IE 492 TERM PROJECT - REPORT

# Profit Maximizing Credit Card Churn Prediction for Campaign Decisions

EYLÜL RANA SARAÇ - 2019402126

MERVE NUR HÜNDÜR - 2019402060

METEHAN YALÇIN - 2018402042

ADVISOR: AYBEK KORUGAN

# CONTENTS

# ÖZ

Günümüz iş ortamında, artan rekabet ve pazarlama maliyetleri nedeniyle müşteri kazanımı her zamankinden daha zor hale gelmektedir. Bu bağlamda, işletmeler için mevcut müşterilerin korunması ve sadakatinin artırılması daha da önem kazanmıştır. Özellikle bankacılık sektöründe, müşteri kaybı ciddi bir endişe kaynağıdır çünkü bu, hem gelir kaybına hem de pazar payı kaybına neden olabilir. Bu nedenle, müşteri kaybını önlemek ve müşteri sadakatini artırmak, bir şirketin karlılığını artırmada kritik faktörler arasındadır.

Bu projede, bir bankayla işbirliği yapılarak kredi kartı müşterilerinin churn olma ihtimalleri analiz edilmiştir. Bankanın mevcut churn modeli incelenmiş ve yeni bir yaklaşım geliştirilmiştir. Önceki modelde, müşterilerin churn olarak etiketlenmesi sadece 3 ay boyunca hiç işlem yapmamalarına dayanıyordu. Ancak geliştirilen projede, müşterilerin sadece aktif olup olmadığı değil, aynı zamanda harcamalarındaki azalmalar da dikkate alınmıştır. Yani harcama davranışı ikili değil sürekli bir değişken olarak ele alınmıştır. Belirli kriterleri sağlayarak harcamalarında azalma gösteren müşteriler churn olarak etiketlenmiş ve bu yeni yaklaşıma göre farklı makine öğrenme modelleri kullanılarak müşterilerin churn olma ihtimalleri hesaplanmıştır.

Seçilen modelin çıktıları, kar maksimizasyonu optimizasyon modelinde girdi olarak eklenerek modelin kampanya kararları alması sağlanmıştır. Bu yaklaşımla, churn olma ihtimali olan müşterilerin önceden belirlenerek doğru aksiyonlarla kayba dönüşmesini engellemek amaçlanmıştır.



**Anahtar Kelimeler:** Müşteri kaybı, bankacılık, müşteri sadakati, makine öğrenimi, kampanya yönetimi, optimizasyon.

# ABSTRACT

In today's business environment, customer acquisition has become increasingly challenging due to rising competition and marketing costs. In this context, preserving and enhancing the loyalty of existing customers has become even more crucial for businesses. Particularly in the banking sector, customer attrition is a significant concern as it can lead to both revenue loss and market share erosion. Therefore, preventing customer churn and increasing customer loyalty are critical factors in enhancing a company's profitability.

This project involved collaborating with a bank to analyze the likelihood of churn among credit card customers. The bank's existing churn model was examined, and a new approach was developed. Unlike the previous model, which relied solely on customers not conducting any transactions for three months in a row to label them as churned, the developed project also considered decreases in customers' spending habits. In other words, spending behavior is treated as a continuous variable rather than a binary variable. Customers who met certain criteria for decreased spending were labeled as churned, and different machine learning models were used to calculate their probabilities of churn based on this new approach.

The outputs of the selected model were incorporated as inputs into a profit maximization optimization model to enable the model to make campaign decisions. With this approach, the aim was to proactively identify customers at risk of churn and prevent their loss through appropriate actions.

**Keywords:** Customer churn, banking, customer loyalty, machine learning, campaign management, optimization.

# 1. Introduction

The main problem that this study aims to solve is the loss of customers who reduce or discontinue the use of credit cards, resulting in a decline in earnings. In order to prevent this loss of profit, it is aimed to predict the churning customers by identifying their decreasing spending trends before they are lost. In this way, it was aimed to increase their credit card usage and loyalty to the bank by taking campaign actions at the right time.

In order to predict the customers who are going to churn, the general spending habits of the customers were analyzed by taking into account their demographic information. In order to perform the necessary data analysis, data cleaning steps such as removing nulls, duplicates and formatting were applied first. With the help of these analyses, various machine learning models were created and the best working algorithm was selected according to their accuracy, precision, recall, and F1-score. The predictions of the selected algorithm were used in the profit maximizing optimization model that chooses which type of campaign to send to which customers.

As a result, the goal is to provide the marketing department the information on which customers should receive a campaign and how much the campaign should cost, using a determined budget. In the light of this information, the marketing department will deliver the campaign at the cost determined by the model to the customers to whom the campaign should be sent, in the appropriate area and in the way it deems appropriate.

According to the results of the model applied on 51169 customers, the bank is projected to generate revenues of 169,539,234 liras in April 2024. According to this projection, the bank will earn 2 million liras more than in the absence of the model.

In the rest of the report, the problem to be solved and the shortcomings of the bank's existing model will be explained in detail, followed by the steps of the solution proposed in this study. Finally, detailed outputs of the model will be shared and assumptions, limitations and further studies will be discussed.

# 2. Problem Definition, Requirements and Limitations

## 2. 1. Problem Definition

Acquiring new customers is costlier than retaining existing ones.[1] In some cases, it can even cost five times more to acquire a new customer than to keep an old one.[2] Although there are many reasons for this situation, the main ones can be listed as follows:

- Targeting new customers is more difficult than targeting existing customers because the company already has data about its existing customers but limited data about the targeted customers.
- Bringing new customers involves additional expenses for onboarding processes and training.

Customer retention is therefore crucial to increase profitability. For this reason, this study aims to predict customers likely to be lost in the future and retain them with targeted campaigns. In fact, there is a method that the bank applies for customer retention, which is, customers who have not used their credit card for three consecutive months are considered to have churned and campaigns are sent to these customers. However there are two main problems with the bank's current churn model:

### a. Inadequate Detection of Decreased Usage:

The bank’s existing churn model labels a customer as churn only when the customer has not used the credit card for three consecutive months. The model fails to consider scenarios where a customer's spending has significantly decreased but not completely stopped. For instance, a customer who previously spent 100k per month but now spends only 10k per month is not labeled as churn, even though this behavior indicates potential disengagement.

---

[1] Dixon, 1999; Floyd, 2000; Slater & Narver, 2000
[2] Invesp - The Importance of Customer Retention

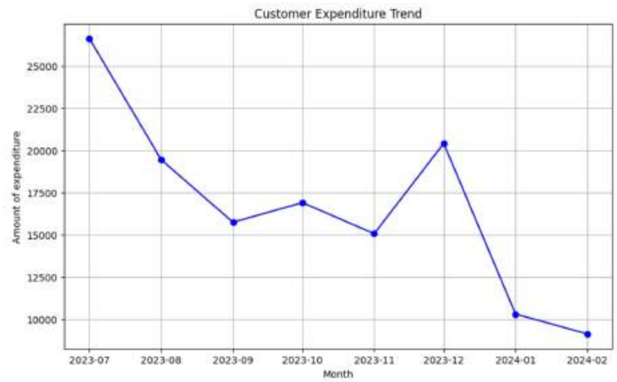




Figure 1

Figure 2

To illustrate visually, Figure 1 shows the monthly spending of a customer labeled as churn in 2024-02 according to the bank's current model. However, the bank's current model does not identify customers with a significant decrease in spending as churn as in Figure 2.

## b. Reactive Approach to Churn Management:

The current model waits for three months of inactivity before taking any action, which means interventions are only initiated at the beginning of the fourth month. It is considerably more difficult to re-engage customers who have not used the card for an extended period (three months). By the time the model identifies them as churn, the likelihood of successfully re-engaging these customers diminishes.

The project aims to address the aforementioned problems by implementing a more sophisticated churn prediction model that detects decreased spending by incorporating changes in spending behavior, such as significant reductions in monthly expenditure, to identify potential churn risks earlier. That is, the project aims to shift from a reactive to a proactive approach by analyzing customers' past behaviors and predicting their future actions.

By using machine learning models to forecast potential churn, this study seeks to label a customer as churn if the prediction indicates a likelihood of churn in the upcoming month, allowing for earlier and potentially more effective intervention, enabling the bank to take preemptive actions to retain customers before they completely disengage.

