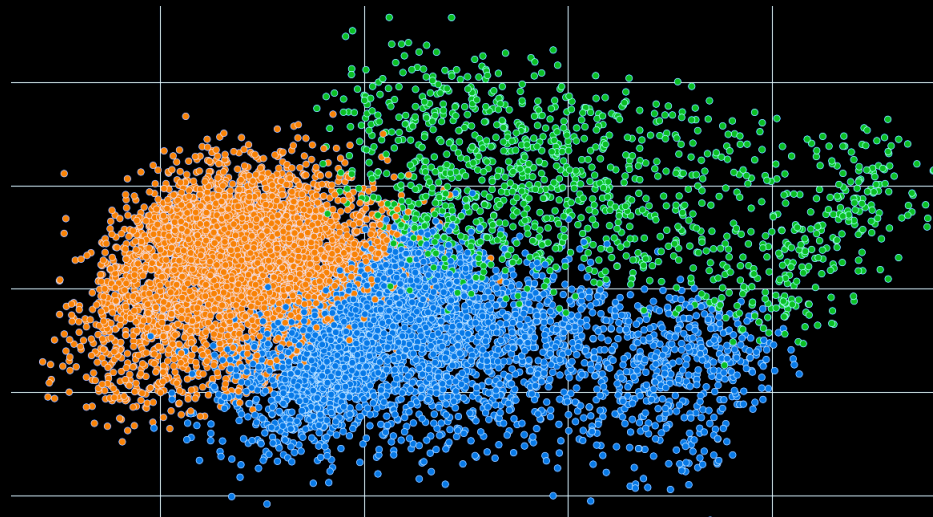


# Bank Churn Prediction using Supervised Machine Learning



Dominik Vukusic



# What's troubling us ?

---

**Credit card portfolio** is facing the problem of **customer churn**

We need to find out **the reason behind** this and leverage the same to **predict** customers who are likely to **drop off**



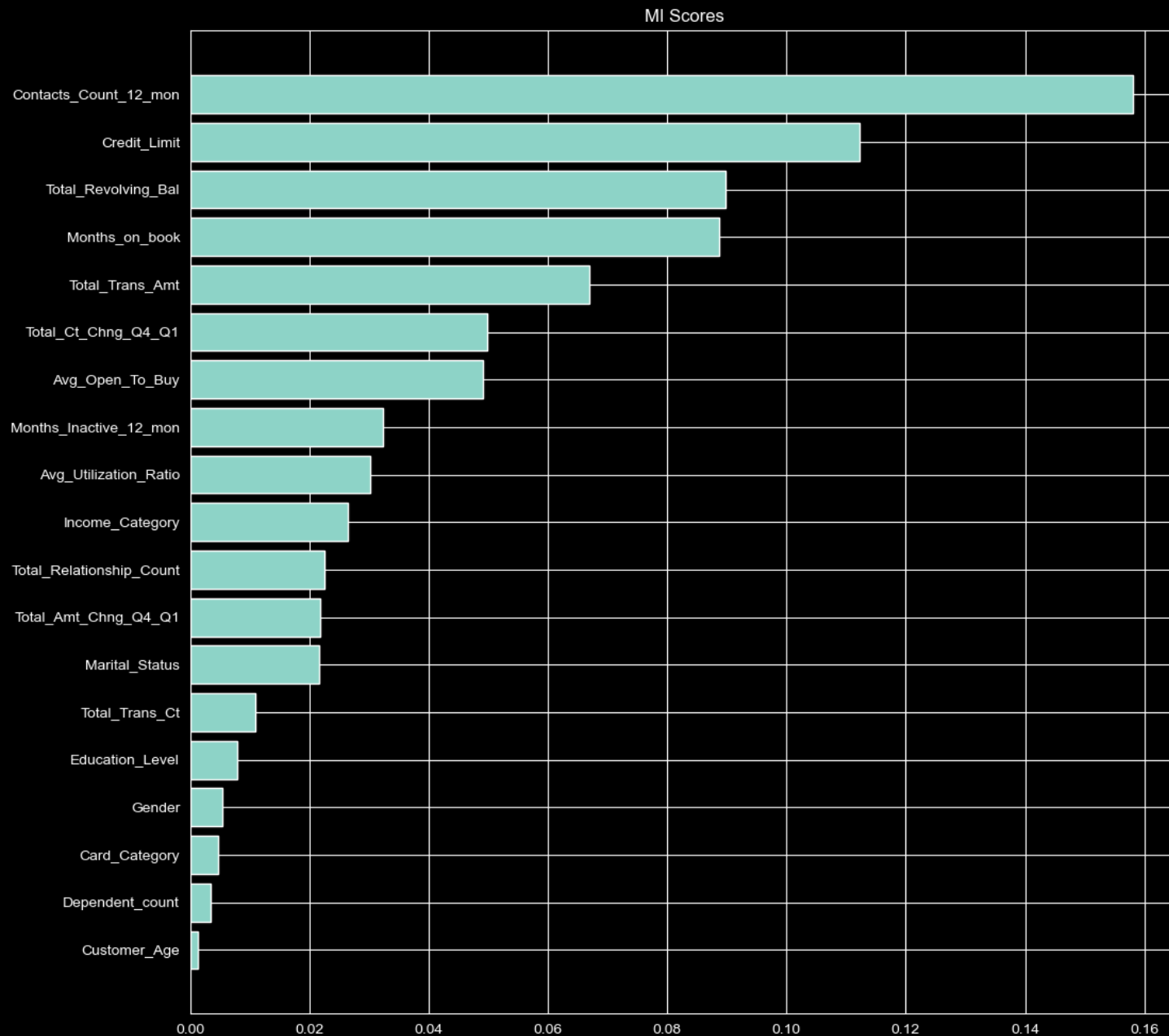
# What data do we have ?

