CSE 473: Introduction to Artificial Intelligence

Homework #2

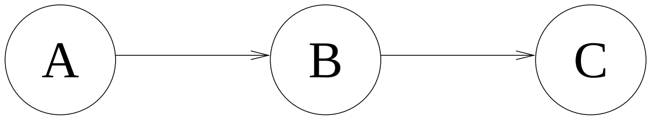
Due May 31st at 9:30 AM in class on paper

1. (20) Consider the following belief network, where all variables are Boolean.

P(A)
0.5

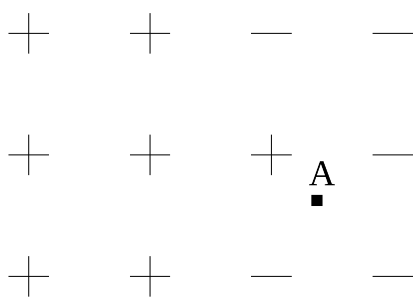
A	P(B)
0	0.2
1	0.8

B	P(C)
0	0.9
1	0.1



What is the probability that A is true given that B is true and C is false?

2. Consider the following training set, where the instance space is the plane, + represents a positive example, and − represents a negative one. Assume Euclidean distance is used throughout.



(a) (20) Draw the Voronoi diagram corresponding to this training set.

(b) (10) What class would the nearest-neighbor algorithm predict for point A? What class would 3-nearest-neighbor predict?

(c) (10) Circle the examples you could remove without changing the frontiers produced by the nearest-neighbor algorithm.

3. (20) Suppose your training set is composed of 200 examples of class Red, 200 examples of class Green, and 400 examples of class Blue. What is the maximum possible information gain of any attribute?

4. (20) Consider the following decision tree, where W, X, Y are the attributes tested at the nodes (all Boolean), A, B, C, D are the classes predicted at the leaves, and the left and right branches of each node correspond respectively to the values True and False. Each leaf is labeled with its accuracy on the validation set, each internal node is labeled with the accuracy on the validation set it would have if it was a leaf, and each branch is labeled with the fraction of examples from the parent node that follow that branch. Which subtrees would you prune?