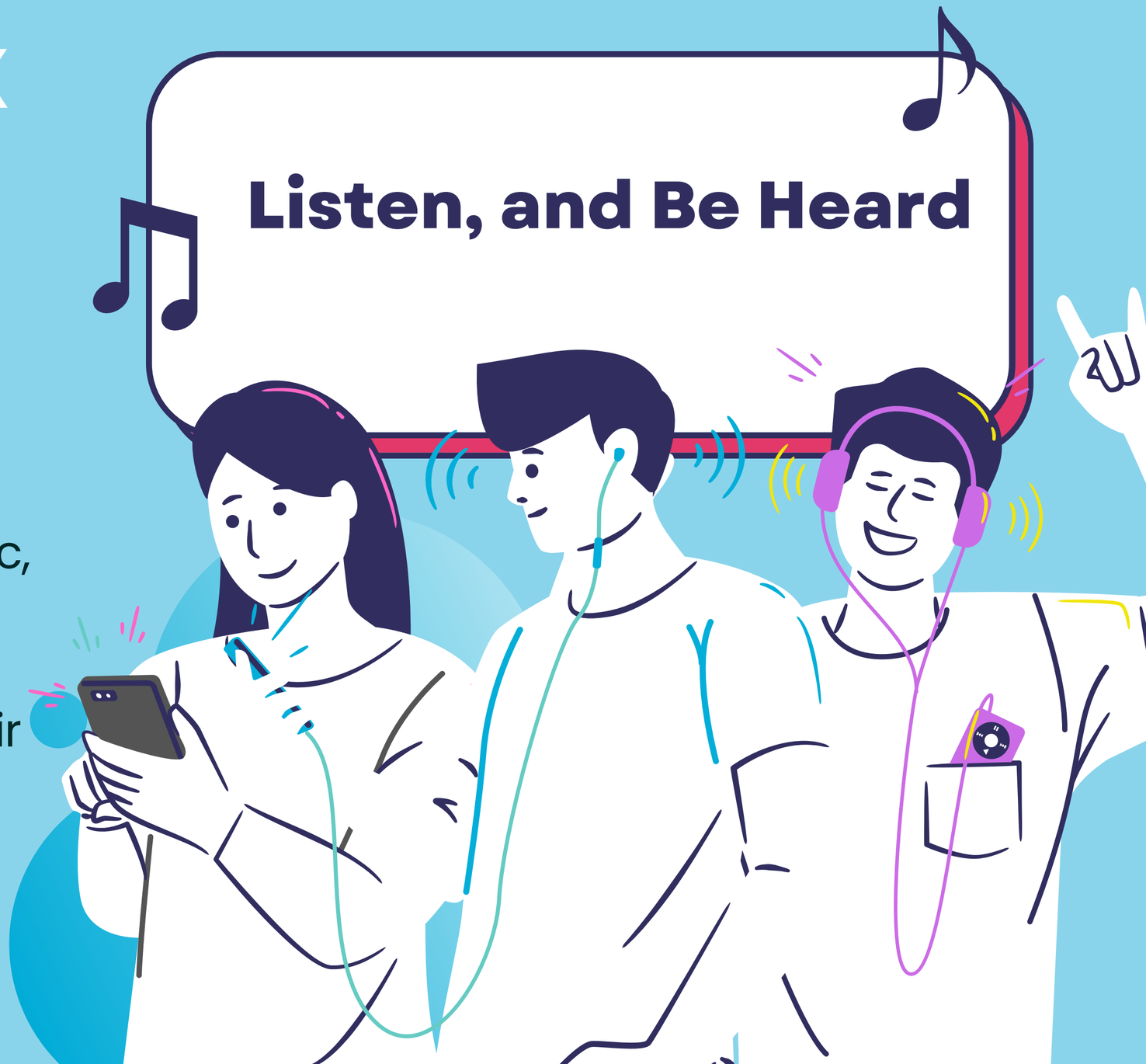
KKbox Churn Rate Prediction

https://www.kkbox.com/intl/

Presented By:
Daniela Magiricu

About KKBox Music App

- KKBox is a Taiwanese company offering a music streaming service to millions of people across Asia.
- Leading provider of Asian Pop music, with over 30 million tracks
- Offering an unlimited version of their service to millions of people supported by advertising and paid subscriptions



Business Problem

Business Question

*How can KKBox improve customer retention and minimize revenue loss with **machine learning**?*

KKBox would like to forecast the likelihood of customers discontinuing their subscriptions in the future.

Their business model is dependent on **accurately predicting the churn of their users.**

Business Value KPI's

Future Expected Revenue

Predicting future churn rate helps KKBox gain a better understanding of **future expected revenue**.

Churn prediction is essential for KKBox to understand what **preventative steps** are necessary to ensure **lost revenue is minimized.**

Customer Retention

Cost of aquiring new customers is **5 times** higher than keeping existing ones

KKBox should focus their efforts towards **retaining existing customers**

Customer Lifetime Value

Churn prediction allows KKBox to **target** individual customers and **prevent** them from discontinuing their subscription.

Churn prediction is also one of the **key components** in determining the lifetime value of customers.