---

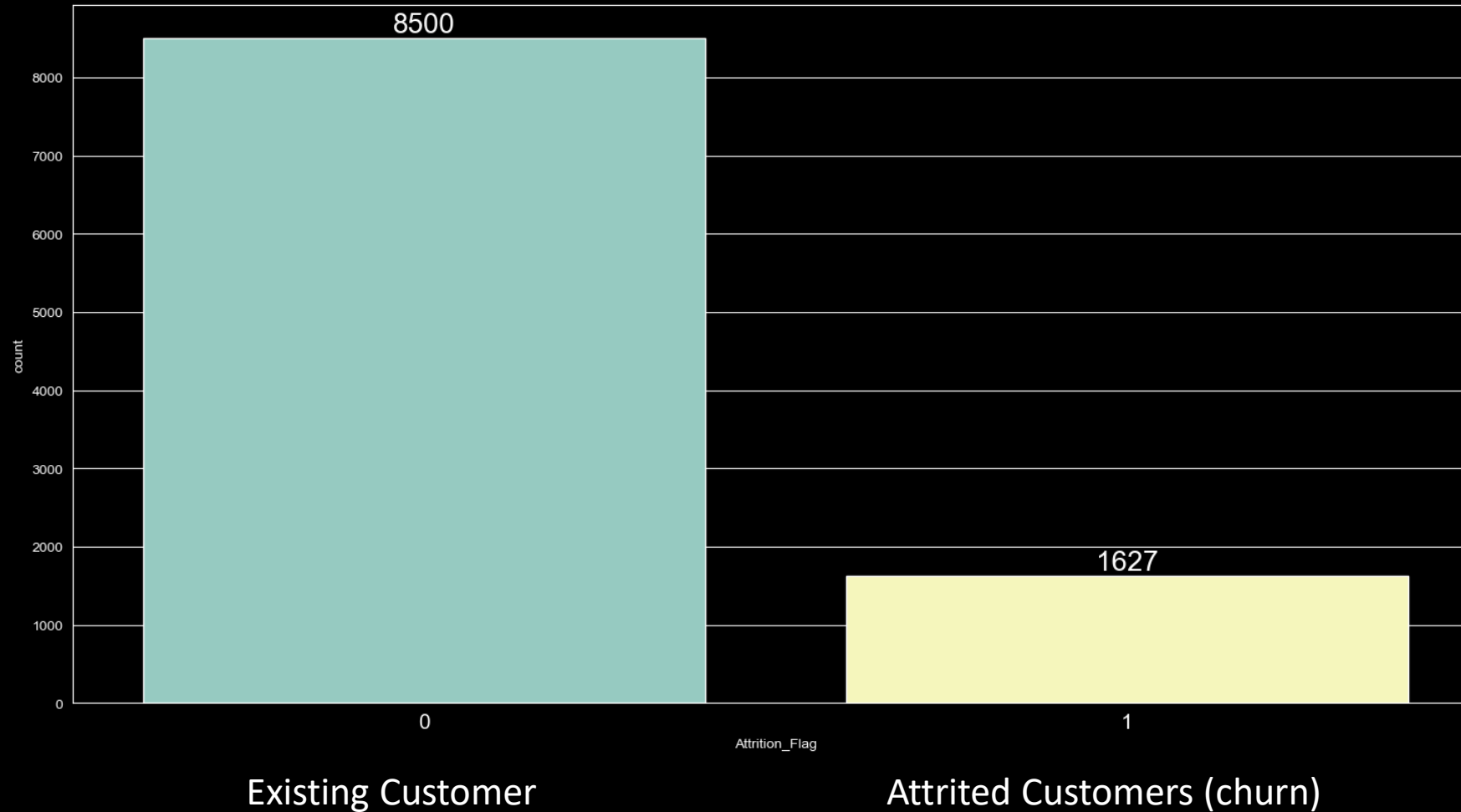
- 23 columns,
- Data on **personal information**,
- User **details**,
- Transaction and spending **statistics** and
- Details of **banking services** that the user has at his disposal


