**Predicting Attrition at D.D.S. Analytics**



**MSDS6306 PROJECT 2**

**Summer 2019**

**8/17/2019**

**By: Justin Howard**

# Abstract

# Data is the buzzword of the twenty-first century. The rapid expansion of predictive analytics to fields outside the finance and manufacturing industries is opening new opportunities in every field. At D.D.S. Analytics, we are embarking on a project that will enhance our ability to provide our clients with a deeper understanding of their human capital flows, beginning with attrition. Every journey begins with a single step, and our first charge is to understand our employees. We focus mainly on attrition and its accompanying trends but will also include a brief but instructive analysis of our current patterns related to compensation.

# We explored the data to find surface-level trends using standard exploratory techniques and used more advanced machine learning techniques to find hidden trends. We conducted a comparative review of three predictive classification models to choose the model that best identified the top three single factors that lead to attrition. Those factors and many other less important predictors had significant variance between their levels. To capture the influence of individual factor levels, we re-tooled the most predictive model to offer a more detailed, level-based analysis to examine which responses were the best predictors of attrition.

The overall conclusions:

The factors contributing to greater estimated odds of attrition are working overtime (odds 4 times greater), being single (odds twice as high), and the number of years between promotions (odds 1.87 times greater for each year).

The largest factors reducing the estimated odds of attrition are attending trainings (81.62% lower odds for each training attended within the last year), choosing Relationship Satisfaction level four (79.45% lower odds) and choosing Environmental Satisfaction level four (72.33% lower odds).

Compensating employees with stock options level 1 (37.06% drop in the odds) and level 2 (31.09% drop in odds) are significant, while stock option level 3 does not significantly increase or decrease the odds of attrition. Also of note is that on average, males are paid $50.03 more each month than females.

Our current estimated compensation model:

MonthlyIncome = -1490.63 + 83.86BusinessTravel -449.56Department -4.24DistanceFromeHome -17.29EnvironmentSatisfaction + 50.03Gender +3.03HourlyRate + 44.45JobInvolvement + 3691.88JobLevel + 50.03JobROle + 8.87NumCompaniesWorked + 60.55TotalWorkingYears + 2.33YearsSinceLastPromotion -23.95YearsWithCurrentManager

# Data Description

The data were collected on a total of 870 employees, identified by an employee ID number. A total of 31 variables are available for analysis.

# Attrition Model Summary

The logistic model without interactions and a lambda penalty of .00943702 achieved an accuracy of 87.74%, a Sensitivity of 76.47% and a Specificity of 88.52%.

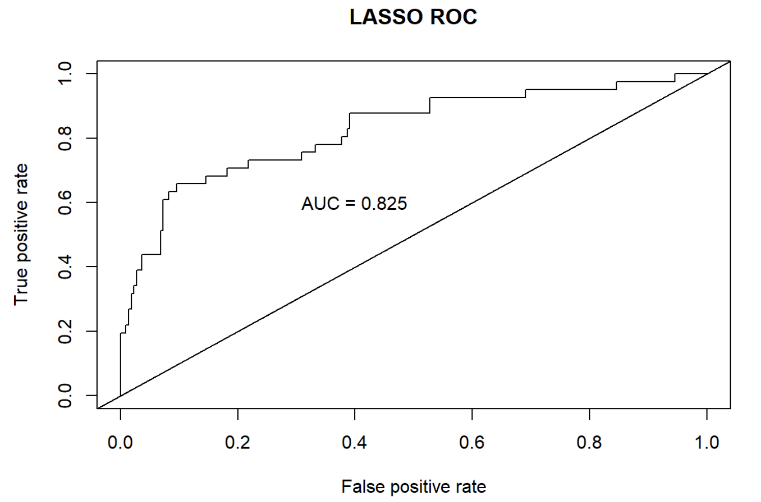


Figure ROC Curve of Logistic Model

|  |  |  |
| --- | --- | --- |
|  | **True “No”** | **True “Yes”** |
| **Predicted “No”** | 216 | 28 |
| **Predicted “Yes”** | 4 | 13 |

