

MACHINE LEARNING

wikipedia.com -

machine learning is the subfield of artificial intelligence that is concerned with the design and development of algorithms that allow computers to improve their performance over time based on data, such as from sensor data or databases

thus machine learning provides the framework allowing statistical methods [mainly, probability] to be easily adapted to the peculiarities of computational [understanding] processes

only a so close relation with probably-the-most active & research-oriented field of science can explain the high degree of autonomy that machine learning has reached in a so short time

1- main algorithm types

- a. supervised learning — **only type under consideration from this point onwards**
it “learns” from a training dataset [inputs-independent variables/outputs-dependent variables], as trendingBot [or a conventional regression] does, in order to predict the value of the function for any valid input object after having seen a number of training examples
- b. unsupervised learning —
these algorithms seek to determine how the data are organised, that is, do not look for establishing specific relationships between labelled elements, but provide some generic solution for a globally-considered set of data - density functions can be considered as their statistical equivalence
- c. reinforcement learning —
they intend to model a “reinforcement situation”, where an event following a response causes an increase in the probability of that response to occur in the future - the exploration [of uncharted territory] vs. exploitation [of current knowledge] trade-off is the most important concern for this type of learning

2- computational learning theory

wikipedia.com -

in theoretical computer science, computational learning theory is a mathematical field related to the analysis of machine learning algorithms

within supervised algorithms, the main concern is the inference principle [used to generalize from limited data] being applicable

main types

1. probably approximately correct (PAC) learning
2. vapnik-chervonenkis (VC) theory
3. bayesian inference
4. algorithmic learning theory [or algorithmic inductive inference]
5. online machine learning

2.1- PAC learning

the learner receives samples and must select a generalisation function [hypothesis] from a certain class of possible functions

with high probability [“probably” part], the selected function will have low generalisation error [“approximately correct” part]

2.2- VC theory

it is closely related to statistical learning theory and to the study of empirical processes

EMPIRICAL PROCESSES —

generalisation of central limit theorem for empirical measurements

glivenko-cantelli theorem - under certain conditions, empirical measures uniformly converge to the probability measure; the rate of convergence is given by the theory of empirical processes

VC framework is based on two main concepts

1. VC dimension [measure of the statistical classification algorithm capacity]
2. structural risk minimisation (SRM) [inductive principle avoiding overfitting]

2.3- bayesian inference

it is a statistical inference in which evidence or observations are used to update or to newly infer the probability that a hypothesis may be true

- see probability [1.statistics/b5] -

2.4- algorithmic learning theory

the underlying concept is “language identification”, whose main objective is having a learner [program in the running computer] capable to develop a hypothesis [another program to determine grammatical correctness] assessing whether any given sentence is “grammatical” or “ungrammatical” [true or false] - note that it does not need to be a natural language [e.g., english], it can be anything known to the tester

2.5- online machine learning

these models learn one instance [variable] at a time in order to predict their corresponding labels [values]

key concept – constant refining process, based on a continuous-feedback system

[once the prediction is made, the true label for the instance is discovered]

3- pattern recognition

wikipedia.com -

pattern recognition is a sub-topic of machine learning. it is "the act of taking in raw data and taking an action based on the category of the data"

it represents a pre-step to the (usually-supervised) learning algorithm, where data is classified into "patterns" [based either on a priori knowledge or on statistical information from the own patterns]

a wide range of algorithms can be applied for pattern recognition, from very simple models [bayesian classifier] to more complex ones [neural networks]

Note that this "pre-set-up" is only required for certain applications like computer-aided diagnostic, automatic speech recognition, automatic recognition of image, etc.

trendingBot point of view

the main statistical drawback [simplifications => user-defined parameters] is common to all these from-traditional-statistics-derived methods

otherwise this theory is, in general, more adequate to solve complex-but-clearly-delimited problems [e.g., overall applicable finding-the-trend algorithm] than its original source [statistics]

supervised learning algorithms might be considered as proper "trend finders", where

precursor → mixture regression analysis [inputs-output]/time series [user-defined parameters]
(calculating) core → probability