Example Only - CSI635 – Mid Term Exam

Worth 30% of Grade

Instructions

Do not take this list of exam questions from the class.

Do not talk or try to communicate during the exam.

Calculators, computers etc. are not permitted.

You are allowed 1 and ten minutes for this exam, including reading time.

Be sure you understand the question before answering it. There are no marks for correctly answering the wrong question.

Show all working, but keep answers short and concise. There are no marks for padding answers.

The Easy Question Section is worth 20 points (2 points per question part)
The Moderate Question Section is worth 25 points (5 points per question part)
The Hard Question section is worth 30 points (15 points per question part)
Easy Questions (2 point each question part) (20 minutes)

1. Consider the following Bayesian Belief Network

We can say that ForestFire is what of Thunder, given knowledge of Lightning, CampFire and Storm?

Conditionally Independent

2. Write the expression to evaluate the probability of the event of ForestFire given we know if the remaining events occurred.

Using the Global Semantic \( P(x_1, x_2 \ldots x_m) = \prod_{i=1 \ldots m} P(x_i \mid \text{Parents}(X_i)) \)

\[
P(\text{Storm}).P(\text{BusGroupTour}).P(\text{Lightning} \mid \text{Storm}).P(\text{Campfire} \mid \text{Storm, BusTourGroup}).P(\text{ForestFire} \mid \text{Campfire, Storm, Lightning})
\]

Moderate Questions (5 point each part) (25 minutes)

3. Using Computational Learning Theory (CLT) we can prove the minimum number of instances, \( m \), required for a Probably Approximately Correct (PAC) learner to find some theory whose generalization error is no more than \( \epsilon \). However, for the exact same problem a neural network or a decision tree may find a model with generalization error no more than \( \epsilon \) using far fewer instances than \( m \), apparently defying the CLT result. What property does the PAC learner have that the neural network and decision tree learner does not?

Saying that a PAC learner is consistent and neural networks and decision trees are not is not strictly true since there are PAC learning formulations for agnostic learning.

The correct answer is that neural networks and decision trees have a search bias/preference meaning that we can find a “good” model (Error < \( \epsilon \)) using less instances. PAC learners have no such bias and hence must eta-exhaust the version space.
**Difficult Questions (15 points each) (25 minutes)**

4. Q-learning is a method of learning a model that can be used to determine what actions to perform when in a particular state. What does the typical Q-learning model look like? Design a neural network to replace the Q-learning model for learning in a deterministic environment that provides immediate (not delayed) rewards. Describe/draw this network fully explaining its input, hidden and output layers. How will you train such a network?

*Typical Q-learning model is a nStates x nActions table.*

A plausible neural network could use one in nStates encoding as input as well one in nActions input (from the previous action). The number of hidden units would need to be experimentally derived. The output layer could be one in nActions encoding.

Note that the network does not explicitly contain any measurement of reward. Training the network would involve creating a random agent that for a given state would try all actions. The training instances would then be for a given state (and the previous action) the action that maximizes the reward.