First Lecture : Machine Learning

lunes, 1 de febrero de 2021 15:08

Lecture of Machine Learning

organisation : i) clustering → Gaussian mixture models

ii) Perceptron, Multi-layer Perceptron } neural-networks

iii) Restricted Boltzmann Machine } generative models

iv) ? . community detection

. Reinforcement learning

I] INTRODUCTION

machine : done by an algorithme without human intervention

learning : we use an iterative method to adjust the parameter

It covers a lot of common task :

ex : regression . set of data points $\vec{x} \in \mathbb{R}^n$

set of target $\vec{y} \in \mathbb{R}^m$

- define a family of functions f_θ when θ is the set of parameter to be learned

- such that $\vec{y} = f_\theta(\vec{x}) + \vec{\varepsilon}$, some random noise

ii) ... the best value θ^*

ML is divided into 3 classes

- supervised learning

both the dataset & the target are provided

$\{\vec{x}_i, \vec{y}_i\}_i$ $\quad \vec{x}_i \in \mathbb{R}^n$, n dim of the data

$\vec{y}_i \begin{cases} \in \mathbb{R}^m \\ \in \mathbb{N}^m \end{cases}$, m ——— target

ex: . classification $y = \{1, 2, \dots, C\}$ where C is the number of classes

. regression

. latent encoding : $\vec{y}_i = \vec{x}_i$

→ finding a non-trivial representation of the dataset.

- unsupervised learning

you only have a set of variables $\{\vec{x}_i\}$

$\vec{x}_i \in \mathbb{R}^n$

Why statistical physics is useful here ?

Bias historical example: $\mathbb{P}$: Ising model → binary variable → constraint satisfaction problem (CSP)

at the beginning of the 20th century

↓

≃ Ising model at $T = 0$

important discovery:

- phase transition (when varying the density of constraint
- design of new algo based on variational approach
- ⋮

What about ML ?

- detect ageing in the learning dynamics
- caracterization of capacity (ex Hopfield) & phase transition between [retrieval phase / Spin-Glass —
- in community detection: phase transition between learnable/not-learnable comm.

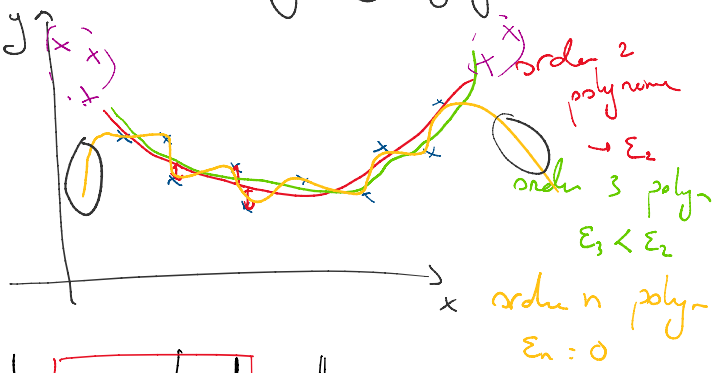
How to find the best value θ^*

. case linear regression : direct optimal θ^* that can be computed

. in more complex situations : iterative procedure

Many issues: i) how to choose the correct family of fct ?

ii) overfitting



$\Rightarrow$ want my fct f_θ to generalized well on new points

$\vec{x}_i \in \mathbb{R}^n$

ex: clustering : group subset of datapt together

- generative model : learn the dist. of dataset → generate new data
  - RBM
  - Generative adversarial network GAN
  - Variational Auto-encoder

☒ Reinforcement Learning

to teach an "agent" how to interact with its environment.

- AlphaGo → game of Go
- AlphaStar → — of starcraft