Model Metrics
Customer Attrition Prediction using ML Algo

Capstone Project
Data Science Career Track, Springboard

Thanks to mentor Julian Jenkins III

Evaluation Metrics

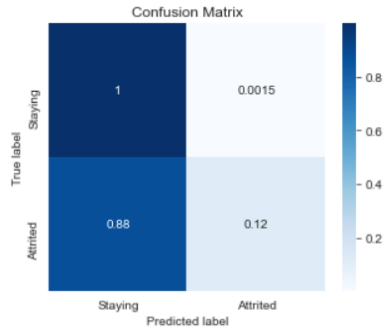
Model used are

1. Logistic Regression
2. Random Forest Classification
3. KNeighbours Classification
4. XGBoost Classification

Logistic Regression: Metrics Evaluation

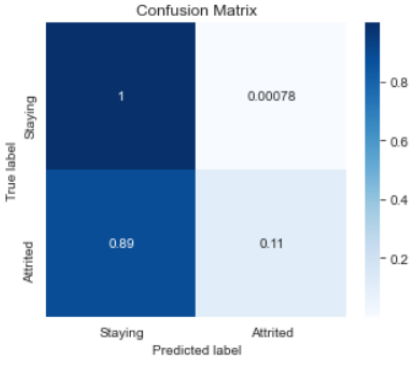
LG classification report on train set

	precision	recall	f1-score	support
Staying	0.86	1.00	0.92	5949
Attrited	0.94	0.12	0.21	1139
accuracy			0.86	7088
macro avg	0.90	0.56	0.57	7088
weighted avg	0.87	0.86	0.81	7088



LG classification report on test set

	precision	recall	f1-score	support
Staying	0.85	1.00	0.92	2551
Attrited	0.96	0.11	0.20	488
accuracy			0.86	3039
macro avg	0.91	0.56	0.56	3039
weighted avg	0.87	0.86	0.81	3039

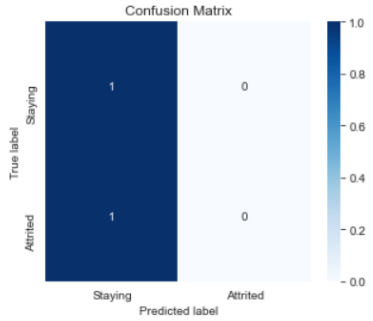


This model is generalized better on train set and test set.However roc_AUC score 0.75, means it can 75% chance of identifying default and non-default class.Recall perfect for defaults, precision is 0.85. only. Indicating this model can not predict the defaults correctly, but non-deaults are predicted as defaults. Meaning those who are actually attrited are predicted as existing. However it is better than naive model.Can identity majority class much better, but still not to the expectation.

Random Forest Classification: Metric Evaluation

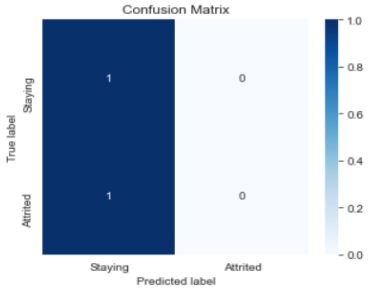
RF classification report on train set

	precision	recall	f1-score	support
Staying	0.84	1.00	0.91	5949
Attrited	1.00	0.00	0.00	1139
accuracy			0.84	7088
macro avg	0.92	0.50	0.46	7088
weighted avg	0.87	0.84	0.77	7088



RF classification report on test set

	precision	recall	f1-score	support
Staying	0.84	1.00	0.91	2551
Attrited	1.00	0.00	0.00	488
accuracy			0.84	3039
macro avg	0.92	0.50	0.46	3039
weighted avg	0.87	0.84	0.77	3039



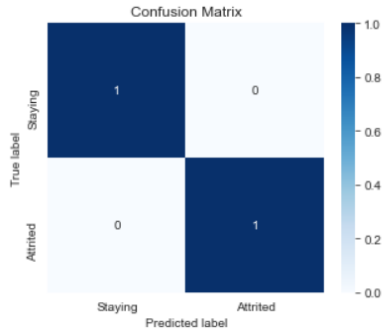
This model performed poorly in comparison to logistic regression. Same as naive model. Grid search may yield better results with hypertuning params, can also try feature importance and apply top 3 or 5 features to get better results. Though recall is 1, +ve predictive rate precision is 0.84. Indicating 84% chance of correct prediction on defaults to non-defaults.

Hyper param tuning did not play the role, it is the same as naive model.

KNeighbours Classification: Metric Evaluation

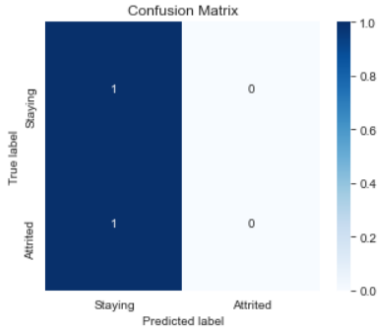
KN classification report on train set

	precision	recall	f1-score	support
Staying	1.00	1.00	1.00	5949
Attrited	1.00	1.00	1.00	1139
accuracy			1.00	7088
macro avg	1.00	1.00	1.00	7088
weighted avg	1.00	1.00	1.00	7088



classification report

	precision	recall	f1-score	support
Staying	0.84	1.00	0.91	2551
Attrited	1.00	0.00	0.00	488
accuracy			0.84	3039
macro avg	0.92	0.50	0.46	3039
weighted avg	0.87	0.84	0.77	3039



