CS7641 ML Practice Quiz

Module SL 1: Decision Trees

Created by Kyle Nakamura with the help of ChatGPT

Question 1

Which of the following statements correctly describe the characteristics of decision trees and neural networks?

A. Decision trees are limited to linear functions, while neural networks can compute both linear and non-linear functions.

B. Neural networks, unlike decision trees, cannot represent Boolean functions such as AND, OR, and XOR.

C. The ID3 algorithm used in decision trees prefers shorter trees with good splits near the top.

D. Decision trees can handle continuous attributes by creating branching factors equal to the number of possible values.

E. Neural networks always require a sigmoid activation function for the neurons.

Question 2

Regarding supervised learning tasks, which of the following statements are true?

A. Classification tasks map inputs to continuous values, while regression tasks map inputs to discrete labels.

B. Regression to the mean is a phenomenon observed in decision trees when dealing with continuous attributes.

C. In regression, variance can be used as a measure instead of information gain when dealing with continuous outputs.

D. Both decision trees and neural networks can be used for classification and regression tasks.

E. Neural networks can only handle discrete inputs and outputs.

Question 3

Which of the following best describe the expressiveness and limitations of decision trees?

A. Decision trees can represent the XOR function using a single node.

B. The n-version of XOR in decision trees has a size exponential in the number of attributes.

C. Decision trees cannot handle continuous attributes without modification.

D. A decision tree for the n-version of OR grows linearly with the number of attributes.

E. Overfitting in decision trees is addressed by increasing the depth of the tree.

Question 4

Concerning the optimization of neural networks, which of the following statements are correct?
A. Gradient descent cannot be used in neural networks as it is not differentiable.
B. Neural networks cannot get stuck in local minima when using gradient descent.
C. Momentum terms in the gradient can help avoid getting stuck in local minima.
D. Neural networks should always be initialized with large random weights to ensure diversity of solutions.
E. Restriction bias in neural networks is influenced by the number of hidden layers and nodes.

Question 5

In the context of the Perceptron Rule and gradient descent in machine learning, which of the following are true?
A. The Perceptron Rule uses unthresholded values for weight updates.
B. The Delta Rule and gradient descent are synonymous, both using threshold outputs.
C. The Perceptron Rule can find a separating line in non-linearly separable datasets.
D. Gradient descent converges only to a global optimum and not local optima.
E. Gradient descent adjusts weights by minimizing the error between activation and target value.

Answer Key

1. C, D
2. C, D
3. B, C, D
4. C, E
5. E