**Jorge Espinosa Campello – Compulsory Assignment Interpretation**

1. The accuracy of predictions for the training and validation data is almost the same due to the tree is so specific that the pruning is not relevant.  
   We can notice that after the pruning the accuracy for the training data decreases because the program only takes care of the validation data accuracy while making the pruning.
2. The rules derived from the tree before and after the pruning are pretty similar in the random example that I tested, so I can determine that the pruning is that not effective in this case because the tree is so specific that if we remove a rule the accuracy will decrease notably and it would be not possible to determine while a mushroom is edible or not with assurance.

|  |  |
| --- | --- |
| Before pruning | After pruning |
| [5, f, 0, p]  [5, a, 0, e]  [5, n, 20, w, 9, b, 0, e]  [5, w, 9, n, 0, e]  [5, w, 9, h, 0, e]  [5, w, 9, p, 0, e]  [5, w, 9, u, 0, e]  [5, w, 9, 19, e, 12, s, 0]  [5, e, 12, k, 0, p]  [5, e, 12, f, 0]  [5, e, 12, y, 0, p]  [5, w, 19, p, 8, n, 0]  [5, p, 8, b, 0, e]  [5, w, 19, l, 0, e]  [5, w, 19, n, 0, e]  [5, w, 19, f, 0, e]  [5, w, 9, g, 0, e]  [5, w, 9, k, 0, e]  [5, w, 9, o, 0, e]  [5, w, 9, y, 0, p]  [5, w, 9, e, 0]  [5, w, 9, r, 0, e]  [5, n, 20, k, 0, e]  [5, n, 20, 0, e]  [5, n, 20, h, 0, e]  [5, n, 20, u, 0, e]  [5, n, 20, b, 0, e]  [5, n, 20, r, 0, p]  [5, n, 20, o, 0, e]  [5, n, 20, y, 0, e]  [5, p, 0]  [5, s, 0, p]  [5, l, 0, e]  [5, y, 0, p]  [5, c, 0, p]  [5, m, 0, p] | [5, f, 0, p]  [5, a, 0, e]  [5, n, 20, w, 9, b, 0, e]  [5, w, 9, n, 0, e]  [5, w, 9, h, 0, e]  [5, w, 9, p, 0, e]  [5, w, 9, u, 0, e]  [5, w, 9, 19, e, 12, s, 0]  [5, e, 12, k, 0, p]  [5, e, 12, f, 0]  [5, e, 12, y, 0, p]  [5, w, 19, p, 8, n, 0]  [5, p, 8, b, 0, e]  [5, w, 19, l, 0, e]  [5, w, 19, n, 0, e]  [5, w, 19, f, 0, e]  [5, w, 9, g, 0, e]  [5, w, 9, k, 0, e]  [5, w, 9, o, 0, e]  [5, w, 9, y, 0, p]  [5, w, 9, e, 0]  [5, w, 9, r, 0, e]  [5, n, 20, k, 0, e]  [5, n, 20, 0, e]  [5, n, 20, h, 0, e]  [5, n, 20, u, 0, e]  [5, n, 20, b, 0, e]  [5, n, 20, r, 0, p]  [5, n, 20, o, 0, e]  [5, n, 20, y, 0, e]  [5, p, 0]  [5, s, 0, p]  [5, l, 0, e]  [5, y, 0, p]  [5, c, 0, p]  [5, m, 0, p] |

1. At first we could think that accuracy is enough to determine whether a mushroom is poisonous or not, but in this example the sensitivity is the best option because it counts the false negative examples compared to the positive and determines with precision of the example given is a negative example or not.
2. Assuming that we know the probability that a poisonous mushroom is predicted as edible, the estimation for p(x) that at least one of them is poisonous given the training and validation data of the example above is p(x)=1, so we can determine that given all this examples at least 1 will be poisonous determined as edible because the tree is not trained as much as we would want to classify all the validation examples.
3. We could use the results of our trained algorithm to classify the mushrooms for example at a farm where they use this kind of food to feed the animals and furthermore we could still train the algorithm and in a future apply it to classify the mushrooms for human people when there were no false predictions in the algorithm.  
     
   We also could use the rules to analyse areas around the world and determine if the population of mushrooms in this area is majority edible or poisonous and warn people to not to go gathering mushrooms there.  
     
   In more general cases, the algorithm could still be useful to set investigation data for teaching purposes, for example to use this data as example for lectures at university or to analyse the performance of a factory training the algorithm to make all the optimizations possible for any purpose.