Assignment : - 01

Machine Learning

Name : - Satyam Kumar

Section :- D

Roll no :- 62

Topic :- Decision tree and random forest

\*Decision tree:-

\* A decision tree, which has a hierarchical structure made up of root, branches, internal, and leaf nodes, is a non-parametric supervised learning approach.

\* it is used for both classification and regression but it is popularly used for classification.

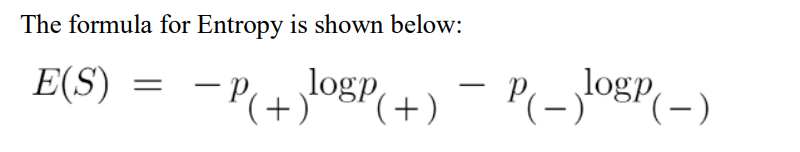
\* decision tree is of three types :-

1. Id3 :- This algorithm measures how mixed up the data is at a node .

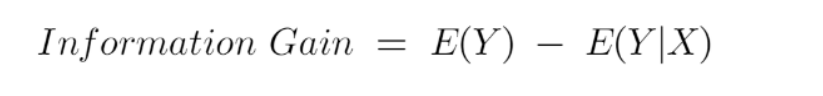
2. C4.5 :- this is a improved version of the id3 and can handle missing data and continuous attributes.

3.CART:- This algorithm uses a different measure called Gini impurity to decide how to split the data. It can be used for both classification and regression

* Pruning: The process of removing or cutting down specific nodes in a decision tree.it is used to prevent overfitting.
* Entropy:- Randomness in data or uncerntainity in dataset.

Information Gain

it is a deciding factor for which attribute should be selected as a decision node or root node.



Random forest:-

Random Forest is a versatile and widely-used ensemble learning algorithm, primarily employed for classification and regression tasks. It builds upon the concept of decision trees and enhances their predictive accuracy and robustness by creating a "forest" of decision trees.

* It is used as a example in Bagging.

thanks