Churn Rate Prediction Impact on Revenue Loss

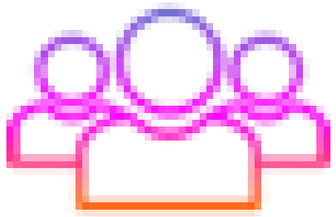
Churn Users: 12,690

Total Users: 970,960

Predicted Churn Rate: 1.3%

Churn Prediction = **$ 1.9 million** in lost revenue within the next 30 days

Data and Cleaning

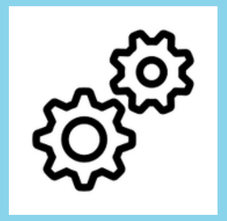


01 Members Data

02 Transactions Data

03 User Logs Data

Modelling Process



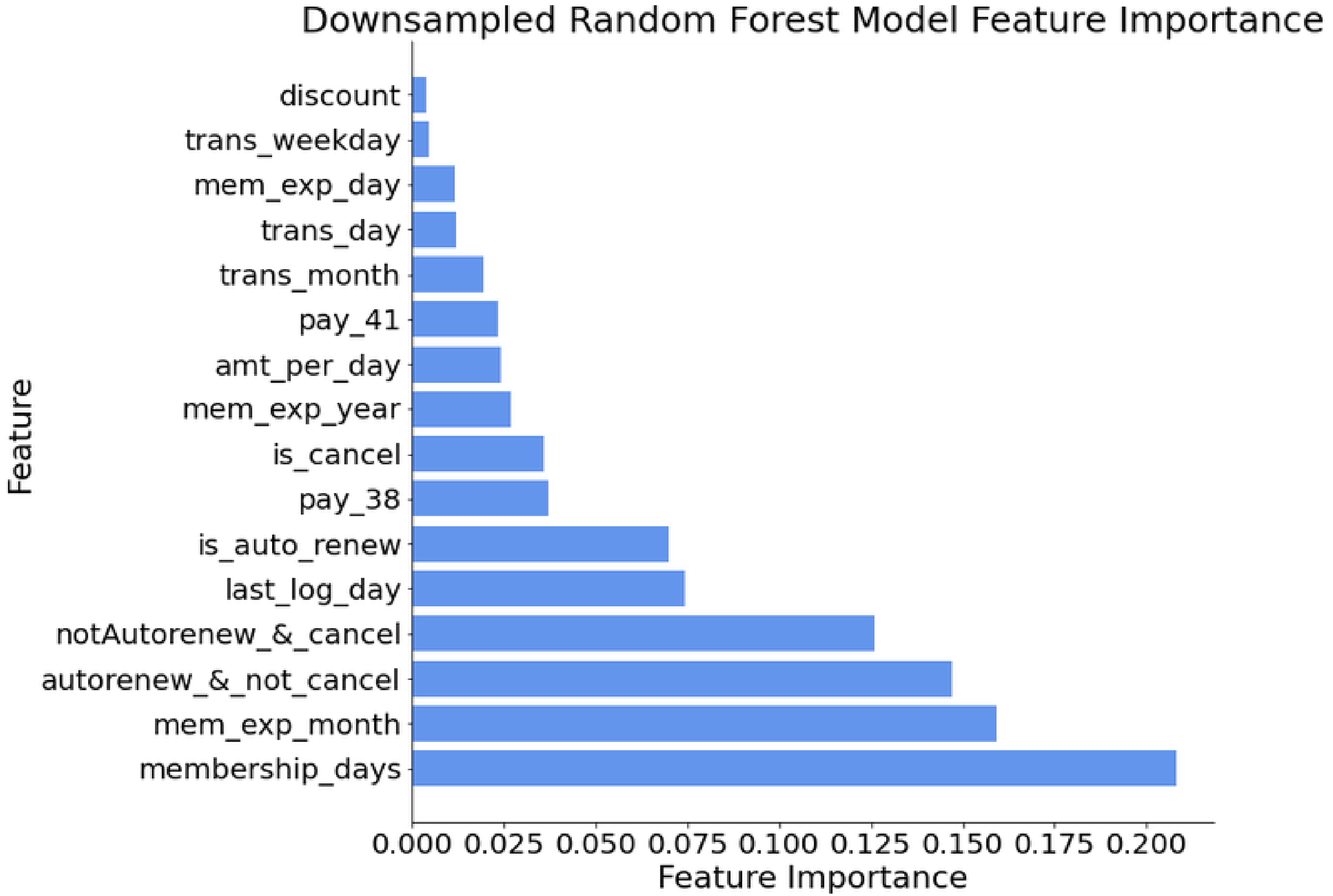
Feature Engineering

Chi-Squared Statistical Correlation Tests for Multicollinearity

