

1. supervised learning types\

- classification (categorical)
- regression (numerical)
- generally
- --input data----->training model and learning patterns----->prediction----->output and reports

algorithms used in supervised learning (self study also)

- Logistic Regression
- K-Nearest neighbour (k-nn)
- Support vector machine (svm)
- Kernel svm
- naive bayes
- Decision tree classification
- Random Forest classification

Random Forest classification\ The random forest is a classification algorithm consisting of many decisions trees. It uses bagging and feature randomness when building each individual tree to try to create an uncorrelated forest of trees whose prediction by committee is more accurate than that of any individual tree.

Decision tree classification\ Decision Tree - Classification. Decision tree builds classification or regression models in the form of a tree structure. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed.\ Decision trees can handle both categorical and numerical data

naive bayes\ In statistics, naive Bayes classifiers are a family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong independence assumptions between the features. They are among the simplest Bayesian network models, but coupled with kernel density estimation, they can achieve higher accuracy levels.

Kernel svm\ SVM algorithms use a set of mathematical functions that are defined as the kernel. The function of kernel is to take data as input and transform it into the required form. Different SVM algorithms use different types of kernel functions. These functions can be different types. For example linear, nonlinear, polynomial, **radial basis function (RBF)**, and sigmoid.

Logistic Regression\ Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is **dichotomous**, which means there would be only two possible classes.\ In simple words, the dependent variable is binary in nature having data coded as either 1 (stands for success/yes) or 0 (stands for failure/no).\ Mathematically, a logistic regression model predicts $P(Y=1)$ as a function of X . It is one of the simplest ML algorithms that can be used for various classification problems such as spam detection, Diabetes prediction, cancer detection etc.

Support vector machine\ In machine learning, support-vector machines (SVMs, also

support-vector networks) are supervised learning models with associated learning algorithms that analyze data for classification and regression analysis. SVMs are one of the most robust prediction methods, being based on statistical learning frameworks

K-Nearest neighbour (k-nn) In statistics, the k-nearest neighbors algorithm (k-NN) is a non-parametric classification method used for classification and regression. In both cases, the input consists of the k closest training examples in a data set

1. **unsupervised Learning/clustering**

- No supervision
 - no teacher
 - self learning
 - find patterns by it self
- --no input data----->training model and learning patterns----->prediction----->output and reports

unsupervised Learning algorithms

- K- Means clustering
- Hierarchical clustering
- Probablistic clustering

Probablistic clustering In probabilistic clustering the assignment of points to clusters is "soft", in the sense that the membership of a data point x in a cluster C_k is given as a probability, denoted by $p_k(x)$. These are subjective probabilities, indicating strength of belief in the event in question.

Hierarchical clustering Hierarchical clustering, also known as hierarchical cluster analysis, is an algorithm that groups similar objects into groups called clusters. The endpoint is a set of clusters, where each cluster is distinct from each other cluster, and the objects within each cluster are broadly similar to each other.

K- Means clustering Kmeans algorithm is an iterative algorithm that tries to partition the dataset into Kpre-defined distinct non-overlapping subgroups (clusters) where each data point belongs to only one group. The less variation we have within clusters, the more homogeneous (similar) the data points are within the same cluster.

1. **SEmi supervised**

- some labelled data and some unlabelled data **mixture of 1 and 2**
- some data is labelled while some is not
- we use both supervised and unsupervised algorithms (usually used for empty values)

1. # **Reinforcement learning** (tukka learning)

- Hit and trial learn
- Learn from mistakes
- reward and punishment rules
- prediction based on reward and punishment
- depends on feedback
- --->input data(but not a learning data)---->training model and learning data by itself

reinforcement Learning algorithms

- Model free reinforcement learning
 1. Policy optimization
 2. Q-learning
- Model based reinforcement learning
 1. learn the model (model based approach)
 2. given the model (system based approach)

Model based reinforcement learning Reinforcement learning systems can make decisions in one of two ways. In the model-based approach, a system uses a predictive model of the world to ask questions of the form "what will happen if I do x?" to choose the best x1.

Model free reinforcement learning In reinforcement learning (RL), a model-free algorithm (as opposed to a model-based one) is an algorithm which does not use the transition probability distribution (and the reward function) associated with the Markov decision process (MDP), which, in RL, represents the problem to be solved. **Policy optimization** Policy optimization methods are centered around the policy, the function that maps the agent's state to its next action. These methods view reinforcement learning as a numerical optimization problem where we optimize the expected reward with respect to the policy's parameters. **Q-learning** Q-learning is a model-free reinforcement learning algorithm to learn the value of an action in a particular state. It does not require a model of the environment (hence "model-free"), and it can handle problems with stochastic transitions and rewards without requiring adaptations.

basic libraries

- pip install scikit-learn