

Problem Statement:

Develop a predictive model using employee data to classify individuals as likely to stay or leave the company. This classification will assist in making informed decisions about employee retention strategies and workplace improvements.

Overview

The dataset contains 900 rows and 15 columns, representing various employee metrics. The data aims to reflect realistic scenarios in a corporate setting, encompassing professional and personal employee metrics. Below is a brief overview of the dataset columns:

1. JobSatisfaction: Employee's job satisfaction level.
2. Performance rating: Performance rating given by the company.
3. YearsAtCompany: Total number of years the employee has been with the company.
4. WorkLifeBalance: Rating of how well the employee feels they balance work and personal life.
5. DistanceFromHome: Distance from the employee's home to the workplace.
6. Monthly Income: The monthly income of the employee.
7. EducationLevel: The highest level of education attained by the employee.
8. Age: The age of the employee.
9. NumCompaniesWorked: The number of companies the employee has worked at before joining the current company.
10. Employee Role: The role or position of the employee within the company.
11. Annual Bonus: Annual bonus received by the employee.
12. Training hours: Number of hours spent in training programs.
13. Department: The department in which the employee works.
14. AnnualBonus_Squared: Square of the annual bonus (a polynomial feature).
15. AnnualBonus_TrainingHours_Interaction: Interaction term between annual bonus and training hours.