Imbalanced Data: Upsampling and Downsampling Techniques

Logistic Model and Random Forest Model

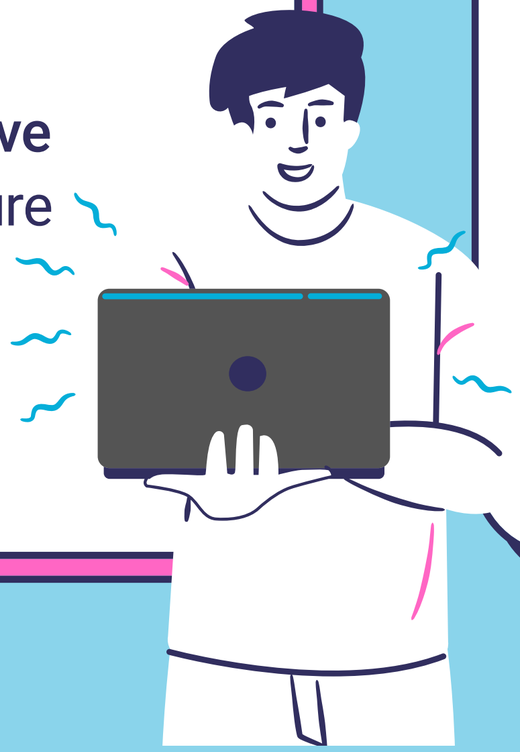
Random Forest Model Top 15 Important Features



Actionable

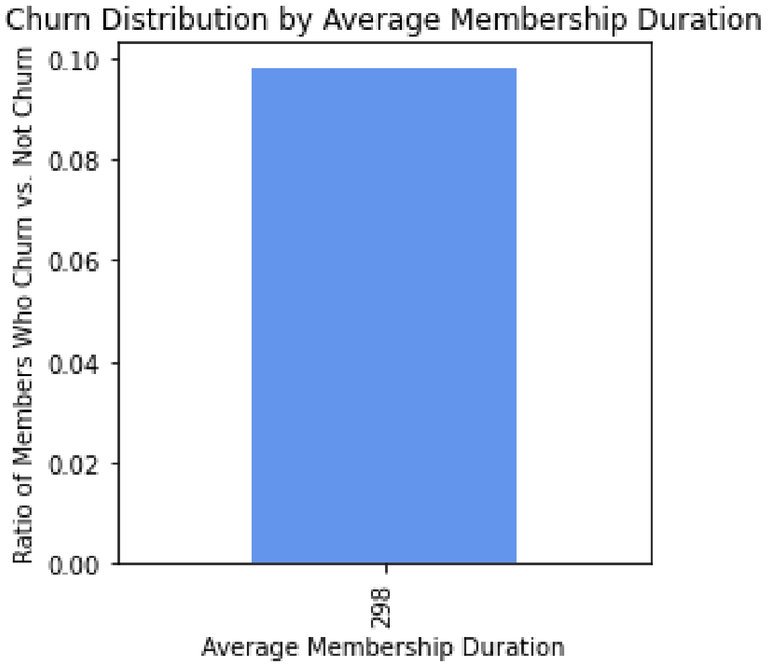
KKBox can use these results to identify and improve upon areas where customer service is lacking

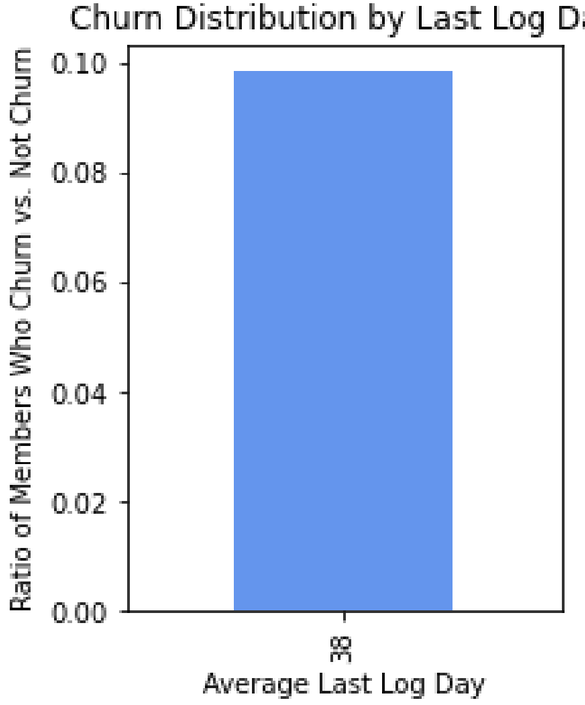
Marketing team can understand what **preventative** steps are necessary to ensure **lost revenue is minimized**

