UniTs - University of Trieste

Faculty of Scientific and Data Intensive Computing

Department of mathematics informatics and geosciences

Probabilistic Machine Learning

Lecturer:
Prof. Luca Bortolussi

Authors:
Andrea Spinelli
Christian Faccio

March 3, 2025

github.com/Spina02 andreaspinelli2002@gmail.com

Abstract

As a student of Scientific and Data Intensive Computing, I've created these notes while attending the **Probabilistic Machine Learning** course.

This course will cover the "*probabilistic side*" of machine learning. In particular, we will focus on the following topics:

- Basics of probability and probabilistic inference
- Pronabilistic formulation of learning (Empirical Risk Minimization and PAC Learning)
- Graphical Models
- Inference with graphical models: belief propagation
- Hidden Markov Models for sequential data
- Bayesian Linear Regression and Classification, Laplace approximation, Model Selection
- Kernel Regression and Kernel functions, Gaussian Processes for regression (hints)
- Monte Carlo sampling
- Expectation Maximization and Variational Inference
- Bayesian Neural Networks
- Generative Modelling: Variational Autoencoders and Diffusion Processes

While these notes were primarily created for my personal study, they may serve as a valuable resource for fellow students and professionals interested in probabilistic machine learning.

Contents

1 Introduction

1.1 Machine Learning

Machine learning is a field of computer science about **learning models**.

Models

Definition: ***Model***

- A ***Model*** is a hypothesis that certain features of a system of interest are well replicated in another, simpler syetem.
- ***Mathematical Model*** is a model where the simpler system consists of a set of mathematical relations between objects (equations, inequalities, etc).
- A ***Stochastic Model*** is a mathematical model where the objects are probability distributions.

All modelling usually starts by defining a family of models indexed by some parameters, which are tweaked to reflect how well the feature of interest are replicated.
Machine learning deals with algorithms for automatic selecrion of a model from observations of the system.

Machine Learning

Definition: ***Machine learning***

Machine learning *explores the study and construction of algorithms that can learn from and make predictions on data.* *Wikipedia*

There are three main types of machine learning:

Generative and Discriminative Learning

- **Generative Learning** im at describing the full probability distribution of inputs x or input/output pairs (x,y).
$$p(x,y) = p(x)p(y|x)$$
- **Discriminative Learning** aims at describing the conditional probability of output given the input, or a statistics/function of such probability
$$p(y|x) \quad or \quad y = f(x)$$

[to fix:]

- **Supervised Learning**: The algorithm learns from labeled data by mapping inputs to outputs.
- **Unsupervised Learning**: The algorithm identifies patterns or structures in unlabeled data.
- **Data Generatiion**: The algorithm generates new data points.

Inference and Estimation

Two central concepts for probabilistic machine learning are:

- **Inference**: Compute marginals and contitionals probability distributions applying the laws of probability.
- **Estimation**: Given data and a family of models, find the best parameters/models that explains the data.

In the Bayesian world: estimation $\approx$ inference.

Probability

Probability is a mathematical theory that deals with **uncertainty**

When a certain problems has to face practical difficulties due to it's complexity, we can use probability to model the ***aleatorical uncertainty***, which is the uncertainty due to the randomness of the system.

More often, we have a limited knowledge of the system, and we can use probability to model the ***epistemic uncertainty***, which is the uncertainty due to the lack of knowledge.

Tip: ***Everything is a probability distribution***

In machine learning **everything is a probability distribution**, even if not explicitly stated.