Machine Learning

1. Supervised Learning
    a. Regression
        i. If the data is labeled and the labeled variable should be numeric, then we use regression algorithm.
        ii. Used in prediction analysis
        iii. Examples--- find house price, find height of the humanbeing
    b. Classification
        i. If the data is labeled, and the data is categorical data then we use classification algorithm.
        ii. Whether student will be placed or not, loan should be given or not, Whether product is expensive, moderate or cheap
2. Unsupervised Learning
   If the data is not labelled, then it is called as unsupervised learning.
    a. Clustering
       The label will not be there, based on common features the data will be clustered,

    b. Dimension reduction
       If the number of features are many, and you want to remove unwanted features, or combine many features and convert into sing feature
       PCA, LDA

Number bath rooms	No of rooms	location	address

    c. Association Rules
       APRIORI algorithms
        i. Finding out patterns from data
    d. Anamoly detection
        i. If any unusual pattern found then it is called as anamoly delection
3. Semisupervised Learning
    a. If the labels are not available for all the data, but given for some data, and based on that machine found similarity and labeled the remaining data.

4. Reinforcement Learning
    a. Machines are learning by themselves based on some perks, like -ve marking or positive marking.

Steps for solving machine learning

1. Data gathering and frame the problem
2. Data preprocessing

    a. Removing missing values
        i. SimpleImputer
    b. Scaling the values
        i. Standardization-StatndardScaler
        ii. Normalization→ MinMaxScaler
    c. Finding outliers
    d. Removing noisy data
    e. Encoding data--→
        i. Labelencoder--- used with label data, for o/p variables
        ii. Onhotencoder----use with nominal categorical data
        iii. Ordinalencoder---used with ordinal categorical data
3. EDA (Exploratory data analysis)
4. Feature Engineering
5. Model generation and evaluation
6. Model deployment
7. Testing and optimization

Data Preprocessing

1. SimpleImputer—it imput/replace the missing values
    a. Mean
    b. Constant value