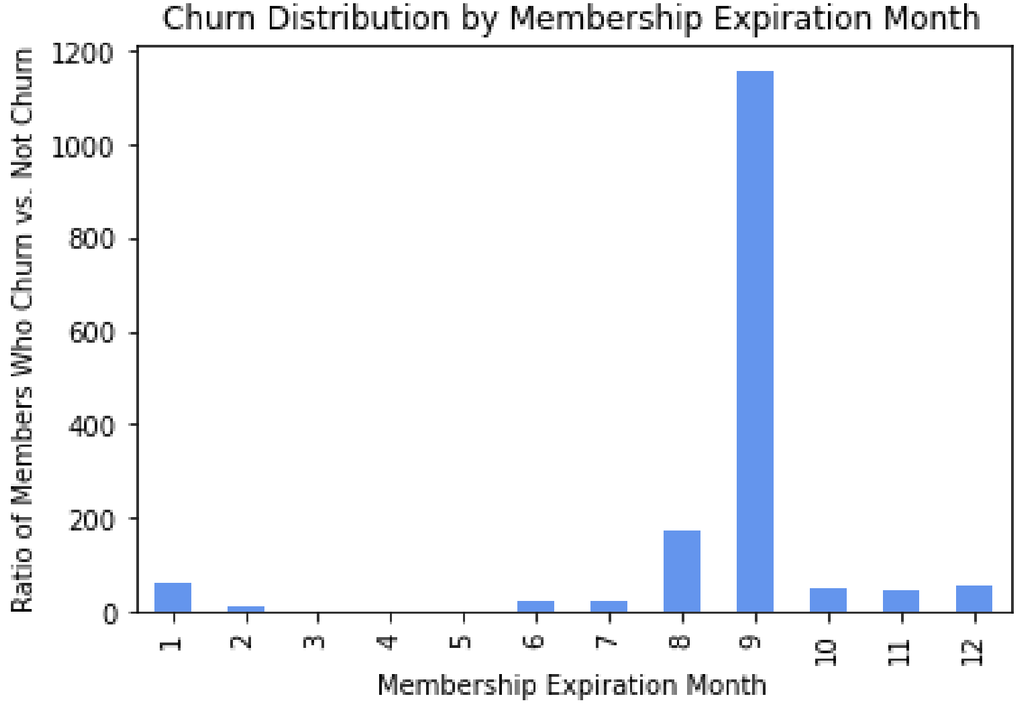


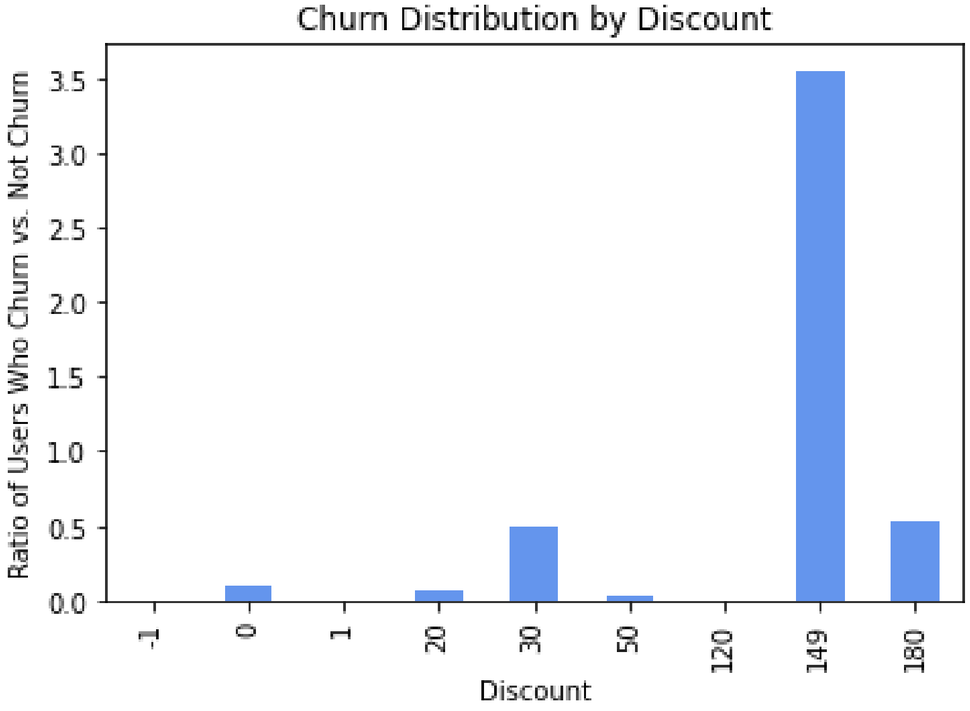
Insights

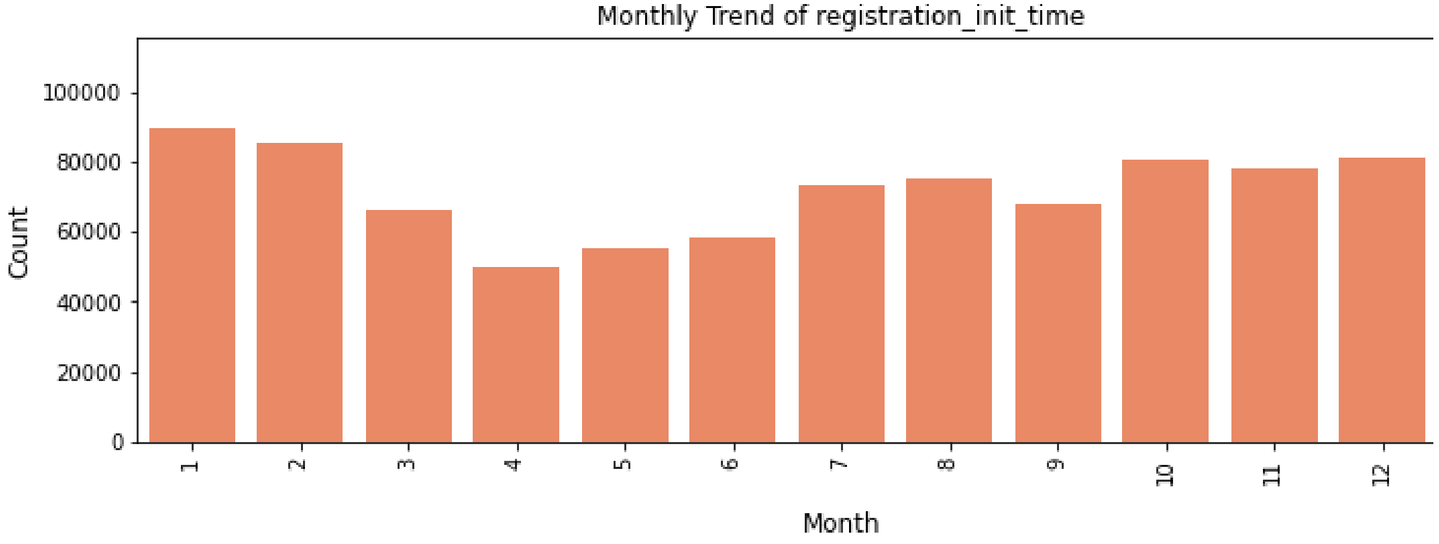
Focus on the **most important features** from our Random Forest Modelling results to make improvements in **customer lifetime value** as well as **customer retention** and **further reduce the churn rate and revenue loss.**

