

Predicting Employee Attrition



What is attrition?

Attrition is a terminology in human resources which refers to the employees leaving the company.



Why do we predict employee attrition?

High attrition is problematic for companies that causes the huge loss. The loss is not only in terms of the money but also the company sometimes loses the skilled employees who are the most valuable assets to the company (Morrison, 2014).

IBM HR Analytics Employee Attrition Dataset

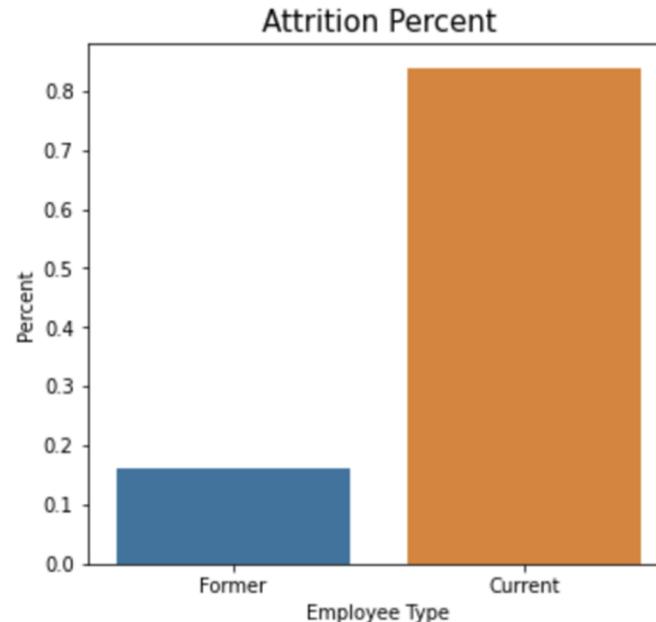
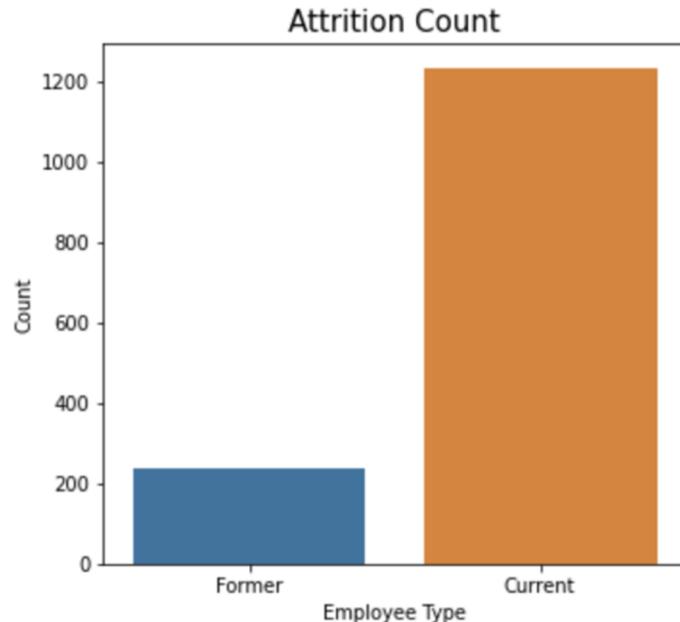
	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField
0	41	Yes	Travel_Rarely	1102	Sales	1	2	Life Sciences
1	49	No	Travel_Frequently	279	Research & Development	8	1	Life Sciences
2	37	Yes	Travel_Rarely	1373	Research & Development	2	2	Other
3	33	No	Travel_Frequently	1392	Research & Development	3	4	Life Sciences
4	27	No	Travel_Rarely	591	Research & Development	2	1	Medical

RelationshipSatisfaction	StandardHours	StockOptionLevel	TotalWorkingYears	TrainingTimesLastYear	WorkLifeBalance	YearsAtCompany	YearsInCurrentRole	YearsSinceLastPromotion
1	80	0	8	0	1	6	4	0
4	80	1	10	3	3	10	7	1
2	80	0	7	3	3	0	0	0
3	80	0	8	3	3	8	7	3
4	80	1	6	3	3	2	2	2



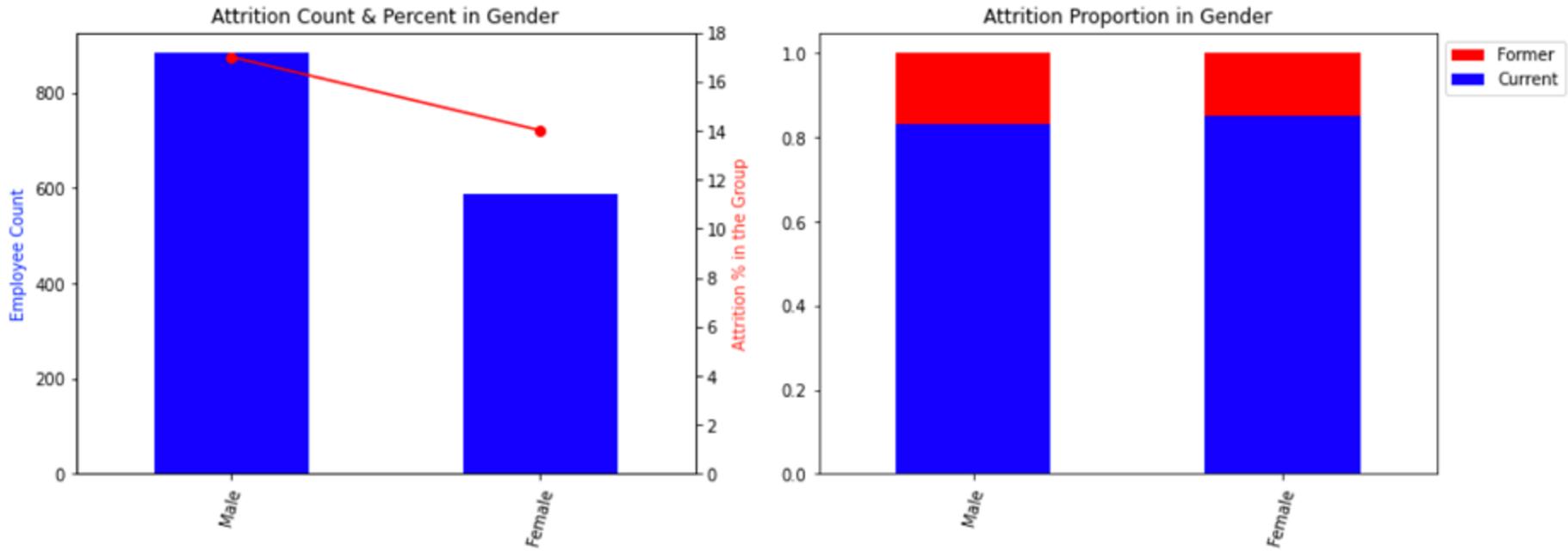
We have imbalance data.

