Print Name:

_____

PART B: Machine Learning

Question 10
What is another name for an *unbiased learning* algorithm?
[4 marks]

It is a rote learner, such as a database.

Question 11
What is the main difference between *concept learning*, *classification* and *regression*?
[6 marks]

Concept learning is when the target attribute (or class variable or dependent variable) is a binary variable.

Classification is when the target attribute (or class variable or dependent variable) is a discrete variable (2 or more distinct values).

Regression is when the target attribute (or class variable or dependent variable) is a continuous variable.

Question 12
How are machine learning algorithms like search algorithms? Give an example of one hypothesis space.
[6 marks]

All machine learning algorithms (except some Bayesian algorithms) are search algorithms. They search through the space of hypotheses looking for the hypothesis that best fits the data (i.e., has the smallest error). An example of a hypothesis space is searching through the space of all decision trees (usually from smaller to larger) or through the space of sets of neural network weights (slowly changing the weighs to reduce error)…..ther are many other examples.
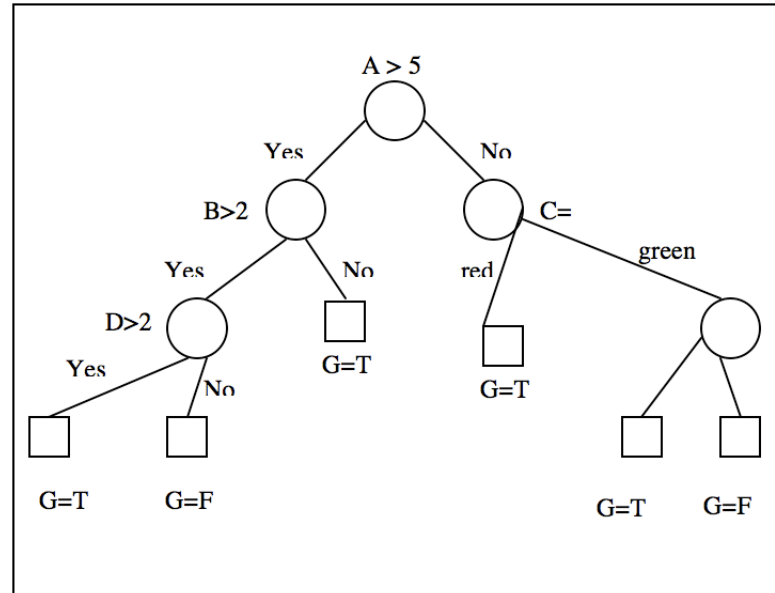
Print Name:

_____

Question 13

How many rules does the decision tree below have? Write out 2 of the rules in propositional logic.

[6 marks]

There are 6 rules.

A>5 ^ B>2 ^ D>2 => G=T

A>5 ^ B>2 ^ D<=2 => G=F

There are 4 other correct rules….

Print Name: _____

Question 14
What is the difference between the *bias* of an algorithm and the *variance* of an algorithm?
[6 marks]

The bias of an algorithm is how far the average error (or accuracy) is from the true error. The true error is based on the whole distribution of data points in the universe (not just the training set).

The variance is a measure of the difference between the different runs of the algorithm from one another.

So if we have two runs with errors of 10 and 20 that would give is an error of 15. Or we could have two runs with errors of 14 and 16. They will both have an average error of 15. But the first will have a much higher bias. People think of this as precision.

The bias is based on the difference between our average error on 15 and the true error (say 5).

Question 15
Given the Hopfield network below what would the stable state be?
[6 marks]
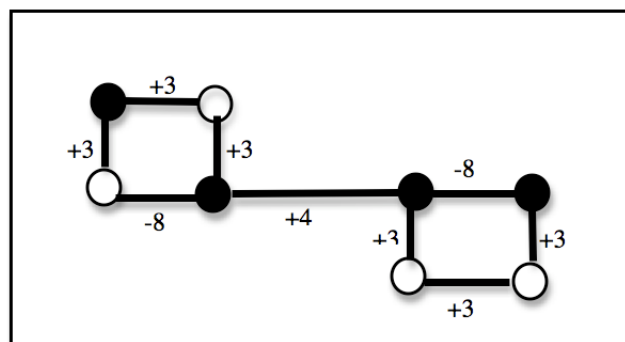
They would all be "on" (black) except for the two nodes in the middle connected by the +4. Also the above is not the starting condition that was on the midterm. I do not have a copy of the midterm that was actually given. But this is the correct answer for the test as it was given.

Print Name:

_____

Question 16
In Genetic Algorithms what is the *hypothesis space* and the *search bias*?
[6 marks]

The hypothesis space is the set of all possible genes. So if we are using binary strings, this is the set of all possible binary strings of the specified size. If we are using tress, this is all the possible trees which could be reached by the GA.

The search bias is complicated. Basically we are searching for genes with higher fitness. It also assumes that parents with higher fitness are more like to have children with higher fitness.

Print Name:
_____

_____