

- Preface of Machine Learning
- Definition of ML by Tom. M. Mitchell
- What is ML?
- Difference between AI and ML
- Understanding ML by real life examples
- Need For Machine Learning
- Machine Learning Definitions
- Machine Learning Process
- How does Machine Learning Work?
- Machine Learning Types
- Supervised Learning
- Unsupervised Learning

- Practical of Linear, Multiple and Polynomial Regression

Lecture-1,2,3,4,5



Supervised Learning

(Introduction to Decision Tree)

Dr. Virendra Singh Kushwah

Assistant Professor Grade-II

School of Computing Science and Engineering

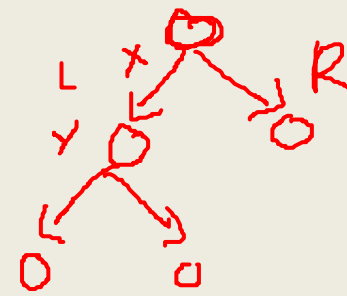
Virendra.Kushwah@vitbhopal.ac.in

7415869616

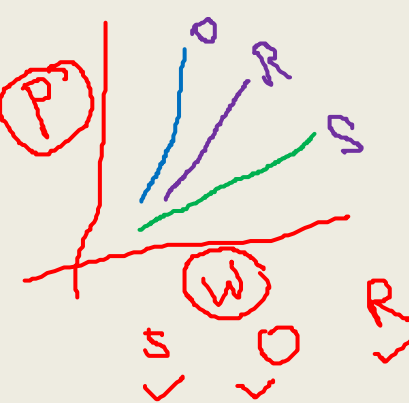
Decision Tree

- A Decision Tree has many analogies in real life and turns out, it has influenced a wide area of Machine Learning, covering both *Classification* and *Regression*.
- In decision analysis, a decision tree can be used to visually and explicitly represent decisions and decision making.

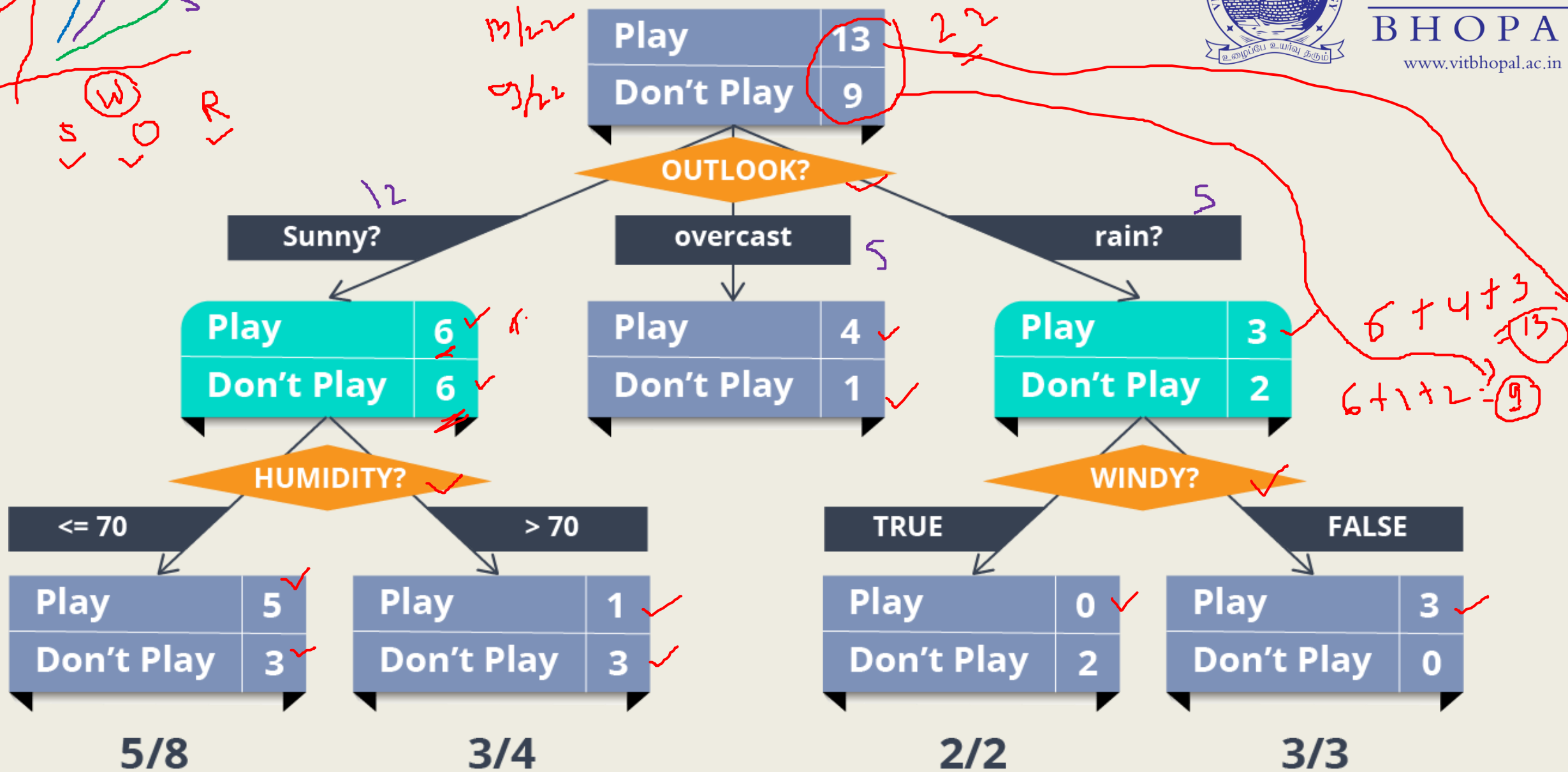
What is a Decision Tree?



