Lecture Notes: 4645: Machine Learning Methods in Empirical Economics

Sebastian Schmidt-Larsen

December 1, 2022

1 Introduction

This is mainly for my own benefit. Any errors are my own. Contact me here on Github if you find any mistakes, or make a push request.

This is based on the lecture slides given for 4645: Machine Learning Methods in Empirical Economic Fall 2023, at Aarhus University BSS. I have tried my best to follow the same notation as given at the lectures. Note that these deviation quite a bit from the curriculum. It should also be noted that I have not included comments from lectures, which have note included derivations. Those have been written in OneNote. Also it should be noted that I don't include equations that are in slides, unless they are needed to provide further information in these notes.

Prerequisites

2 Prediction

Literature Chapters: ISL 2, 5; ESL 2, 7

2.1 Introduction

Goal of machine learning is to automatically discover patterns from the data. the learning part is given af the learning is happening automatically by the computer. We typically use these discovered pattens for prediction.
For prediction we focus on supervised learning, where we have information on the outcome variable. These methods are designed for labelled data (we have observations on the outcome variable).
Important distinction here. We refer to continuous variable prediction as prediction and discrete variable prediction as classification. Exception is causal machine learning where ML is used to discover causal relationships.

2.1.1 Big Data

Machine learning is well suited to handle large datasets. Big data tern not precisely defined but often a relative term. We don't care about the definition, but just note that our methods can handle big data.
High dimensional problems is where we have many parameters to estimate relative to number of observations. High dimensional problems are normally also characterized by an unknown form. Digitalization has also made new types of data available.

2.2 What is Statistical learning

General formula. We observe a quantitative response Y and p different predictors, $X_1, X_2, \ldots, X_p$. We assume that there is some relationship between Y and $X = (X_1, X_2, \ldots, X_p)$, which can be written in the very general form:

$$Y = f(X) + \epsilon \tag{1}$$

Where f is some fixed but unknown function of $X_1, \ldots, X_p$, and ϵ is a random error term which is is independent of X and has mean zero. Here we can say that f represent the systematic information that X provides Y [**james2013introduction**].

3 Prediction

Define outcome (target, response) variable as Y and predictors as $\mathbf{X} \in \mathbb{R}^p$. Assume that Y and $\mathbf{X}$ are random data. Data is a $(p+1)$-dimensional random vector $(Y, \mathbf{X}')'$

4 Prediction Versus Inference

Applied econometrics usually has the goal of statistical inference and in that context Y wil be the dependent variable and $\mathbf{X}$ the regressors. We sometimes want establish a causal relationship between Y and $\mathbf{X}$. In machine learning we are focussed on prediction an we do not care if the relationships are statistically significant or not. Our goal in predictive modelling is whether the model predict new observations well. Consider a simple linear regression model

$$Y = \beta_0 + \beta_1 X + \varepsilon \tag{2}$$

If we want to predict Y given X then our end-goal is Y. If we instead want to consider the causal effect of X on Y our end-goal is β_1. Not that we don't observe β_1 but we observe Y, so we can take advantage of this fact in predictive modelling. Note: that we also normally do prediction in time series econometrics.

9 High-dimensional linear regression

Literature Chapters: ISL 6; ESL 3, 7

10 Network analysis

Literature: Only articles

11 Neural Networks

Literature Chapters: ISL 10; ESL 11; CASI 11

12 Classification methods

Literature Chapters: ISL 4, 9; ESL 4.3; CASI 19

13 Tress and random forests

Literature Chapters: ISL 8; ESL 8.7, 9.2, 10, 15; CASI 17

14 Non-linear models

Literature Chapters: ISL 3.5, 7; ESL 5,1-5.7, 9.1

15 Causal inference

Selected articles.

16 Unsupervised learning

Literature Chapters: ISL 6.3,6.5.3, 12; ESL 3.5-3.6, 14.1-14.3, 14.5.