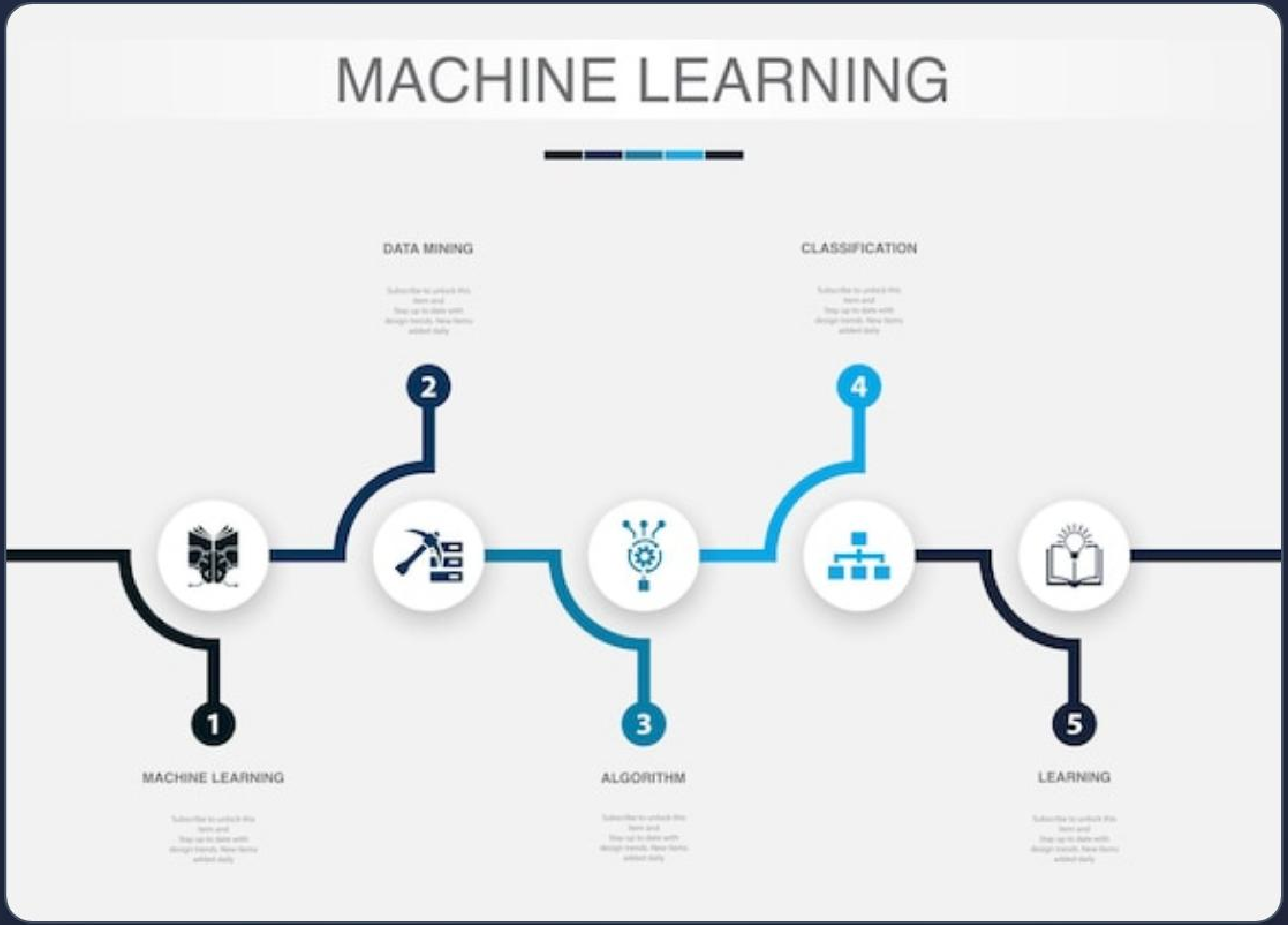
MACHINE LEARNING: CLASSIFICATION & DECISION TREES

From Theory to Implementation

Presenter Name
Date / Event

What is Classification?