- A decision tree is a map of the possible outcomes of a series of related choices. It allows an individual or organization to weigh possible actions against one another based on their costs, probabilities, and benefits.
- As the name goes, it uses a tree-like model of decisions. They can be used either to drive informal discussion or to map out an algorithm that predicts the best choice mathematically.
- A decision tree typically starts with a single node, which branches into possible outcomes. Each of those outcomes leads to additional nodes, which branch off into other possibilities. This gives it a tree-like shape.



Dependent variable: PLAY



PROJECT MANAGEMENT DECISION HELPER

SHOULD I
START THIS
PROJECT?

YES

NO

Ask not to be involved

Is it important
for the
company?

YES

NO

Is it too risky?

Is it important
for my career?

YES

NO

Consider not
to be involved

Start the
project

Start the
project

Do not start it

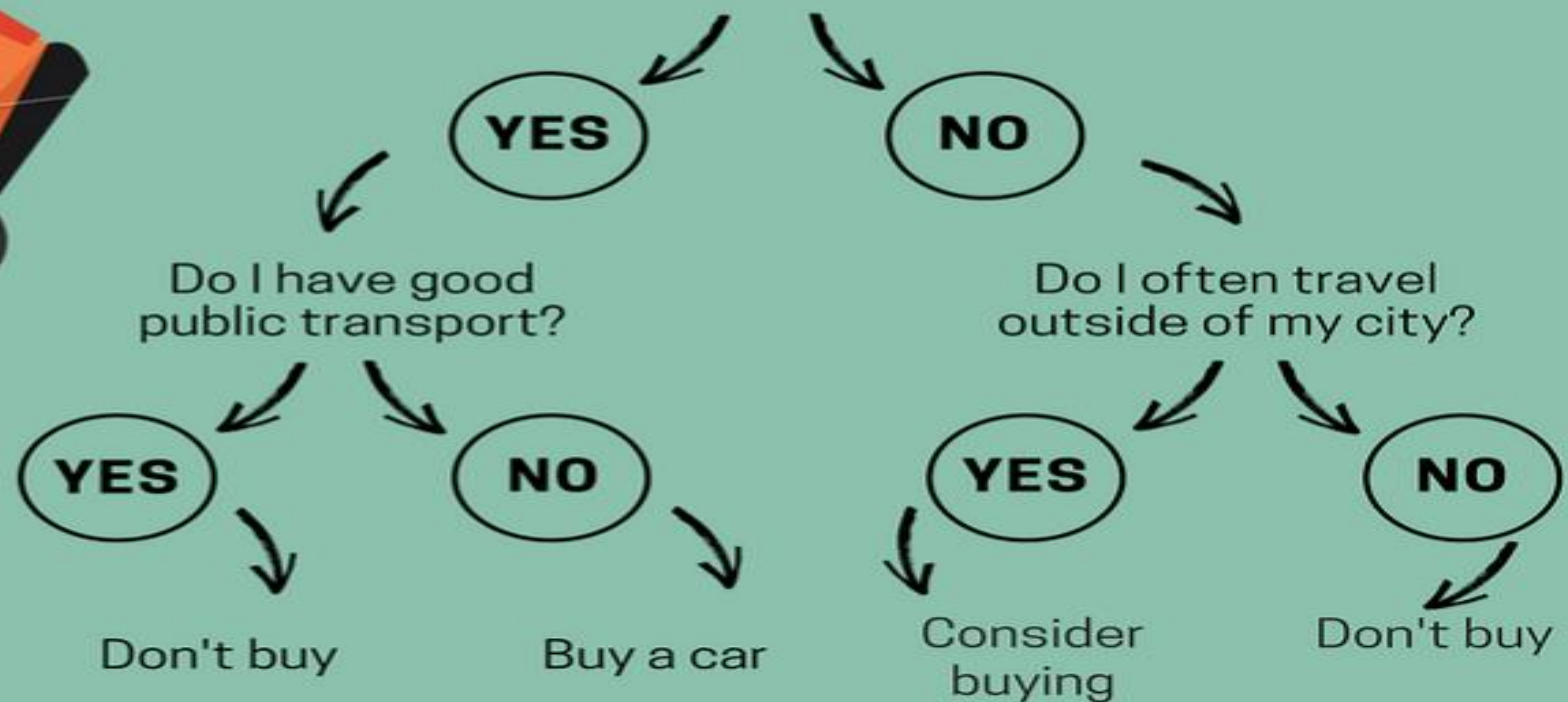


Classification

CAR BUYING DECISION HELPER



SHOULD I BUY A CAR?





$$y = m_1x_1 + m_2x_2 + m_3x_3 - m_4x_4 + c$$



Why Decision Tree Algorithm?



- Decision Tree is considered to be one of the most useful Machine Learning algorithms since it can be used to solve a variety of problems. Here are a few reasons why you should use Decision Tree:

$$y = m_1x_1 + m_2x_2$$



1. It is considered to be the most understandable Machine Learning algorithm and it can be easily interpreted. ✓
2. It can be used for classification and regression problems.
3. Unlike most Machine Learning algorithms, it works effectively with non-linear data.
4. Constructing a Decision Tree is a very quick process since it uses only one feature per node to split the data.

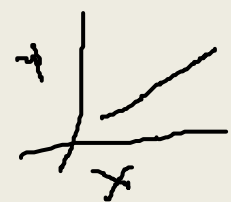
Linear data vs

non-linear data

polynomial
degree = n

linear?

① $y = mx + c$



② $y = m_1x_1 + m_2x_2 + \dots + m_nx_n$



$y = m_1x_1^1 + m_2x_1^2 + m_3x_1^3 + \dots + m_nx_1^n + c$

always
degree = 1

degree = n

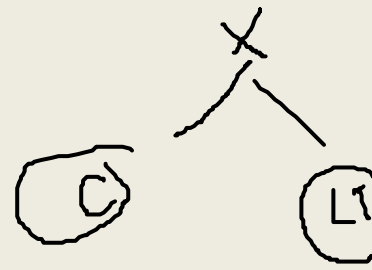


Advantages & Disadvantages of Decision Trees

- Advantages

- ✓ Decision trees generate understandable rules.
- ✓ Decision trees perform classification without requiring much computation.
- ✓ Decision trees are capable of handling both continuous and categorical variables.
- ✓ Decision trees provide a clear indication of which fields are most important for prediction or classification.