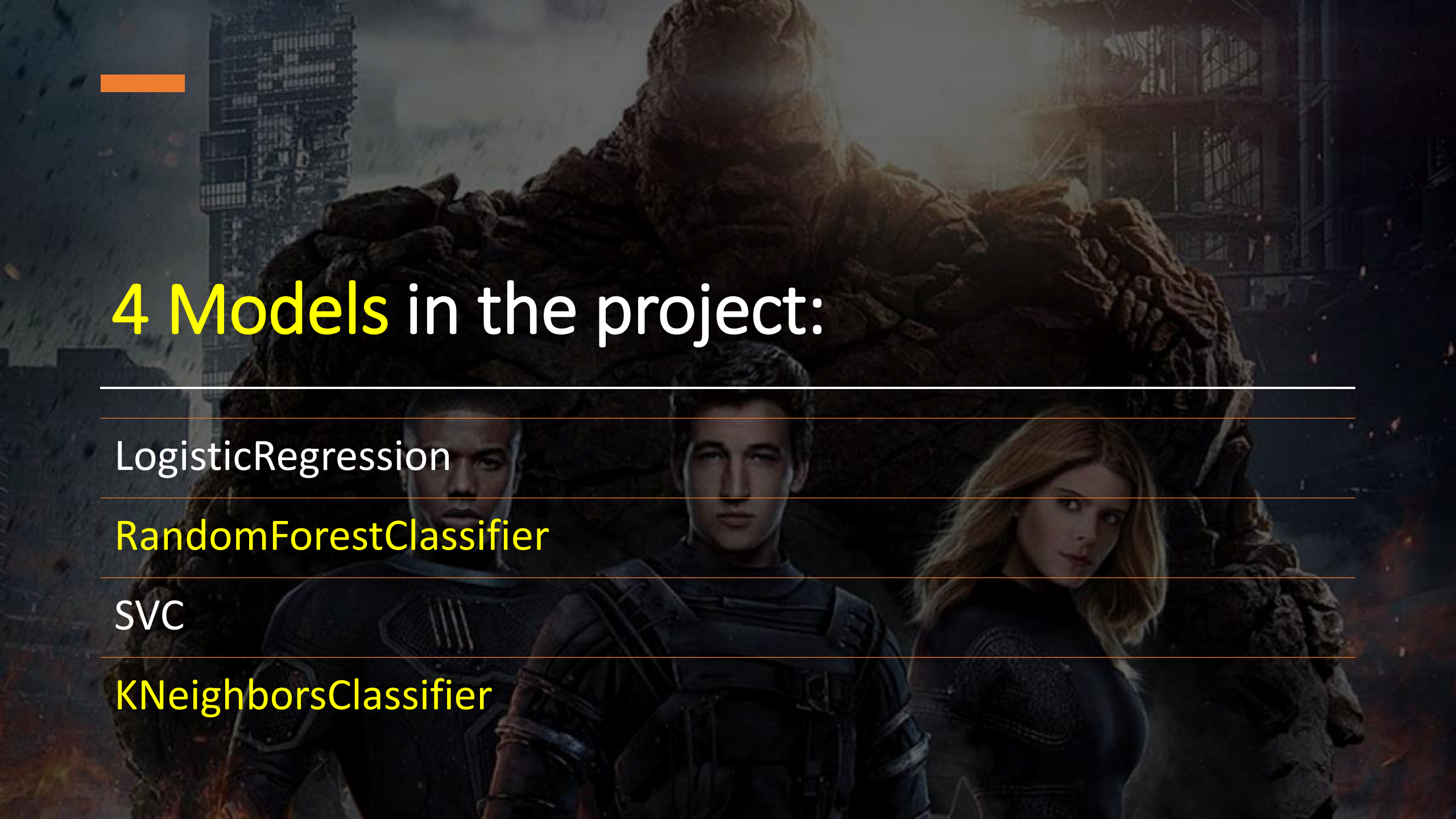
# MI scores

derived from spearman  
correlation



# Attrition Flag distribution in BankChurners dataset





# 4 Models in the project:

---

LogisticRegression

---