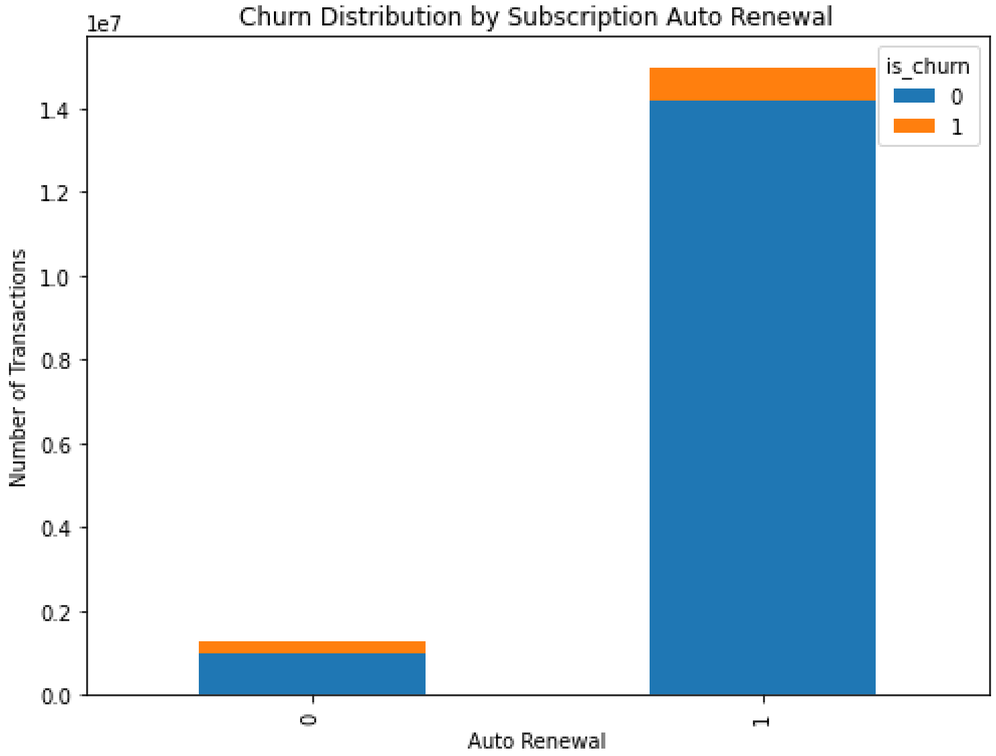

Churn Distribution by Average Membership Duration
Ratio of Members Who Churn vs. Not Churn
0.00
0.02
0.04
0.06
0.08
0.10
298
Average Membership Duration


Churn Distribution by Last Log Day
Ratio of Members Who Churn vs. Not Churn
0.00
0.02
0.04
0.06
0.08
0.10
38
Average Last Log Day


Churn Distribution by Membership Expiration Month
Ratio of Members Who Churn vs. Not Churn
0
200
400
600
800
1000
1200
1
2
3
4
5
6
7
8
9
10
11
12
Membership Expiration Month