1233 employees stayed, 237 employees left.





17% of Male and 14% of Female Leave the Company.





Gender Equal Opportunity

Null Hypothesis:

H_0 : There is no difference in the number of males and females in the company.

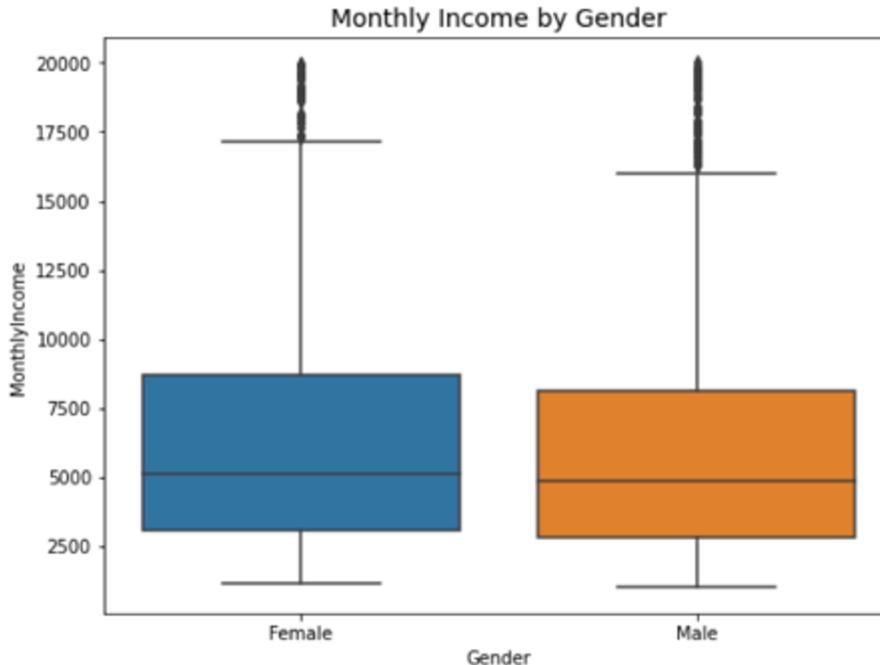
Alternative Hypothesis:

H_A : There is difference in the number of males and females in the company.

Category	Observed	Expected	Residual = Obsereved - Expected	$(\text{Residual})^2$	$(\text{Residual})^2/\text{Expected}$	Chi Square Score	Degree of Freedom	P Value
Female	588	735	-147	21609	29.4	58.8	1	< 0.05
Male	882	735	147	21609	29.4			

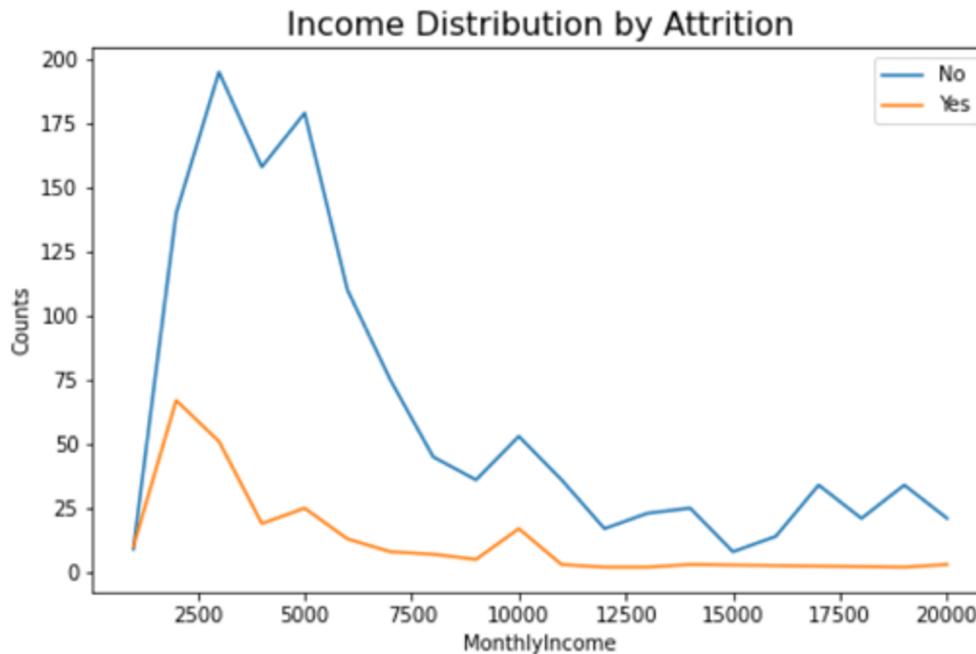


There is no gender disparity in income.



