HRA
Turnover Intention at BenzInfotech Assignment

Submitted by:

PALASH BAJPAI

B180759CS

Submitted to:

Dr. Sreejith SS

15-11-2021

TABLE OF CONTENTS

Turnover Intention at BenzInfotech

1. Case Analysis

1.1 Case Summary

BenzInfotech is a company that offers IT solutions for its clients. They've started to experience an increase in attrition rates from the past two quarters. BenzInfotech was not able to identify the reason behind this higher turnover rate. They offer benefits such as Employee Awards, Onsite Opportunities, Flexible Working hours, etc.

1.2 Decision Summary

The company wants to know the reason for the high attrition rates. For this reason, they have engaged an MBA intern from a third-party organization to conduct a fair survey among employees. The company wants to know factors that they can work on to decrease the turnover rate. It is proposed to develop a regression model to know factors affecting intentions to stay or leave the company.

1.3 Tools for Data Analysis

Jamovi (1.6.23 Solid version) is an open-source software used for the analysis.

1.4 Dataset

The survey has compiled the data of about 361 employees of BenzInfotech. The variables are shown in table below.

Table 1: Variables

SL . No	Variable Name	Variable Type	Variable Description
1.	Sl. No	ID	The number provided to each entry in serial order
2.	Gender	Nominal (Categorical)	Gender of the employee
3.	Boss Gender	Nominal (Categorical)	Gender of the Boss for each corresponding employee
4.	Tenure	Nominal	Years in organizations
5.	Account	Nominal	The regional account where they are associated at (APAC- Asia Pacific; MEA – Middle East and Africa; UKNA- United Kingdom and North America; LA – Latin America; EU – European Union)
6.	Level	Nominal	Designation Level
7.	Performance Rating	Nominal	Previous Year Performance Score (scale of 1-5; 5 being highest)
8.	Job Satisfaction	Nominal	The satisfaction level in their job (from another survey on a scale of 1-5; 5 being highest)
9.	Engagement	Nominal	The employee engagement level (from another survey on a scale of 1-5; 5 being highest)
10.	LeavesPreYear	Nominal	The number of leaves availed by employees during previous one year
11.	OnsiteOppurtunity	Nominal	Whether or not the employees have got an onsite opportunity
12.	Awards	Nominal	Whether or not the employee received any award from BenzInfotech

13.	Flexihours	Nominal	A benefit offered to employees – whether or not employees are permitted to work during flexible hours
14.	Intention	Nominal	The turnover intention (Yes indicates, the employee is thinking to move out)

Table 2: Descriptive Statistics

Descriptives

	Sl No	Gender	BossGender	Tenure	Account	Level	PerfRating
N	361	361	361	361	361	361	361
Missing	0	0	0	0	0	0	0
Mean	181			4.25			2.73
Median	181			4			3
Mode	1.00[a]			2.00			1.00
Standard deviation	104			2.35			1.40
Minimum	1			1			1
Maximum	361			8			5

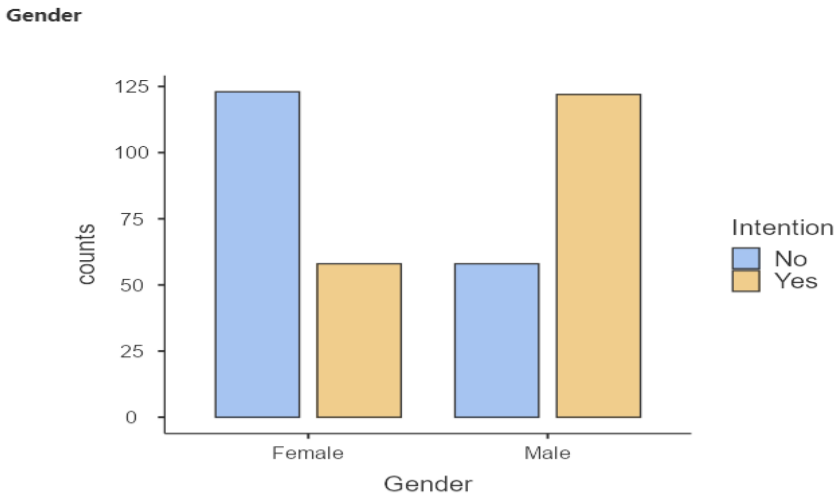
[a] More than one mode exists, only the first is reported