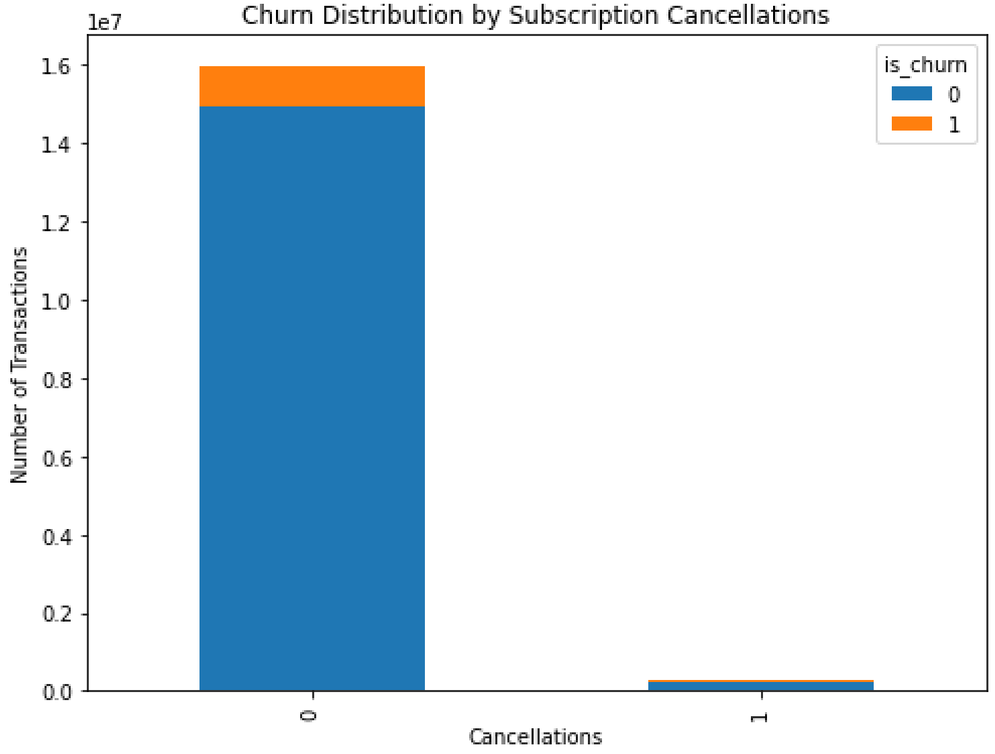

Churn Distribution by Discount
Ratio of Users Who Churn vs. Not Churn
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
-1
0
1
20
30
50
120
149
180
Discount


Monthly Trend of registration_init_time
Count
0
20000
40000
60000
80000
100000
1
2
3
4
5
6
7
8
9
10
11
12
Month



	No Churn	Churn	Ratio
Not Autorenew	964,662	31,7794	0.329436
Autorenew	14,199,457	773,709	0.054489

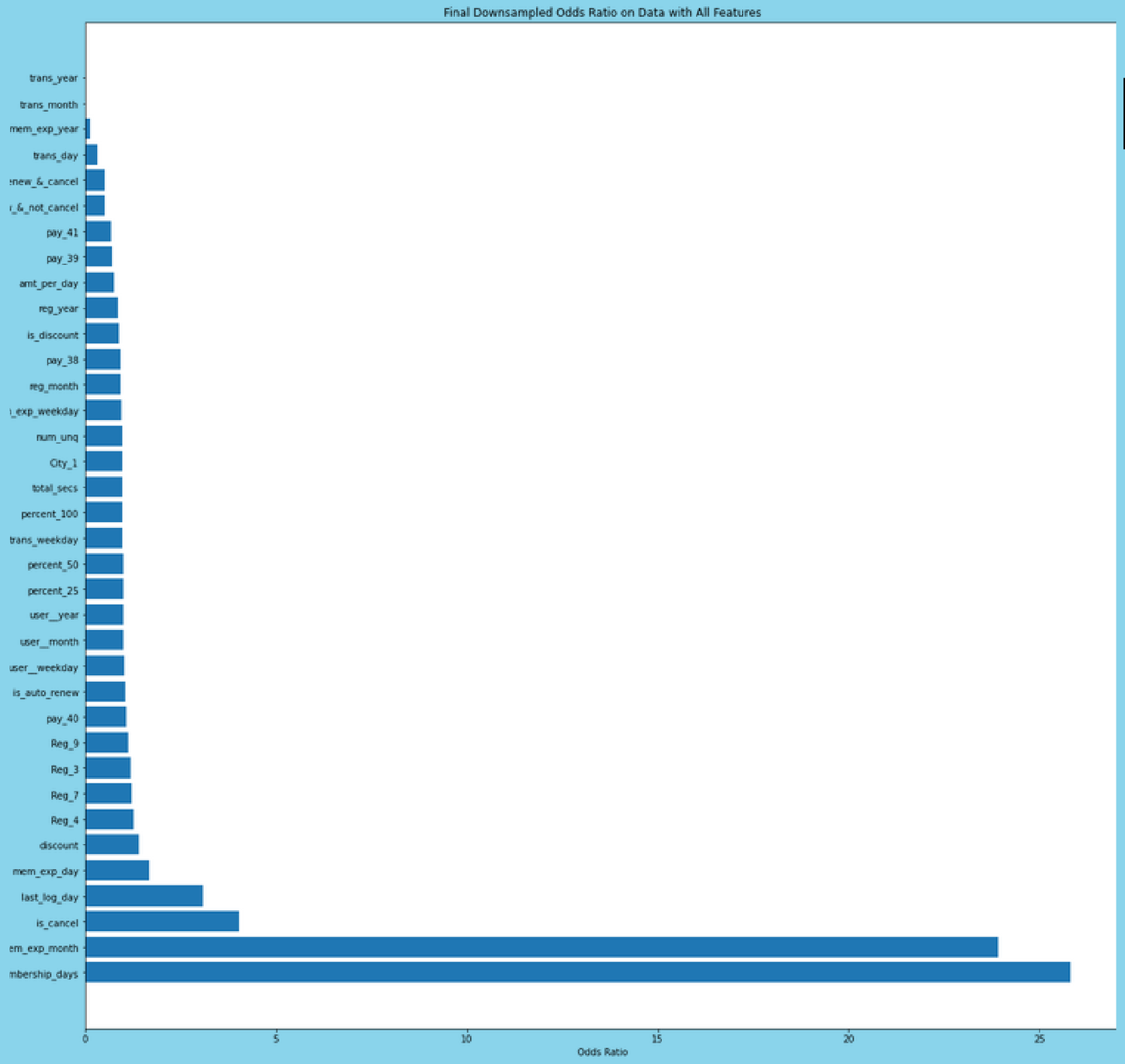
	No Churn	Churn	Ratio
Not Cancel	14930693	1043633	0.069898
Cancel	233426	47870	0.205076