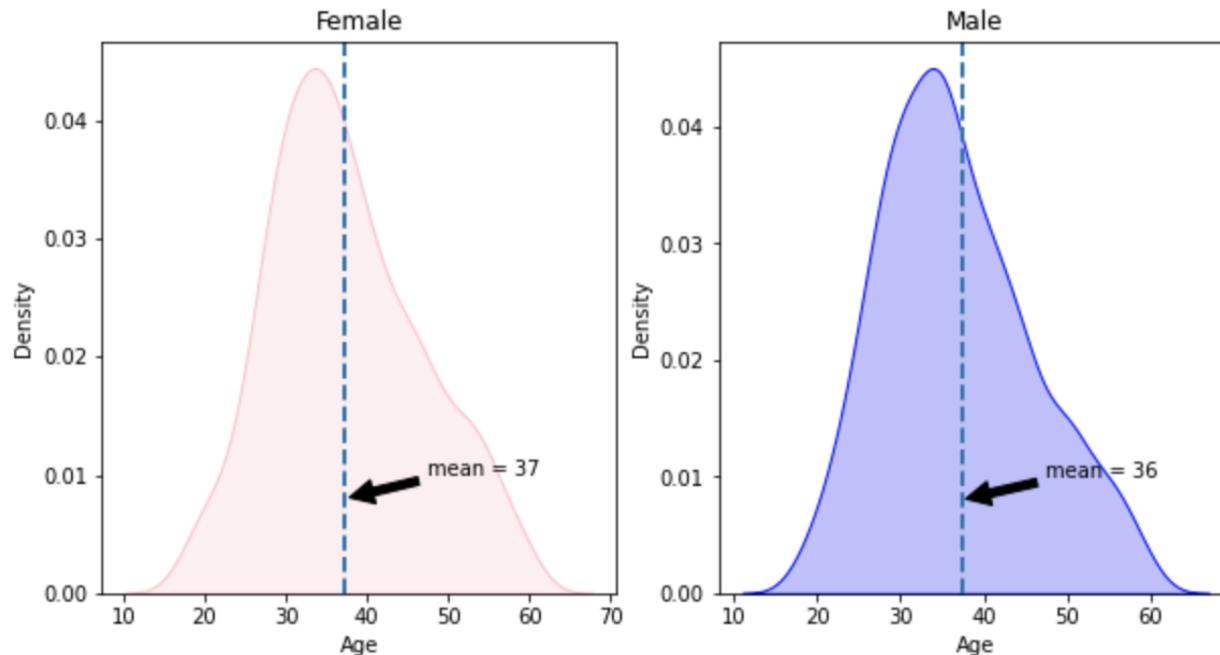


Attrition rate is high at low income level.



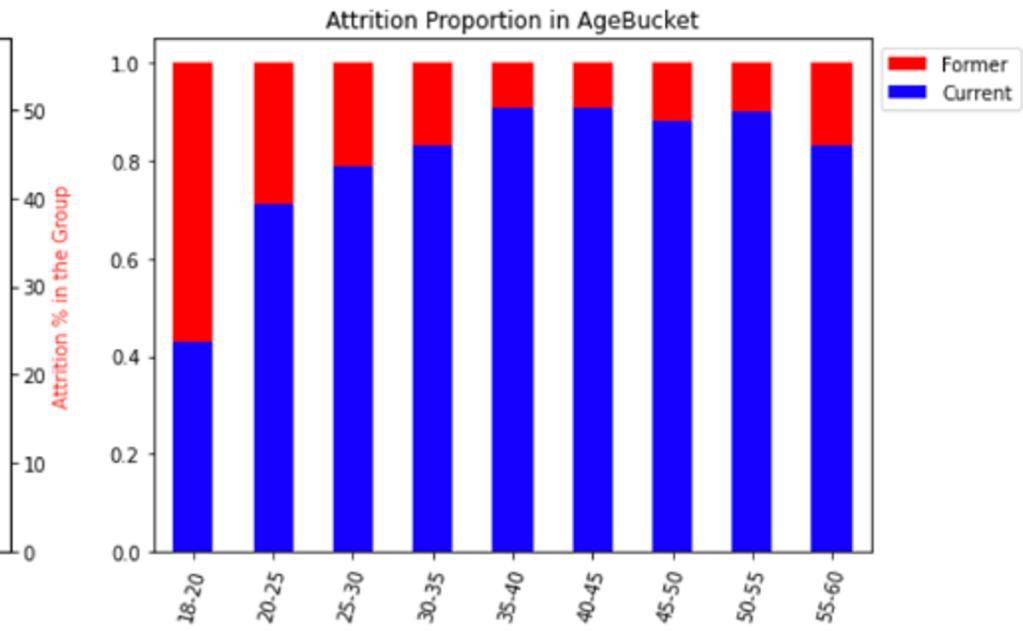
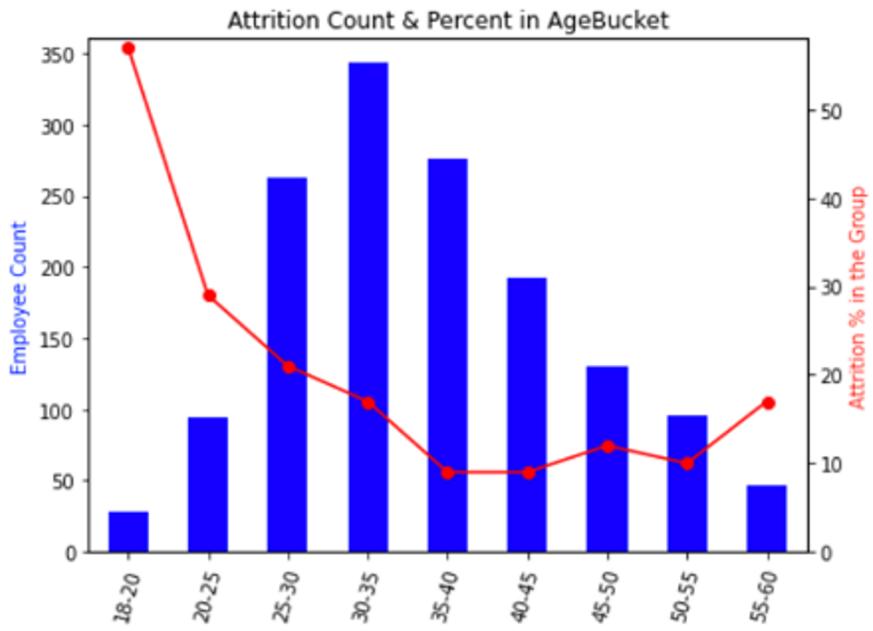


