Introduction-Summary Sheet

ML@LSE

2018/2019

Abstract

Objectives: Understand the basics concepts and notions underpinning machine learning theories and techniques.

Requirements: Basic definitions of random variable, expectation and variance. Basic knowledge of Linear Algebra may be useful.

Keywords: Dataset, Number of Observations, Dimensionality, Machine learning techniques, High dimensional statistics, Statistical pattern, Supervised Learning, Unsupervised learning, Learning Function, Inputs and Outputs, Training Data, Test Data, Irreducible Error, Regression, Classification, Loss, Risk, Empirical Risk, MSE, MER, Testing Errors, Overfitting, Generalization, Training vs. Testing Errors, Bias- Variance Trade-off.

A *DATA STRUCTURE*

A.1 What is Data?

a Intuitive example

b Datapoint as a Vector

A.2 Datasets

a Datasets: n vs. p

b Large n, large p and associated techniques

B *MACHINE LEARNING AND PATTERNS*

B.1 Supervised vs. Unsupervised learning

a What are patterns?

b Two kinds of patters and two machine learning fields

B.2 More on supervised learning

a The learning function and the training data

b Regression vs. Classification

C *ASSESSING THE MODEL ACCURACY*

C.1 Measuring the quality of fit to the data

a Risk of a learning function

b Empirical Risk: MSE and MER

C.2 The Bias-Variance trade-off

a Overfitting vs. Generalization

b Cross-Validation