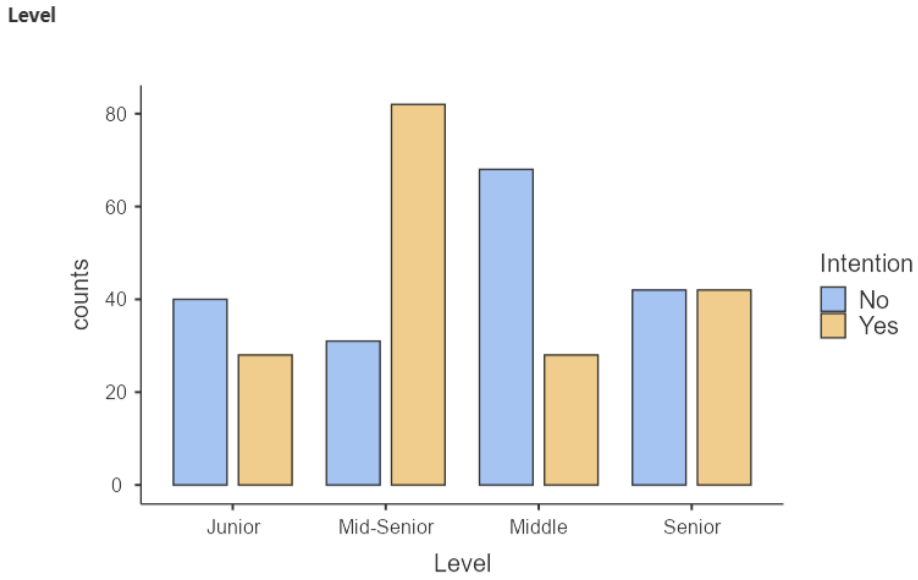
Descriptives

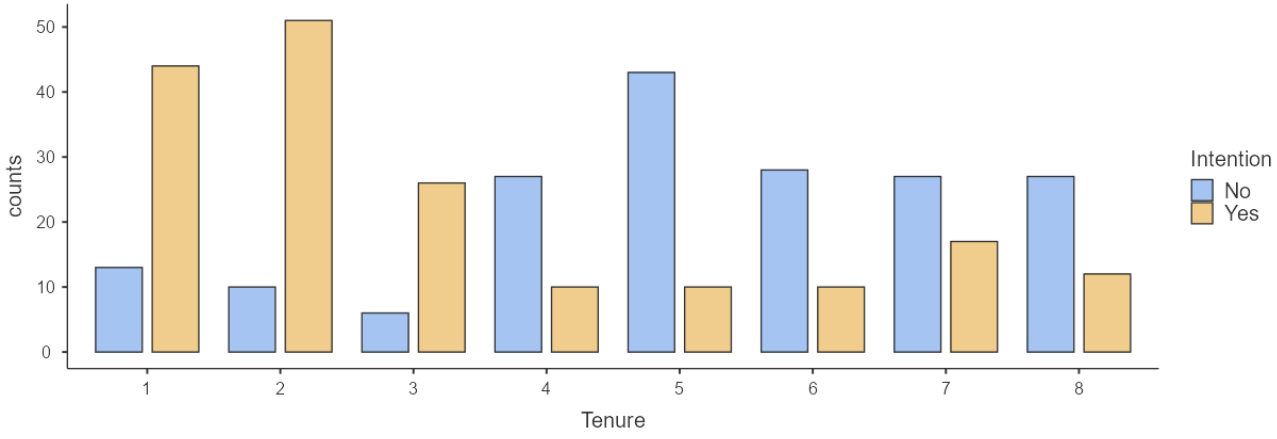
	JobSatisfaction	Engagement	LeavesPrevYear	OnsiteOpportunity	Awards	Flexihours	Intention
N	361	361	361	361	361	361	361
Missing	0	0	0	0	0	0	0
Mean	3.10	3.00	7.43				
Median	3	3	7				
Mode	4.00	4.00	7.00[a]				
Standard deviation	1.35	1.41	4.37				
Minimum	1	1	0				
Maximum	5	5	16				