Average age between genders



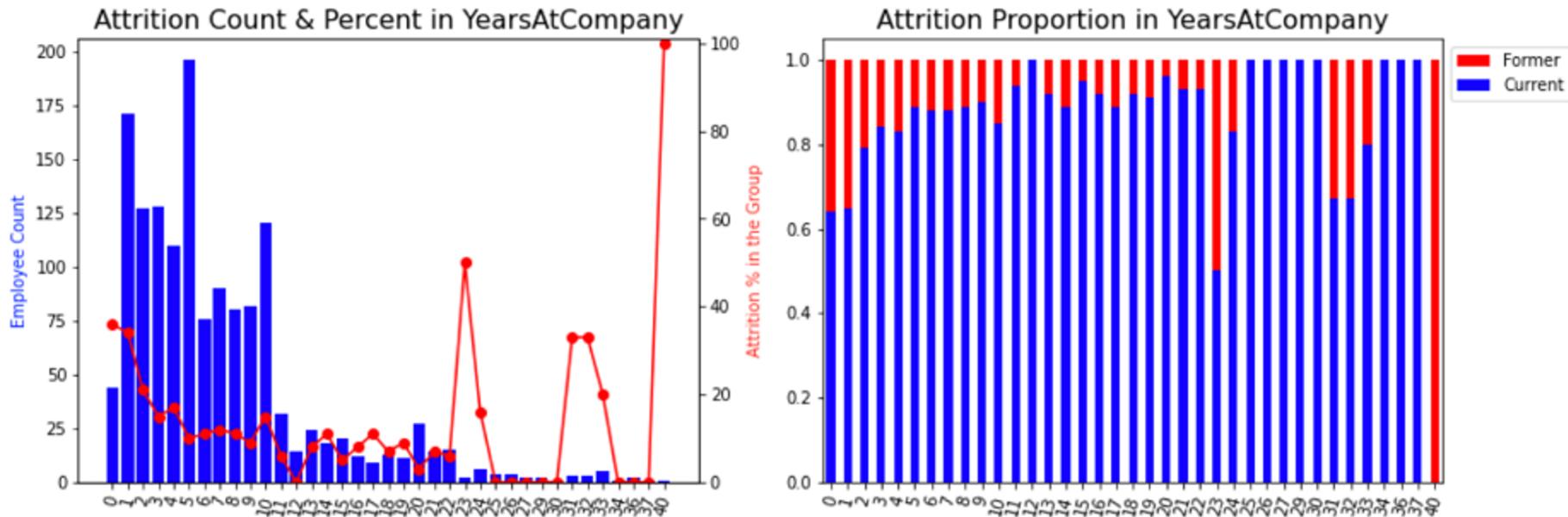


Age



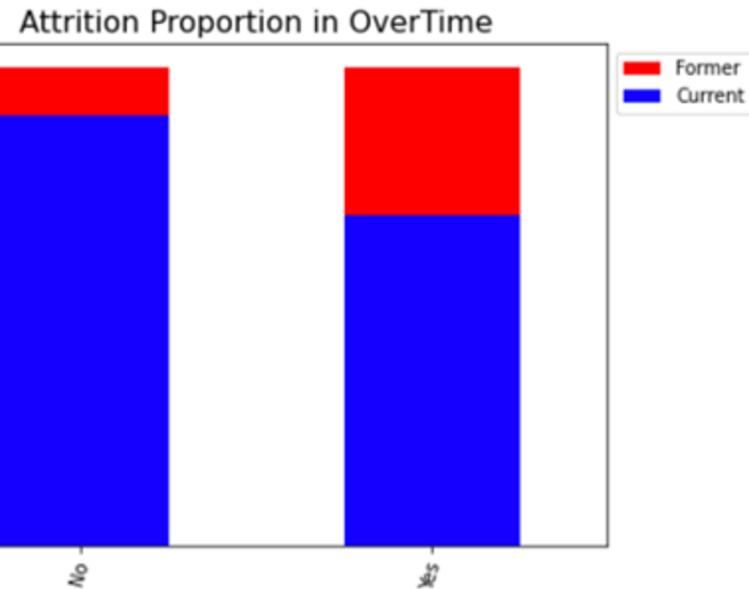
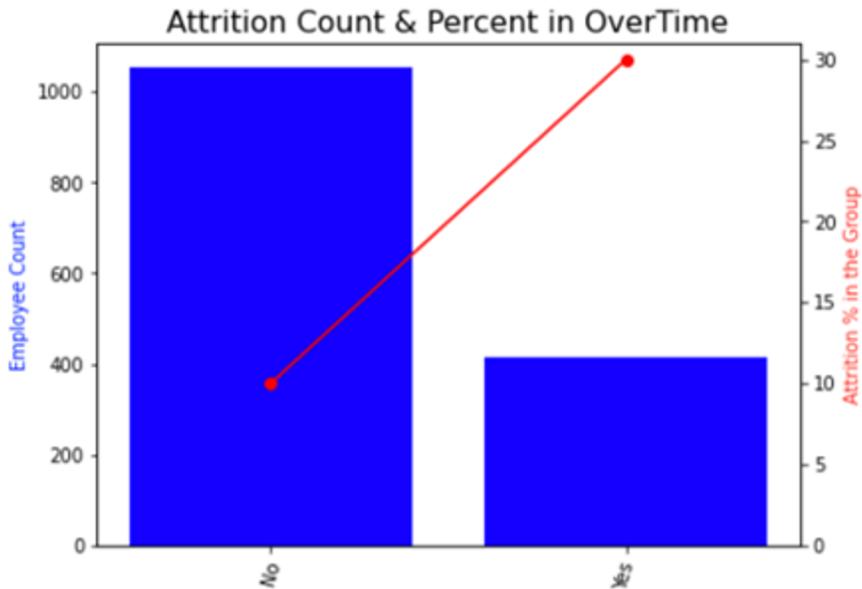


Years at company.