## 2.2. About the Data

Card spending data of 50.000 customers consisting of X rows and Y columns was worked with, as well as customer demographic data provided by the bank. Below is the explanation of the features in the dataset:

### Demographic Data

- **Customer_id**: Unique identifier for the customer
- **Accommodation**: Accommodation information
- **BranchName**: Name of the bank branch
- **Citizenship**: Citizenship status of the customer
- **Gender**: Gender of the customer
- **MaritalStatus**: Marital status of the customer
- **ProfessionCode**: Code representing the profession of the customer
- **ValueSegment**: Segment value of the customer
- **RecordingTime**: The date when the customer was recorded as a bank customer
- **sesLevel**: Socio-economic status level of the customer
- **EducationLevelPI**: Education level of the customer
- **Age**: Age of the customer
- **EmploymentType**: Type of employment of the customer
- **EmployerSector**: Sector of the employer
- **JobTitle**: Job title of the customer
- **City**: City where the customer resides
- **County**: County where the customer resides

### Card Spending Data

- **Customer_id**: Unique identifier for the customer
- **trxdate**: Transaction date
- **max_taksit**: Maximum installment amount in study period

- **qty**: Total number of transaction in study period
- **TotalAmountLC**: Total amount spent in transaction date in local currency
- **total_amount_market**: Total amount spent on market transactions
- **total_qty_count_market**: Total quantity count of market transactions
- **total_amount_diger**: Total amount spent on other transactions
- **total_qty_count_diger**: Total quantity count of other transactions
- **total_amount_giyim**: Total amount spent on clothing transactions
- **total_qty_count_giyim**: Total quantity count of clothing transactions
- **total_amount_saglik_kozmetik**: Total amount spent on health and cosmetics transactions
- **total_qty_count_saglik_kozmetik**: Total quantity count of health and cosmetics transactions
- **total_amount_sigorta**: Total amount spent on insurance transactions
- **total_qty_count_sigorta**: Total quantity count of insurance transactions
- **total_amount_teknoloji**: Total amount spent on technology transactions
- **total_qty_count_teknoloji**: Total quantity count of technology transactions
- **total_amount_ulasim**: Total amount spent on transportation transactions
- **total_qty_count_ulasim**: Total quantity count of transportation transactions
- **total_amount_yemek**: Total amount spent on food transactions
- **total_qty_count_yemek**: Total quantity count of food transactions

## 2.3. Limitations and Constraints

Thanks to working with a large bank, the data set involved customers from all cities of Turkey, from various age and occupational groups. Although the data was real and up-to-date data from a very large population, there were some challenges and limitations with it.

### a. Data Sharing Issue

Due to the large data size and bank protection policies, the entire dataset could not be shared with all team members. Only one team member had access to the full data, and small samples were provided to the rest of the team. After writing the code[3] using the sample, the code was run with the full dataset, which was time-consuming and required significant effort.

### b. Bias Related to Past Campaigns

The data did not include information on whether customers had been sent campaigns before. Therefore, if customers tended to increase their loyalty after the campaign was sent, this could not be included in data analysis.

### c. Limited Card Spending Information

Although the credit card expenditure information in the data was useful because it included the category and amount of the expenditure, it was monthly rather than continuous. Therefore, it was not possible to conduct analysis on a continuous or weekly basis; therefore only an analysis on a monthly basis could be conducted.

It was also only 9 months of data, so seasonality could not be taken into account in the analysis as there was no data for the past few years. And the success of the forecasts was limited by the scarcity of these data points.

## Churn Definition

First, an understanding of the data was aimed for by deriving metrics such as changes in total customer spending month-to-month and the spending patterns of the top 80% of customers. Individual customer behaviors were also analyzed to determine if different churn behaviors were captured by our definitions. It was known that diminishing spending behavior should be identified as churn, but uncertainty remained about the rate of decrease that should be considered as such. Since customer spending fluctuates, a threshold for the speed of decrease was needed. To establish this threshold, the average rate of decrease was calculated, and one criterion was set: if a customer's spending decreases faster than average, they might be considered at risk of churn.

[3] The link that you may find all relative code files is given in the appendix section.

However, it was observed that even when a customer's spending decreased faster than average, their spending amount for that month could still be relatively high compared to their past purchases. Therefore, a second criterion was added: if a customer's spending decreases, the amount spent in the observed month should also be lower than the average of the previous months.

As seen in Figure 3 below, the customer makes high spending in month 3 and decreases their spending in month 4. However, the spending amount in month 4 is still higher than the average of months 1 and 2 (when investigating churn at month i, the average of the months up until month (i-1) should be compared with the amount of spending in month i). Therefore, the decrease from month 3 to month 4 is not accepted as churn.

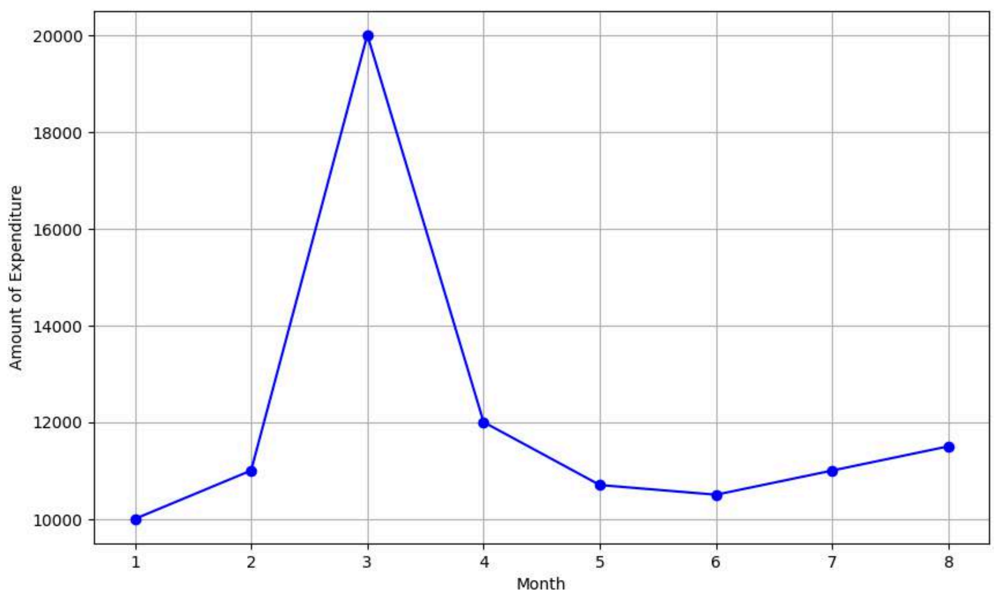
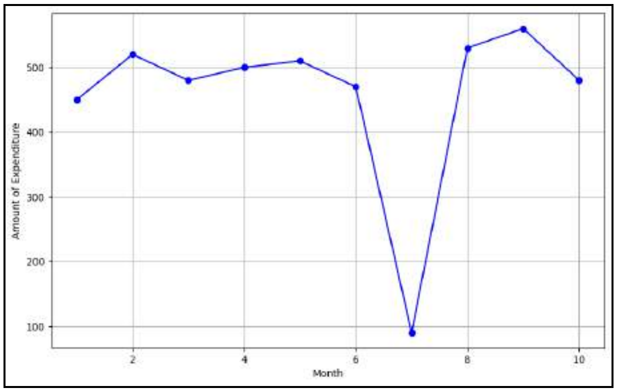
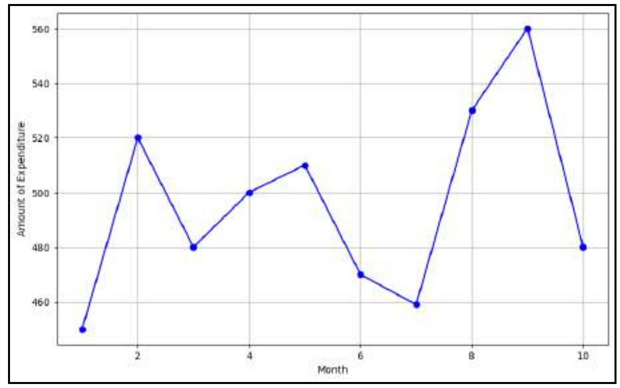



Figure 3

Another observation was made that some customers had unexpectedly high or low purchases in particular months but then returned to their normal behavior. These months were identified as outliers and excluded from the analysis to avoid mistakenly classifying a customer as churned due to an atypical spending spike followed by a return to normal behavior. In our study, a threshold of 2 times the standard deviation was utilized, and spendings that were more than 2 standard deviations above or below the mean were considered as outliers. After the removal of those points, they were replaced with the average of the data to prevent the loss of a data point.




(Before removing the outlier)




(After removing the outlier)

## Data Cleaning

After churn data points were determined, it was needed to merge churn results with first card spending and then demographic data so that we can use that merged data in machine learning models. However, a requirement of the study was the fact that Machine Learning models do not accept null values and data types rather than floats or integers. So, we needed to start with cleaning the dataset.

Starting with null values here were the columns with null values:

| | |
|---|---|
| total_amount_market | 123443 |
| total_qty_count_market | 123443 |
| total_amount_diger | 109923 |
| total_qty_count_diger | 109923 |
| total_amount_giyim | 286029 |
| total_qty_count_giyim | 286029 |
| total_amount_saglik_kozmetik | 316018 |
| total_qty_count_saglik_kozmetik | 316018 |
| total_amount_sigorta | 407822 |
| total_qty_count_sigorta | 407822 |

| | |
|---|---|
| total_amount_teknoloji | 338836 |
| total_qty_count_teknoloji | 338836 |
| total_amount_ulasim | 252845 |
| total_qty_count_ulasim | 252845 |
| total_amount_yemek | 271016 |
| total_qty_count_yemek | 271016 |
| MaritalStatus | 38 |
| sesLevel | 2 |
| EmploymentType | 805 |
| EmployerSector | 3765 |
| JobTitle | 2869 |
| City | 38129 |
| County | 38148 |
| Slope | 42121 |
| Moving_Average | 42121 |

For columns [total_amount_market, total_qty_count_market, total_amount_diger, total_qty_count_diger, total_amount_giyim, total_qty_count_giyim, total_amount_saglik_kozmetik, total_qty_count_saglik_kozmetik, total_amount_sigorta, total_qty_count_sigorta, total_amount_teknoloji, total_qty_count_teknoloji, total_amount_ulasim, total_qty_count_ulasim, total_amount_yemek, total_qty_count_yemek], null values indicate that the customer had no transactions in that month for this category. Therefore, for those columns, the null cells are replaced with "0".

