Monday, April 8, 2019 Review for COSC 4368 Midtem2 Exam

1) Classification and Supervised Learning in General

a) What is overfitting? What are the characteristics of overfitting? What can be done in the context of decision trees to battle overfitting?

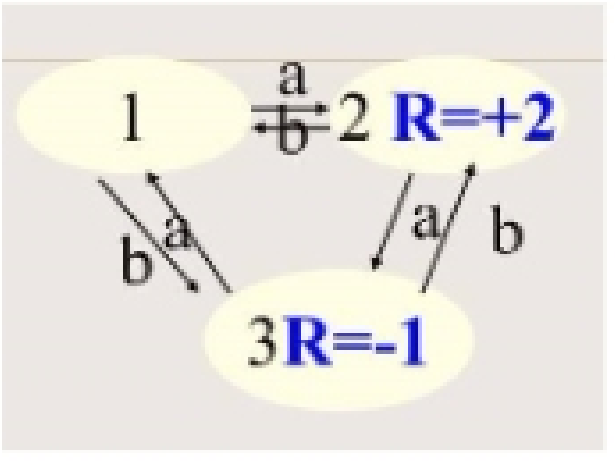
b) What is the key contribution of the backpropagation algorithm? What problems does it solve?

c) What is the purpose of training, test and validation sets in Supervised Learning? What else can be said about their relationship?

2) Reinforcement Learning

a) Assume you have a policy that always selects the action that leads to the state with the highest expected utility. Present arguments that this is usually not a good policy by describing scenarios in which this policy leads to suboptimal behavior of the agent!

b) Assume the following world is given:



Moreover, the current Q-table contain the following entries:

Assume the agent is currently in state 2 and her policy always applies action b in every state. How does the updated Q-Table look like after the agent has applied action b the fourth time assuming when Q-Learning is used? Assume that the learning rate α and the discount rate γ are both 0.5. Do not only report the updated value, but also give the formulas for the four Q-table updates. Assume that the q-table entries are initially set to 0.

c) How does SARSA differ from Q-learning?

3) SVMs [9]

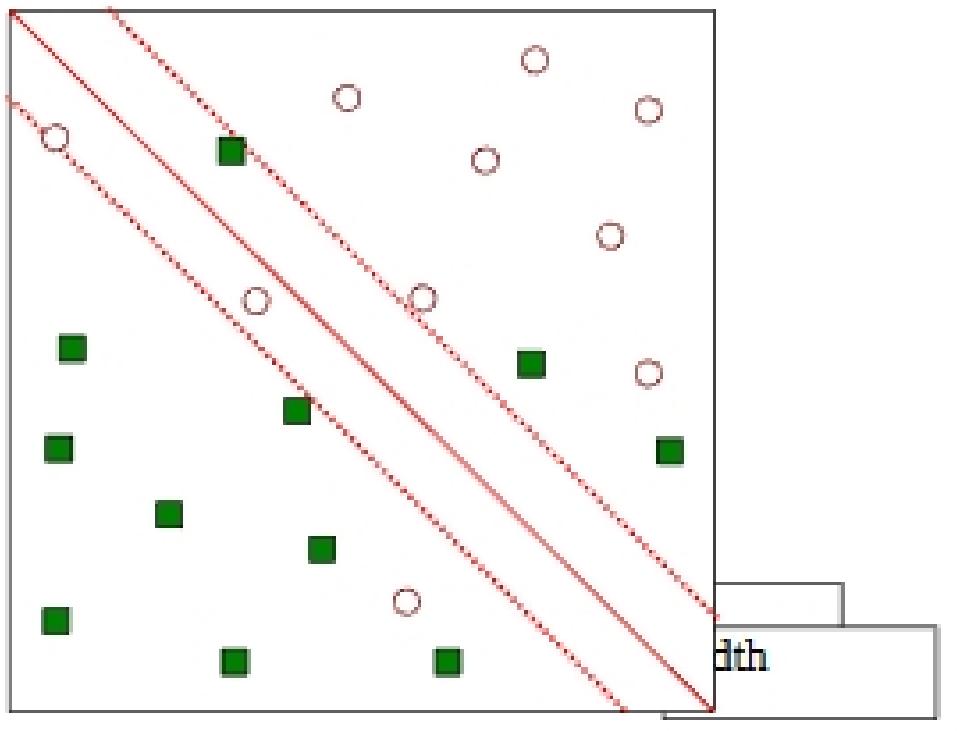
a) What are the characteristics of hyperplanes that support vector machines learn from a training set? [3]

b) The soft margin support vector machine solves the following optimization problem:

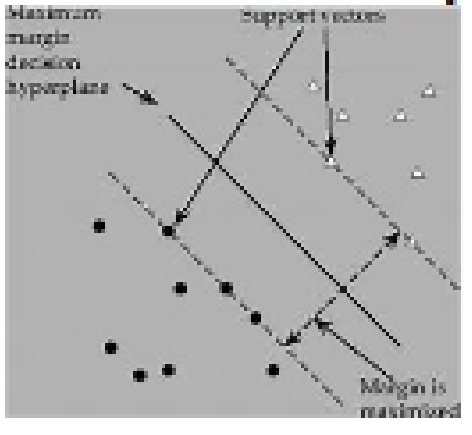
$$\text{argmin}\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_i \xi_i \qquad \text{subject to } c_i(\mathbf{w}\cdot\mathbf{x}_i - b) \geq 1 - \xi_i \quad 1 \leq i \leq n.$$

What does the first term minimize? Depict all non-zero ξ_i in the figure below! Depict all support vectors in the figure below---if example j is a support vector what is its value for ξ_j. What is the advantage of the soft margin approach over the linear SVM approach?

All other points have ξ_i values of 0!



What is a support vector?



Support vectors are the **data points nearest to the hyperplane**, the points of a data set that, if removed, would alter the position of the dividing hyperplane. Because of this, they can be considered the critical elements of a data set.

c) Explain how examples are classified by SVMs!

4) Expressing Natural Language Statement as FOPL Formulas

a) "No vegetarian eats fish"

b) "There is a student who is enrolled in COSC 4368 who got a grade of A in all exams of the course"

5) Neural Networks

a) How are activation functions used in neural network computations? What is neural network learning all about?

b) What is the main purpose of the back propagation algorithm?