Statistical Machine Learning

Luca Bortolussi

Department of Mathematics and Geosciences
University of Trieste

Office 328, third floor, H2bis
lbortolussi@units.it

Trieste, DSSC, Spring Semester

Machine Learning

What is Machine Learning all about?

Machine Learning

What is Machine Learning all about?

Learning Models

Machine Learning

What is a Model?

Discuss among you 5 minutes and then answer...

My own answer

- A model is a hypothesis that certain features of a system of interest are well replicated in another, simpler system.

My own answer

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

My own answer

- A model is a hypothesis that certain features of a system of interest are well replicated in another, simpler system.
- A *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.

MY OWN ANSWER

- A model is a hypothesis that certain features of a system of interest are well replicated in another, simpler system.
- A *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 is captured.

My own answer

- A model is a hypothesis that certain features of a system of interest are well replicated in another, simpler system.
- A *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 is captured.
- Machine learning deals with algorithms for automatic selection of a model from observations of the system.

A rough classification

Machine learning explores the study and construction of algorithms that can learn from and make predictions on data. [source: wikipedia]

- Supervised learning: learn a model from input-output data. The goal is to predict a the (most-likely) output value for a new, unobserved, input. We distinguish
  - Regression (continuous output)
  - Classification (binary/ discrete output)
- Unsupervised learning: extract information/ learn a model from input-only data
- Reinforcement Learning: find suitable actions to take in a given situation in order to maximize a reward.

Generative and Discriminative models

- **Generative models** aim at describing the full probability distribution of inputs x or input/output pairs x, y:

$$p(x, y) = p(x)p(y \mid x)$$

- **Discriminative learning** aims at describing the conditional probability of output given the input, or a statistics/ function of such probability:

$$p(y \mid x) \text{ or } y = f(x)$$

MACHINE LEARNING

- Supervised learning: learn $p(y \mid x)$ or the best prediction (discriminant) $y = f(x)$
- Unsupervised learning: learn $p(x)$ or some properties of it (e.g. clusters)
- Data generation: learn a model of $p(x)$ and sample new elements from it (or from $p(x \mid y)$).

Inference and Estimation

Two central concepts for probabilistic machine learning

- Inference: compute marginals and conditional probability distributions applying the laws of probability.
- Estimation: Given data and a family of models, find the best parameters/models for the data.

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

Course plan

- Recall of basic notions of probability and probabilistic inference
- Graphical models
- Inference with graphical models: belief propagation and Monte Carlo sampling
- Hidden Markov Models for sequential data
- Probabilistic estimation and Bayesian statistics primer
- Bayesian Linear Regression and Classification, Laplace approximation, Model Selection;
- Kernel Methods: Gaussian Processes for Regression and Classification
- Variational Inference, Mixtures of Gaussians and Expectation Maximisation (guest lecturer: Guido Sanguinetti)
- Neural Networks and Deep Learning.

Lab

Lab

The Lab will account for a good fraction of the course. In the Lab, we will experiment with Machine Learning in Python...

Bring your own laptop...

Lab will be learn by doing, with a lot of self learning. Working in groups is ok. You can also work on your own data and problems (from Kaggle, from your past courses).

Exercises and Exam

Exercises

During the course you will get exercises for homework, both pen-and-paper and Python based. If you complete them successfully and hand them in time, you can get a bonus at the exam.

Exam

- Final team project, with presentation - possibly on datasets coming from companies, or from Kaggle (typically chosen by the team).
- Teams are chosen by me according to some undisclosed probabilistic model.
- An individual oral colloquium with questions on stuff seen during the course, including theory.

Coordinates

Moodle

There is a moodle page of the course. Register, it is where you will get all the material.

Where can you find me?

- Room 328, 3rd floor - email me first at `lbortolussi@units.it`.

Other stuff

- question time at the end of each lecture
- Requests?