Thank You

Music is the color for the world and everything in it

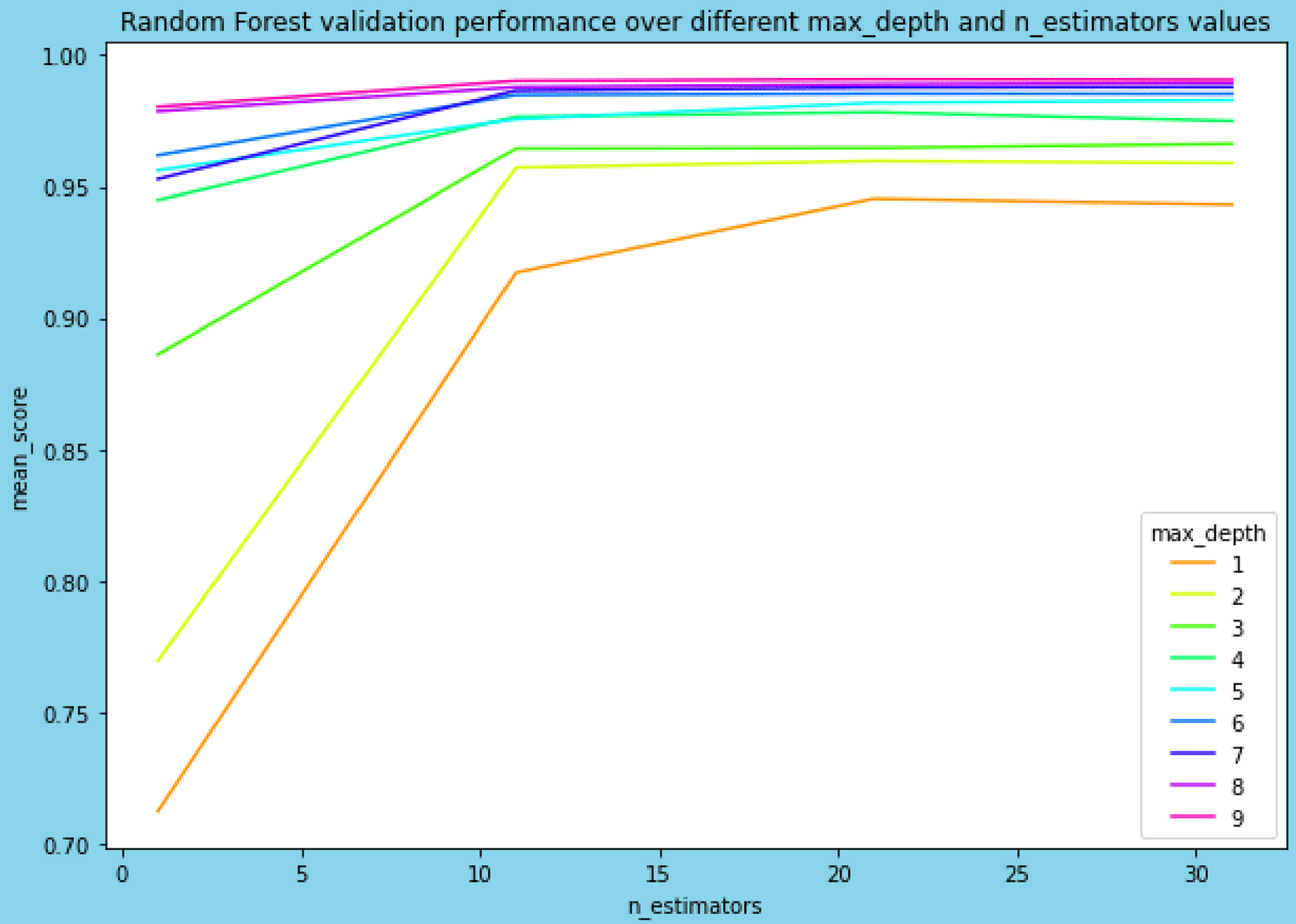
https://www.kaggle.com/competitions/kkbox-churn-prediction-challenge/data