- **Disadvantages**



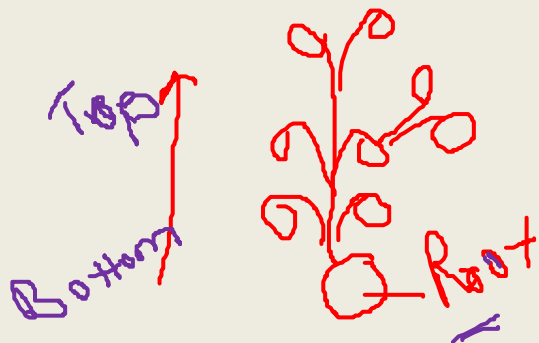
min max



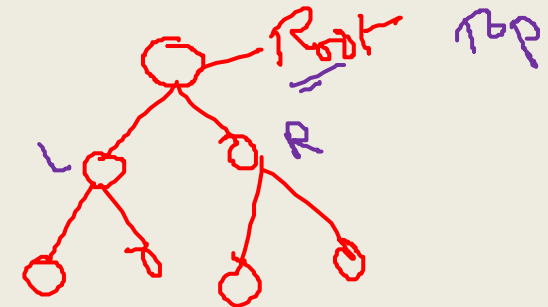
- Decision trees are less appropriate for estimation tasks where the goal is to predict the value of a continuous attribute.
- Decision trees are prone to errors in classification problems with many class and a relatively small number of training examples.
- Decision trees can be computationally expensive to train. The process of growing a decision tree is computationally expensive. At each node, each candidate splitting field must be sorted before its best split can be found. In some algorithms, combinations of fields are used and a search must be made for optimal combining weights. Pruning algorithms can also be expensive since many candidate sub-trees must be formed and compared.

What Is A Decision Tree Algorithm?

- A Decision Tree is a Supervised Machine Learning algorithm which looks like an inverted tree, wherein each node represents a predictor variable (feature), the link between the nodes represents a Decision and each leaf node represents an outcome (response variable).

