



# DAY 21 — Decision Trees

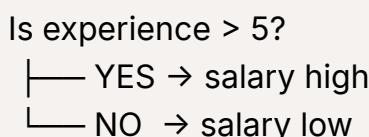
**Goal : Understand tree-based decision making**

## 1 What is a Decision Tree ?

A Decision Tree :

- Splits data using **if-else rules**
- Each split asks a **question**
- Leaves give the final decision

Example:



Trees mimic human decision logic.

## 2 Components of a Decision Tree

| Team          | Meaning              |
|---------------|----------------------|
| Root node     | First split          |
| Internal node | Decision rule        |
| Leaf node     | Final prediction     |
| Depth         | How deep the tree is |

## 3 How does a Tree Decide Where to Split?

Decision trees use **impurity measures**.

**For Classification:**

- Gini Impurity (default)
- Entropy (information gain)

**Gini Intuition:**

"How mixed are the classes?"

- Gini = 0 → pure node
- High Gini → mixed classes

Tree chooses split that **reduces impurity the most.**

## 4 Information Gain (Entropy)

Entropy measures **uncertainty**.

High entropy → uncertain

Low entropy → confident

Best split = **max information gain**

## 5 Overfitting in Decision Trees

Decision Trees:

- Can perfectly fit training data
- Learn noise
- Fail on Test data **✗**

Unrestricted trees almost always overfit.

## 6 How to Control Overfitting

Key hyperparameters:

| Parameter         | Effect                   |
|-------------------|--------------------------|
| max_depth         | Limits tree depth        |
| min_samples_split | Minimum samples to split |
| min_samples_leaf  | Minimum samples in leaf  |

These act as **regularization** for trees.

## 7 Decision Trees Need NO Scaling

Unlike KNN or Logistic Regression:

- Trees do NOT use distance

- Feature scale does not matter

This is a big advantage.

---

## 8 Decision Trees for Regression

Trees can:

- Predict categories (classification)
- Predict numbers (regression)

Regression trees:

- Predict mean value in leaf
- 

## 9 Interpretability

Decision Trees are:

- Visual
- Explainable
- Easy to debug

This is why they're used in:

- Finance
  - Healthcare
  - Rule-based systems
-