CMPT 310, Spring 2019, Written Assignment - Replacement

Due date: No submission required

Problem 1. In this question we consider decision trees with *continuous* input attributes $A_1, \ldots, A_n$ and a *Boolean* output attribute Y. In such trees, the test at each internal node is an inequality of the form $A_k > c$, where c, the *split point* may be any real number (to be chosen by the learning algorithm. The value at each leaf is *true* or *false.* In a *test-once* tree, each attribute may be tested at most once on any path in the tree; in a *test-many* tree, each attribute may be tested more than once on a path.

Suppose we are given the four examples (also shown in the Figure 1).

A_1	A_2	Y
3	3	*false*
6	13	*true*
15	14	*true*
14	22	*false*

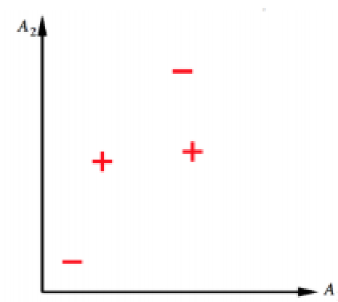

Figure 1: Examples used in Problems 1 and 2

(a) Draw a test-once decision tree that classifies the examples correctly.

(b) Write down the information gain of your root test and the child test (your answer may contain logs; numerical evaluation not required).

Problem 2. Considering the Figure 1 and the Problem 1 description:

(a) Can every non-noisy training set can be correctly classified by a test-once decision tree? Why or why not?

(b) Can every non-noisy training set can be correctly classified by a test-many decision tree?Why or why not?

(c) Now consider the four training sets below (Figure 2). Which are correctly classifiable by test-many decision trees with a reasonably few nodes?

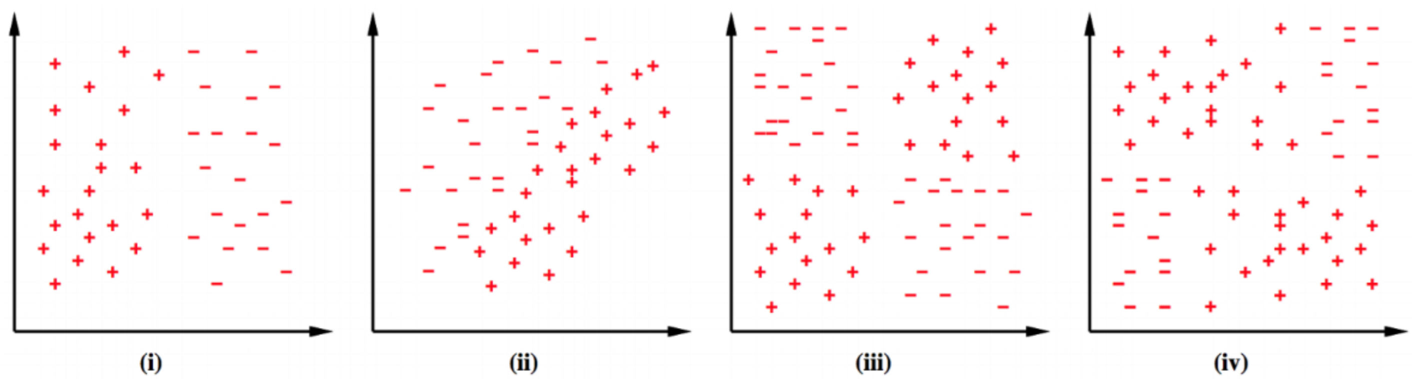

Figure 2: Training sets for Problem 2.d and 2.e.

(d) Which are correctly classifiable by a single-layer perceptron (Figure 2)?

Problem 3. Consider the class of neural networks with inputs are either 0 or 1 and where g is a step function.

(a) Describe how to specify a network that computes the majority function on n inputs. That is, it should output 1 if at least half the inputs are 1.

(b) Draw a decision tree that represents the disjunction of five inputs.

(c) Suppose you're training a neural network in a genuinely nondeterminstic domain. The training set consists of N copies of the same example, a fraction $p > 0.5$ of which are positive and a fraction $1 - p$ of which are negative. Suppose we use the *absolute error* function $E = \sum_{i=1}^{N} |T_i - O|$ where T_i is the correct value for example i and O is the network's output for this example. Suppose that O must also be in the range [0,1]. By writing out an expression for the error in terms of O, find the value of O that minimizes the error.

Problem 4. The Surprise Candy company makes candy in two flavors: 70% are strawberry flavor and 30% are anchovy flavor. Each new piece of candy starts out with a round shape; as it moves down the production line, a machine randomly selects a certain percentage to be trimmed into a square; then, each piece is wrapped in a wrapper whose color is chosen randomly to be red or brown. 80% of strawberry candies are round and 80% have a red wrapper, while 90% of the anchovy candies are square and 90% have a brown wrapper. All candies are sold individually in sealed, identical black boxes.

Now you, the customer, have just bought a Surprise candy at the store but have not yet opened the box. Consider these three Bayes nets shown in Figure 3.

(a) Which network(s) can correctly represent $P(\text{Flavor}, \text{Wrapper}, \text{Shape})$? Why?

(b) Which network is the best representation for this problem? Why?

(c) *True/False*: Network (i) asserts that $P(\text{Wrapper}|\text{Shape}) = P(\text{Wrapper})$.

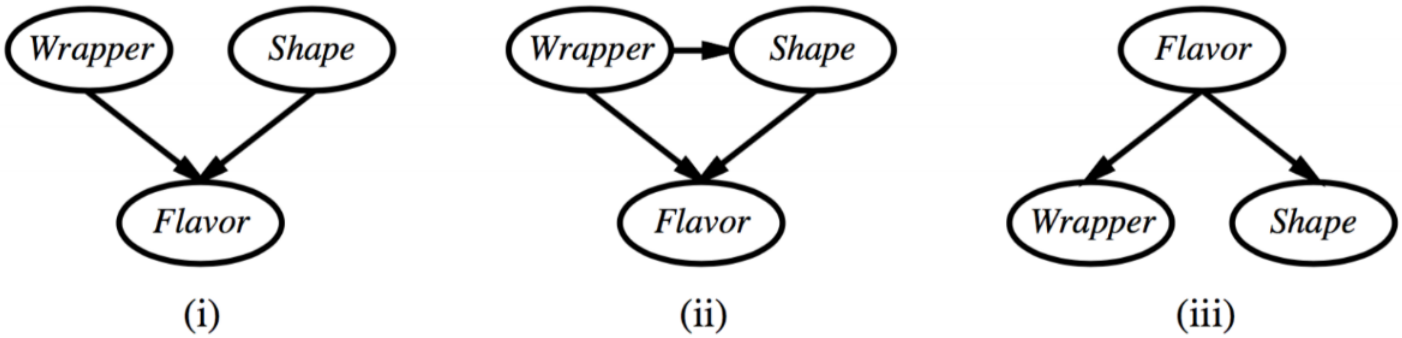

Figure 3: Bayes Nets used in Problems 4 and 5

Problem 5. Considering the Bayes Nets in Figure 3 and the situation described in Problem 4, answer the following items.

(a) What is the probability that your candy has a red wrapper?

(b) In the box is a round candy with a red wrapper. The probability that its flavor is strawberry is ?

(c) An unwrapped strawberry candy is worth s on the open market and an unwrapped anchovy candy is worth a. Write an expression for the value of an unopened candy box.