[a] More than one mode exists, only the first is reported

Figure 1: Descriptive Plots







2 ANALYSIS

2.1 Binomial Logistic Regression Model

Log (yes/1 − yes) = a + b1 (x1) + ⋯ + bn(xn)

2.2 Generic/Full Model

Intention = a + b1 (Gender) + b2 (Boss Gender) + b3 (Tenure) + b4 (Account) + b5 (Level) + b6 (Performance Rating) + b7 (Job Satisfaction) + b8 (Engagement) + b9 (LeavesPerYear) + b10(Onsite Opportunity) + b11(Awards) + b12(Flexihours)

2.2.1 Binomial Logistic Regression: Model

Table 3: Generic Model Fit Measures

Results

Binomial Logistic Regression

Model Fit Measures

Model	Deviance	AIC	R^2_{McF}	R^2_{CS}	R^2_N
1	238	274	0.525	0.517	0.689

It can be observed from the given table that the AIC (Akaike information criterion) value is 274. A lower AIC value indicates a better fit. The pseudo-R^2 values, i.e., McFadden's R^2 is 0.525, Cox and Snell's R^2 is 0.517, Nagelkerke's R^2 is 0.689.

Table 4: Generic Model Coefficients

Model Coefficients - Intention

Predictor	Estimate	SE	Z	p
Intercept	4.2639	1.0913	3.9071	< .001
Tenure	−0.4887	0.1000	−4.8854	< .001
PerfRating	−0.5209	0.1323	−3.9380	< .001
JobSatisfaction	−0.5247	0.1336	−3.9269	< .001
Engagement	−0.3962	0.1204	−3.2910	< .001
LeavesPrevYear	0.1715	0.0400	4.2857	< .001
Gender:				
Male – Female	1.5492	0.3636	4.2611	< .001
BossGender:				
Male – Female	−1.7592	0.4237	−4.1520	< .001
Account:				
EU – APAC	0.0486	0.5729	0.0848	0.932
LA – APAC	0.4378	0.5577	0.7850	0.432
MEA – APAC	−0.1343	0.5649	−0.2378	0.812
UKNA – APAC	0.7062	0.5873	1.2026	0.229
Level:				
Mid-Senior – Junior	1.7332	0.5834	2.9709	0.003
Middle – Junior	−0.2731	0.5246	−0.5206	0.603
Senior – Junior	1.3924	0.6314	2.2053	0.027
OnsiteOpportunity:				
Yes – No	0.0541	0.3529	0.1533	0.878
Awards:				
Yes – No	0.4013	0.3408	1.1776	0.239
Flexihours:				
Yes – No	0.2264	0.3377	0.6703	0.503

Note. Estimates represent the log odds of "Intention = Yes" vs. "Intention = No"

It can be observed from the above table that the p value at 95% confidence level is not significant for 4 variables namely Account, Awards, OnsiteOpportunity, Flexihours. The p values for other variables are significant.

Interpretation of estimate values for the significant variables is as follows,

• If the sign of estimate value for a variable is positive, it indicates that more the value of variable more is chance of having the intention to leave. In brief, these variables affect intention positively.

• If the sign of estimate value for a variable is negative, it indicates that the more the value of the variable lesser is chance of having the intention to leave. In brief, these variables affect intention negatively.

