

ML_Theory

January 19, 2022

1 Machine Learning

1.0.1 Data Driven Decesion Making

1.1 Types of Machine learning

1. Supervised
2. Un-Supervised (Clustring)
3. Semi-Supervised
4. Reinforcement

1.2 1- Supervised

- Works under supervision
- Teacher teaches
- Prediction
- Outcome

Input data ==> Training models and learning patterns of data ==> Prediction ==> Output and Reports

1.2.1 Types

- Classification
 - For catagories
- Regression
 - For numerical data

1.2.2 Supervised learning algorithms

- Logistic Regression
 - Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable.
- K- Nearest Neighbours (K-NN)
 - The k-nearest neighbors (KNN) algorithm is a supervised machine learning algorithm that can be used to solve both classification and regression problems. It's easy to implement and understand, but has a major drawback of becoming significantly slows as the size of that data in use grows.
- Support Vector Machines (SVM)

- Support Vector Machine” (SVM) is a supervised machine learning algorithm that can be used for both classification or regression challenges.
- 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.
- Naive Bayes
 - Naïve Bayes Classifier is effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.
- Decision Tree Clasification
 - Decision Trees are a type of Supervised Machine Learning (that is you explain what the input is and what the corresponding output is in the training data) where the data is continuously split according to a certain parameter. The tree can be explained by two entities, namely decision nodes and leaves.
- Random Forest Classification
 - Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.

1.3 2- Un-Supervised (Clustering)

- No supervision
- No Teacher
- Self learning
- No labelling of data
- Find patterns by itself

Training models and learning patterns of data by making different clusters ==> **Prediction**
 ==> **Output and Reports**

1.3.1 Supervised learning algorithms

- 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.
- 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.
- Probabilistic 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.

1.4 3- Semi-Supervised

- Mixture of 1 and 2
- Some data is labelled but most is not

Input data ==> Training models and learning patterns of data ==> Prediction ==> Output and Reports

1.5 4- Reinforcement

- Hit and trial method
- Learn from mistakes
- Reward and punishment rule
- Prediction based on reward and punishment
- Depends on feedback

Input data ==> Training models and learning patterns of data [WELL-DONE] ==> Prediction ==> Output and Reports

1.5.1 Reinforcement learning algorithms

1. 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
 - Q-Learning

2. Model-Based Reinforcement learning

- Model-based Reinforcement Learning refers to learning optimal behavior indirectly by learning a model of the environment by taking actions and observing the outcomes that include the next state and the immediate reward.
 - Learn the model
 - Given the model