



HR ANALYTICS CASE STUDY

Submitted By:

AAQIB MUSHTAQ KHAN

PREETHI DAMODARAN

SHIRIN LOURDES



ABSTRACT

A company called XYZ, is facing the problem of employee attrition. Every year around 15% of the employees are leaving the company and requires replacement. This attrition is bad for the company since it results in :

- Project delay and difficulty to meet the deadlines.
- A department needs to be maintained for the recruitment.
- The new employees need to be provided with training.

They need to obtain the factors that can help them curb attrition.

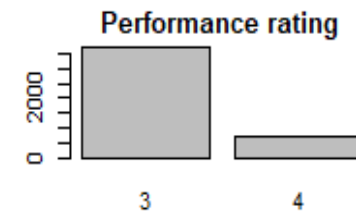
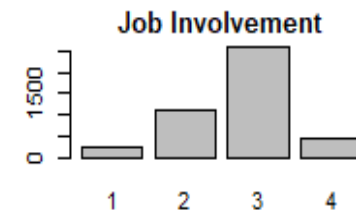
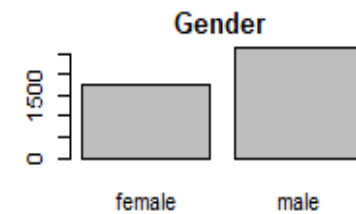
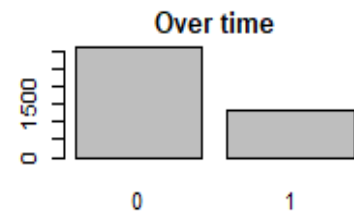
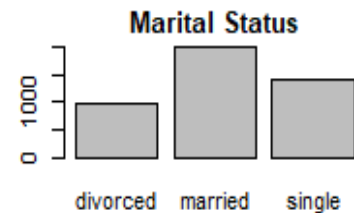
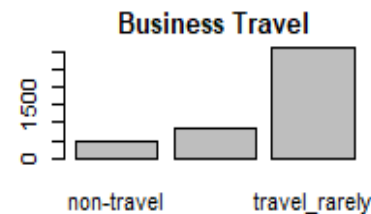
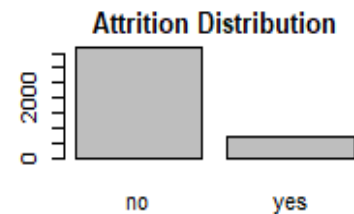


PROBLEM SOLVING METHODOLOGY

- First clean the datasets to remove any unnecessary variables or values.
- Calculating the average working hours of each employee using the time datasets provided.
- Merge all the datasets together to a single one (hr_analytics).
- Exploratory Data analysis is done and the factors affected are identified.
- Build a model of the dataset using logistic regression to find the factors affecting the attrition.

What do the plots show?

- Attrition distribution is very low.
- Employees are mostly satisfied but there are some who are not satisfied (rating 1 and 2).
- Most of the employees travel less.
- In gender distribution, males are double in numbers compared to females.
- Most of them are married.
- Job involvement is less than the highest.
- Most of the employees serve overtime.
- Performance rating is mostly between 3 and 4.





Approach for Logistic Regression

For creating Train and test datasets from final data set:

- Fixed seed to 100
- Used split ratio of 0.7 for training dataset and remaining data has been assigned to test dataset
- Initial model has been conceived with *glm* function, then *StepAIC* has been applied to arrive at standard model which yielded on iterative predictor selection with out major reduction in AIC Score (2077.1 after 6 scoring iterations).
- Then based on VIF (variance inflation factor) and P value (with significance) predictors have been filtered and after 23 more iterations we could achieve our final model with almost all predictors being significant with lowest VIFs.



Final Logistic Regression Model

The final model contains the following variables which effect the attrition rate for company XYZ:

-> *Age*

-> *Job satisfaction*

-> *Number of Companies worked*

-> *Years with current Manager*

-> *Total working years*

-> *Over time*

-> *Years since last promoted*

-> *Business Travel*

-> *Environment Satisfaction*

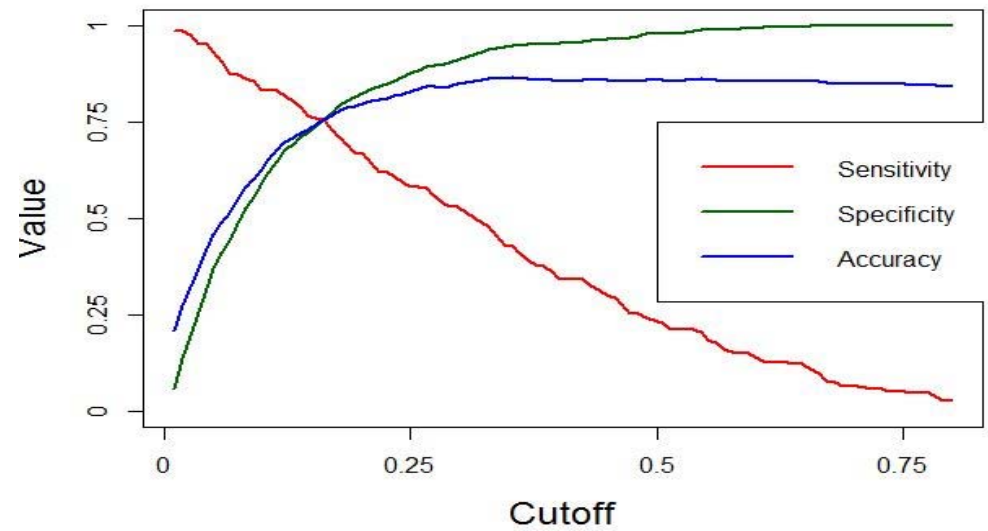
-> *Department*

-> *Marital Status*

Confusion Matrix and Statistics

The final model gives the following statistical values :