2.3 Reduced / Working Model

Intention = a + b1 (Gender) + b2 (Boss Gender) + b3 (Tenure) + b4 (Performance Rating) + b5 (Job Satisfaction) + b6 (Engagement) + b7 (LeavesPerYear) +b8 (Level)

Intention = 4.661 + 1.673 (Gender) -1.816 (Boss Gender) – 0.476 (Tenure) -0.509 (Performance Rating) – 0.503(Job Satisfaction) - 0.392(Engagement) + 0.170 (LeavesPerYear) + 1.653(Level)

Model Specific Results Model 1

Model Coefficients - Intention

Predictor	Estimate	SE	Z	p
Intercept	4.661	0.9410	4.953	< .001
Tenure	−0.476	0.0967	−4.920	< .001
PerfRating	−0.509	0.1307	−3.899	< .001
JobSatisfaction	−0.503	0.1259	−3.995	< .001
Engagement	−0.392	0.1150	−3.411	< .001
LeavesPrevYear	0.170	0.0392	4.335	< .001
Gender:				
Male – Female	1.673	0.3525	4.746	< .001
BossGender:				
Male – Female	−1.816	0.4146	−4.381	< .001
Level:				
Mid-Senior – Junior	1.653	0.5648	2.926	0.003
Middle – Junior	−0.263	0.5118	−0.514	0.608
Senior – Junior	1.410	0.6191	2.277	0.023

Note. Estimates represent the log odds of "Intention = Yes" vs. "Intention = No"

2.4 Model Comparison

Table 5: Model Fit Measure (Model 1 and 2)

Model Fit Measures

Model	Deviance	AIC	R^2_{McF}	R^2_{CS}	R^2_N
1	242	264	0.516	0.511	0.681
2	238	274	0.525	0.517	0.689

Model Comparisons

Comparison			χ^2	df	p
Model		Model			
1	-	2	4.40	7	0.733

In Table 5, Model 1 represents the reduced/working model and Model 2 represents the generic/full model. It can be observed that the AIC (Akaike information criterion) value is 264 for Model 1. The AIC value obtained for the reduced model is lower than that of the generic model. A lower AIC value indicates a better fit. The pseudo-R^2 values. McFadden's R^2 is 0.516, Cox and Snell's R^2 is 0.511, Nagelkerke's R^2 is 0.681. The Nagelkerke's R^2 value for Model 1 is only slightly less than the value of Model 2, hence, the reduced model with 7 variables holds good for prediction.

From the Model Comparisons table, we can see that p-value indicates that the difference between the 2 models is not significant. Therefore model 1 with 7 variables is good to predict intentions.

2.5 Model Specific Results

2.5.1 Model 1 (Reduced / Working Model)

Intention = 4.661 + 1.673 (Gender) -1.816 (Boss Gender) – 0.476 (Tenure) -0.509 (Performance Rating) – 0.503(Job Satisfaction) - 0.392(Engagement) + 0.170 (LeavesPerYear) + 1.653(Level)

1. Model Cofficients

Model Coefficients – Intention

Predictor	Estimate	SE	Z	P
Intercept	4.661	0.9410	4.953	< .001
Tenure	-0.476	0.0967	-4.920	< .001
PerfRating	-0.509	0.1307	-3.899	< .001
JobSatisfaction	-0.503	0.1259	-3.995	< .001
Engagement	-0.392	0.1150	-3.411	< .001
LeavesPrevYear	0.170	0.0392	4.335	< .001
Gender:				
Male – Female	1.673	0.3525	4.746	< .001
BossGender:				
Male – Female	-1.816	0.4146	-4.381	< .001
Level:				
Mid-Senior – Junior	1.653	0.5648	2.926	0.003
Middle – Junior	-0.263	0.5118	-0.514	0.608
Senior – Junior	1.410	0.6191	2.277	0.023

Note. Estimates represent the log odds of "Intention = Yes" vs. "Intention = No"

2. Assumption Check

We see Tenure is having more VIF (Variation Inflation Factor) value than other variables hence it affects intention more in comparison of others. For ideal case tolerance should be less than 1 and VIF should be less than 3.

