AI & ML Course
Quiz 2(Apr 3, 2024)

Time: 30 minutes

Instructions

- Answer all questions
- See upload instructions in the form

Question:	1	2	3	4	5	Total
Points:	3	1	1	1	4	10
Score:						

1. Consider the Expectation Maximization(EM) algorithm applied to a Dataset $\mathcal{D}$ of observations from a mixture of Gaussians. Let Θ^* be the parameters of the mixture model.
   (a) (1 point) At each iteration of the EM algorithm the log-likelihood of $\mathcal{D}$
      A. decreases
      B. increases
      C. unable to determine
      D. At each step the likelihood is not guaranteed to increase but over many iterations likelihood is guaranteed to increase.
   (b) (1 point) Parameters estimated through the EM algorithm started from an arbitrary initial point will
      A. always converge to the Maximum Likelihood solution
      B. always converge to a local maximum of the maximum Likelihood
      C. always converge to the Maximum likelhood solution for mixture models
   (c) (1 point) The EM algorithm determines $\Delta\Theta$ using the current estimate of Θ to produce the new estimate
$$\Theta^{(new)} = \Theta + \Delta\Theta.$$

   ___

2. (1 point) Consider learning BM1 and BM2, two Boltzmann Machines, using the maximum likelihood approch. BM1 has 100 hidden units and BM2 has 10 hidden units. They have the same number of visible units.
   A. BM1 is 10 times more expensive then BM2 computationally.
   B. BM1 is 10^3 times more expensive then BM2 computationally.
   C. BM1 is 10^9 times more expensive then BM2 computationally.
   D. BM1 is 2^{90} times more expensive then BM2 computationally.

3. (1 point) To specify a hidden markov model over N states and M observables, the number of parameters required are
   A. $N^2 + M$
   B. $N(1 + N + M)$
   C. $N^2 + NM$
   D. $N + M$

4. (1 point) For the above model, to compute the likelihood of an observation sequence of length T the number of latent random variables we need to marginalize over is
   A. T
   B. M
   C. N
   D. NM

5. Consider an Ising model over N sites(as discussed in class).
   (a) (2 points) Let $E(s)$ be the energy for a specific configuration $s \in \{-1, 1\}^N$. Suppose it is given that $E(s^{(1)}) = 9, E(s^{(2)}) = 1$ where $s^{(1)} = [1, s_2, \ldots, s_N]^\top, s^{(2)} = [-1, s_2, \ldots, s_N]^\top$.The value of $P(S_1 = 1 | S_2 = s_2, \ldots, S_N = s_N)$
      A. is 0.9
      B. is 0.1
      C. cannot be determined

D. can be determined

(b) (2 points) Let $F_i = \sum_{j=1}^{N} w_{ij} S_j + b_i$. Consider the random variable $S_i|F_i = f$. It can be deduced that $P(S_i = 1|F_i = f) = g(f)$ where $g : \mathbb{R} \to [0, 1]$. Mark all the correct choices

A. $g(f) = \frac{1}{1+e^{\ 2f}}$

B. $g(f) = \frac{1}{2}$

C. $g(f) = min(e^{(-f)}, 1)$

D. $g(f)$ is a linear function between 0 and 1.