

# CECS 551 Homework Assignment 1

## Exercises

1. Draw a minimum-sized decision tree for the three-input XOR function which produces a 1 iff an odd number of the inputs evaluate to one.
2. Provide decision trees to represent the following Boolean functions: A and not B, A or (B and C), A xor B, (A and B) or (C and D).
3. Construct a decision-tree learning model based on the following dataset.

Weight	color	texture	classification
medium	orange	smooth	orange
heavy	green	smooth	melon
medium	green	smooth	apple
light	red	bumpy	berry
medium	orange	bumpy	orange
light	red	bumpy	berry
heavy	green	rough	melon
medium	red	smooth	apple
heavy	yellow	smooth	melon
medium	yellow	smooth	orange
medium	red	smooth	apple
medium	green	smooth	apple
medium	orange	rough	orange

4. A test for the AIDs virus has a false positive rate of 1 out of 10000 (i.e. on average one out of every ten-thousand people who do not have AIDs will test positive), and a false negative rate of 3 out of every thousand (i.e. on average, three out of every one-thousand AIDs virus carriers will test negative). Finally, it is estimated that 0.4% of the population carry the AIDs virus. Now suppose a person is selected at random, and tests positive for the virus, what is the probability that this person does in fact have the virus?
5. A test for a disorder has the following properties. If you have the disorder, the probability that the test returns positive is 0.999. If you don't have the disorder the probability that the test returns positive is 0.004. Assume that 3% of the population has the disorder. If a person is randomly chosen from the population and tested, and the test returns positive, what is the probability that the person has the disorder?

6. Using the data from the previous problem, suppose that a particular region of the country wanted to estimate the percentage of people who have the disorder. Based on a five-year study conducted at a regional hospital, it was found that 60% of all people who tested positive for the disorder actually had the disorder. Assuming a statistically significant number of tests, what can we conclude about the percent of people in the region who have the disorder?
7. When observing a blue vehicle driving by on a dimly-lit night, there is a 25% chance that its true color is green. The same can be said when observing a green vehicle under the same conditions: there is a 25% chance that its true color is blue. Suppose that 90% of taxis on the streets have blue color, while the other 10% have green color. On a dimly-lit night you see a taxi randomly drive by and observe its color as blue. What is the probability that its true color is blue? Hint: condition the event that the true color is blue on the event that you've observed it to have blue color.
8. Develop a Naive Bayes learning model for the Exercise-3 dataset, and use it to classify a green piece of fruit of medium size, and having a smooth texture.
9. An artificial neuron with discrete activation function consists of an  $n$ -dimensional weight vector  $\bar{w}$  and real-valued activation threshold  $b$ . For input vector  $\bar{x}$ , the neuron outputs 1 if  $\bar{w} \cdot \bar{x} \geq b$ , and outputs 0 otherwise. Provide an appropriate two-dimensional weight vector and threshold that yields an artificial neuron that functions as a i) AND gate, ii) OR gate, iii) NOT gate (assume a one-dimensional vector).
10. Provide a neural network consisting of three neurons that computes the binary XOR function.