RandomForestClassifier

---

SVC

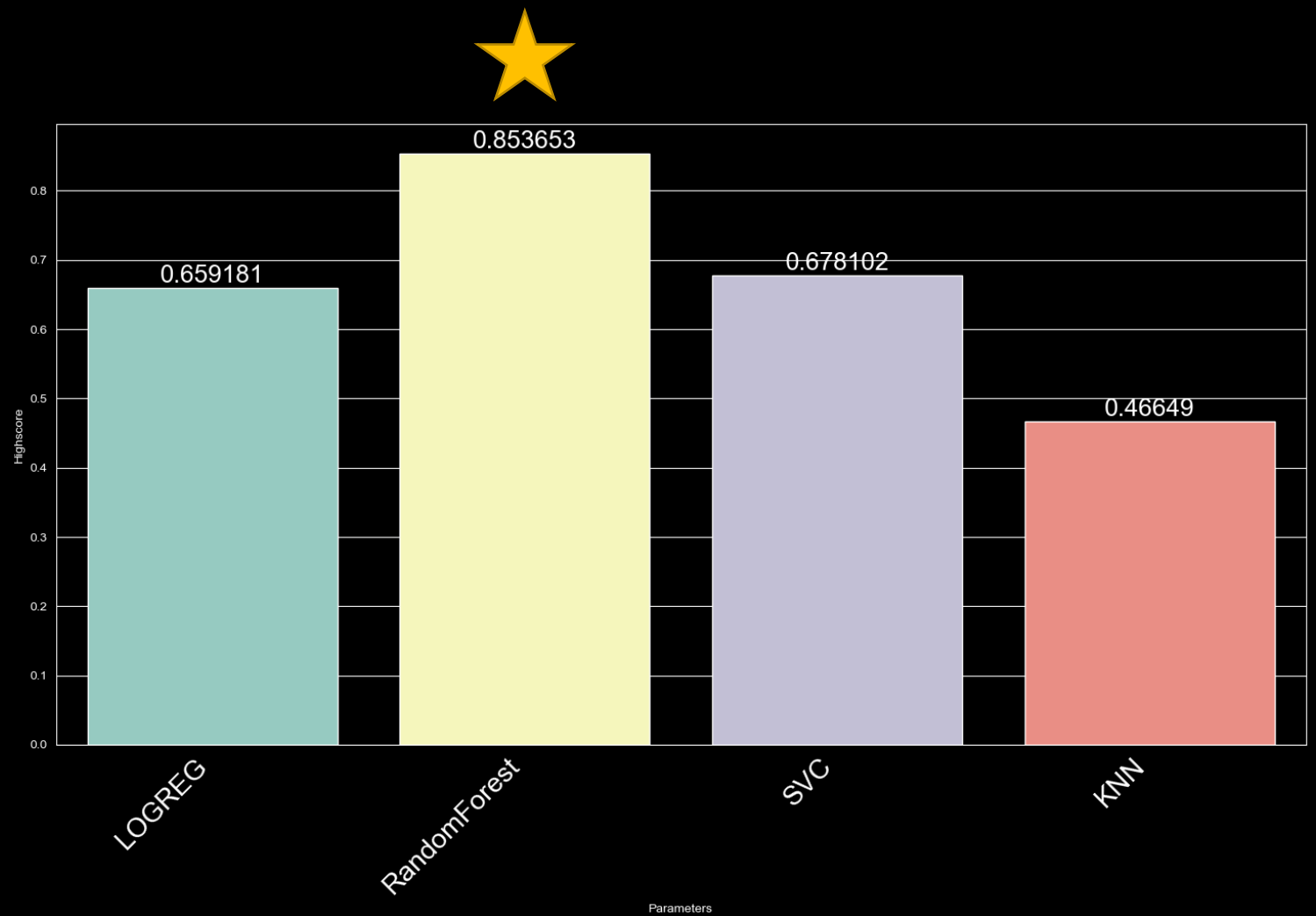
---

KNeighborsClassifier

## train\_models\_results

The  $F_1$  Score is given by :

$$F_1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$



