

Customer Churn Prediction – Raport

1. Goal

The goal of this analysis is to predict if a customer will **churn** (leave the service). This helps the bank understand customer behavior and reduce churn risk.

2. Dataset

- **Number of customers:** 1000
 - **Features:** Age, Gender, MaritalStatus, IncomeLevel, TotalSpent, AvgSpent, NumTransactions, NumInteractions, ResolvedRatio, AvgLoginFreq, MostUsedPlatform
 - **Target:** ChurnStatus (0 = stay, 1 = churn)
-

3. Model

- **Algorithm:** Random Forest Classifier
 - **Train/Test Split:** 80% train, 20% test
 - **Preprocessing:**
 - Missing values filled with median
 - One-hot encoding for categorical features
 - CustomerID removed
-

4. Model Performance

Metric	Value
Accuracy	0.745
Precision	0.0
Recall	0.0
F1-score	0.0

Confusion Matrix:

```
[[149   1]
 [ 50   0]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.75	0.99	0.85	150
1	0.00	0.00	0.00	50
accuracy			0.74	200
macro avg	0.37	0.50	0.43	200
weighted avg	0.56	0.74	0.64	200

Note: The model predicts class **0 (stay)** well, but it fails to predict **1 (churn)**. This happens because there are fewer churn cases in the data.

5. Top Features (Importance)

Feature	Importance
AvgSpent	0.1736
AvgLoginFreq	0.1631
TotalSpent	0.1553
Age	0.1487
NumTransactions	0.0825
NumInteractions	0.0437
ResolvedRatio	0.0430
Gender_M	0.0292
MostUsedPlatform_Online Banking	0.0269
IncomeLevel_Medium	0.0248

Explanation:

- AvgSpent, AvgLoginFreq, and TotalSpent are the most important features.
 - Bank can focus on these features to reduce churn.
-

6. Recommendations

1. Solve **class imbalance** problem (use SMOTE, undersampling, or class weights).
2. Add more features to detect churn customers better.
3. Train the model again to increase recall and F1-score.
4. Use top features to make business strategies (loyalty programs, special offers).