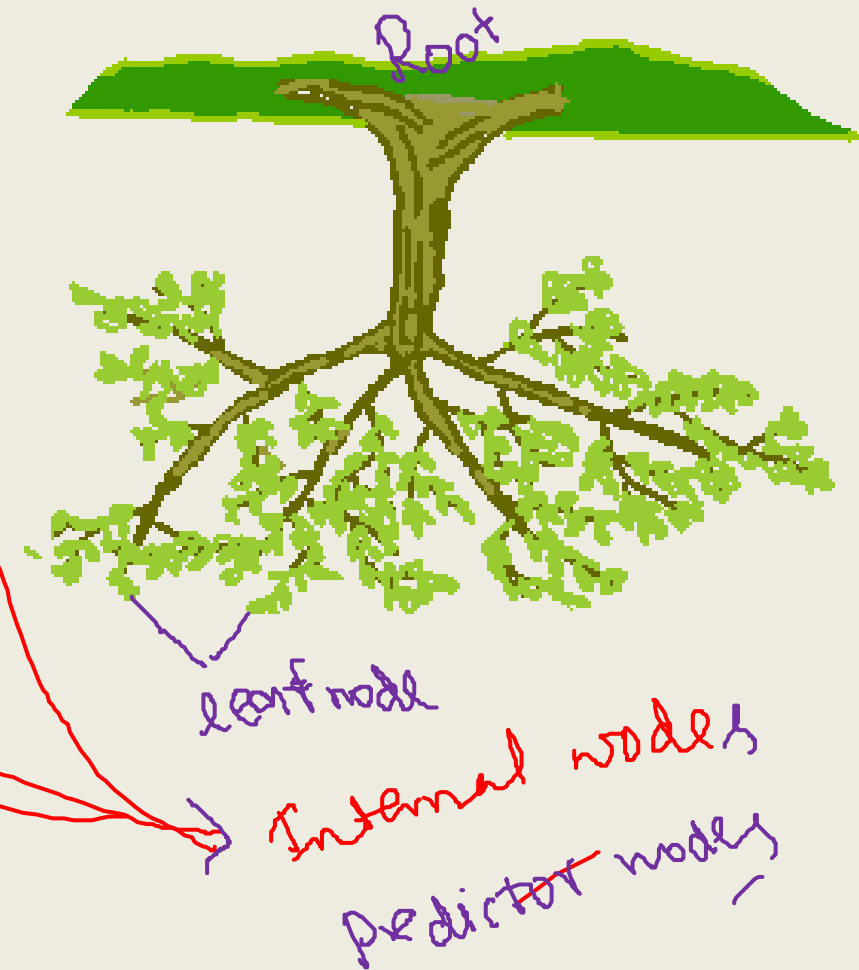
