CS 8750

HW #5: Learning with Hidden Variables, HMMs and Filters (10 points)

Spring 2017
(Due 4/4, Tuesday, midnight)

Part I (8 points)

1. (4 points) Given 2 binary variables A and B, estimate their joint distribution from the following dataset with missing data (represented as H). Assume data are missing at random.

A	0	0	1	1	1
B	1	0	1	H	1

   a) Ignore the missing data, 1) give the ML (maximum likelihood) estimation of the joint distribution parameters; 2) Based on the ML estimation, calculate the log likelihood of the data given the model, i.e. $\log P(D|M)$.

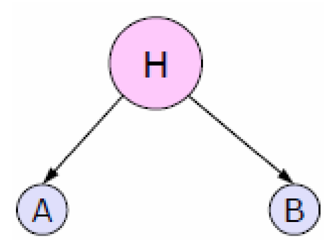
   b) Fill in the missing value H with the value that makes the likelihood of the actual data biggest, then 1) give the ML (maximum likelihood) estimation of the joint distribution parameters; 2) Based on the ML estimation, calculate the log likelihood of the data given the model, i.e. $\log P(D|M)$.

   c) Fill in the missing value H with distribution (0.5, 0.5), then 1) give the ML (maximum likelihood) estimation of the joint distribution parameters; 2) Based on the ML estimation, calculate the log likelihood of the data given the model, i.e. $\log P(D|M)$.

   d) Continue from c). Use the joint distribution estimation in c) to get a new distribution over H. Fill in the missing value H with this distribution, then 1) give the ML (maximum likelihood) estimation of the joint distribution parameters; 2) Based on the ML estimation, calculate the log likelihood of the data given the model, i.e. $\log P(D|M)$.

2. (2 points) EM for BN. Given the following BN with a hidden cause H and 2 observable variables A and B. Given the following training data.

A	0	0	1	1	1	1
B	1	0	1	0	1	1



   Start EM with this initial model:

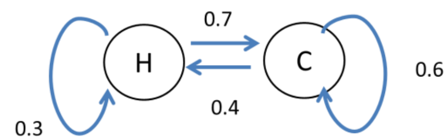
$$P(H) = 0.4, P(A|H) = 0.55, P(A|{\sim}H) = 0.6, P(B|H) = 0.42, P(B|{\sim}H) = 0.52.$$

   Show the result of 2 iterations of EM. In each iteration, a) fill in H data, b) re-estimate CPT parameters, and c) calculate the log likelihood of the training data given the model.

3. (1 point) Given the following Markov chain of the weather condition (H: hot, C: cold) and it is cool on day 0, i.e. $P(C_0) = 1$ and $P(H_0) = 0$.

   a) What is the probability of hot on day 1, $P(H_1)$, and day 2, $P(H_2)$?

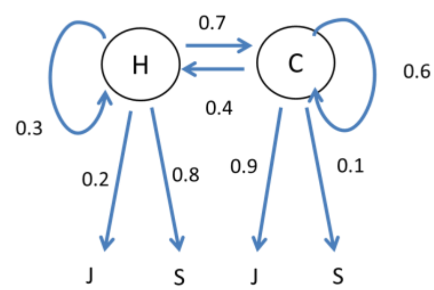
   b) What is $\lim_{t \to \infty} P(C_t)$?



4. (1 point) Consider the following Hidden Markov Model of the weather condition (H: hot, C: cold). You may observe students wear jacket(J) or shorts(S) under different weather condition. Suppose on day 0, $P(C_0) = 1$ and $P(H_0) = 0$, and on day 1, you observe students wear shorts.

   a) What is $P(C_1|S_1)$?

   b) What is $P(C_2|S_1)$?



Part II (2 points)

Write a proposal for a potential final project (a 3-week project) for up to 2 pages and prepare a 2-minute presentation to pitch your project ideas to the class on Tuesday, 4/4.

The proposal should contain motivation, project goals, problem definition, technical approach, possible algorithms, datasets to be used, and planned implementation and experiments.