For MaritalStatus, null values are likely indicative of missing information about the customer's marital status. Customers with null values for MaritalStatus were identified, and all columns associated with those customers were deleted. This process ensures data integrity, as solely deleting the rows with null values for MaritalStatus could potentially lead to manipulated data if a customer had transactions but only the row with null MaritalStatus was deleted. The same procedure was applied to sesLevel, as there are a small number of null values in that column, making deletion a reasonable option.

For Count Column, While the County column is important for reflecting the socio-economic standing of the customer's location, sesLevel already serves as an indicator of the customer's socio-economic level. Therefore, the County column has been deleted.

For EmployerSector, null values corresponding to EmploymentType values such as ['Çalışmıyor', 'Emekli Çalışan', 'Serbest Meslek Sahibi'] are replaced with 'Sektör Yok'. Null values associated with EmploymentType values like ['İşveren', 'Tacir/Esnaf'] are replaced with 'Özel Sektör'. Any remaining null values are replaced with “Sektör Yok”.

For EmploymentType, here is the professionCodes of the customer whose EmploymentType is null:

[17: Ev Hanımı, 83:Mobilyacı, 32: Mimar, 39: Öğretmen, 38: Öğrenci, 25: İşletmeci, 30: Memur, 24: İşçi, 60: Nakliyeci, 55: Veteriner Hekim, 34: Mühendis, 50: Teknisyen]

Those customers can be attained to one of Employment types ['Serbest Meslek Sahibi' 'Emekli' 'Emekli Çalışan' 'Öğrenci' 'Ücretli' 'Çalışmıyor' 'İşveren' 'Tacir/Esnaf']

Ev Hanımı, Öğrenci - Çalışmıyor

Mobilyacı, İşletmeci : Tacir/Esnaf

Mimar,Öğretmen, Memur, İşçi, Nakliyeci, Veteriner Hekim, Mühendis: Ücretli

For null City values, the BranchName values of customers are utilized to determine the city where the branch is located, and this corresponding city value is then filled.

For null JobTitle values, the ProfessionCodes of those customers are examined. Since certain professions, such as 'Ev Hanımı', do not have JobTitles, they are replaced with the value 'Unvan Yok'.

For Slope and Moving_Average columns, the null values stem from the fact that slope and moving average are calculated based on past data points. For the data points at the beginning of the period, we cannot calculate the moving average or the slope. Therefore, we simply set null slope and moving average values to 0 as initiation values.

Continuing with datatypes here are the columns with data types other than float or integer :

| | |
|---|---|
| trxdate | object |
| ValueSegment | object |
| RecordingTime | datetime64[ns] |
| EmploymentType | object |
| EmployerSector | object |
| JobTitle | object |
| City | object |

The Trxdate column will be deleted before machine learning, so it can be ignored. For the ValueSegment, EmploymentType, and JobTitle columns, one-hot encoding method is used to convert these columns into integers.

For the City column, cities are first divided into regions, and then one-hot encoding is applied.

Lastly, the RecordingTime indicates the date when the customer became a bank's customer, so by subtracting the recording time from today, the Tenure, which is the number of years the customer has been enrolled, is calculated.

After these steps, the data consists of card spending data, churn status, and demographic data, all ready to be used as input for machine learning models.

# 3. Analysis for Solution/Design Methodology

## 3.1. Literature Overview

The main purpose of this study is to predict customers who identify as churn according to the defined churn in this study. There are many methods to predict future values. A literature search has been made. Initially, a selection is made between forecasting and machine learning models. In this case, in the banking industry & churn prediction related papers indicate that supervised machine learning models are mostly used. Large datasets with defined attributions are common in these kinds of studies, therefore supervised classification methods are used. The literatures are also suggesting a feature selection, which will be mentioned.

In the study, according to the literature search, 5 commonly and successfully used machine learning models are selected:

- Logistic Regression
- Decision Tree
- Random Forest
- Support Vector Machine
- Neural Networks

These models' results will be used in the optimization model. The probabilities of the target attribution are embedded within the models.

## 3.2. Alternative Solution/Design Approaches

30 different subsets of the main dataset are trained and tested with 5 different ML models, with and without feature selection. In order to find the churn rate that should be in the data on which the model was trained, we trained the models with a dataset containing many different rates of churn-labeled customers and the rate that gave the best results was used in the model.

## 3.3. Assumptions

- Since we do not have information about the months before the first month, the initial status of the customers is considered not churn.
- Monthly productivity of non-churn customers is reflected as profit to the bank.
- All customer segments return to the campaign at the same rate.
- If the customer accepts the campaign, he/she is considered as a retention.
- If a campaign is sent a campaign to a customer who would not churn would not churn, this campaign was completely wasted (the benefit to be obtained is negligible).
- Customers who are slowly reducing their spending are not labeled as churn because trying to identify them at the same time as customers who are falling fast results in labeling everyone as churn. Detecting slow churners is the subject of a different study.
- Assuming that the SesLevel value corresponds to the information we will receive from the county, the county has been deleted. Similarly, branch name is also deleted.
- Original MCCPI categories have been edited. Gasoline and Fuel Stations and Car Rental and Vehicle Rental-Sales/Service/Spare Parts categories were brought together under one roof and the category was named Vehicle Expenditures and Rental.

## 3.4. Brief Overview of the Selected Approach(es)

### a. Linear Regression:

It is a statistical method that models the relationship between two or more variables and its goal is to predict how a dependent variable varies with an independent variable. It is based on the formula $y = mx + b$ where:

y: Predicted dependent variable (churn, non-churn)

x: Independent variable

m: Slope (regression coefficient)

b: Point that intersects the Y axis (constant term)

### b. Decision Tree:

Decision tree is a model used in machine learning that performs classification or regression analysis (in the study: classification) by categorizing data sets. It is briefly shown below, how it works:

The decision tree consists of the root node, internal nodes and leaf nodes. Starting from the root node, the data set is separated into branches, and every branch represents a decision made on a feature.

Splitting: Starting from the root node, the feature that best divides the data set is selected and the data set is separated into subgroups through this feature.

Prediction: When a new data point arrives, it is directed to the appropriate branches, starting from the root node in the decision tree, and eventually reaches the leaf node. The value at the leaf node represents the predicted class or value.

### c. Random Forest:

It is an ensemble method of decision trees (models working together). Essentially, it aims to create a stronger and more useful model by combining multiple decision trees.

Ensemble Method: Each tree of the random forest is trained independently, and the final prediction is obtained by combining the predictions of these trees.

Bootstrap Sampling: Different samples are selected from the training data set repeatedly. Each tree is trained on one of these examples. This increases the independence of the trees and the diversity of the model.

Combining Predictions: In classification problems, each tree makes a class prediction and the class with the most votes is selected as the final prediction.

### d. Support Vector Machine (SVM):

It is a powerful machine learning algorithm used for classification analysis. Essentially, SVMs aim to find the best boundary separating data points. SVM tries to find the best hyper-plane (decision boundary) to classify the data.

### e. Neural Networks:

It is a machine learning model inspired by the working principles of the human brain and often used for classification, regression and more complex tasks. In its structure, there are layers and neurons just as in the human neural system. Just as a neural system, it has a constant feedback system, called “Feedforward” and “Backpropogation”.

## 3.5. Integrated IE Skills

- Data Analysis and Statistical Analysis: Statistical analysis and data analysis capabilities were used to analyze customers' past spending and behavior to predict future spending.
- Prediction and Modeling: Prediction and modeling techniques were used to predict customers' spending trends and label churn, which requires knowledge of machine learning and data science.
- Optimization Techniques: Optimization techniques were used to get the best result by analyzing costs and benefits when deciding whether to send a campaign or not.
- System Design and Simulation: System design and simulation capabilities were used during the creation and implementation of the model. This is important to test the accuracy and effectiveness of the model.
- Decision Support Systems: Capabilities were used to develop and implement decision support systems to be used when making the decision to send a campaign.
- Project Management: Project management capabilities were used to plan, implement and monitor the project.
- Process Improvement and Quality Control: Process improvement and quality control methods were used to continuously monitor and improve the performance of the model.

# 4. Development of Alternative Solutions

There is an enormous dataset that limits the efficiency of the workflow. Therefore, 30 different subsets of this dataset are to be tried with 5 different ML models. The apparent result was as the dataset is bigger, the obtained results are better. Also, when the number of churn and non-churn lines are closer to each other, the results were better. According to this knowledge, oversampling is done on the churned customers since it has a lower number of rows.

There are 51169 unique customers between the 2023/07-2024/04-time interval, 45906 of them identified as non-churn customers and 5263 of them identified as churn customers. There are 420755 non-churn lines and 44413 churn lines, it is made equal by oversampling.

