Predictive Analytics using Machine Learning

Can we predict employee attribution?

10Alytics
data & strategy

Problem Overview

Employee Attrition is a HR Concept that refers to the inferences generated by HR when they try to explain reasons for events, the behavior of employees, and their own behavior. Attributions may be internal (dispositional), based on something within a person, or external (situational), based on something outside a person. In other words it helps HR understand the personality of an employee, hence how likely the employee may respond to certain stimuli or event.

In this Case Study, we will attempt to Uncover the factors that can determine employee attribution. We will then build a predictive model that will be capable of classifying employed attribution.

In the second scenario, we will attempt to build a predictive model that predicts the salary of employees given employee level of education, attribution etc amongst other features present in the data. We will also be identifying the factors that determine an employee's salary

10Alytics
data & strategy

OBJECTIVES

As a Data Scientist, you will not only be analyzing datasets the data to determine the driving factors Deliverables, but you will also be building predictive and forecast models. In this case study, we will be applying Predictive analytics to two different case scenarios. To predict employee attribution and to also build a salary prediction model

GOAL

1. They have hired you as a Data Scientist, shared the data of a company XYZ, and asked to analyze the data to determine the driving factors for employee attribution , and build a predictive model that will be able to predict wether or not an employee will attribute (this is a classic classification problem)
2. In the Second scenario, a the same company XYZ desires a predictive model that can accurately predict the salary of an employee based of the features from the data (this is a classical regression problem)

10Alytics
data & strategy

Here Your are expected to use Machine Learning Techniques to perform some predictive Analytics on the data set.

Specifically we are going to implement both Regression and Classification (Supervised ML techniques) to solve our objectives

10Alytics
data & strategy

Data Dictionary

1. Age
2. Attrition
3. BusinessTravel
4. DailyRate
5. Department
6. HourlyRate
7. TrainingTimesLastYear
8. WorkLifeBalance
9. YearsAtCompany
10. YearsInCurrentRole
11. Years SinceLastPromotion
12. YearsWithCurrManager

1. DistanceFromHome
2. Education
3. EducationField
4. EmployeeCount
5. EmployeeNumber
6. EnvironmentSatisfaction
7. Gender
8. JobInvolvement
9. JobLevel
10. JobRole
11. JobSatisfaction

1. MaritalStatus
2. MonthlyIncome
3. MonthlyRate
4. NumCompaniesWorked
5. Over18
6. OverTime
7. PercentSalaryHike
8. PerformanceRating
9. RelationshipSatisfaction
10. StandardHours
11. StockOptionLevel
12. TotalWorkingYears

To obtain the full data description, please visit
Data Source: https://www.kaggle.com/datasets/rohitsahoo/employee

10Alytics
data & strategy

10Alytics
data & strategy

Create a detailed analysis and upload on your Github.

(Tag @10Alytics)