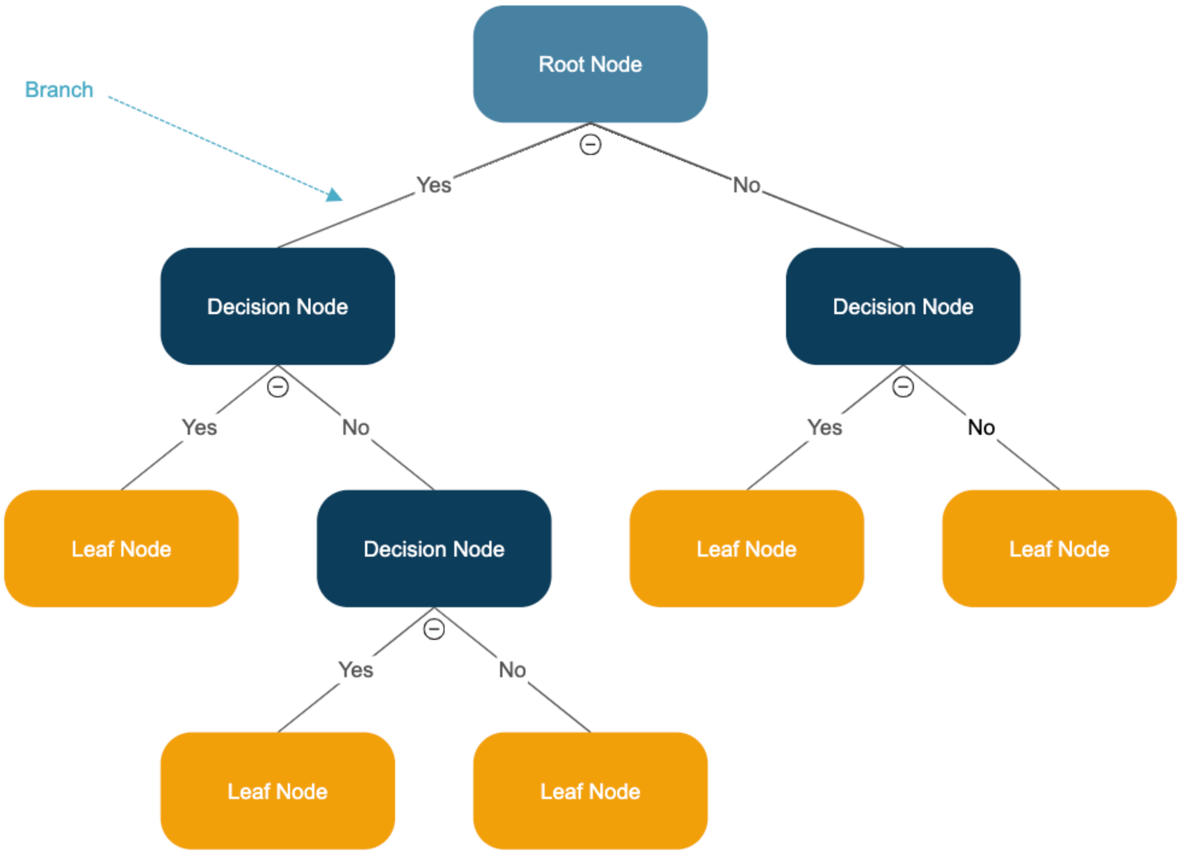
Decision Tree

A
Decision Tree is a machine learning model that helps in making decisions by breaking down a problem into a series of simple questions or conditions. It works like a flowchart, where each decision point (called a "node") asks a question, and based on the answer, it moves to the next step until it reaches an outcome (called a "leaf node")

Simple Explanation:

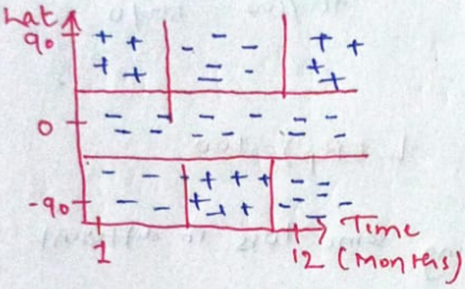
- Imagine a tree where each branch represents a choice or decision, and each leaf is a final decision or outcome.
- The model asks questions at each step, and based on the answer (like "yes" or "no"), it moves to the next question or final decision.