The dataset in the model is initially separated into three parts, 70% of the data as train, 15% of the data as validation, 15% of the data as test. The accuracy, precision, recall and F1-score results of those did not differ as much, which means the model is valid, the validation could be terminated in order to give train and test more data. 80% of the data is given to train, 20% of the data is given to test. Additionally, the results indicate that there is no overfitting in the oversampling since the train and test results do not differ by a significant amount.

The literature suggests a feature selection. Models such as decision tree and random forest include feature importance within them. They are used to implement and visualize feature selection and feature importance. Afterwards, they have been tried with every subset and the last subset. Their results are compared to the results without feature selection. The results of models with or without feature selection did not differ significantly, however, those that do not include feature selection were better in terms of the accuracy, precision, recall and F1-score results.

These were the most significant features: Tenure, Age, ProfessionCode, TotalAmountLC, EducationLevelPI.

# 5. Comparison of Alternatives and Recommendation

## 5.1. Numerical studies or evaluation procedure

The evaluation of the machine learning methods is based on accuracy, precision, recall and F1-score.

Accuracy: Measures the aggregate accuracy of the model.

Precision: It shows how many of the positive values are classified as actual positives.

Its calculation: TP/(TP+FP)

Recall: It shows how many of the actual positive values are classified correctly.

Its calculation: TP/(TP/FN)

F1-Score: It is the weighted average of precision and recall. It provides a balanced evaluation of these evaluation metrics.

Its calculation: 2*Precision*Recall/(Precision+Recall)

Importance of Recall:

In this project, the prediction of the churned customers correctly is the main focus. Additionally, the number of customers that churn is much lower than the total number of customers. Therefore, a high score of recall would imply that the model is viable to use.

## 5.2. Solution Proposal and Justification

All machine learning methods are evaluated in terms of the predefined metrics. Three of them could not demonstrate valuable performance even if they were widely used techniques in the literature search regarding this project's subject. Those are: Logistic Regression, Support Vector Machine, and Neural Networks.

The dataset used in this project has too many dimensions. Also, the data is real data, it is not generated. Naturally, it has many noises. Support Vector Machine does not perform well under these circumstances.

Neural Networks has comparably better results than these two, however, they were not enough. Neural Networks method is very sensitive to the dataset's size. If it is big, it could overfit. If it is small, it does not perform well.

These two models have been brought up until some point, however, because of the data availability and library issues of the computers used in this study, It could not be preceded further. They were not as promising as the other models until to the point where they left when the other models were at the same level.

The best performing models in the project are Random Forest and Decision Tree.

**Results of the Best Decision Tree Iteration**

```
Churn Samples: 39374, Non-Churn Samples: 91872
Churn Ratio: 30.0% | Max Depth: 10
Validation Metrics:
Accuracy: 0.8896733885305024
Precision: 0.7866070050918068
Recall: 0.8659758790555461
F1 Score: 0.824385510996119

Test Metrics:
Accuracy: 0.8897241834713262
Precision: 0.7904440006145337
Recall: 0.8644153225806451
F1 Score: 0.8257764224380066

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.94      0.96     45852
           1       0.63      0.83      0.72      5317

    accuracy                           0.93     51169
   macro avg       0.80      0.89      0.84     51169
weighted avg       0.94      0.93      0.94     51169

Confusion Matrix for April 2024:
[[43250  2602]
 [  887  4430]]
```

**Results of the Best Random Forest Iteration**

```
Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.96      0.96     45852
           1       0.67      0.76      0.71      5317

    accuracy                           0.94     51169
   macro avg       0.82      0.86      0.84     51169
weighted avg       0.94      0.94      0.94     51169

Confusion Matrix for April 2024:
[[43829  2023]
 [ 1264  4053]]
```

Results of the Best Logistic Regression Iteration

```
Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.96      0.58      0.72     45852
           1       0.17      0.77      0.28      5317

    accuracy                           0.60     51169
   macro avg       0.57      0.67      0.50     51169
weighted avg       0.87      0.60      0.67     51169

Confusion Matrix for April 2024:
[[26484 19368]
 [ 1218  4099]]
```

According to the results, Decision Tree method is chosen with 30% Churn rate data and 10 max depth for the prediction of the churned customers prediction. Their embedded probability in the model will be used in the optimization model.

The confusion matrix of the model is as follows:

| | | 1 | 0 | | | |
|---|---|---|---|---|---|---|
| | Observed/Predicted | TRUE | FALSE | SUM | Percentage Correct (%) | Error (%) |
| 1 | TRUE | 43250 | 2602 | 45852 | 0.845238328 | 0.0508511 |
| 0 | FALSE | 887 | 4430 | 5317 | 0.086575856 | 0.01733471 |
| | SUM | 44137 | 7032 | 51169 | 0.931814184 | 0.06818582 |

Limitations:

Limited time interval, a longer time interval would strengthen the prediction models.

The data used in the project has monthly monthly information. Daily information would be better.

Data availability is limited for all the contributors.

## 5.3. Further Assessment of the Recommended Solution

Each month, changing the efficiency values used by the model according to that month's efficiency expectations is sufficient for monthly evaluation. There is even the flexibility to differentiate the model by changing the campaign costs and number. In this way, the model is designed to be sustainable.

In addition, even different companies can adapt the model for themselves, and it can be applied for different customer bases. The only requirement of the model is the knowledge of customer value, segment, churn probability, and likelihood of accepting the campaign.

The proposed solution method leaves no room for ambiguity in customer selection by determining which customer to send the campaign to. In addition to this, a very robust solution is obtained by determining the appropriate cost for the campaign for a specific customer, offered according to the campaign options given to it.

---

# 6. Suggestions for a Successful Implementation

## 6.1. Model Monitoring

Each model has a lifecycle, from inception to deployment and eventual decommissioning, and can experience drift over time. Consequently, ongoing monitoring of model performance at specific intervals is necessary. As part of this process, monthly churn analysis should be conducted. Each month, the latest model predictions should be revisited and compared to actual data, with the confusion matrix recorded for detailed analysis.

One of the most significant metrics in this analysis is recall, as it measures the ratio of correctly predicted churn customers to the total number of actual churn customers; this is why ensuring high recall is critical for taking proactive measures to retain these customers. To maintain the model's effectiveness, a threshold for recall will be established. This threshold will act as a benchmark for acceptable performance. If the recall metric falls below this threshold, it will indicate that the model is not performing adequately in identifying churn customers.

By implementing this comprehensive monitoring system, deviations in model performance can be promptly addressed. This proactive approach will ensure that the model continues to perform effectively over time, providing accurate predictions and supporting informed decision-making.

## 6.2. Optimization

After comparing several machine learning models and selecting the most successful one, the data obtained from this model was used in the optimization step. In this step, the churn probability obtained by machine learning was used together with the campaign acceptance probability of each segment, which varies according to the campaigns, to determine which type of campaigns should be sent to which customers.

In the bank's current categorization, customers are divided into 5 segments from A1 to D, A1 with the highest return and D with the lowest return. For each segment, the bank determines a monthly productivity value. These values are calculated by taking into account factors such as the amount and frequency of customer spending and other expenses of the bank. These productivity values given by the bank are used in the optimization equation as they represent the money that the bank would earn if the customers do not churn in the predicted month.

3 different campaign types with 3 different costs were created and the 2% value given by the bank as the average campaign acceptance rate was used as the acceptance rate for the average-cost campaign, C2. This rate was used in the model with small differences according to segments, which are seen in the Base Acceptance Rate column in Table 1. The other campaigns were found by multiplying the acceptance probability of the C2 campaign by 0.6 or 1.7 times.

| Campaign Types | Cost | Effect Rate |
|---|---|---|
| **C1** | 75 | 0.6 |
| **C2** | 125 | 1 |
| **C3** | 250 | 1.8 |

**Table 2**

| Segment | Monthly Value | Base Acceptance Rate |
|---|---|---|
| **A1** | 22558.55 | 0.015 |
| **A2** | 1894.69 | 0.018 |
| **B** | 824.30 | 0.020 |
| **C** | 215.28 | 0.025 |
| **D** | 27.79 | 0.030 |
| **Table 1** | | |

## Optimization Model

**Sets:**

I: Set of customers

C: Set of campaigns

**Parameters:**

Pi: Probability of customer i to churn in the following month (derived from machine learning)

Aic: Probability of customer i to accept campaign c (Ai * Ec in the code, Ec being the effect of campaign c to acceptance probability)

Cc: Cost of campaign c

Vi: Monthly value of customer i

M: Marketing Budget

B: probability of a customer labeled as churn is actually churn

D: probability of a customer labeled as non-churn is actually non-churn

**Decision Variables:**

Xic = 1 If campaign c is send to customer i

0 otw

**Objective Function:**

max z = ΣI ΣC Vi [(Pi · B · 0.5) · ((Xic · Aic) - (1 - Xic)) + (1 - Pi) · D] - Xic · Cc · Aic