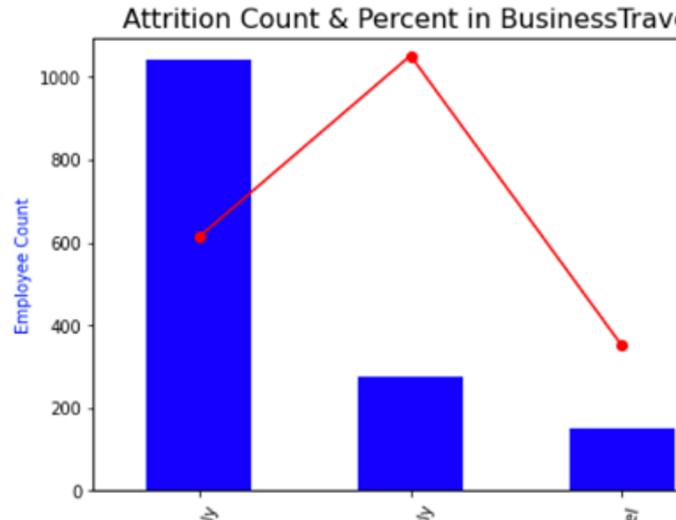
Over Time



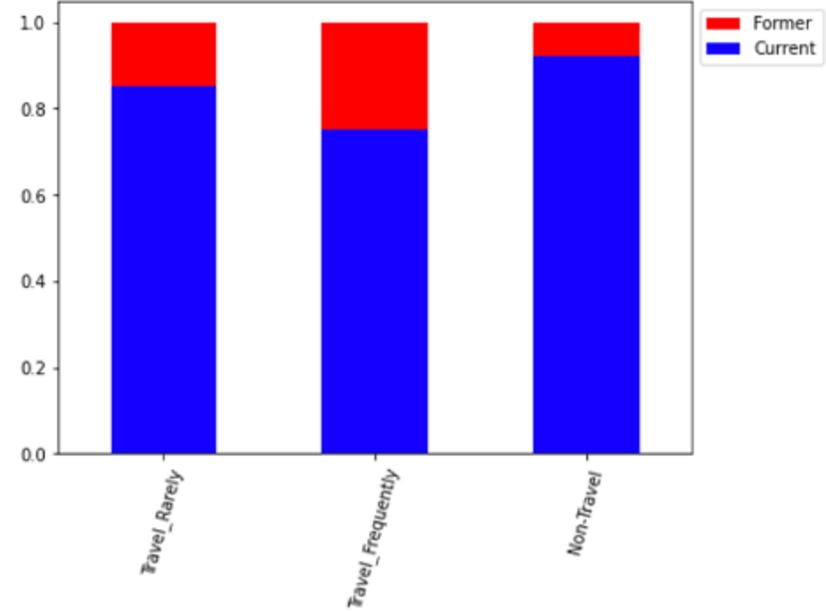
Former
Current



Business Travel



Attrition Proportion in BusinessTravel

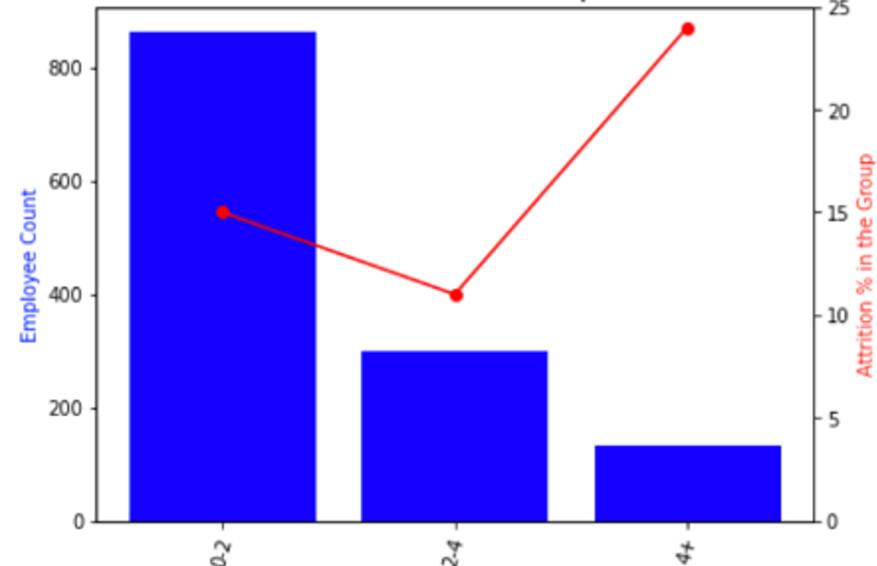


Former
Current

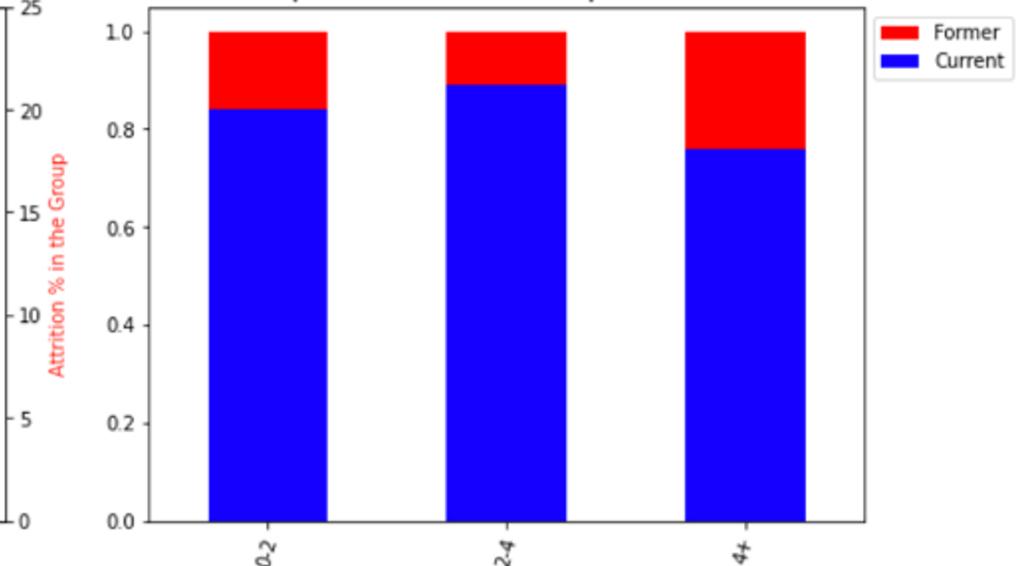


Number of companies worked before.

