CPSC 633 Exam #2 (4/26/2011, Tue)[1]

Last name: _____, First name: _____

Time: **9:35am-10:50am (1 hour 15 minutes)**, Total Points: **100**

Subject	Score
Genetic Algorithms	/20
Dimensionality Reduction	/20
Local Models	/20
Bayesian Learning	/30
Kernel Machines	/20
Sensorimotor Learning	/10
Total	/120
Adjusted final score	min(,100)

- The total adds up to 120. You may solve all problems, or solve problems that add up to 100 points. In case your score is above 100, your final score will be 100. This will allow you to skip some problems that you are not confident with.
- Total 8 pages (12 questions), including the cover sheet.
- Be as **succinct** as possible. Usually, one sentence is enough to answer one specific question. If your statement includes both correct and incorrect answers, you will get 0 points.
- Read the questions carefully to see what kind of answer is expected (*explain blah* **in terms of ...** *blah*).
- If you feel that the question is not specific enough, please do ask. Be careful not to mention any hint, or suggest alternative answers.
- This is a closed-book, open-note (the one-sheet note you brought) exam.

[1] Instructor: Yoonsuck Choe.

1 Genetic Algorithms

Question 1 (10 pts): Suppose there are two schemas s_1 and s_2. If s_1 has higher average fitness than s_2 and if all other conditions are equal, which schema has a higher expected number of instances in the next generation? Explain your intuition.

Question 2 (10 pts): What are the two selection strategies that overcome the problem of crowding that occurs with fitness proportionate selection (probability of selection proportional to $\frac{\text{fitness}}{\text{total fitness}}$)?

2 Dimensionality Reduction

Question 3 (10 pts): In PCA, encoding and decoding can be done with the eigenvector matrix and the transpose of the same eigenvector matrix. (1) Explain why this is possible, in terms of the property of the eigenvectors, and (2) why this property arises naturally.

Question 4 (10 pts): In Isomap, (1) explain how the distance between two arbitrary data points are calculated and (2) explain what aspect of this distance calculation relates to the concept of manifolds.

3 Local Models

Question 5 (10 pts): What is the main difference between k-means clustering and Adaptive resonance Theory? Explain in terms of the number of clusters.

Question 6 (10 pts): Explain why learning the weight between RBF units and the output unit can be done in one shot without iterative learning.

4 Bayesian Learning

Question 7 (10 pts): In Naive Bayes Classifier, what happens when the conditional independence assumption is violated? Explain in terms of the consistency of the outcome (i.e., is the output the same as the target value).

Question 8 (10 pts): Does h_{MAP} always give the same result as a Bayes optimal classifier? Explain why or why not.

Question 9 (10 pts): (1) Explain how Gibbs sampling works and (2) explain its property (in terms of the error bound).

5 Kernel Machines

Question 10 (10 pts): SVM usually projects the input vector into higher dimensional space. (1) What is the disadvantage in doing this (explain in terms of VC dimension)? (2) What is the advantage of doing this (explain in terms of a certain theorem). (3) How does SVM overcome the above disadvantage in (1)?: explain in terms of VC dimension.

Question 11 (10 pts): How does SVM solve the computational overhead of calculating the inner product between projected feature vectors?

6 Sensorimotor Learning

Question 12 (10 pts): Is grounding possible based on just the neural spikes (green lights), given no other capability? (1) Explain why or why not and (2) show how it can be done in either case.