using the Chow-Liu Bayesian structure learning algorithm.

Naïve Bayes:

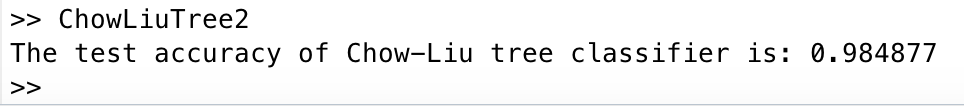
I used **NaiveBayes1.m** to train a naïve Bayes classifier to predict whether or not the mushroom is poisonous using the data. It can be executed in matlab like this:

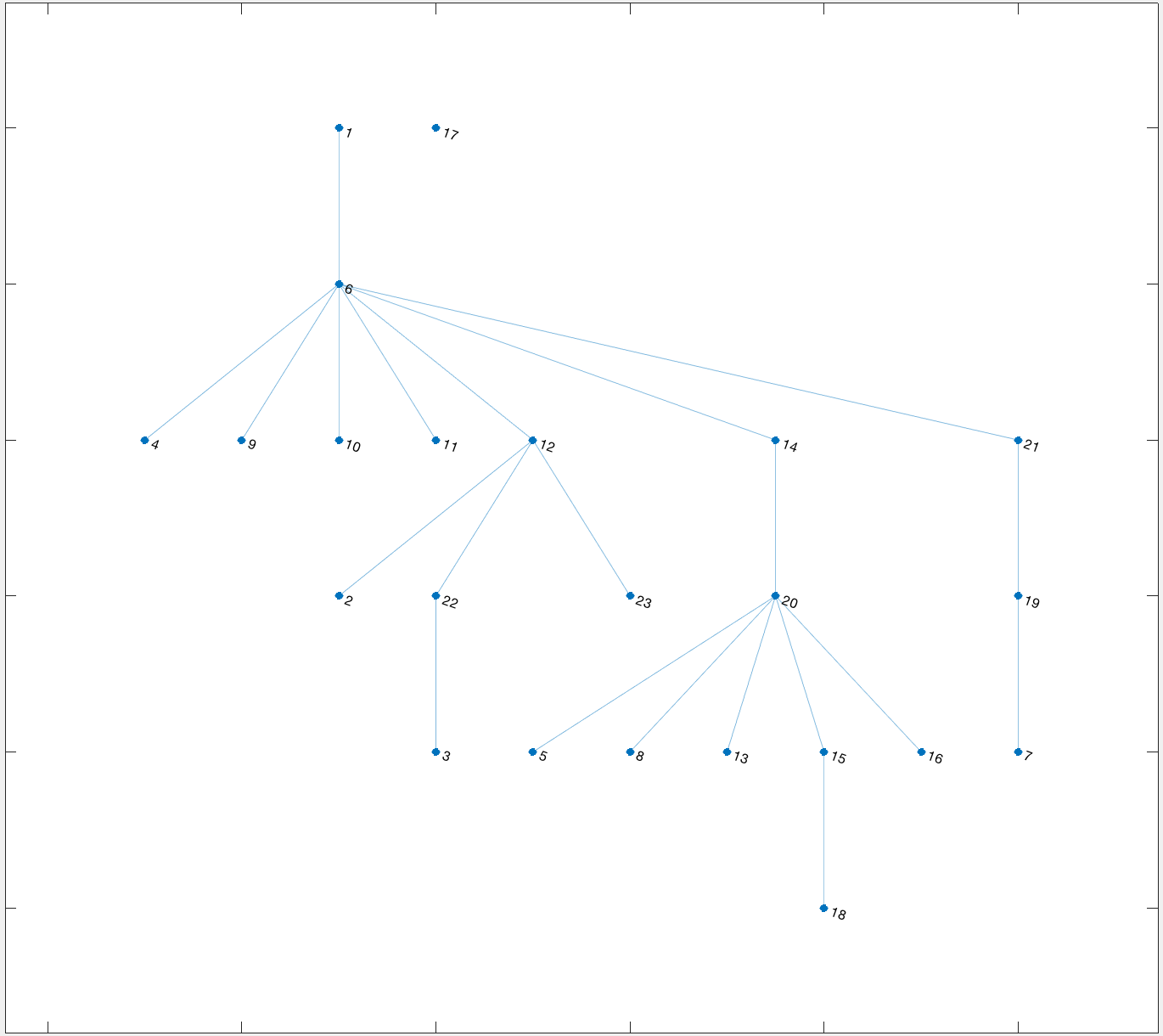


the test accuracy of Naïve Bayes classifier is 0.998110. The performance is reasonable. Because the data set is simple and Naïve Bayes can almost perfectly classify the test data.

Structure Learning

I used **ChowLiuTree2.m** to learn a discriminative Bayesian network for the class label (the first column) by using the Chow-Liu Bayesian structure learning algorithm. It can be executed in matlab like this:



I also use plot the print the tree, which like this:

the test accuracy of Naïve Bayes classifier is 0.984877. The performance is also as good as Naïve Bayes even it keeps only edge 1-6 and discards other edges. It can build a simplified model with faster prediction process while still keeps good performance.