Attrition Count & Percent in NumCompaniesWorkedBucket



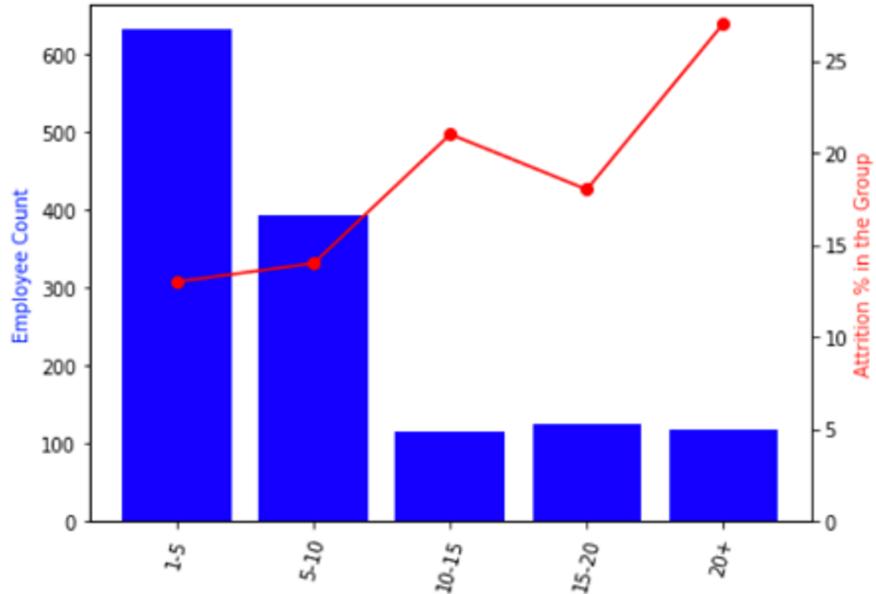
Attrition Proportion in NumCompaniesWorkedBucket



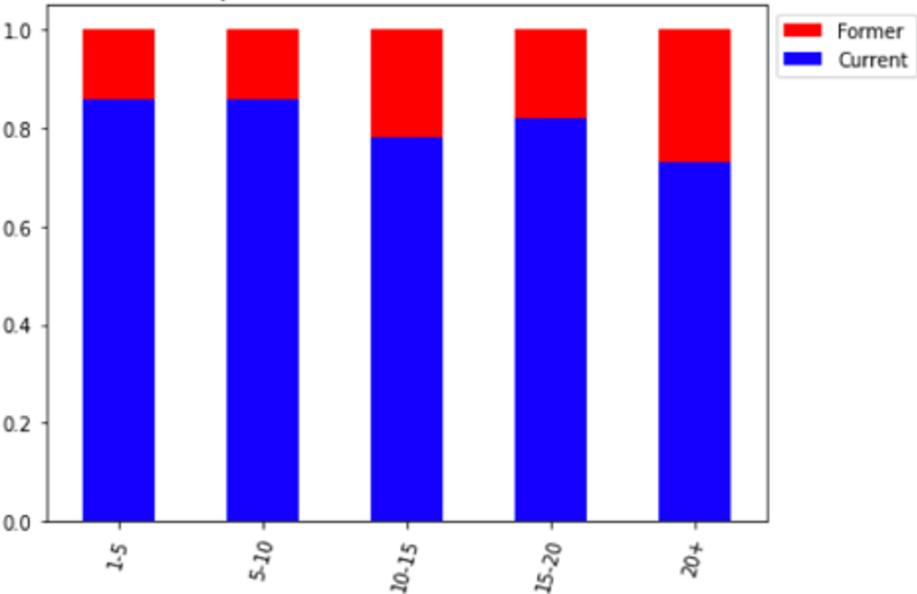


Distance from home

Attrition Count & Percent in DistanceFromHomeBucket

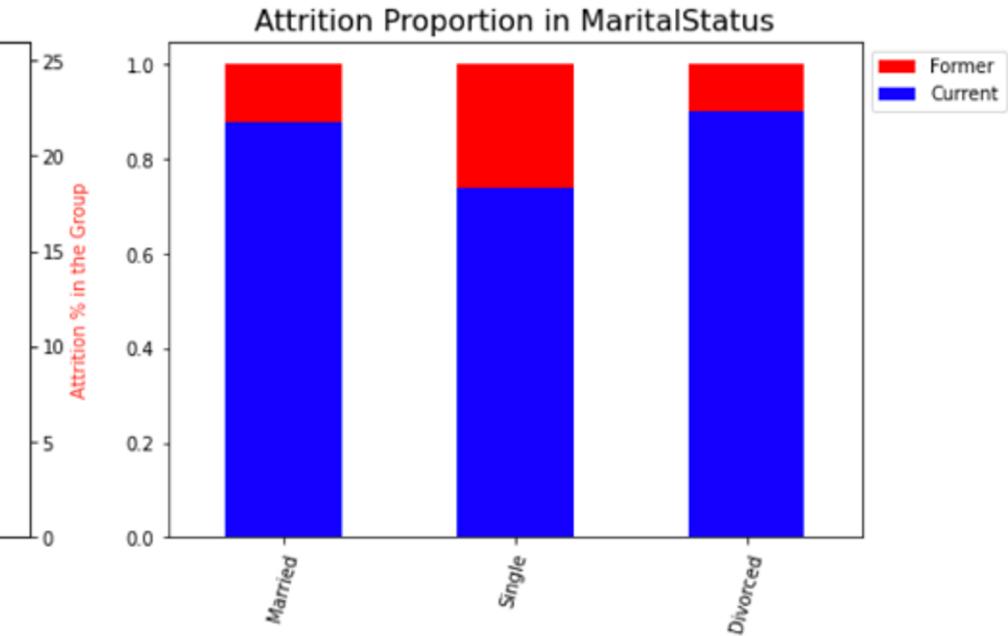
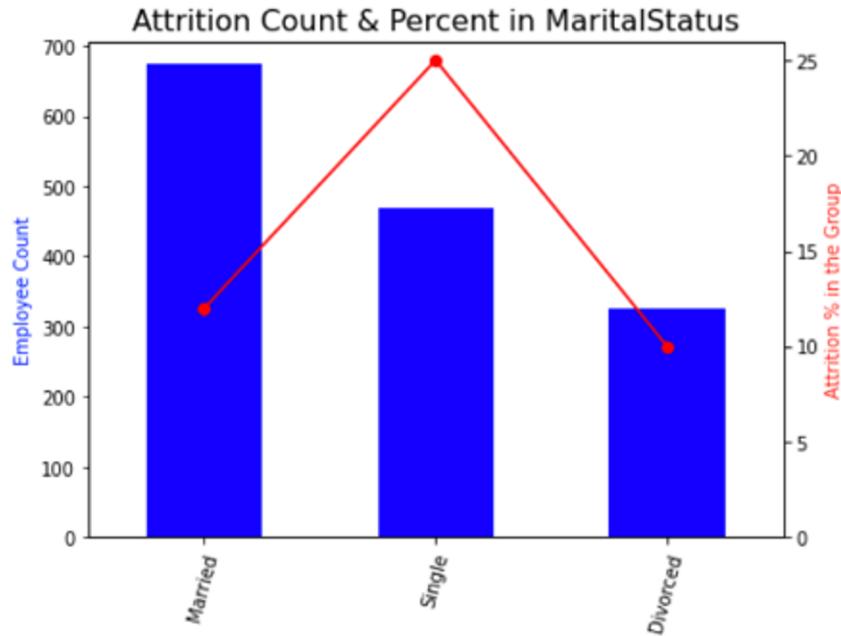


Attrition Proportion in DistanceFromHomeBucket





Single people tend to leave.





