

Bank Customer Churn Prediction – Case Study

1. Problem Statement

Predict which customers are likely to leave the bank (churn) based on demographic and account-related features.

2. Dataset Overview

- 10,000 customers
- Features: Credit Score, Geography, Gender, Age, Tenure, Balance, NumOfProducts, IsActiveMember, EstimatedSalary
- Target: Churn (0 = No, 1 = Yes)

3. Data Cleaning

- Removed duplicates
- Standardized target values
- Filled missing values
- Encoded categorical columns
- Scaled numerical features

4. EDA Insights

- Germany customers churn more
- Older customers churn more
- Low activity increases churn
- Single-product customers churn more

- Balance has weak correlation with churn

5. Model Building

Trained models:

- Logistic Regression
- Decision Tree
- Random Forest (best baseline)

Best Model: Random Forest

6. Feature Importance (Permutation)

Top influential features:

- IsActiveMember
- Age
- NumOfProducts
- TenureGroup
- HasCrCard

7. Dashboard Summary

Excel dashboard includes:

- Churn distribution
- Geography churn

- Gender churn
- Tenure analysis
- Balance analysis
- Activity impact

8. Conclusion

Key drivers of churn:

- Low activity
- Higher age
- Fewer products
- Country differences

Business actions:

- Targeted retention campaigns
- Personalized offers
- Customer activity engagement programs