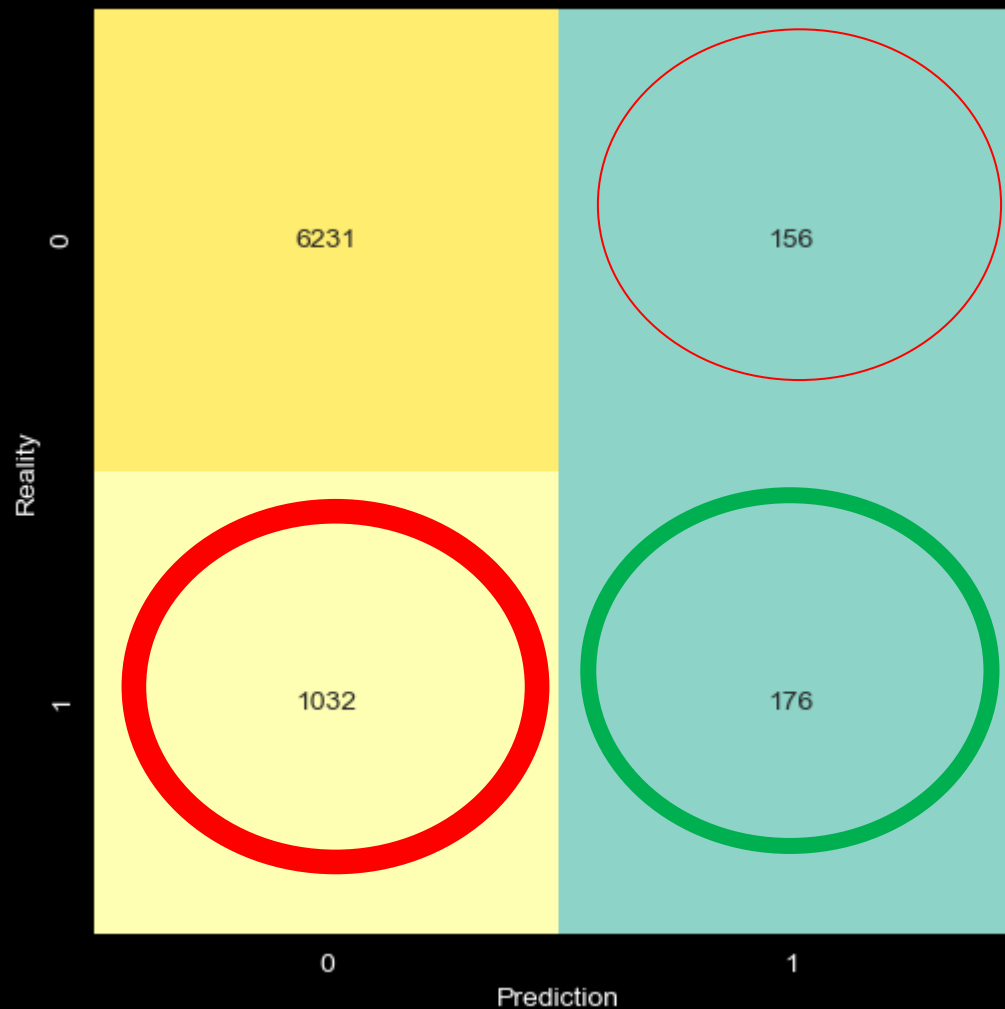
# confusion\_matrix

## train\_data : RandomForest

Best f1\_score: 84.56%

Best params

```
{  
'max_depth': None,  
'min_samples_leaf': 1,  
'n_estimators': 50  
}
```