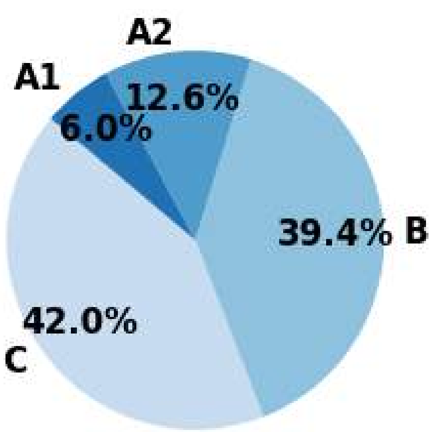
**Constraints:**

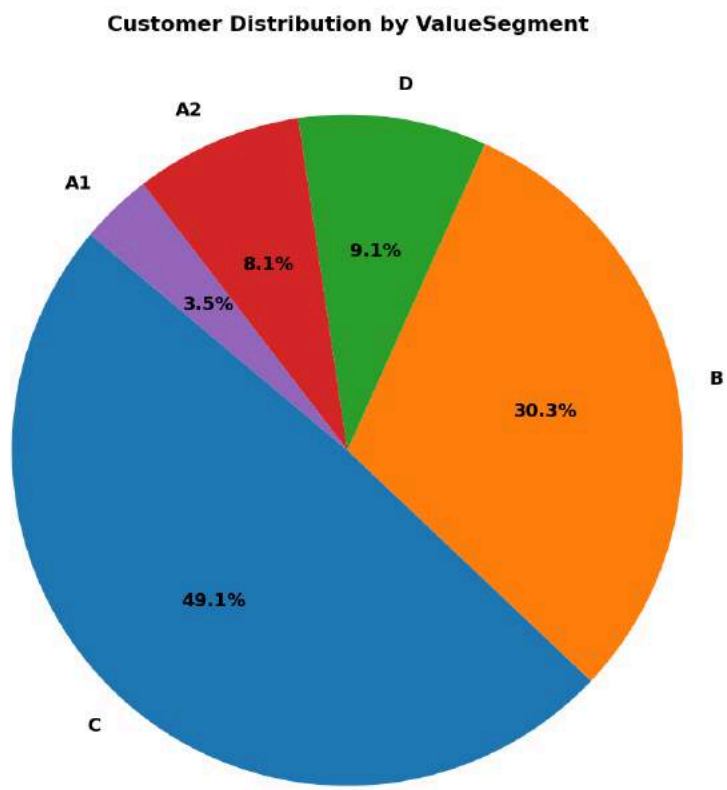
ΣI ΣC Xic * Cc ≤ M (budget constraint)

ΣC Xic ≤ 1 , ∀i ∈ I (at most one campaign to one customer)

Xic ∈ {0,1}, ∀i ∈ I, ∀c ∈ C

## Results



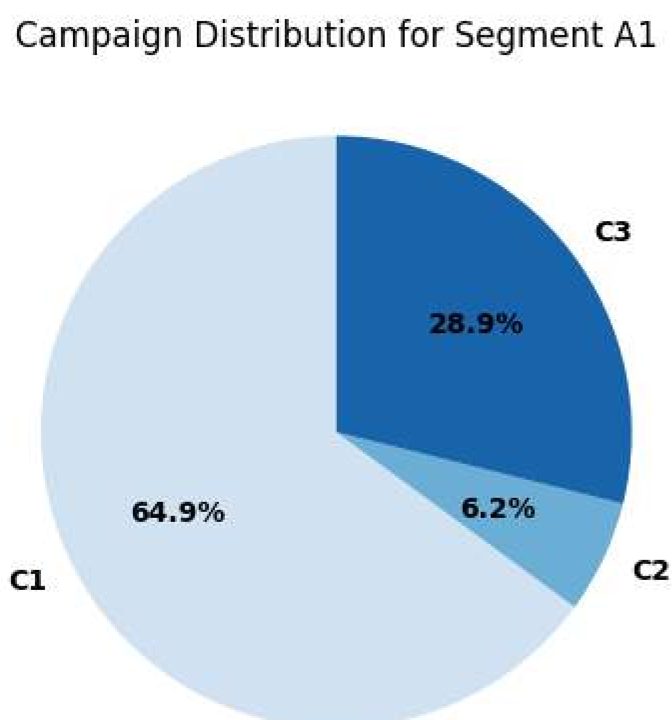




The optimization model was run on 51.169 of the customers due to computer processing speed constraints. When the pie chart is analyzed which shows the segments of these customers, it is seen that approximately half of the customers belong to C class. 30% of them belong to class B. It can be said that a2 and d classes have similar proportions and the least common class is A1 with 3%.

When the model was used on a data set of exactly 51.169 customers from 5 segments, the results were in line with expectations. Using the entire budget of 400,000 liras given by the bank, the model decided to send campaigns to a total of 5113 customers. The objective value is 169,539,234.66, which is nearly 2 million liras more than expected profit of 167,768,220 if the model is not used. That is, 1% increase in total earnings of the bank from credit card usage.

When we look at the customers to whom campaigns were sent, the first thing that draws attention is that no campaign was sent to any customer from the D segment. This is not surprising as the cost of even the lowest cost campaign is about three months of expected profit for a D segment customer.

It is also noteworthy that only the lowest cost campaign was sent to the A2, B and C segments. Higher cost campaigns were only sent to customers from the A1 segment. Again, this is to be expected, as the expected monthly return from the A1 segment is much higher than all other segments. While the expected value of other segments is close to each other, the expected return of the A1 segment is about 12 times higher than even the closest segment, A2.



# 7. Conclusions and Discussion

The project achieved its objectives successfully by utilizing a comprehensive array of Industrial Engineering (IE) tools, techniques, and methods. Initially, rigorous data analysis and preprocessing were undertaken to handle missing values, normalize data, and employ one-hot encoding for categorical variables. This step ensured the dataset was clean and ready for analysis. Subsequently, various machine learning models were deployed to predict customer churn. After comparing these models, the most successful one was selected based on its performance metrics.

The selected model's output was then integrated into the optimization step. This step utilized the churn probability obtained from the machine learning model along with the campaign acceptance probability of each customer segment to determine the most suitable type of campaign for each customer. The bank's current categorization divides customers into five segments (A1 to D), with A1 having the highest return and D the lowest. Monthly productivity values for each segment, provided by the bank, were incorporated into the optimization equation. These values represent the potential earnings if customers do not churn in the predicted month.

Three different campaign types with varying costs were designed, and the base acceptance rate was established using the average acceptance rate provided by the bank for the mid-cost campaign (C2) and the acceptance probabilities for other campaigns were adjusted accordingly. The optimization model was formulated with specific sets, parameters, and decision variables. The objective function aimed to maximize the expected return, considering the probabilities of churn, campaign acceptance rates, and associated costs, while adhering to budget constraints. The decision variables determined whether a campaign should be sent to a particular customer.

To implement the model, a subset of 12,000 customers was used due to computational limitations. The results indicated that the model effectively utilized the entire marketing budget of 350,000 liras, sending campaigns to 4,512 customers and achieving an objective value significantly higher than the expected profit without the model. The model strategically avoided sending campaigns to the D segment due to their low expected return and allocated higher-cost campaigns to the A1 segment, which had the highest expected return. Overall, the integration of data preprocessing, machine learning, and optimization techniques was crucial in processing the dataset, accurately predicting customer churn, and effectively optimizing marketing strategies to enhance customer retention and maximize returns.

The design presented in this project showcased several merits and held significant implications. Firstly, the integration of multiple machine learning models and optimization techniques led to highly accurate predictions of customer churn, significantly enhancing the bank's ability to retain valuable customers. The design also proved to be cost-effective; by accurately targeting the right customers with appropriate marketing campaigns, the bank could utilize its marketing budget more efficiently, thus maximizing returns. This innovation represents a proactive approach to predicting customer churn, shifting from a reactive to a proactive customer retention strategy, which is a significant advancement in customer relationship management.

Moreover, the design has several notable impacts across economic, environmental, ethical, and societal dimensions. Economically, it contributes to increased revenue by retaining more customers and optimizing marketing expenditures, leading to significant cost savings. Environmentally, the optimized marketing campaigns reduce waste by minimizing ineffective marketing efforts, thereby lowering both paper and digital waste. Societally, the design enhances customer satisfaction through better-targeted marketing campaigns, improving the overall customer experience. Furthermore, by improving business performance, the bank can contribute to economic stability and growth within the community. Overall, the design not only provides economic benefits to the bank but also positively impacts societal well-being.

As a further implication, campaign costs can be fully customized for each customer. The assumption that the customer who accepts the campaign is really not a churned customer can be based on a solid foundation by making the campaign spending amounts in a way that prevents them from falling below the level that needs to fall for the customer to be a churn when he accepts it.

