PR Final Project Description

Project Title - **Employee Attrition Prediction in Large Organizations**

Team Members:

1. Venkata Pavan Surya Kanakam
2. Ganga Sai Krishna Battu
3. Pavan Thalla
4. Anuhya Paturu
5. Goutham Pulivarthi

Link to the Dataset: https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset

Description:

The **Synthetic Employee Attrition Dataset** a real HR dataset and contains **10000 rows and 31 columns.** Each row represents an individual employee's details, and each column captures a unique attribute. The dataset is designed to predict **attrition status** (Attrition column) based on various factors that typically influence turnover in large organizations.

Five Sample Rows:

Employee ID	Age	Attrition	BusinessTrav	DailyRate	Department	DistanceFro	Education	EducationFie	Environment	Gender	HourlyRate	JobInvolvme	JobLevel	JobRole	JobSatisfacti	MaritalStatu	MonthlyInco	NumCompan	OverTime	PercentSalar
1	18	No	Travel_Frequ	365	Research & D	13	4	Other	1	Female	70	2	4	Sales Execut	1	Divorced	2591	6	Yes	13
2	48	No	Travel_Rarely	564	Research & D	20	3	Life Sciences	3	Male	68	2	3	Manufacturir	3	Divorced	6074	0	Yes	15
3	59	No	Travel_Rarely	1111	Sales	18	2	Technical De	2	Female	64	1	4	Laboratory Te	2	Divorced	12478	9	Yes	17
4	20	No	Travel_Frequ	811	Human Reso	8	2	Human Reso	1	Female	36	1	4	Manufacturir	2	Divorced	1821	9	No	23
5	44	No	Travel_Frequ	787	Research & D	25	2	Human Reso	1	Male	82	4	1	Sales Repres	2	Single	11105	7	No	11
6	56	No	Non-Travel	1124	Sales	27	4	Technical De	4	Female	97	3	1	Research Sci	2	Divorced	8681	6	No	15

Performance	Relationship	StockOptionL	TotalWorking	TrainingTime	WorkLifeBala	YearsAtComp	YearsInCurre	YearsSinceLa	YearsWithCurrManager
4	2	3	27	2	3	12	1	8	12
4	4	0	2	5	1	35	8	11	1
3	4	0	15	4	2	9	19	13	19
3	2	1	16	1	2	35	1	12	16
3	4	0	38	4	4	14	16	0	14
3	1	0	38	1	2	7	14	12	5

Dataset Attributes

Below are the key columns with descriptions for each:

1. Employee ID: Employee ID
2. **Age**: Employee's age.
3. **Attrition** (Target Variable): Indicates if an employee left (Yes) or stayed (No).
4. **BusinessTravel**: Frequency of travel required for the job (Travel_Rarely, Travel_Frequently, Non-Travel).
5. **DailyRate**: Daily salary rate of the employee.
6. **Department**: Department of the employee (Sales, Research & Development, Human Resources).
7. **DistanceFromHome**: Distance between employee's home and workplace (in miles).
8. **Education**: Education level (1-4, where 1 = Below College, 4 = Doctorate).
9. **EducationField**: Field of education (e.g., Life Sciences, Medical, Marketing).
10. **EnvironmentSatisfaction**: Employee's satisfaction with the work environment (1-4 scale).
11. **Gender**: Gender of the employee (Male, Female).
12. **HourlyRate**: Hourly wage rate of the employee.
13. **JobInvolvement**: Employee's involvement in their job (1-4 scale).
14. **JobLevel**: Level of the job position within the organization (1-5).
15. **JobRole**: Job role (e.g., Sales Executive, Research Scientist, Manager).
16. **JobSatisfaction**: Job satisfaction rating (1-4 scale).
17. **MaritalStatus**: Marital status of the employee (Single, Married, Divorced).
18. **MonthlyIncome**: Monthly income of the employee.
19. **NumCompaniesWorked**: Number of previous companies the employee has worked for.
20. **OverTime**: Indicates if the employee works overtime (Yes, No).
21. **PercentSalaryHike**: Percentage increase in salary.
22. **PerformanceRating**: Performance rating (3 = Good, 4 = Excellent).
23. **RelationshipSatisfaction**: Satisfaction with relationships at work (1-4 scale).
24. **StockOptionLevel**: Level of stock options granted (0-3).
25. **TotalWorkingYears**: Total years the employee has worked.
26. **TrainingTimesLastYear**: Number of training sessions attended last year.
27. **WorkLifeBalance**: Work-life balance rating (1-4 scale).
28. **YearsAtCompany**: Number of years the employee has been with the company.
29. **YearsInCurrentRole**: Years the employee has been in their current role.
30. **YearsSinceLastPromotion**: Years since the employee's last promotion.
31. **YearsWithCurrManager**: Number of years the employee has worked with their current manager.

Project Goal

The primary objective of this study is to predict attrition, or the likelihood that an employee would leave, using the variables in the dataset. With the use of this data, we might be able to identify patterns that contribute to employee turnover and eventually create a model that would assist HR departments in proactively managing attrition. The information gathered from this study can be applied to the creation of focused staff engagement and retention plans.