

DECISION TREE

WHAT IS DECISION TREE?

A DECISION TREE IS A NON-PARAMETRIC SUPERVISED LEARNING ALGORITHM. IT HAS A HIERARCHICAL, TREE STRUCTURE, WHICH CONSISTS OF A ROOT NODE, BRANCHES, INTERNAL NODES AND LEAF NODES.

TERMINOLOGIES

Root Nodes — It is the node present at the beginning of a decision tree from this node the population starts dividing according to various features.

Decision Nodes — the nodes we get after splitting the root nodes are called Decision Node

Leaf Nodes — the nodes where further splitting is not possible are called leaf nodes or terminal nodes

Branch/Sub-tree — just like a small portion of a graph is called sub-graph similarly a sub-section of this decision tree is called sub-tree.

Pruning — is nothing but cutting down some nodes to stop overfitting.

TYPES OF DECISION TREES

1. **CART** (Classification and Regression Trees) → uses ***Gini Index(Classification)*** as metric.
2. **ID3** (Iterative Dichotomiser 3) → uses ***Entropy function*** and **Information gain** as metrics.

WHY?

There are various algorithms in Machine learning. Below are the two reasons for using the Decision tree:

- Decision Trees usually mimic human thinking ability while making a decision, so it is easy to understand.
- The logic behind the decision tree can be easily understood because it shows a tree-like structure.

REAL WORLD EXAMPLES

- Medical Diagnosis – Predict disease based on symptoms (fever, cough, BP, etc.)
- Loan Approval – Decide loan eligibility using income, credit score, debt history
- Spam Detection – Classify emails as spam or not spam
- Customer Segmentation – Group customers based on age, purchase history
- Fraud Detection – Identify suspicious credit card transactions
- Heart Disease Prediction – Predict risk using health parameters

THANK YOU