Applies Machine Learning ALgorithms

- Logistic Regression
- Random Forest Classifier
- Kth Nearest Neighbor (KNN)
- Support Vector Machine
- Gradient Boosting
- ADA Boosting

**Right evaluation metric is f1-score
for this business problem.**

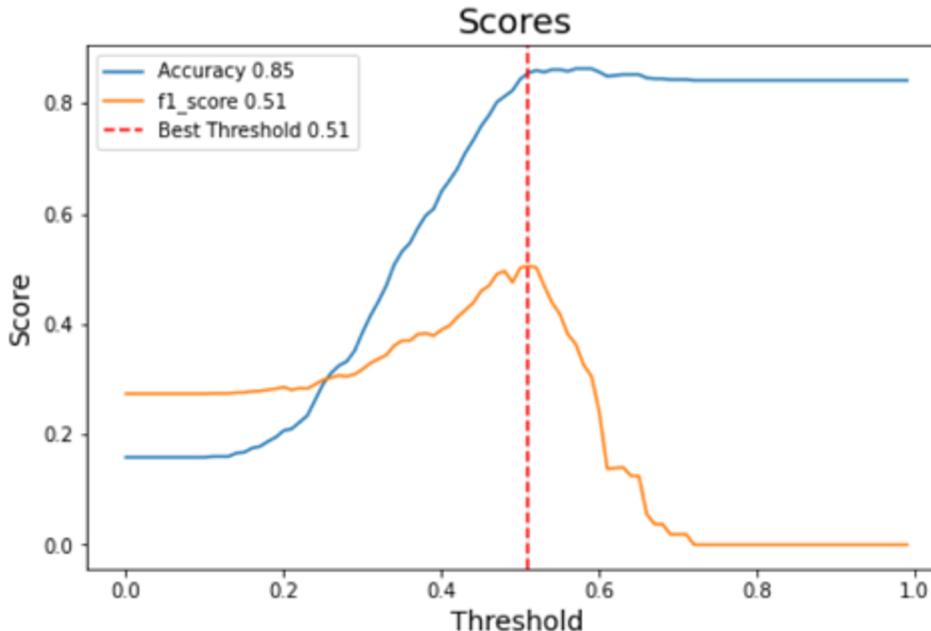


Logistic regression gives higher F1 score

	default_model	grid_search_model	moved_threshold
Logistic Regression	0.4694	0.5024	0.5052
Random Forest	0.3014	0.4096	0.4889
KNN	0.2818	0.3305	0.3580
SVM	0.3380	0.3920	0.4107
Gradient Boosting	0.4118	0.4118	0.4608
ADA Boosting	0.3700	0.4615	0.4615

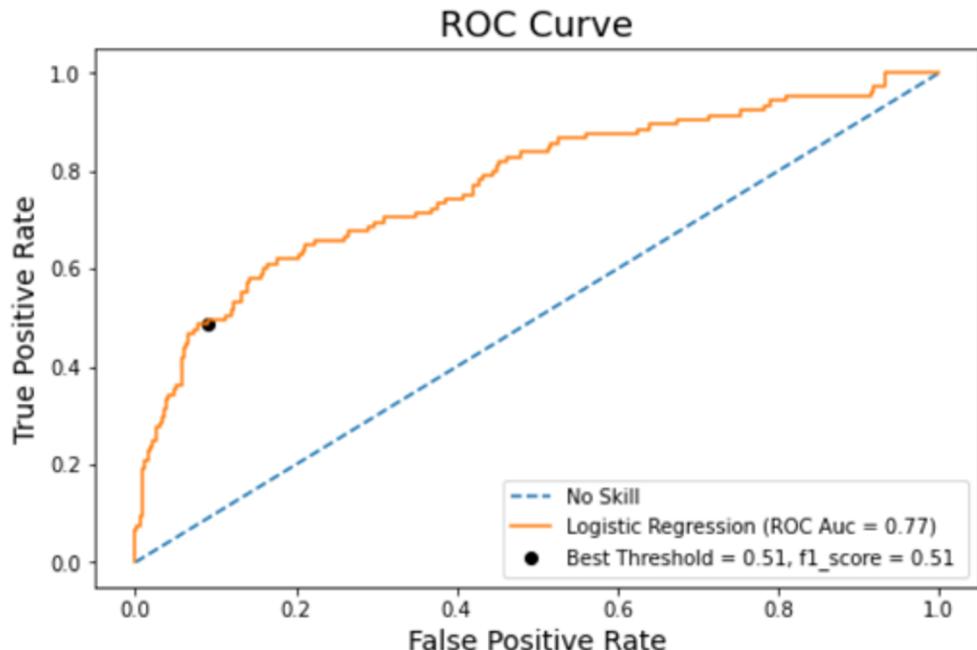


F1 score is high when threshold is 0.51.