- A **supervised machine learning technique**.
- Predicts **categorical class labels** (discrete values).
- **Examples:**
  - Spam vs. Not Spam
  - Yes/No
  - Animal: Cat, Dog, Bird
- Learns from labeled training data to classify new data.



Types of Classification Algorithms

Decision Tree

Tree-like model of decisions.

Random Forest

Ensemble of multiple decision trees.

Naïve Bayes

Based on Bayes' theorem, assumes feature independence.

KNN

Classifies based on nearest data points.

Classification Use Cases

Email Filtering: Spam detection

Medical Diagnosis: Disease prediction

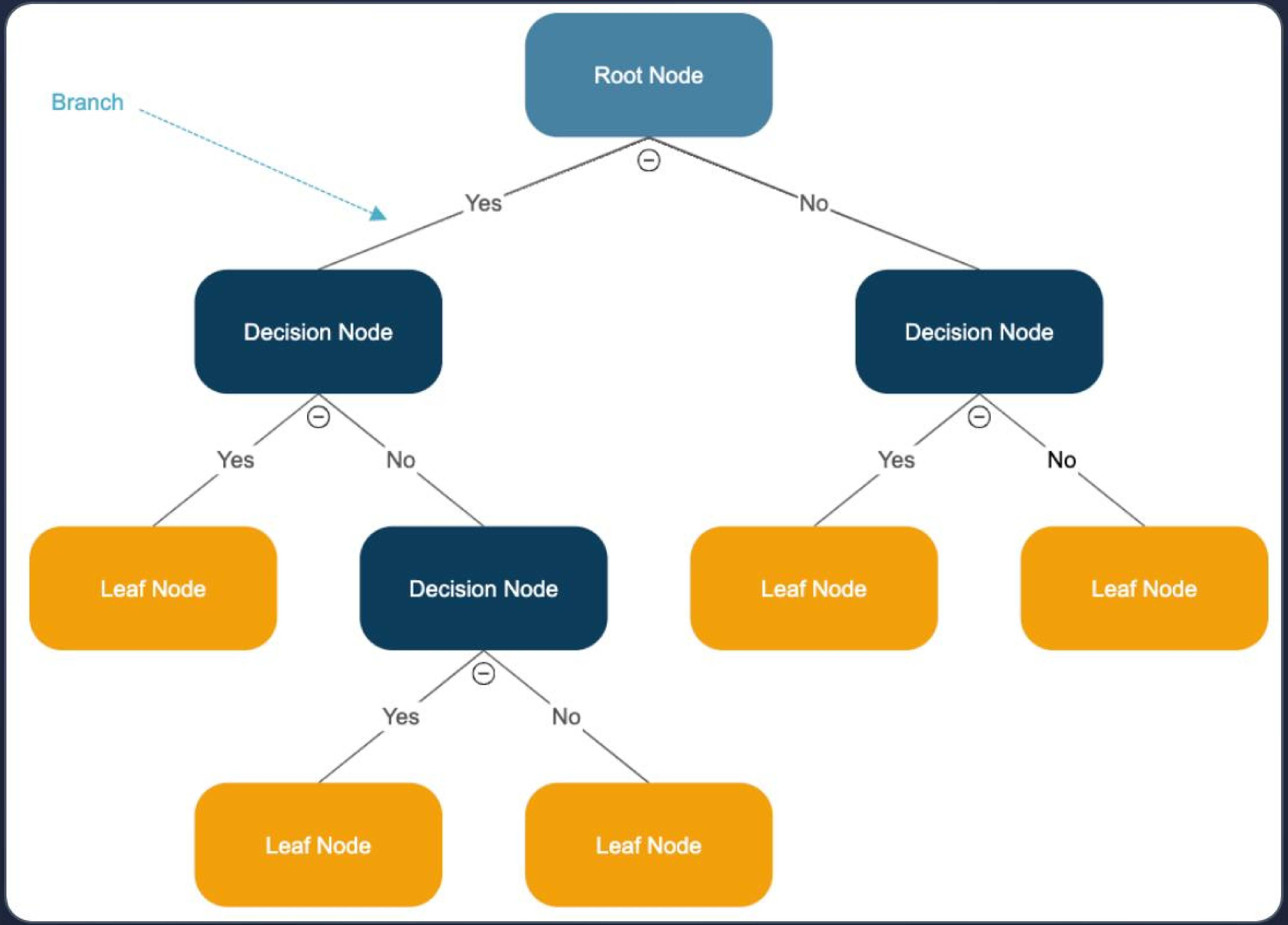
Banking: Loan approval, fraud detection

Customer Behavior: Churn prediction

Image Recognition: Object classification

What is a Decision Tree?

- Flowchart-like tree structure.
- **Internal nodes:** Features/attributes.
- **Branches:** Decision rules.
- **Leaf nodes:** Outcomes/class labels.
- **Root node:** Topmost decision point.
- Intuitive, easy to interpret, mimics human decision-making.



Decision Tree Terminology

Core Components

- **Root Node:** First feature to split data.
- **Decision/Internal Node:** Sub-node that further splits.
- **Leaf/Terminal Node:** Final output (no further split).

Actions & Structure

- **Splitting:** Dividing nodes into sub-nodes.
- **Pruning:** Removing unnecessary branches to avoid overfitting.
- **Branch/Sub-Tree:** Section of the tree.

Visualizing a Decision Tree

Example: Should I play tennis?