Collinearity Statistics

	VIF	Tolerance
Tenure	1.30	0.769
PerfRating	1.07	0.936
JobSatisfaction	1.02	0.979
Engagement	1.02	0.983
LeavesPrevYear	1.02	0.979
Gender	1.10	0.913
BossGender	1.03	0.971
Level	1.10	0.911

3. Prediction

Classification Table – Intention

	Predicted		
Observed	**No**	**Yes**	**% Correct**
No	155	26	85.6
Yes	29	151	83.9

Note. The cut-off value is set to 0.5

The observed or actual intention has been shown in above table using the working model. The correctness percentage of no intention is 85.6%, this is the specificity. The correctness percentage of intention yes is 83.9%, this is the sensitivity.

Predictive Measures

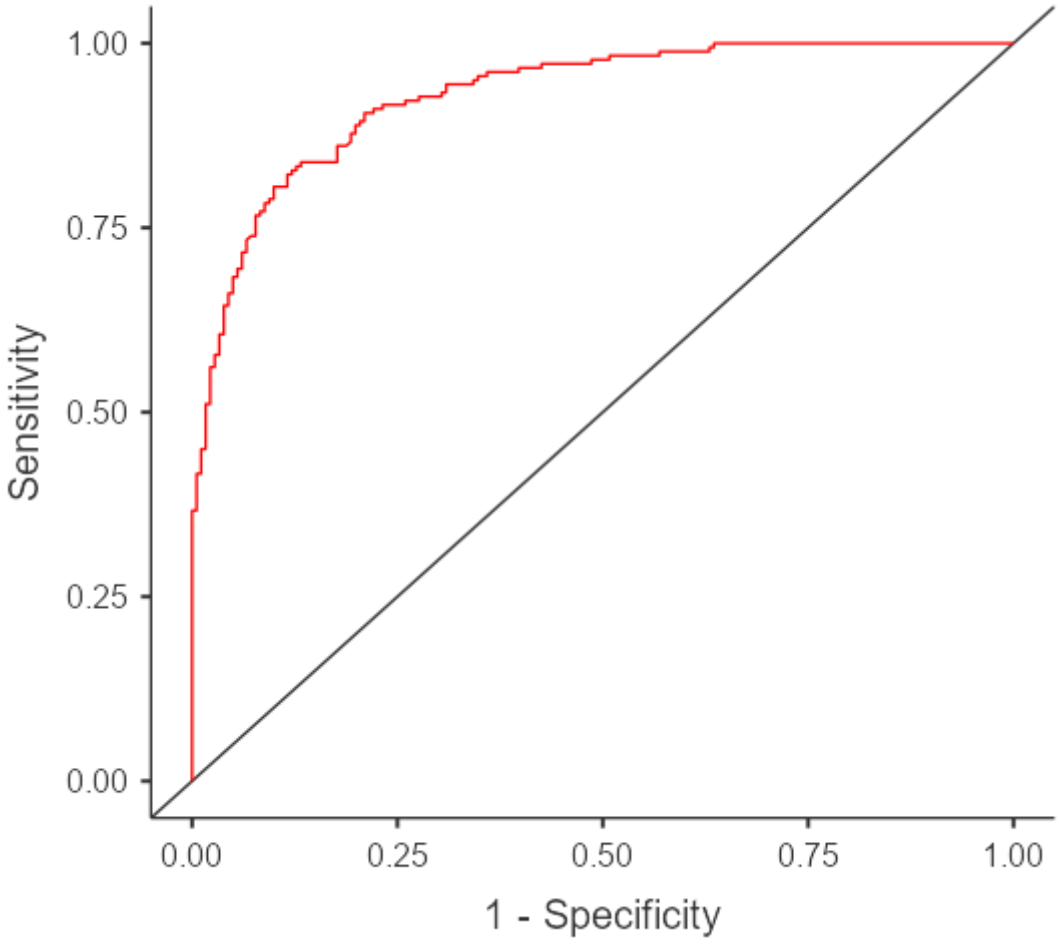
Accuracy	Specificity	Sensitivity	AUC
0.848	0.856	0.839	0.930