KNeighbours is performing same as Random Forest on Test set. However Max roc_auc is perfect 1 on train set,meaning model can identify between all the defaults and the

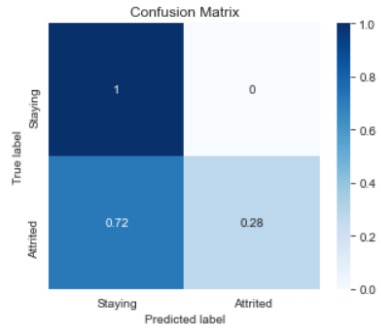
non-defaults points correctly. This is supported with confusion matrix of train set. However on applying to test set it is visible False -ve is 1, and true -ve is 0. can be seen with precision 0.84 Hyper param tuning did not play the role, it is same as naive model.

It is potential that all may be predicted as true +ves meaning, there is attrition, Actually customers are leaving.

XGBoost Classification: Metric Evaluation

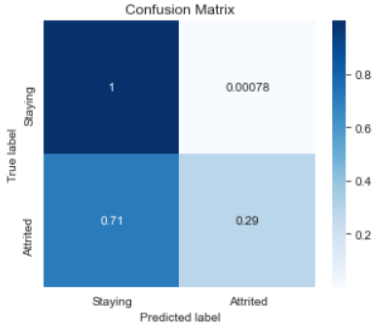
classification report

	precision	recall	f1-score	support
Staying	0.88	1.00	0.94	2551
Attrited	0.99	0.29	0.45	488
accuracy			0.89	3039
macro avg	0.93	0.64	0.69	3039
weighted avg	0.90	0.89	0.86	3039



classification report

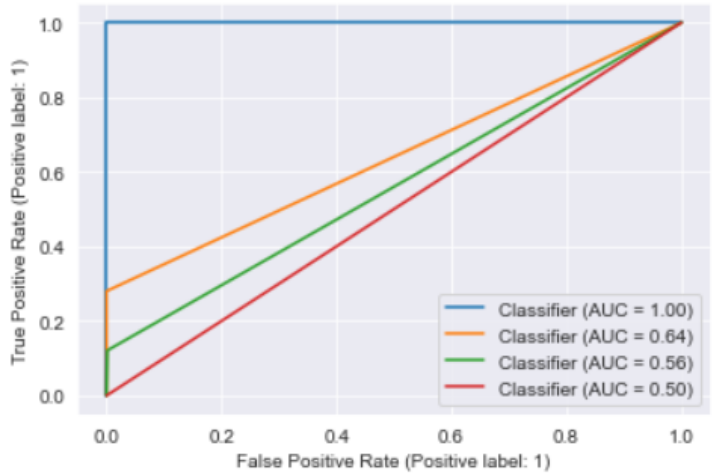
	precision	recall	f1-score	support
Staying	0.88	1.00	0.94	2551
Attrited	0.99	0.29	0.45	488
accuracy			0.89	3039
macro avg	0.93	0.64	0.69	3039
weighted avg	0.90	0.89	0.86	3039



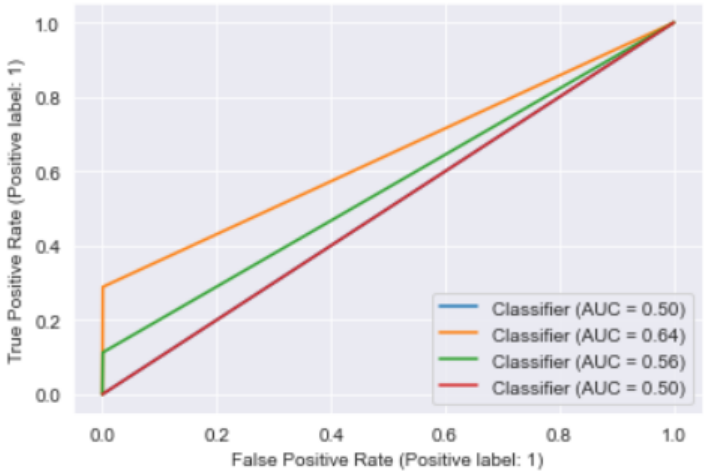
This is generalized better compared to all modeland performed fairly well, as false -ve rate is 0.7 compared with logistic and random forest on both on train and test set. And I can see true +ve rate stayed 1, with recall 1 and improved precision 88%.generalized. This brings a model that can fairly make mistakes in unseen data, in comparison to other models.

ROC_AUC Curve Evaluation

On train set



On test set



- Logistic regression is performing well on both training and test sets. With a score of 0.55 This is generalized better.
- KNeighbors performed very badly on train is 1.0 and on test data with roc_auc 0.71. It is overfitting.Meaning it can perfectly identify between all the Positive and the Negative class points correctly on train set.However on test data there is a chance KN identify the defaults from the non-defaults class values
- RandomForest has both train and test roc_auc is 0.5. It is the same as naive. Model does not have the ability to predict defaults and non-defaults.
- XGboost has done fairly well in terms of roc_auc performance with slight difference between train and test roc_auc, those are 0.64 and .64.scores are almost close.Same as Logistic regression, generalized better.