Table Confusion Matrix of Logistic Model

Table Interpretation of top ten coefficients

|  |  |  |
| --- | --- | --- |
| **Factor** | **Model Coefficient** | **Interpretation** |
| **Working Overtime** | 1.367 | The estimated odds of attrition for an employee who works over-time are 3.92 times greater than for those who do not work over-time. |
| **Marital Status** | .824 | With each level increase in Marital Status, the estimated odds of attrition increase by 2.27 times the previous level. |
| **Level of Job Involvement** | -.629 | With each level increase in Job Involvement, the odds of attrition decrease by 53.31% |
| **Work-Life Balance** | .308 | With each level increase in Work-Life Balance, the odds of attrition decrease by 73.49%. |
| **Job Satisfaction** | .248 | With each level increase in Job Satisfaction, the odds of attrition decrease by 72.12%. |
| **Years Since Last Promotion** | .183 | For each year after an employee’s last promotion, the estimated odds of attrition rise by 34%. |
| **Environment Satisfaction** | .168 | With each level increase in Environment Satisfaction, the odds of attrition are 78.11% less. |
| **Job Level** | .129 | With each level increase in Job Level, the estimated odds of attrition rise by 28.9% . |
| **Number of Companies Worked For** | .114 | For every company an employee has worked for previously, the estimated odds of attrition rise by 12.07% |
| **Relationship Satisfaction** | .111 | For every company an employee has worked for previously, the estimated odds of attrition rise by 11.73% |

Table

|  |  |
| --- | --- |
| **Variable** | **Description** |
| [**ID**](#id) | Employee ID Number |
| [**Age**](#age) | Employee Age |
| [**Attrition**](#attrition) | "Yes" or "No" if employee left company |
| [**BusinessTravel**](#businesstravel) | Non-Travel, Travel\_Frequently, Travel\_Rarely |
| [**DailyRate**](#dailyrate) | Daily pay |
| [**Department**](#department) | Human Resources, Research & Dev., Sales |
| [**DistanceFromHome**](#distancefromhome) | Commute distance |
| [**Education**](#education) | Levels of education from 1-5 |
| [**EducationField**](#educationfield) | Degree field |
| [**EmployeeCount**](#employeecount) | Number of employees |
| [**EmployeeNumber**](#employeenumber) | Employee number |
| [**EnvironmentSatisfaction**](#environmentsatisfaction) | Satisfaction rating from 1-4 |
| [**Gender**](#gender) | Male or Female |
| [**HourlyRate**](#hourlyrate) | Hourly pay |
| [**JobInvolvement**](#jobinvolvement) | Employee's expressed involvement in job |
| [**JobLevel**](#joblevel) | Employee rank |
| [**JobRole**](#jobrole) | Job title |
| [**JobSatisfaction**](#jobsatisfaction) | Employee's satisfaction rating from 1-4 |
| [**MaritalStatus**](#maritalstatus) | Marital status from Divorced, Married, Single |
| [**MonthlyIncome**](#monthlyincome) | Monthly salary |
| [**MonthlyRate**](#monthlyrate) | Monthly rate of pay |
| [**NumCompaniesWorked**](#numcompaniesworked) | Number of companies worked for in the past |
| [**Over18**](#over18) | Yes for over 18 |
| [**OverTime**](#overtime) | Yes or No for having worked overtime |
| [**PercentSalaryHike**](#percentsalaryhike) | Percentage of last raise in pay |
| [**PerformanceRating**](#performancerating) | Job skill rating from 1-4 |
| [**RelationshipSatisfaction**](#relationshipsatisfaction) | Employee satisfaction with their manager |
| [**StandardHours**](#standardhours) | Number of hours worked per week |
| [**StockOptionLevel**](#stockoptionlevel) | Vestager in company stock |
| [**TotalWorkingYears**](#totalworkingyears) | Total number of working years for employee |
| [**TrainingTimesLastYear**](#trainingtimeslastyear) | Number of trainings attended within 1 year |
| [**WorkLifeBalance**](#worklifebalance) | Employee rating from 1-4 |
| [**YearsAtCompany**](#yearsatcompany) | Number of years at D.D.S. Analytics |
| [**YearsInCurrentRole**](#yearsincurrentrole) | Number of years in current position |
| [**YearsSinceLastPromotion**](#yearssincelastpromotion) | Number of years since last promotion |
| [**YearsWithCurrManager**](#yearswithcurrmanager) | Number of years working with manager |