**HEADING : DATA ANALYTICS PHASE 2**



**Problem Definition of Phase 2 :**

\* The problem definition is to find the prediction of how many customers leave according to their own perspectives.

\* To protect the exact reason for the customer living the market is done through some models.

\*These algorithms analyze various factors such as customer demographics, purchase behavior, and interactions with the company.

Innovation:

\*The predictive modelling techniques we have used here are:

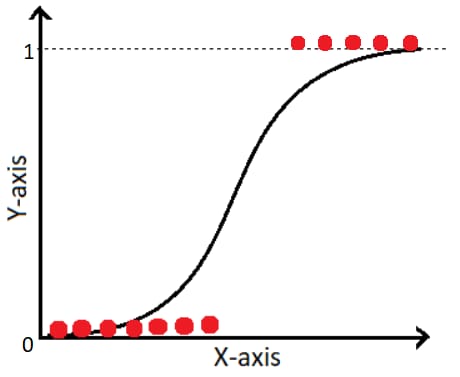
1. Logistic Regression : suitable for binary classification problems, predicting whether a customer is churn or not.

2. Decision tree and random forest:

Effective for capturing complex relationships in the data.

**Visual Representation of Models:**

1. Logistic Regression :

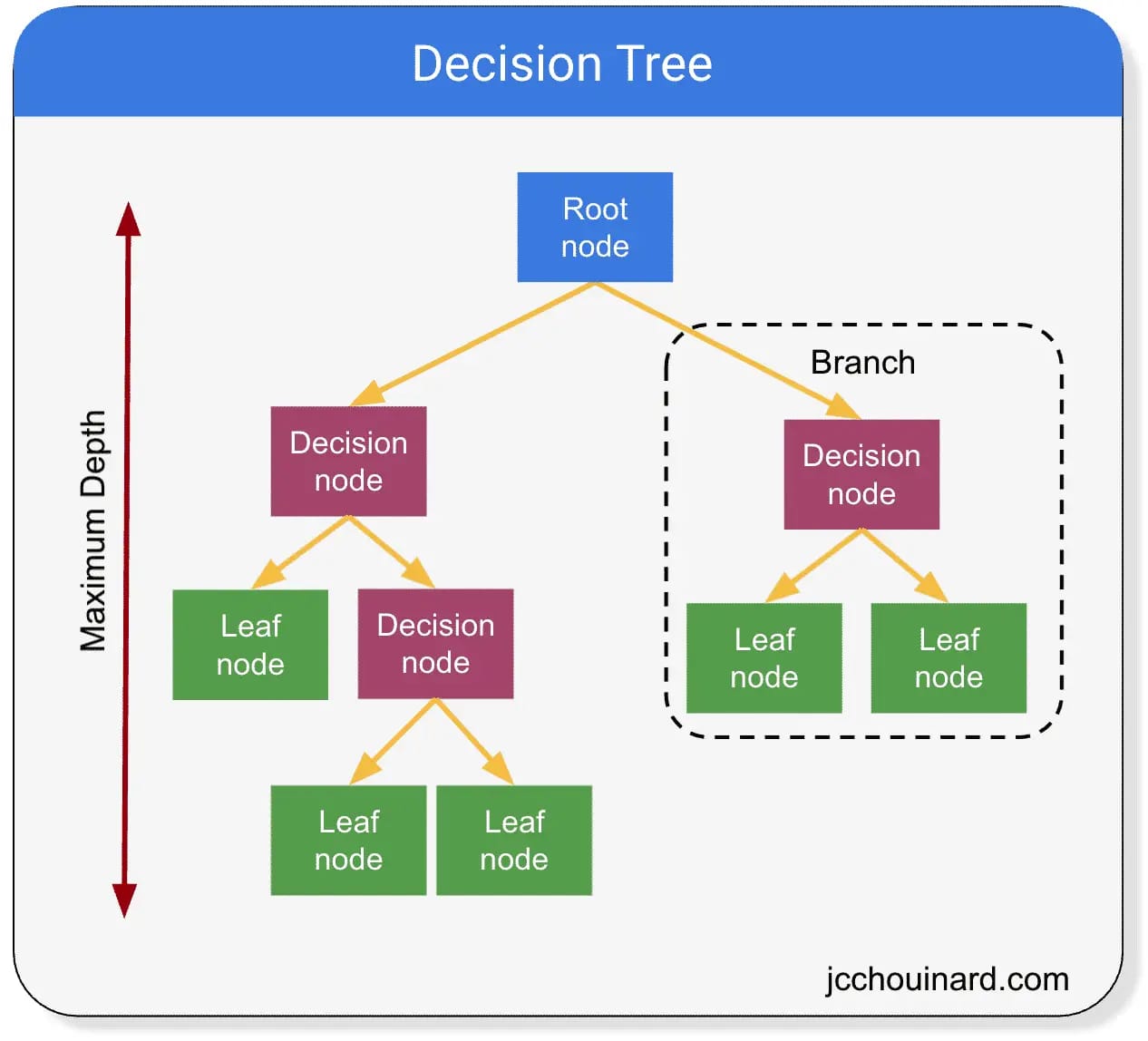


1. Random forest :

A diagram of a sample

Description automatically generated

1. Decision tree:



**Description on Model:**

\* Logistic regression calculates probabilities of a customer churning. Unlike linear regression, which predicts a continuous outcome, logistic regression models the probability of an event occurring, which is suitable for predicting binary outcomes like churn.

\*Random Forest models often achieve high accuracy in predicting customer churn, especially when there are a large number of diverse features. This accuracy is crucial for businesses aiming to identify potential churners accurately and take timely actions to retain customers.

**Conclusion:**

At last this phase has done with the decision of choosing perfect models for the prediction of costumer churn.