Employee Attrition Project – Feature Importance Guide

This document summarizes the relation of different HR dataset columns with employee attrition. It helps in deciding which columns are useful for predictive modeling, which are optional, and which can be dropped. This guide is useful for data preprocessing, feature selection, and explaining model decisions in your GitHub repository.

Column	Relation Strength	Notes
Attrition	Target	Yes/No – main target variable.
Age	Strong	Younger employees leave more frequently.
BusinessTravel	Strong	Frequent travel = higher attrition.
Department	Strong	Attrition varies across departments.
DistanceFromHome	Strong	Longer distance increases attrition risk.
JobLevel	Strong	Lower job levels show more attrition.
JobRole	Strong	Certain roles (e.g., Sales Rep, Lab Tech) have higher attrition.
JobSatisfaction	Strong	Low satisfaction strongly drives attrition.
EnvironmentSatisfaction	Strong	Poor environment → higher attrition.
OverTime	Strong	Overtime is one of the strongest predictors of attrition.
MonthlyIncome	Strong	Lower income = higher attrition.
PercentSalaryHike	Strong	Low hike = higher attrition.
TotalWorkingYears	Strong	Mid-career employees tend to leave more.
YearsAtCompany	Strong	Short tenure = higher attrition.
YearsSinceLastPromotion	Strong	Long gap since promotion increases attrition.
WorkLifeBalance	Strong	Poor balance = higher attrition.
Education	Medium	Some effect, but not strong.
EducationField	Medium	Slight differences across fields.
JobInvolvement	Medium	Moderate relation to attrition.
RelationshipSatisfaction	Medium	Small effect on attrition.
StockOptionLevel	Medium	Fewer stock options = higher attrition.
YearsInCurrentRole	Medium	Stagnation in same role increases attrition.
YearsWithCurrManager	Medium	Weak, but some influence.
NumCompaniesWorked	Medium	More job changes = more likely to leave.
TrainingTimesLastYear	Medium	Very weak relation, but addable.
PerformanceRating	Medium	Most employees rated 'Good' – weak predictor.
MaritalStatus	Medium	Single employees slightly higher attrition.
Gender	Medium	Very weak relation; mainly for analysis, not prediction.
DailyRate	None	No strong correlation.
HourlyRate	None	No strong correlation.

Column	Relation Strength	Notes
MonthlyRate	None	No strong correlation.
EmployeeNumber	None	Just an ID – drop.
EmployeeCount	None	Constant value – drop.
StandardHours	None	Constant value – drop.