Root Node
Branch
Yes
No
Decision Node
Decision Node
Yes
No
Yes
No
Leaf Node
Decision Node
Leaf Node
Leaf Node
Yes
No
Leaf Node
Leaf Node

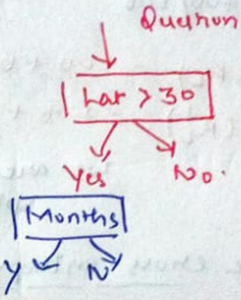
Day 3: **Decision Tree**

A decision tree is a non-parametric supervised learning algorithm, which is utilized for both classification and regression tasks. It has a hierarchical tree structure, which consist of a root node, branches, internal nodes and leaf nodes.

Decision Tree



Greedy, Top-Down Recursive Partitioning.



whether You can ski or not

{Region R_{parent}

$\Rightarrow$ Looking for a Split S_p

$$S_p(j,t) = \left(\{X \mid X_j < t,\ X \in R_p\} \xrightarrow{} R_1,\ \{X \mid X_j \geq t,\ X \in R_p\} \xrightarrow{} R_2 \right)$$

Q. How to choose Split.

Define $L(R)$: Loss on R

Given C classes, define

$\hat{p}_c$ to be proportion of example in R that are of class c.

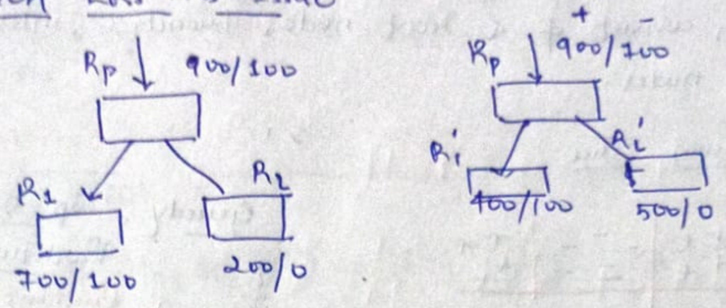
$$\boxed{L_{(misclassification)} = 1 - \max \hat{p}_c}$$

our Aim is to minimize loss

$$\text{s.t. } \min \underbrace{L(R_p)}_{\text{parent loss}} - (\underbrace{L(R_1) + L(R_2)}_{\text{children loss}})$$

\# Misclassification Loss Has Issue

ex



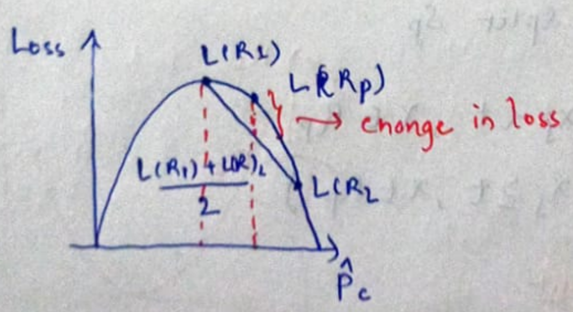
$L(R_1) + L(R_2) = 100 + 0$ $\quad$ $L(R_p) = 100$

$L(R_1') + L(R_2') = 100 + 0$

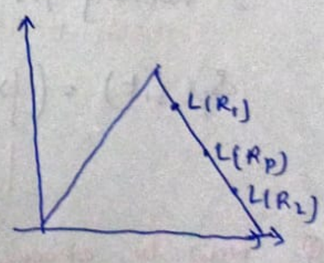
we are getting same loss in different split

* Instead, define cross-entropy loss.

$$\boxed{L_{(cross)} = \sum_{class} \hat{p}_c \log_2 \hat{p}_c}$$



cross Entropy



Misclassification

$$L(R_p) = \frac{L(R_1) + L(R_2)}{2}$$