## References


- Glady, N., Baesens, B., & Croux, C. (2009). Modeling churn using customer lifetime value. European Journal of Operational Research, 197(1), 402-411.
- Cramér, H. (1946). Mathematical Methods of Statistics. Princeton. Press, Princeton, NJ, 367-369.
- Nicolas Glady a,*, Bart Baesens a,b,c, Christophe Croux, 2008
- Wouter Buckinx, Dirk Van den Poel, 2002
- AL-Najjar D, Al-Rousan N, and AL-Najjar H. (2022). Machine Learning to Develop Credit Card Customer Churn Prediction. Journal of Theoretical and Applied Electronic Commerce Research. 17(4), 1529-1542.
- Ullah I., Raza B., Malik A. K., Imran M., Islam S. U., and Kim S. W., (2019) "A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector," in IEEE Access, vol. 7, pp. 60134-60149.
- J. Runge, P. Gao, F. Garcin and B. Faltings, (2014) "Churn prediction for high-value players in casual social games," 2014 IEEE Conference on Computational Intelligence and Games, Dortmund, Germany, pp. 1-8.


## Appendix

Here you may find the relative files of the project : [https://github.com/eylulrana/IE492](https://github.com/eylulrana/IE492)
Note: Due to the large size of the real dataset and the restriction that it cannot be taken out of the bank's laptop, only a sample of the data is shared on GitHub.

Random Forest

```
Churn Ratio: 60.0% | n_estimators: 50
Validation Metrics:
Accuracy: 0.8907853296759118
Precision: 0.8797970187123375
Recall: 0.9459505541346973
F1 Score: 0.9116752937309999

Test Metrics:
Accuracy: 0.8924217797643235
Precision: 0.8801894238358327
Recall: 0.9489448604492853
F1 Score: 0.9132749160592908

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.90      0.94     45852
           1       0.51      0.88      0.64      5317

    accuracy                           0.90     51169
   macro avg       0.75      0.89      0.79     51169
weighted avg       0.94      0.90      0.91     51169

Confusion Matrix for April 2024:
[[41320  4532]
 [  638  4679]]
```

```
Churn Ratio: 60.0% | n_estimators: 100
Validation Metrics:
Accuracy: 0.8882454536218632
Precision: 0.8749016522423289
Recall: 0.947996589940324
F1 Score: 0.9099836333878887

Test Metrics:
Accuracy: 0.8890694839496139
Precision: 0.877763739734681
Recall: 0.9458815520762424
F1 Score: 0.9105504587155964

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.90      0.94     45852
           1       0.50      0.88      0.64      5317

    accuracy                           0.90     51169
   macro avg       0.74      0.89      0.79     51169
weighted avg       0.93      0.90      0.91     51169

Confusion Matrix for April 2024:
[[41207  4645]
 [  638  4679]]
```

```
Churn Ratio: 60.0% | n_estimators: 200
Validation Metrics:
Accuracy: 0.8920044701818551
Precision: 0.8791850915982312
Recall: 0.9491901108269395
F1 Score: 0.9128474214970895

Test Metrics:
Accuracy: 0.8926249492076391
Precision: 0.8804673930206853
Recall: 0.9489448604492853
F1 Score: 0.9134245228929478

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.99      0.90      0.94     45852
           1       0.51      0.88      0.65      5317

    accuracy                           0.90     51169
   macro avg       0.75      0.89      0.80     51169
weighted avg       0.94      0.90      0.91     51169

Confusion Matrix for April 2024:
[[41417  4435]
 [  629  4688]]
```

```
Churn Ratio: 50.0% | n_estimators: 50
Validation Metrics:
Accuracy: 0.8824077209617338
Precision: 0.8625530525628469
Recall: 0.9061910478477105
F1 Score: 0.8838337375595885

Test Metrics:
Accuracy: 0.8767459578430543
Precision: 0.8620080321285141
Recall: 0.8998826094247862
F1 Score: 0.880538234328848

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.93      0.95     45852
           1       0.58      0.82      0.68      5317

    accuracy                           0.92     51169
   macro avg       0.78      0.88      0.82     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42748  3104]
 [  944  4373]]
```

```
Churn Ratio: 50.0% | n_estimators: 100
Validation Metrics:
Accuracy: 0.8838469353200136
Precision: 0.8655517297917692
Recall: 0.905333561996227
F1 Score: 0.8849958088851635

Test Metrics:
Accuracy: 0.8811478879200088
Precision: 0.866067127027461
Recall: 0.9044105316116049
F1 Score: 0.8848236259228875

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.93      0.96     45852
           1       0.59      0.83      0.69      5317

    accuracy                           0.92     51169
   macro avg       0.78      0.88      0.82     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42766  3086]
 [  911  4406]]
```

```
Churn Ratio: 50.0% | n_estimators: 200
Validation Metrics:
Accuracy: 0.8844395529969522
Precision: 0.8646309601567603
Recall: 0.9080775167209741
F1 Score: 0.8858218318695107

Test Metrics:
Accuracy: 0.8811478879200088
Precision: 0.8650120096076861
Recall: 0.9059198390072112
F1 Score: 0.8849934469200524

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.93      0.96     45852
           1       0.59      0.83      0.69      5317

    accuracy                           0.92     51169
   macro avg       0.78      0.88      0.82     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42766  3086]
 [  914  4403]]
```

```
Churn Ratio: 40.0% | n_estimators: 50
Validation Metrics:
Accuracy: 0.8709109380291229
Precision: 0.8513676242903836
Recall: 0.8260724419963278
F1 Score: 0.838529312097594

Test Metrics:
Accuracy: 0.8748476229175132
Precision: 0.8480280455740579
Recall: 0.8314143323595119
F1 Score: 0.8396390142311698

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.95      0.96     45852
           1       0.66      0.76      0.71      5317

    accuracy                           0.93     51169
   macro avg       0.81      0.86      0.83     51169
weighted avg       0.94      0.93      0.94     51169

Confusion Matrix for April 2024:
[[43751  2101]
 [ 1273  4044]]
```

```
Churn Ratio: 40.0% | n_estimators: 100
Validation Metrics:
Accuracy: 0.8736200474094141
Precision: 0.8573904002772483
Recall: 0.8259055249540979
F1 Score: 0.8413535113076008

Test Metrics:
Accuracy: 0.8761343627251794
Precision: 0.850861765740415
Recall: 0.8314143323595119
F1 Score: 0.8410256410256409

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.96      0.96     45852
           1       0.66      0.76      0.71      5317

    accuracy                           0.93     51169
   macro avg       0.82      0.86      0.84     51169
weighted avg       0.94      0.93      0.94     51169

Confusion Matrix for April 2024:
[[43791  2061]
 [ 1267  4050]]
```

```
Churn Ratio: 40.0% | n_estimators: 200
Validation Metrics:
Accuracy: 0.8736200474094141
Precision: 0.8582595101615424
Recall: 0.8247371056584877
F1 Score: 0.841164453524004

Test Metrics:
Accuracy: 0.876540701611811
Precision: 0.8523809523809524
Recall: 0.8305550781921293
F1 Score: 0.8413264862041954

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.96      0.96     45852
           1       0.67      0.76      0.71      5317

    accuracy                           0.94     51169
   macro avg       0.82      0.86      0.84     51169
weighted avg       0.94      0.94      0.94     51169

Confusion Matrix for April 2024:
[[43829  2023]
 [ 1264  4053]]
```

Logistic Regression

```
Churn Ratio: 50.0%
Validation Metrics:
Accuracy: 0.6352014900101591
Precision: 0.5923319582625576
Recall: 0.8372491853884411
F1 Score: 0.693810843459106

Test Metrics:
Accuracy: 0.6380259036654533
Precision: 0.6014187808103884
Recall: 0.8388395103136005
F1 Score: 0.7005602240896358

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.96      0.58      0.72     45852
           1       0.17      0.77      0.28      5317

    accuracy                           0.60     51169
   macro avg       0.57      0.67      0.50     51169
weighted avg       0.87      0.60      0.67     51169

Confusion Matrix for April 2024:
[[26484 19368]
 [ 1218  4099]]
```

```
Churn Ratio: 40.0%
Validation Metrics:
Accuracy: 0.6009481882831019
Precision: 0.6386554621848739
Recall: 0.03805708562844266
F1 Score: 0.0718336483931947

Test Metrics:
Accuracy: 0.6099146688338074
Precision: 0.5896656534954408
Recall: 0.03333906169444922
F1 Score: 0.06310995445673339

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.90      0.99      0.94     45852
           1       0.35      0.03      0.05      5317

    accuracy                           0.89     51169
   macro avg       0.62      0.51      0.50     51169
weighted avg       0.84      0.89      0.85     51169

Confusion Matrix for April 2024:
[[45558   294]
 [ 5162   155]]
```

```
Churn Ratio: 30.0%
Validation Metrics:
Accuracy: 0.7009193884289124
Precision: 0.48
Recall: 0.0020383896721589945
F1 Score: 0.0040595399188092015

Test Metrics:
Accuracy: 0.6976177172753594
Precision: 0.48484848484848486
Recall: 0.002688172043010753
F1 Score: 0.005346700083542189

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.90      1.00      0.95     45852
           1       0.45      0.00      0.01      5317

    accuracy                           0.90     51169
   macro avg       0.67      0.50      0.48     51169
weighted avg       0.85      0.90      0.85     51169

Confusion Matrix for April 2024:
[[45825    27]
 [ 5295    22]]
```

```
Churn Ratio: 20.0%
Validation Metrics:
Accuracy: 0.8002031832035218
Precision: 0.0
Recall: 0.0
F1 Score: 0.0

Test Metrics:
Accuracy: 0.8010903796010972
Precision: 0.0
Recall: 0.0
F1 Score: 0.0
```

```
              precision    recall  f1-score   support

           0       0.90      1.00      0.95     45852
           1       0.00      0.00      0.00      5317

    accuracy                           0.90     51169
   macro avg       0.45      0.50      0.47     51169
weighted avg       0.80      0.90      0.85     51169

Confusion Matrix for April 2024:
[[45852     0]
 [ 5317     0]]
```

