Machine Learning

Assignment 1

by **Himanshu Sardana**

TABLE OF CONTENTS

Assignment 1

Question 1

Explain how KNN algorithm works. What are the advantages and disadvantages of KNN?

Solution

K Nearest Neighbors (KNN) is a simple, yet powerful, supervised machine learning algorithm used for classification and regression tasks. It works by finding the 'k' nearest data points in the feature space to a given query point and making predictions based on the majority class (for classification) or average value (for regression) of those neighbors.

Advantages of KNN:
1. Simple and easy to implement.
2. No training phase, making it fast for small datasets.

Disadvantages of KNN:
1. Computationally expensive for large datasets, as it requires calculating distances to all points.
2. Sensitive to irrelevant features and the choice of 'k'.

Question 2

What is the difference between supervised and unsupervised learning? Provide examples of each.

Solution

1. **Supervised Learning** involves training a model on labeled data, where the input features and corresponding output labels are known. Examples include classification tasks like spam detection and regression tasks like predicting house prices.
2. **Unsupervised Learning** involves training a model on unlabeled data, where only input features are available. The model learns patterns or structures in the data without explicit labels. Examples include clustering tasks like customer segmentation and dimensionality reduction tasks like PCA.