Aimee Co
MA189 SP24
Reading Summary 1

Summary of Murphy Chapter 1

Machine learning can be used to automatically detect patterns in data, which can then be used for prediction and decision making. Machine learning is categorized into three main types:

1. Supervised learning: Mapping from inputs to outputs based on a labeled dataset. It includes classification (predicting categorical outcomes) and regression (predicting continuous outcomes).
2. Unsupervised learning: Finding interesting patterns or structures in data without the guidance of a specific output variable. Examples include clustering (grouping similar instances) and dimensionality reduction (simplifying data while retaining its structure).
3. Reinforcement learning: Learning how to act or behave based on rewards or punishments.

There are various applications of machine learning, from document classification and spam filtering to more complex tasks like image classification, face detection, and financial predictions. This chapter also discusses parametric and non-parametric models. Parametric models are quicker but assume more about data distribution, whereas non-parametric models are more flexible but computationally intensive. Some models discussed are: K-Nearest Neighbors (KNN), which is a non-parametric classifier that predicts the class of a point based on the majority class among its K nearest neighbors. Logistic Regression is an extension of linear regression for binary classification problems. It models the probability of a binary outcome using the logistic function to ensure outputs between 0 and 1.
Some potential issues lie in overfitting and model selection, where a model learns noise in the training data as if it were a true signal, leading to poor generalization.