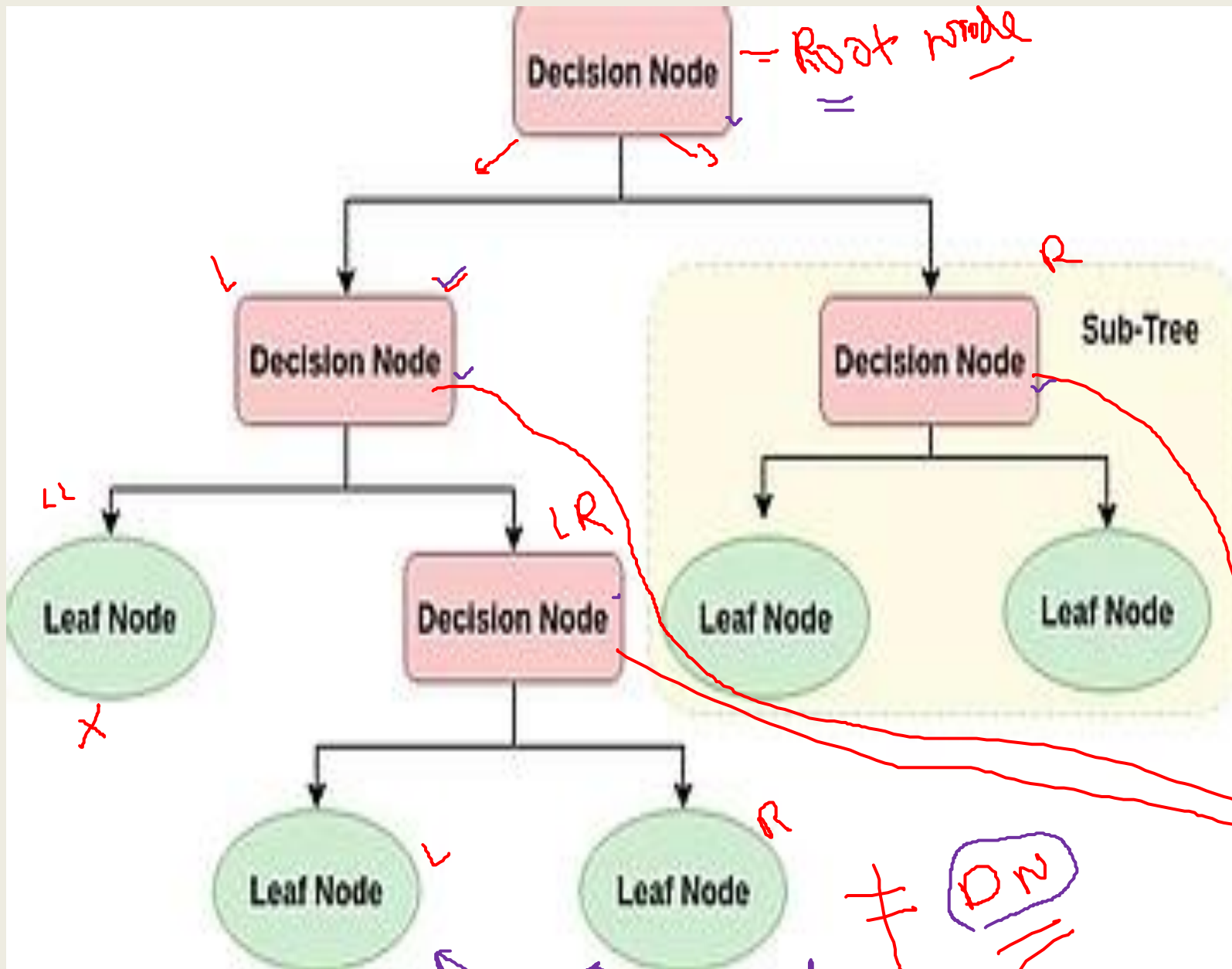


