Homework 5
CS 6375: Machine Learning

Spring 2012

Due date: Wednesday, May 2, 11:59 p.m.

1 Learning Theory [40 points, 10 points each]

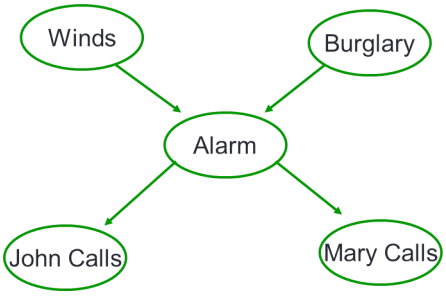
- Mitchell 7.2
- Mitchell 7.3
- Mitchell 7.4
- Mitchell 7.5

2 Bayesian networks [40 points]

- (**30 points**) Consider the following Bayesian network: $A \rightarrow B \rightarrow C$. And the following data table, with entries '?1' and '?2' missing at random:

A	B	C
F	F	F
F	F	?1
F	T	F
T	T	T
T	?2	T
T	F	T

  - Use the data to estimate initial parameters for this network, using maximum likelihood estimation for simplicity.
  - Apply the EM algorithm (by hand) to estimate the values of the missing data, reestimate the parameters, etc. until convergence. Show your calculations.
  - How many iterations does EM take to converge? Will this always be the case? Explain.



- (**10 points**) Consider the Bayesian network given above. It has five variables: {Windy (W), Burglary(B), Alarm(A), John Calls(J), Mary Calls (M) }.
  - Is J independent of M?
  - Is B independent of W given A?
  - Is M independent of W given A?
  - Is A independent of B given W?
  - Is B independent of J given A?

  Hint: Use the d-separation test.

3 Hidden Markov models [20 points]

Consider an HMM with three states, three outputs, and the following transition $P(X_{t+1}|X_t)$ and sensor $P(E_t|X_t)$ models. Assume a uniform distribution for the initial state, X_0.

X_t	X_{t+1}	a	b	c
a		0.5	0.4	0.1
b		0.1	0.5	0.4
c		0.4	0.1	0.5

X_t	E_t	p	q	r
a		0.7	0.1	0.2
b		0.2	0.7	0.1
c		0.1	0.2	0.7

1. Compute the most likely sequence of hidden states for the observed sequence, (p, p, r, r, q, r) by stepping through the Viterbi algorithm by hand. Show your work.
2. Use the forward-backward algorithm to compute the probability distribution over states at position 3. Show your work.