Decision Tree

```
Churn Samples: 39374, Non-Churn Samples: 39374
Churn Ratio: 50.0% | Max Depth: 10
Validation Metrics:
Accuracy: 0.899339654588554
Precision: 0.8592879256965944
Recall: 0.9519807923169268
F1 Score: 0.9032625498332112

Test Metrics:
Accuracy: 0.8961313806822991
Precision: 0.8586791881248107
Recall: 0.9506959584101962
F1 Score: 0.9023477914842817

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.99      0.92      0.95     45852
           1       0.56      0.93      0.70      5317

    accuracy                           0.92     51169
   macro avg       0.78      0.92      0.83     51169
weighted avg       0.95      0.92      0.93     51169

Confusion Matrix for April 2024:
[[41998  3854]
 [  372  4945]]
```

```
Churn Samples: 39374, Non-Churn Samples: 39374
Churn Ratio: 50.0% | Max Depth: 15
Validation Metrics:
Accuracy: 0.8880799187267185
Precision: 0.8622850715089185
Recall: 0.9202538158120391
F1 Score: 0.890326862452298

Test Metrics:
Accuracy: 0.8876661305341573
Precision: 0.8635508155583438
Recall: 0.9233607244675499
F1 Score: 0.8924548180565687

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.99      0.92      0.95     45852
           1       0.57      0.90      0.70      5317

    accuracy                           0.92     51169
   macro avg       0.78      0.91      0.83     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42291  3561]
 [  557  4760]]
```

```
Churn Samples: 39374, Non-Churn Samples: 39374
Churn Ratio: 50.0% | Max Depth: 15
Validation Metrics:
Accuracy: 0.8880799187267185
Precision: 0.8622850715089185
Recall: 0.9202538158120391
F1 Score: 0.890326862452298

Test Metrics:
Accuracy: 0.8876661305341573
Precision: 0.8635508155583438
Recall: 0.9233607244675499
F1 Score: 0.8924548180565687

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.99      0.92      0.95     45852
           1       0.57      0.90      0.70      5317

    accuracy                           0.92     51169
   macro avg       0.78      0.91      0.83     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42291  3561]
 [  557  4760]]
```

```
Churn Samples: 39374, Non-Churn Samples: 39374
Churn Ratio: 50.0% | Max Depth: 25
Validation Metrics:
Accuracy: 0.8692854724009482
Precision: 0.8629974597798475
Recall: 0.8739495798319328
F1 Score: 0.8684389911383776

Test Metrics:
Accuracy: 0.8699737577245408
Precision: 0.8649629018961253
Recall: 0.879758510816703
F1 Score: 0.8722979714000665

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.93      0.95     45852
           1       0.58      0.86      0.69      5317

    accuracy                           0.92     51169
   macro avg       0.78      0.89      0.82     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42546  3306]
 [  761  4556]]
```

```
Churn Samples: 39374, Non-Churn Samples: 39374
Churn Ratio: 50.0% | Max Depth: 30
Validation Metrics:
Accuracy: 0.8651371486623772
Precision: 0.8620984278879016
Recall: 0.8652032241468016
F1 Score: 0.8636480356072926

Test Metrics:
Accuracy: 0.8700584102260221
Precision: 0.8659504132231405
Recall: 0.8785846050645648
F1 Score: 0.8722217597602597

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.93      0.95     45852
           1       0.58      0.85      0.69      5317

    accuracy                           0.92     51169
   macro avg       0.78      0.89      0.82     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42554  3298]
 [  811  4506]]
```

```
Churn Samples: 39374, Non-Churn Samples: 39374
Churn Ratio: 50.0% | Max Depth: 35
Validation Metrics:
Accuracy: 0.8650524889942431
Precision: 0.863316755273538
Recall: 0.863316755273538
F1 Score: 0.863316755273538

Test Metrics:
Accuracy: 0.8672648776771353
Precision: 0.8664332166083042
Recall: 0.8713734697300016
F1 Score: 0.8688963210702342

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.93      0.95     45852
           1       0.58      0.85      0.69      5317

    accuracy                           0.92     51169
   macro avg       0.78      0.89      0.82     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42567  3285]
 [  809  4508]]
```

```
Churn Samples: 39374, Non-Churn Samples: 39374
Churn Ratio: 50.0% | Max Depth: 40
Validation Metrics:
Accuracy: 0.8623433796139519
Precision: 0.8629380286552736
Recall: 0.8573143543131538
F1 Score: 0.8601169993117687

Test Metrics:
Accuracy: 0.865995090154914
Precision: 0.865975935828877
Recall: 0.8690256582257253
F1 Score: 0.8674981166820122

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.93      0.95     45852
           1       0.58      0.84      0.69      5317

    accuracy                           0.92     51169
   macro avg       0.78      0.89      0.82     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42599  3253]
 [  843  4474]]
```

```
Churn Samples: 39374, Non-Churn Samples: 59060
Churn Ratio: 40.0% | Max Depth: 10
Validation Metrics:
Accuracy: 0.8877074161869285
Precision: 0.8286060973760049
Recall: 0.9118678017025539
F1 Score: 0.8682453909726637

Test Metrics:
Accuracy: 0.8894080996884736
Precision: 0.8225955610357584
Recall: 0.9171678982643066
F1 Score: 0.8673112862598521

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.99      0.93      0.96     45852
           1       0.59      0.89      0.71      5317

    accuracy                           0.92     51169
   macro avg       0.79      0.91      0.83     51169
weighted avg       0.95      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42512  3340]
 [  568  4749]]
```

```
Churn Samples: 39374, Non-Churn Samples: 59060
Churn Ratio: 40.0% | Max Depth: 10
Validation Metrics:
Accuracy: 0.8877074161869285
Precision: 0.8286060973760049
Recall: 0.9118678017025539
F1 Score: 0.8682453909726637

Test Metrics:
Accuracy: 0.8894080996884736
Precision: 0.8225955610357584
Recall: 0.9171678982643066
F1 Score: 0.8673112862598521

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.99      0.93      0.96     45852
           1       0.59      0.89      0.71      5317

    accuracy                           0.92     51169
   macro avg       0.79      0.91      0.83     51169
weighted avg       0.95      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42512  3340]
 [  568  4749]]
```

```
Churn Samples: 39374, Non-Churn Samples: 59060
Churn Ratio: 40.0% | Max Depth: 20
Validation Metrics:
Accuracy: 0.8736200474094141
Precision: 0.8322861285645239
Recall: 0.8622934401602403
F1 Score: 0.8470241023118544

Test Metrics:
Accuracy: 0.8721386970066368
Precision: 0.8206885299396313
Recall: 0.864409692387008
F1 Score: 0.841981921660529

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.93      0.96     45852
           1       0.60      0.84      0.70      5317

    accuracy                           0.92     51169
   macro avg       0.79      0.89      0.83     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42860  2992]
 [  870  4447]]
```

```
Churn Samples: 39374, Non-Churn Samples: 59060
Churn Ratio: 40.0% | Max Depth: 25
Validation Metrics:
Accuracy: 0.8650186251269895
Precision: 0.829976890062727
Recall: 0.8392588883324987
F1 Score: 0.8345920823304839

Test Metrics:
Accuracy: 0.8662467831504809
Precision: 0.8212972250083584
Recall: 0.8443031448702526
F1 Score: 0.8326413015846115

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.94      0.96     45852
           1       0.60      0.82      0.69      5317

    accuracy                           0.92     51169
   macro avg       0.79      0.88      0.82     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42914  2938]
 [  949  4368]]
```

```
Churn Samples: 39374, Non-Churn Samples: 59060
Churn Ratio: 40.0% | Max Depth: 30
Validation Metrics:
Accuracy: 0.8631899762952929
Precision: 0.8314137873476882
Recall: 0.8314137873476882
F1 Score: 0.8314137873476882

Test Metrics:
Accuracy: 0.8633346877962887
Precision: 0.8225012726964195
Recall: 0.8329609898608008
F1 Score: 0.8276980874316939

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.94      0.96     45852
           1       0.60      0.81      0.69      5317

    accuracy                           0.92     51169
   macro avg       0.79      0.87      0.82     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[42983  2869]
 [ 1017  4300]]
```

```
Churn Samples: 39374, Non-Churn Samples: 59060
Churn Ratio: 40.0% | Max Depth: 35
Validation Metrics:
Accuracy: 0.8631899762952929
Precision: 0.8340905266700319
Recall: 0.8274077783341679
F1 Score: 0.8307357130886542

Test Metrics:
Accuracy: 0.8614384396586753
Precision: 0.8218734004436103
Recall: 0.8278054648565045
F1 Score: 0.8248287671232877

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.94      0.96     45852
           1       0.60      0.80      0.69      5317

    accuracy                           0.92     51169
   macro avg       0.79      0.87      0.82     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[43011  2841]