division node



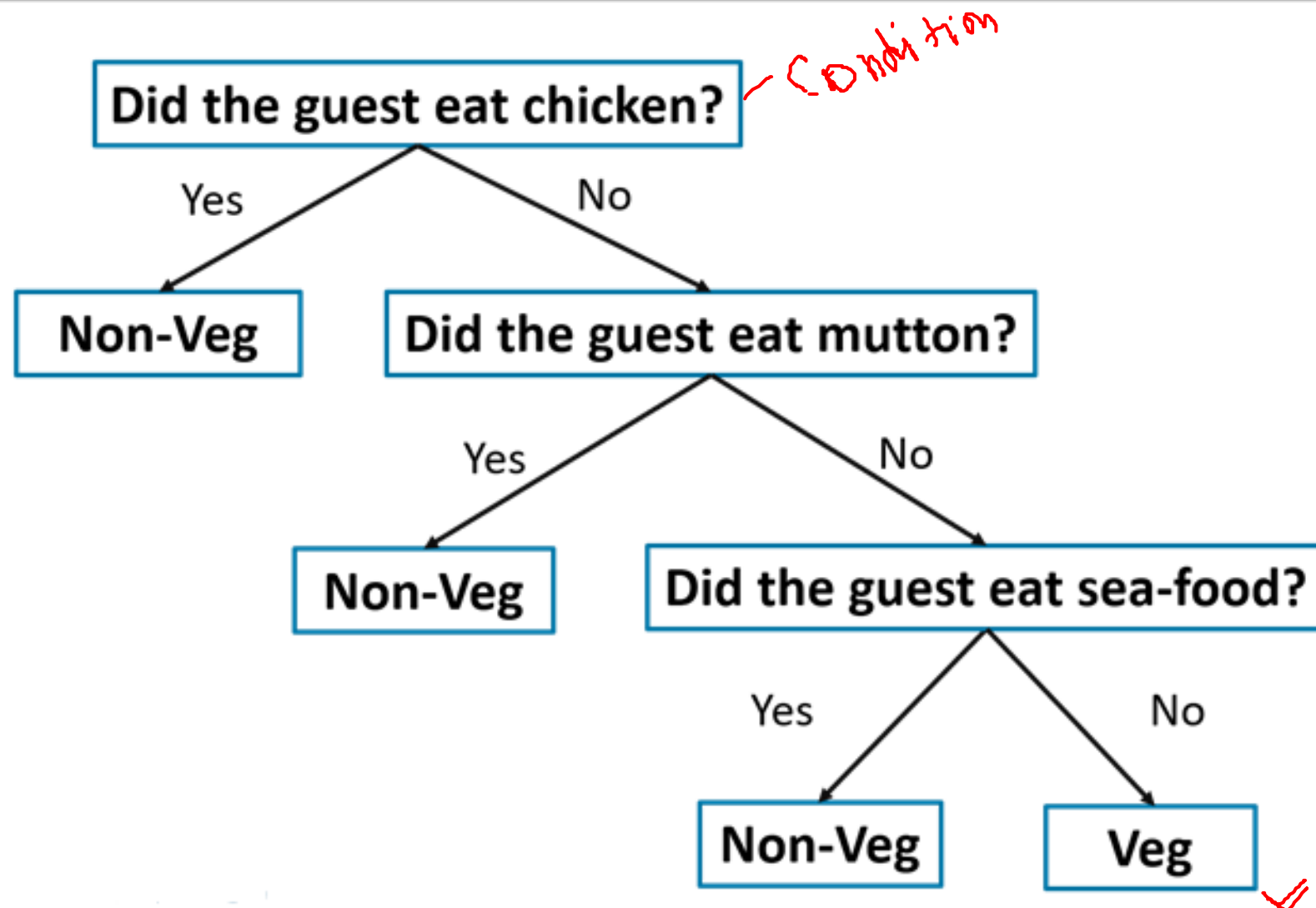
- A Decision Tree has the following structure:

- Root Node: The root node is the starting point of a tree. At this point, the first split is performed.
- Internal Nodes: Each internal node represents a decision point (predictor variable) that eventually leads to the prediction of the outcome.
- Leaf/ Terminal Nodes: Leaf nodes represent the final class of the outcome and therefore they're also called terminating nodes.
- Branches: Branches are connections between nodes, they're represented as arrows. Each branch represents a response such as yes or no.