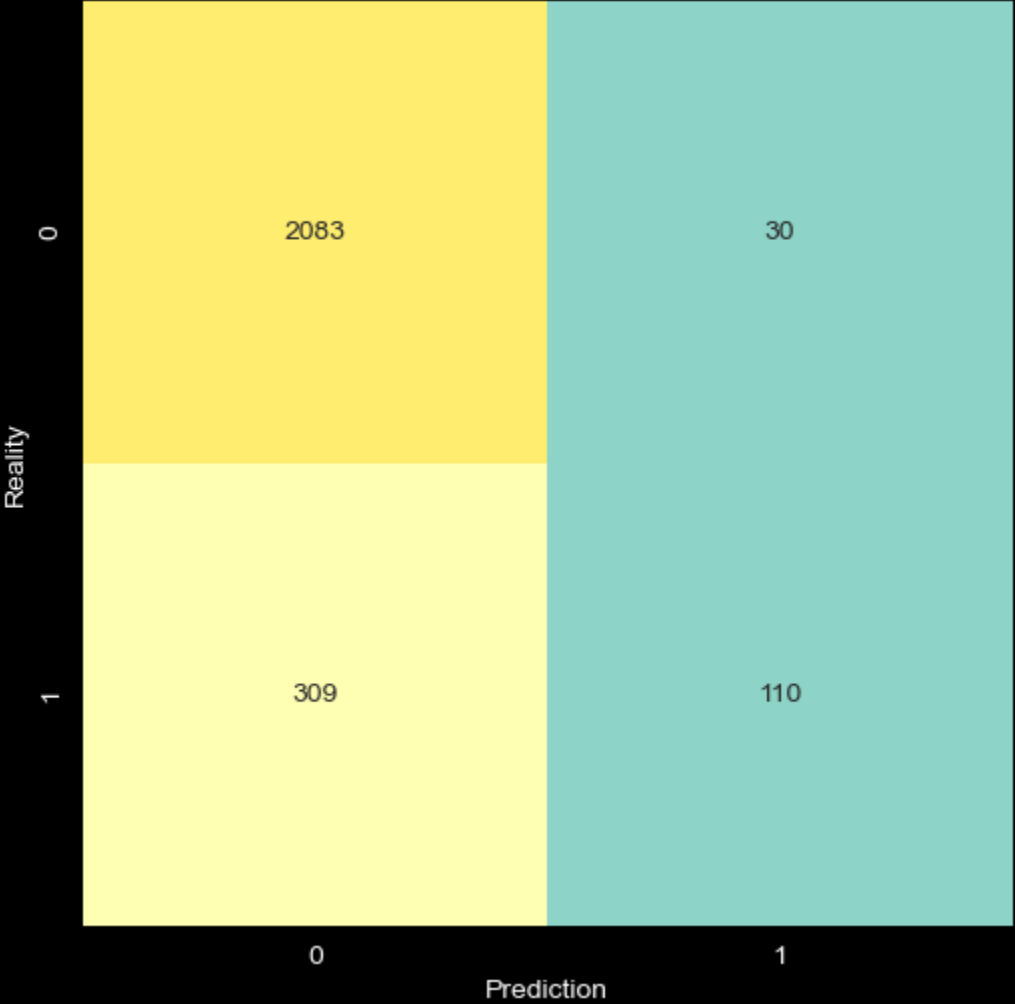






# test\_RandomForest

f1\_Score: 83.71%



## What now?

---



Potential tactics to prevent churns:

- 1.) Increase the limit on credit cards for targeted users
- 2.) Make customer service a priority of your business
- 3.) Research the competition

Questions?