```
Outlook? ├── Sunny → Humidity? │ ├── High → No │ └── Normal → Yes ├── Overcast → Yes └── Rain → Wind? ├── Strong → No └──
```

Visual tools: `sklearn.tree.plot_tree`, Graphviz

Writing a Classifier from Scratch

Key Steps:

1. Define the Node structure.
2. Choose best split (using Gini/Entropy).
3. Build tree recursively.
4. Make predictions by traversing tree.
5. (Optional) Add pruning, depth limit

Simple Code Structure Preview:

```
class DecisionNode: def __init__(self, feature=None, threshold=None, left=None, right=None, value=No
```

Summary

Prediction

Classification predicts discrete categories based on input data.

Intuition

Decision trees are intuitive, interpretable models that mimic human logic.

Fundamentals

Building from scratch enhances understanding of core ML fundamentals.

?

Questions?

Contact: [Your Email/Handle]

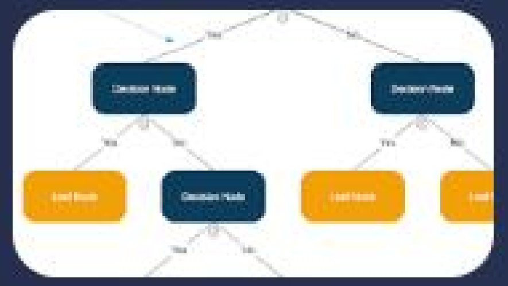
Resources:

- Scikit-learn Documentation
- "Introduction to Statistical Learning"
- GitHub: Example code repository link

Image Sources

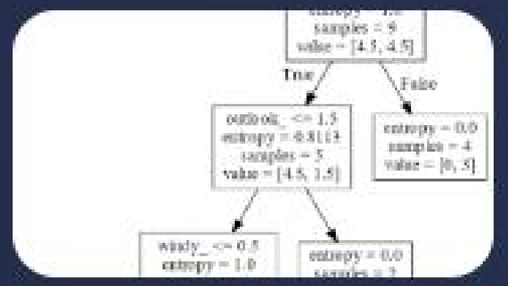
https://img.freepik.com/premium-vector/machine-learning-data-mining-algorithm-classification-learning-icons-infographic-design-template-creative-concept-with-5-steps_159242-13150.jpg

Source: www.freepik.com



https://waz.smartdraw.com/decision-tree/img/structure-of-a-decision-tree.png?bn=15100111939

Source: www.smartdraw.com



https://i.sstatic.net/IGS70.png

Source: stackoverflow.com