The HR team stored data of the promotion cycle last year, which consists of details of all the employees in the company working last year and also if they got promoted or not, but every time this process gets delayed due to so many details available for each employee - it gets difficult to compare and decide.

So this time the HR team wants to utilize the stored data to make a model that will predict if a person is eligible for promotion or not.

Need to come up with a model that will help the HR team to predict if a person is eligible for promotion or not.

Objective

Explore and visualize the dataset.

Build a classification model to predict if the customer has a higher probability of getting a promotion

Optimize the model using appropriate techniques

Generate a set of insights and recommendations that will help the company

Data Dictionary

• employee id: Unique ID for the employee

• department: Department of employee

• region: Region of employment (unordered)

• education: Education Level

• gender: Gender of Employee

• recruitment channel: Channel of recruitment for employee

• no\_ of\_ training: no of other training completed in the previous year on soft skills, technical skills, etc.

• age: Age of Employee

• previous\_ year\_ rating: Employee Rating for the previous year

• length\_ of\_ service: Length of service in years

• awards\_ won: if awards won during the previous year then 1 else 0

• avg\_ training\_ score: Average score in current training evaluations

• is\_promoted: (Target) Recommended for promotion

**Protocols:**

1. All participants should adhere to agreed timelines and timelines will not be extended.
2. All the documentation – Final presentation and R/python code to be submitted before the final presentation day.
3. All the participants must attend review meetings.