Q/A.

Over View o
Machine lea

1. Linear Regression

It is Used to Predict real Values, (or) Predicting Output Variable is "Continous Data" a relationship is established between independent and dependent Variable by fitting a best fit line. (or) Assuming relation between x & y is "Linear". The best fit Line is known as "regression line". represented as

$$y = mx + b$$

- m → slope
- b → intercept
- y → dependent variable
- x → independent Variable

2. Logistic Regression :

it is a classification algorithm used to estimate "Discrete Data Values". based on given set of independent Variables. its Output Value lies between [0 and 1] or [Yes / No]. Uses The sigmoid function. Probability

$$\frac{e^y}{1+e^y}$$

3. Decision tree

It is Supervised learning algorithm Used for classification problems. which Works For both Categorical and Continous dependent Variables.

In this algorithm "population is Split" into two (or) More homogenous sets. according to our Given Conditions.

it Uses different Techinques Like

* Gini impurity
* Information Gain
* Chi-Square Entropy.

4. Random Forest

A Collection of Decision trees is called as "random Forest." To clasify a new object based on its attributes, Each Tree is classified and Take majority of votes for That class. and classifies them. (Bagging Technique)

applicable for both Classification & Regression

5. K-Nearest Neighbors

K-Nearest Neighbors is a Simple algorithm that stores all available cases and classifies New Cases by a majority Vote of its "K-Neighbors".

it is Measured by distance Function

* Manhattan Distance $|x_2-x_1| + |y_2-y_1|$
* Euclidean Distance $\sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$

6. Naive Bayes Theorm (algorithm)

Conditional Probability

it is Supervised learning algorithm. Which is based on "Bayes theorm". and used for Solving classification problems.

Bayes theorm is used to determine Probability of a hypothesis with Prior knowledge. and Likelihood.

$$P(A/B) = \frac{P(B/A) * P(A)}{P(B)}$$

* Naive :- it assumes that Each input variable is independent.

7. SVM (Support Vector Machine)

it is Supervised Machine learning algorithm Used for both classification & Regression. Main objective of SVM to find a hyplane in a "N-dimensional space that distinctly classifies data points

(or)

SVM to find best line in two dimensions (or) The best hyperplane in more than two dimension in order to help us seperate our space into classes.

* Linear Seperable
* Non-Linear Seperable (or) Kernal SVM

8. Boosting Techniques

Ensemble learning (or) boosting Techniques it is combination of "Weak Learners". that implies Combination of estimators (models) with an applied coefficient could act as effective Ensemble Estimator.

Boosting Formula :

$$F_T(x) = \sum_{t=1}^{T} f(x) = \alpha_t \, h(x)$$

Sumation of t=1 to "T"

Each model

α_t → learning rate

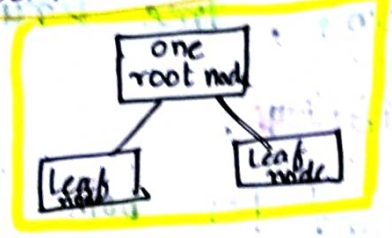
9. AdaBoost Algorithm

AdaBoost (adaptive Boosting) Ensemble method.

it assigns weights to observations which are incorrectly predicted. Next model works to predict these values correctly which are wrongly predicted before.

* Each "Weak Learner" is called as "stump"

- Depth = 1



consists of
- 1 root node
- 2 leaf nodes

10. Gradient Boost

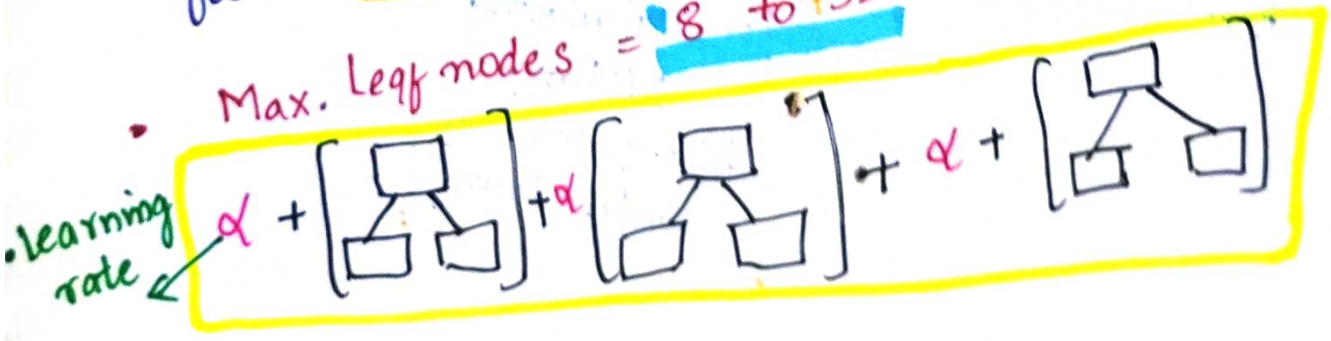
Gradient boost is used to predict both Continuous and categorical variables. it is same as boosting techinques. Here we calculate The Errors and minimize "loss function" (single training Error) in a sequential order.

↓

(Sum of error) min

and Trees are adjusted after multiple times.

- Max. Leaf nodes = 8 to 32



11. XGBoost

XGBoost also called as "Regularized Boosting" Like L_1 & L_2 regularization. it is "10" times faster Than Gradient boost. it implement "parallel processing" it reduces The overfitting, Handle Missing Data, Inbuilt "Cross validation", Tree pruning using depth approach; High Flexibility.

it Can Used For both "classification" and Regression problems".

12. K-Means algorithm

it is "Un Supervised" learing algorithm. (or) Unlabel data. K-means algorithm "identifies 'k' no. Of Centroids, and then allocates every data point to nearest Cluster, while keeping Centroids small as possible, it uses "Distance formula" same as K-nearest Neighbours.

Based on Assumption.

↓

For calculating Distance blw points & cluster them.

13. **Hierarchial clustering**

it combines Data points which are very close by distance. That Groups Similar objects into Groups called clusters. This clustering involves creating clusters that are predetermined Ordering from Top to bottom.

These are two of hierarchial clustering

* Agglomerative (From Top to bottom)
* Divisive. (From bottom to Top)

14. **DB-Scan clustering**

DB-scan (Density Based Spatial Clustering of application with noice")

DB-Scan clustering locates regions of "high densi and Seperate Outliers.

(or)

DB Scan is to find the neighborhoods of data Points Exceeds Certain Density threshold.

* two parameters.
* R (radius of neighborhood)
* m (min. no. of neighbours)

15. **Gradient descent algorithm**

Gradient Descent is an "optimization" algorithm for Finding a Local minima (minimum Point of a differentiable function.) (or) it Gradually reducing Weights and making Errors as Minimum Value.

(or) Ex:- By applying diffirent "x" values and identifying The perfect "x" value, till we get error as "0" Zero is called gradient "descent" minize Cost Functi

Update rule formula

$$y = m \times b$$

16. **Stochastic Gradient Descent**

it has both * "Local Minima" and "Global minima".

Entire data We divide into parts (or) batches Each part we are identifying The "local minima" and from that Over all identify The "Global minima" is called "stochastic Gradient Descent"

Ex:- State ranker (local Minima)
All india ranker (Global Minima)