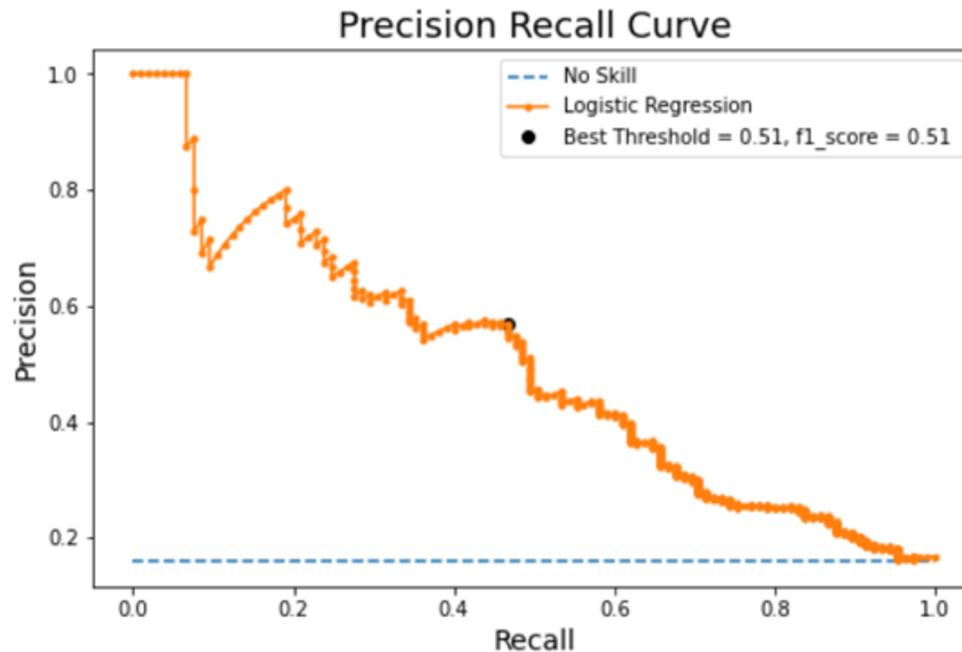


Area under curve is 0.77.



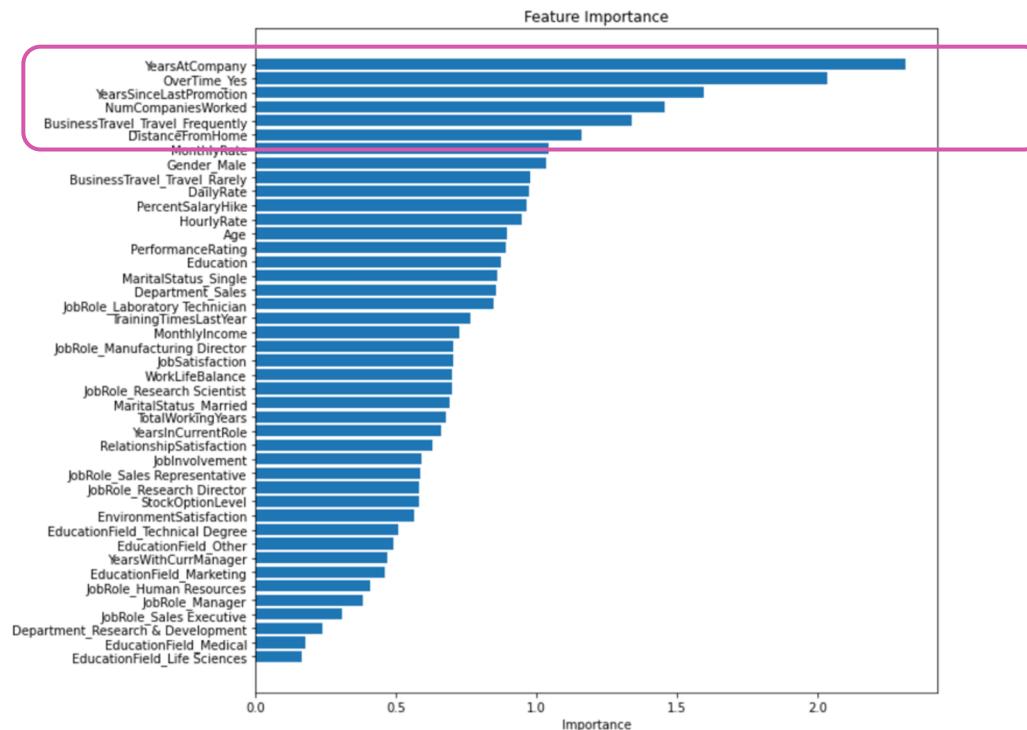
The Machine Learning Predictive Model has been successful in effectively classifying 77% unknown (Validation Set) examples correctly.

Precision Recall curve is pretty different than no skill line.





Feature Importance





Top 6 Important Features

- Years at Company
- Overtime
- Years Since Last Promotion
- Number of Companies Worked
- Business Travel
- Distance from Home

Recommendations



Years at Company

23% of employees have two years or less working experience in the company and 30% of those employees leave the company. The model shows that number of years at company is the strongest factor in attrition. Employees who are in their first years are more likely to leave the company. Employees who have gained working experience prefer to stay in the company. Therefore, the company should understand why their new employees leave the company.



Overtime

28% of employees works overtime in the company and 30% of those employees leave the company. The model also shows that overtime work is the second strongest factor in attrition. Therefore, the company should understand the reason why they are working overtime. Is it for too high workload or are employees' qualifications not enough to complete the scheduled tasks on time? There might be some other reasons behind that. Our recommendation will be to understand the reason(s) for overtime with detail research and take appropriate measures to reduce the factors behind this attrition factor.



Years since last promotion

There are 1097 employees who have less than 2 years since last promotion and 16% of those employees leave the company. Attrition rate increases if the number of years increases since the last promotion. Because of this reason, the company should review the promotion policy. They can make promotion requirements clear to all employees how and when they can be promoted.



Number of companies worked before

9% of employees worked in 4 or more companies before this company and 24% of those employees leave the company. The company should understand why those employees are leaving the company. The company can question the applicants why they want to quit the previous job during interview process. The company can even get in touch with previous company.



Business travel

19% of the employees travels frequently and 70% of those employees leave the company. This is one of the strongest factors in employee attrition. The company should understand if a high number of business travels are really necessary. They can adjust the frequency. If they can't, they can give some extra incentives to motivate those employees.



Distance from home

24% of the employees live more than 10 miles away from the company and 22% of those employees leave the company. The company should understand how distance is affecting their work. They can support their employees to move closer areas to the company by providing moving expenses. If they can't move closer due to their mandatory things, the company can adjust their work shift to prevent employees wasting their time from rush hours. The company can enable some days to working from home if they can.

Thank you!