CS7641 ML Practice Quiz

Module SL 4: Instance Based Learning

Created by Kyle Nakamura with the help of ChatGPT

Question 1

Which of the following are true about instance-based learning and k-nearest neighbors (k-NN)?
A. Instance-based learning stores all training data and discards it after learning the function.
B. In k-NN, the concept of distance is critical for determining similarity between instances.
C. k-NN assumes that all features in the dataset have equal importance.
D. The curse of dimensionality does not significantly impact instance-based learning methods like k-NN.
E. Locality in k-NN means that near points are assumed to be similar to each other.

Question 2

What are some characteristics and considerations of k-NN?
A. k-NN can be used for both classification and regression tasks.
B. The choice of distance metric (Euclidean, Manhattan) has no impact on k-NN's performance.
C. Adding more features or dimensions always improves k-NN's accuracy.
D. k-NN is a lazy learner, deferring the learning process until a query is made.
E. In k-NN, a large value of k always results in more accurate predictions.

Question 3

Regarding the curse of dimensionality in instance-based learning, which statements are correct?
A. The curse of dimensionality refers to the increase in computation time with more dimensions.
B. As the number of dimensions increases, the number of data points needed grows exponentially.
C. Adding more dimensions to a dataset always leads to better performance of instance-based learning algorithms.
D. The curse of dimensionality does not affect the selection of distance functions in instance-based learning.
E. Strategies for dealing with the curse of dimensionality include selecting relevant features and reducing dimensionality.

Question 4

What implications does the preference bias in k-nearest neighbors (k-NN) have?
A. k-NN assumes that functions behave smoothly and similarly across different regions of the dataset.
B. The importance of features bias in k-NN indicates that the algorithm gives more weight to less relevant features.
C. The smoothness assumption in k-NN means that distant points are more related than nearby points.
D. k-NN inherently considers some features more important based on their squared values or other transformations.
E. Preference bias in k-NN is not influenced by the choice of distance metric (e.g., Euclidean or Manhattan).

Question 5

Regarding the computational aspects of instance-based learning, which of the following are true?
A. One nearest neighbor requires significant learning time due to its complexity.
B. In nearest neighbor methods, querying has a linear time complexity when the data is sorted.
C. Linear regression requires more space than nearest neighbor methods.
D. Nearest neighbor methods are eager learners, processing and learning from data immediately.
E. The balance between learning time and querying time is a critical consideration in choosing instance-based learning algorithms.

Answer Key

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