Note. The cut-off value is set to 0.5

The Accuracy of the working model is 0.848. Thus our model with 8 variables can predict with 84.8% accuracy whether the employee has the intention to leave the company or not. Specificity measures the proportion of true negatives, i.e., 0.856. Sensitivity measures the proportion of true positives that are correctly identified, i.e., 0.839. The AUC (Area Under the Curve) value is 0. 930.The higher the AUC i.e., close to 1, the better the performance of the model.

The ROC (Receiver Operating Characteristic) curve shows the relationship between sensitivity and 1-specificity for possible cut-off.

ROC Curve



2.5.2 Model 1 (Generic / Full Model)

1. Model Cofficients

Model Coefficients – Intention

Predictor	Estimate	SE	Z	p
Intercept	4.2639	1.0913	3.9071	< .001
Tenure	-0.4887	0.1000	-4.8854	< .001
PerfRating	-0.5209	0.1323	-3.9380	< .001
JobSatisfaction	-0.5247	0.1336	-3.9269	< .001
Engagement	-0.3962	0.1204	-3.2910	< .001
LeavesPrevYear	0.1715	0.0400	4.2857	< .001
Gender:				
Male – Female	1.5492	0.3636	4.2611	< .001
BossGender:				
Male – Female	-1.7592	0.4237	-4.1520	< .001
Level:				
Mid-Senior – Junior	1.7332	0.5834	2.9709	0.003
Middle – Junior	-0.2731	0.5246	-0.5206	0.603
Senior – Junior	1.3924	0.6314	2.2053	0.027
Account:				
EU – APAC	0.0486	0.5729	0.0848	0.932
LA – APAC	0.4378	0.5577	0.7850	0.432
MEA – APAC	-0.1343	0.5649	-0.2378	0.812
UKNA – APAC	0.7062	0.5873	1.2026	0.229
OnsiteOpportunity:				
Yes – No	0.0541	0.3529	0.1533	0.878
Awards:				
Yes – No	0.4013	0.3408	1.1776	0.239
Flexihours:				
Yes – No	0.2264	0.3377	0.6703	0.503

Note. Estimates represent the log odds of "Intention = Yes" vs. "Intention = No"

2. Assumption Check

We see Tenure is having more VIF (Variation Inflation Factor) value than other variables hence it affects intention more in comparison of others.

For ideal case tolerance should be less than 1 and VIF should be less than 3.

Collinearity Statistics

	VIF	Tolerance
Tenure	1.34	0.747
PerfRating	1.07	0.932
JobSatisfaction	1.07	0.935
Engagement	1.05	0.954
LeavesPrevYear	1.03	0.971
Gender	1.12	0.894
BossGender	1.04	0.957
Level	1.13	0.885
Account	1.05	0.951
OnsiteOpportunity	1.09	0.921
Awards	1.05	0.954
Flexihours	1.04	0.963

3. Prediction

Classification Table – Intention

	Predicted		
Observed	No	Yes	% Correct
No	158	23	87.3
Yes	29	151	83.9

Note. The cut-off value is set to 0.5

The observed or actual intention has been shown in above table using the working model. The correctness percentage of no intention is 85.6%, this is the specificity. The correctness percentage of intention yes is 83.9%, this is the sensitivity.