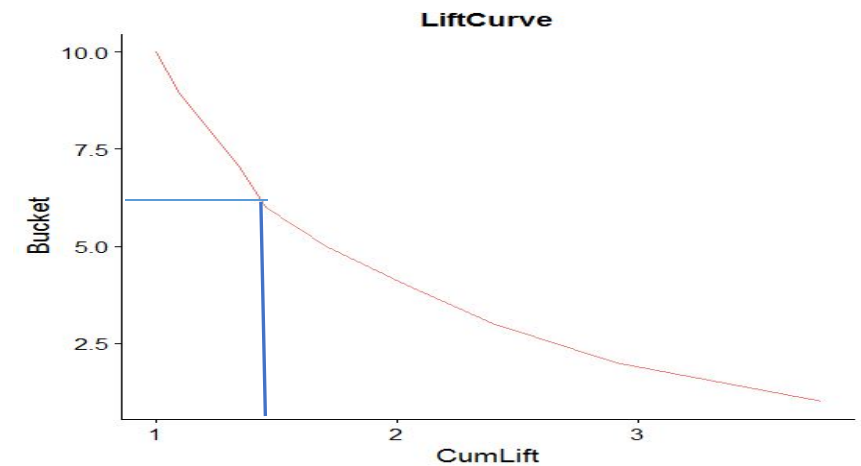
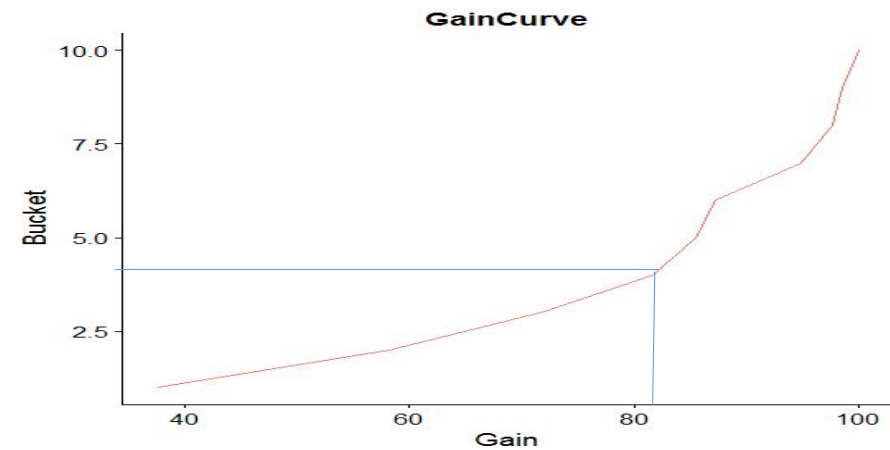
Prediction	Reference		
	No	Yes	Total
No	839	52	891
Yes	271	161	432
Total	1110	213	1323



The final model's confusion matrix shows that, it is 75% accurate, sensitivity and specificity is also 75% which shows this model is good for analysis.

Model Assessment (GAIN & LIFT charts)

bucket	total	totalresp	Cumresp	Gain	Cumlift
1	133	80	80	37.55869	3.755869
2	132	44	124	58.21596	2.910798
3	132	29	153	71.83099	2.394366
4	133	21	174	81.69014	2.042254
5	132	8	182	85.44601	1.70892
6	132	4	186	87.32394	1.455399
7	133	16	202	94.83568	1.354795
8	132	6	208	97.65258	1.220657
9	132	2	210	98.59155	1.095462
10	132	3	213	100	1





Model Assessment Summary

- The model has an increasing Gain and a decreasing Lift.
- The Model predicts more than 80% of the attritions within the 4th Decile with 75% accuracy.
- The KS statistic shows that the model is very good in distinguishing between employees who will leave the company and employees who won't.



Recommendations



Observation	Recommendation
Mostly young, middle aged single employees leave the company. It can be due to not having much family dependency.	They should be given more challenging roles and responsibilities in order to feel the need to stay.
Employees also leave due to working over time but very much satisfied with the environment and job	Incase of overtime, they should be compensated for it separately or given a week off on some other day. They could also be given incentives. Some fun activities could be planned once a month in order to boost the morale of the employees
Employees leave because of working under the same Manager for a long time and don't get to travel for Business abroad.	More opportunities should be created in terms of projects and assignments , so that the employees can work under different managers and also get to travel for business once in a while