

Exercise Guide for *The Elements of Statistical Learning (2nd Ed.)* by Hastie, Tibshirani, and Friedman

Muthu Chidambaram

Last Updated: July 10, 2019

Contents

1	Introduction	3
2	Overview of Supervised Learning	4
2.1	Exercise 2.1	4

About

"I do not fear computers. I fear the lack of them." - Isaac Asimov

This book is popular, and there are a lot of solution guides available - the first result was this one. What follows are my own solutions to some of the problems.

1 Introduction

Not much to say here; included so that section numbering would be correct.

2 Overview of Supervised Learning

2.1 Exercise 2.1

This question is poorly worded. We assume that the question means there exists some model that, given x , predicts a vector \hat{y} such that \hat{y}_k is the probability that x belongs to class k . We then have that

$$\begin{aligned}\arg \min_k \|t_k - \hat{y}\| &= \arg \min_k \sum_{i=1}^K ((t_k)_i - \hat{y}_i)^2 \\ &= \arg \min_k (1 - \hat{y}_k)^2 + \sum_{i=1, i \neq k}^K \hat{y}_i^2 \\ &= \arg \min_k (1 - \hat{y}_k)^2 + \sum_{i=1, i \neq k}^K \hat{y}_i^2 - \sum_{i=1}^K \hat{y}_i^2 \\ &= \arg \min_k (1 - \hat{y}_k)^2 - \hat{y}_k^2 \\ &= \arg \min_k 1 - 2\hat{y}_k \\ &= \arg \max_k \hat{y}_k\end{aligned}$$