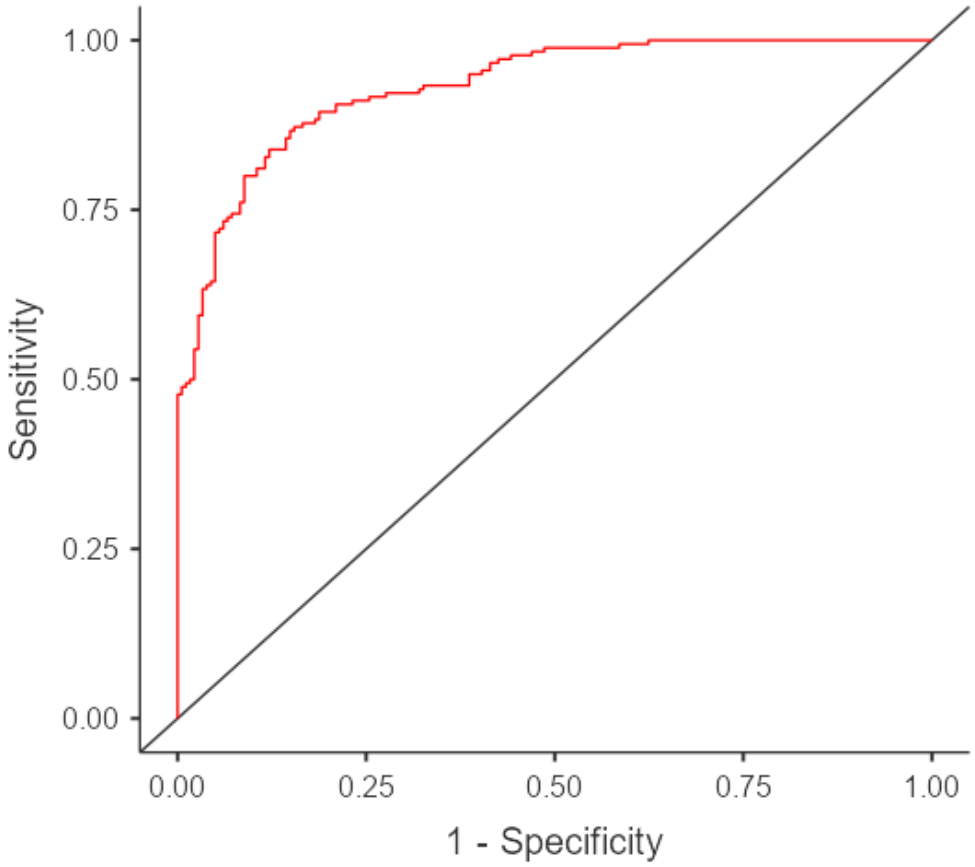
Predictive Measures

Accuracy	Specificity	Sensitivity	AUC
0.856	0.873	0.839	0.932

Note. The cut-off value is set to 0.5

The Accuracy of the generic model is 0.856. Thus our model with 12 variables can predict with 85.6% accuracy whether the employee has the intention to leave the company or not. Specificity measures the proportion of true negatives, i.e., 0.873. Sensitivity measures the proportion of true positives that are correctly identified, i.e., 0.839. The AUC (Area Under the Curve) value is 0. 932.The higher the AUC i.e., close to 1, the better the performance of the model.

ROC Curve



3. Conclusions

We have developed our model with 8 variables that affect the intention most. For this prediction, we applied Binomial Logistic Regression on the dataset collected through a survey on Employees of BenzInfotech.

Intention = 4.661 + 1.673 (Gender) -1.816 (Boss Gender) – 0.476 (Tenure) -0.509 (Performance Rating) – 0.503(Job Satisfaction) - 0.392(Engagement) + 0.170 (LeavesPerYear) + 1.653(Level)

Observations :-

- Male employees tend to have more turnover intention than females.
- Employees with Female bosses tend to have lesser turnover intentions.
- Employees with more leaves per year tend to have more turnover intentions.
- Employees with lesser tenure, less performance rating, less job satisfaction, less engagement tend to have more turnover intentions.

Strategies to reduce turnover intentions:-

- The company should include some more policies to help employees to perform better in their work. More performance rating lesser will be turnover intentions.
- Should train male bosses to learn how female bosses behave with employees working under them.
- Bring changes in working culture to improve job satisfaction among employees .