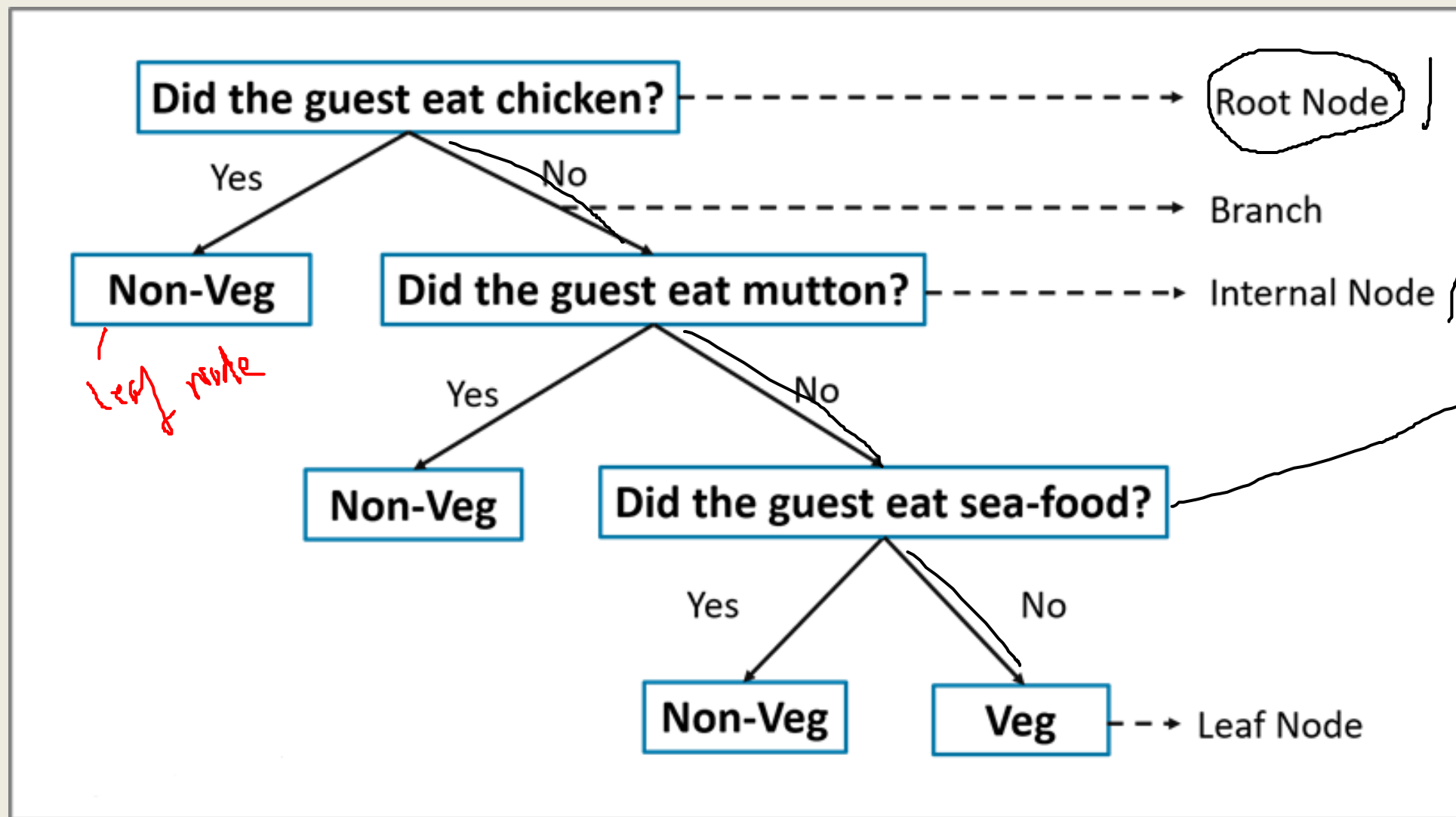


Let us see an example

- Let's say that you hosted a huge party and you want to know how many of your guests were non-vegetarians. To solve this problem, let's create a simple Decision Tree.
- In the above illustration, I've created a Decision tree that classifies a guest as either vegetarian or non-vegetarian. Each node represents a predictor variable that will help to conclude whether or not a guest is a non-vegetarian. As you traverse down the tree, you must make decisions at each node, until you reach a dead end.



Structure Of A Decision Tree



Root Node

Branch

Internal Node

Leaf Node