Logistic Model Odds Ratio

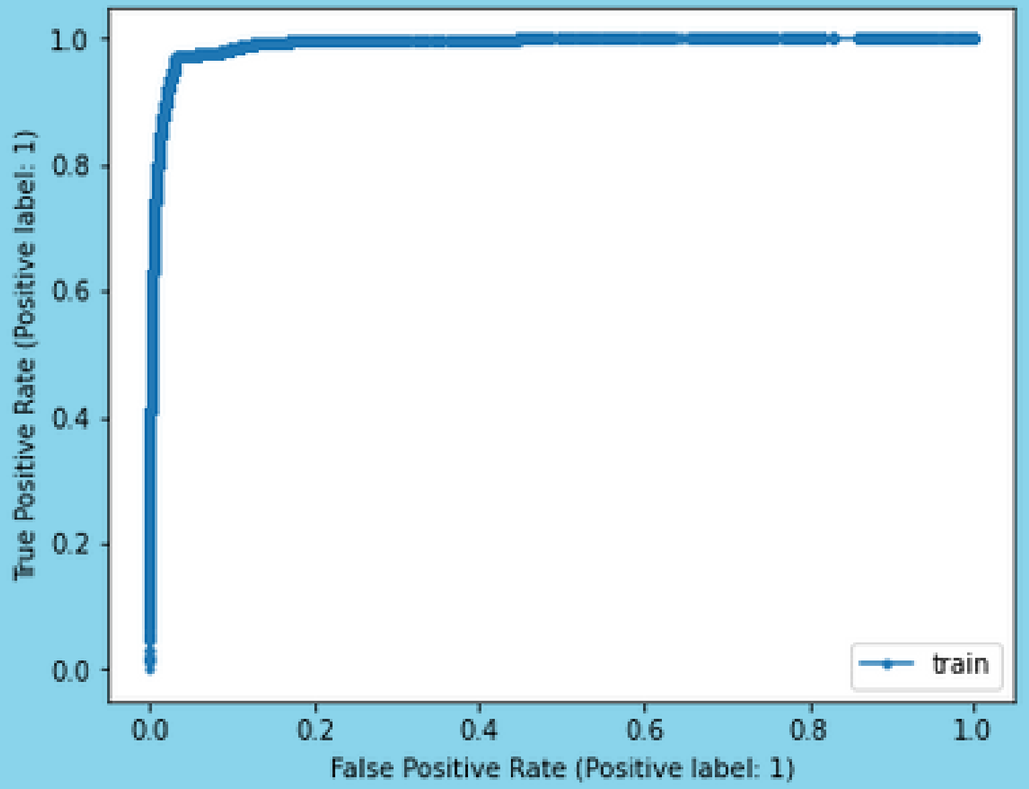
- **Membership Duration**
- Membership Expiration Month
- Cancellation
- Last Log Day

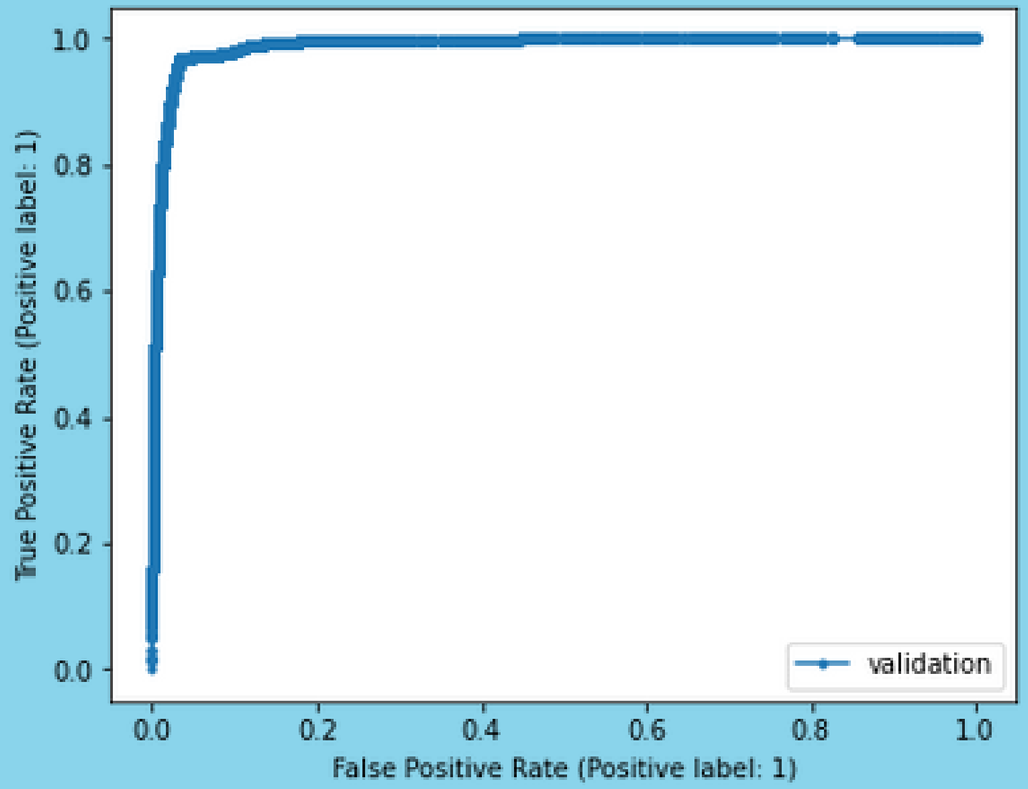
Random Forest Model Hyperparameter Optimization



- Max depth 9
- 10 estimators

Random Forest Model: ROC-AUC Curve





Train AUC Score: 0.99

Validation AUC Score:0.99