 [ 1051  4266]]
```

```
Churn Samples: 39374, Non-Churn Samples: 59060
Churn Ratio: 40.0% | Max Depth: 40
Validation Metrics:
Accuracy: 0.8602099559769726
Precision: 0.8316162810336092
Recall: 0.8218995159405775
F1 Score: 0.8267293485560779

Test Metrics:
Accuracy: 0.8585940674522552
Precision: 0.8212502152574479
Recall: 0.8195566248496305
F1 Score: 0.8204025460175468

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.94      0.96     45852
           1       0.60      0.80      0.68      5317

    accuracy                           0.92     51169
   macro avg       0.79      0.87      0.82     51169
weighted avg       0.94      0.92      0.93     51169

Confusion Matrix for April 2024:
[[43027  2825]
 [ 1081  4236]]
```

```
Churn Samples: 39374, Non-Churn Samples: 91872
Churn Ratio: 30.0% | Max Depth: 10
Validation Metrics:
Accuracy: 0.8896733885305024
Precision: 0.7866070050918068
Recall: 0.8659758790555461
F1 Score: 0.824385510996119

Test Metrics:
Accuracy: 0.8897241834713262
Precision: 0.7904440006145337
Recall: 0.8644153225806451
F1 Score: 0.8257764224380066

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.94      0.96     45852
           1       0.63      0.83      0.72      5317

    accuracy                           0.93     51169
   macro avg       0.80      0.89      0.84     51169
weighted avg       0.94      0.93      0.94     51169

Confusion Matrix for April 2024:
[[43250  2602]
 [  887  4430]]
```

```
Churn Samples: 39374, Non-Churn Samples: 91872
Churn Ratio: 30.0% | Max Depth: 15
Validation Metrics:
Accuracy: 0.8854574084421192
Precision: 0.7877059569074778
Recall: 0.8445727874978767
F1 Score: 0.8151487826871053

Test Metrics:
Accuracy: 0.8877431807791943
Precision: 0.7979299363057325
Recall: 0.8419018817204301
F1 Score: 0.8193263570961413

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.94      0.96     45852
           1       0.63      0.81      0.71      5317

    accuracy                           0.93     51169
   macro avg       0.80      0.88      0.83     51169
weighted avg       0.94      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43265  2587]
 [ 1005  4312]]
```

```
Churn Samples: 39374, Non-Churn Samples: 91872
Churn Ratio: 30.0% | Max Depth: 20
Validation Metrics:
Accuracy: 0.8753492152181642
Precision: 0.7800620003263176
Recall: 0.812128418549346
F1 Score: 0.7957723035952065

Test Metrics:
Accuracy: 0.8773302179102961
Precision: 0.7841876908243612
Recall: 0.8198924731182796
F1 Score: 0.8016427104722792

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.94      0.96     45852
           1       0.62      0.80      0.70      5317

    accuracy                           0.93     51169
   macro avg       0.80      0.87      0.83     51169
weighted avg       0.94      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43253  2599]
 [ 1088  4229]]
```

```
Churn Samples: 39374, Non-Churn Samples: 91872
Churn Ratio: 30.0% | Max Depth: 20
Validation Metrics:
Accuracy: 0.8753492152181642
Precision: 0.7800620003263176
Recall: 0.812128418549346
F1 Score: 0.7957723035952065

Test Metrics:
Accuracy: 0.8773302179102961
Precision: 0.7841876908243612
Recall: 0.8198924731182796
F1 Score: 0.8016427104722792

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.98      0.94      0.96     45852
           1       0.62      0.80      0.70      5317

    accuracy                           0.93     51169
   macro avg       0.80      0.87      0.83     51169
weighted avg       0.94      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43253  2599]
 [ 1088  4229]]
```

```
Churn Samples: 39374, Non-Churn Samples: 91872
Churn Ratio: 30.0% | Max Depth: 25
Validation Metrics:
Accuracy: 0.8693554122009448
Precision: 0.7735599933982505
Recall: 0.7961610327841006
F1 Score: 0.7846978067972543

Test Metrics:
Accuracy: 0.8703205160765988
Precision: 0.7795690080605363
Recall: 0.7962029569892473
F1 Score: 0.7877981880142964

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.94      0.96     45852
           1       0.62      0.78      0.69      5317

    accuracy                           0.93     51169
   macro avg       0.80      0.86      0.82     51169
weighted avg       0.94      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43293  2559]
 [ 1193  4124]]
```

```
Churn Samples: 39374, Non-Churn Samples: 91872
Churn Ratio: 30.0% | Max Depth: 30
Validation Metrics:
Accuracy: 0.8654949966983289
Precision: 0.7743520243943758
Recall: 0.7764565992865636
F1 Score: 0.7754028837998305

Test Metrics:
Accuracy: 0.866866460100574
Precision: 0.7787447698744769
Recall: 0.7817540322580645
F1 Score: 0.7802464995388614

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.95      0.96     45852
           1       0.62      0.77      0.68      5317

    accuracy                           0.93     51169
   macro avg       0.80      0.86      0.82     51169
weighted avg       0.94      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43341  2511]
 [ 1245  4072]]
```

```
Churn Samples: 39374, Non-Churn Samples: 91872
Churn Ratio: 30.0% | Max Depth: 35
Validation Metrics:
Accuracy: 0.8652410219942094
Precision: 0.7744399185336049
Recall: 0.7750976728384577
F1 Score: 0.7747686560828593

Test Metrics:
Accuracy: 0.8660537410473916
Precision: 0.7792754844144903
Recall: 0.7770497311827957
F1 Score: 0.7781610162362246

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.95      0.96     45852
           1       0.62      0.76      0.68      5317

    accuracy                           0.93     51169
   macro avg       0.80      0.85      0.82     51169
weighted avg       0.94      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43369  2483]
 [ 1268  4049]]
```

```
Churn Samples: 39374, Non-Churn Samples: 91872
Churn Ratio: 30.0% | Max Depth: 40
Validation Metrics:
Accuracy: 0.8633108142429015
Precision: 0.7712152070604209
Recall: 0.7718702225242059
F1 Score: 0.7715425757704389

Test Metrics:
Accuracy: 0.8655457916391528
Precision: 0.7807032444368949
Recall: 0.7721774193548387
F1 Score: 0.7764169271053298

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.95      0.96     45852
           1       0.62      0.76      0.68      5317

    accuracy                           0.93     51169
   macro avg       0.79      0.85      0.82     51169
weighted avg       0.93      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43363  2489]
 [ 1281  4036]]
```

```
Churn Samples: 39374, Non-Churn Samples: 157496
Churn Ratio: 20.0% | Max Depth: 10
Validation Metrics:
Accuracy: 0.9030816119200813
Precision: 0.7586852861035422
Recall: 0.7550847457627119
F1 Score: 0.7568807339449541

Test Metrics:
Accuracy: 0.9026446784734685
Precision: 0.7560184394741335
Recall: 0.7538304392236976
F1 Score: 0.754922853976643

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.96      0.96     45852
           1       0.69      0.72      0.70      5317

    accuracy                           0.94     51169
   macro avg       0.83      0.84      0.83     51169
weighted avg       0.94      0.94      0.94     51169

Confusion Matrix for April 2024:
[[44124  1728]
 [ 1514  3803]]
```

```
Churn Samples: 39374, Non-Churn Samples: 157496
Churn Ratio: 20.0% | Max Depth: 15
Validation Metrics:
Accuracy: 0.9012868269556383
Precision: 0.7490405473051894
Recall: 0.7608474576271187
F1 Score: 0.7548978390649963

Test Metrics:
Accuracy: 0.8982763875249737
Precision: 0.7388482023968043
Recall: 0.7557030983997276
F1 Score: 0.7471806093250295

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.96      0.96     45852
           1       0.67      0.74      0.70      5317

    accuracy                           0.93     51169
   macro avg       0.82      0.85      0.83     51169
weighted avg       0.94      0.93      0.94     51169

Confusion Matrix for April 2024:
[[43897  1955]
 [ 1408  3909]]
```

```
Churn Samples: 39374, Non-Churn Samples: 157496
Churn Ratio: 20.0% | Max Depth: 20
Validation Metrics:
Accuracy: 0.8898747036911615
Precision: 0.7176915488326209
Recall: 0.7398305084745763
F1 Score: 0.728592889334001

Test Metrics:
Accuracy: 0.8906911381260371
Precision: 0.7207943925233645
Recall: 0.7352740892066735
F1 Score: 0.7279622450699477

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.96      0.96     45852
           1       0.65      0.72      0.68      5317

    accuracy                           0.93     51169
   macro avg       0.81      0.84      0.82     51169
weighted avg       0.93      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43830  2022]
 [ 1509  3808]]
```

```
Churn Samples: 39374, Non-Churn Samples: 157496
Churn Ratio: 20.0% | Max Depth: 25
Validation Metrics:
Accuracy: 0.8848628513376228
Precision: 0.7057951926243003
Recall: 0.7266101694915255
F1 Score: 0.7160514447970603

Test Metrics:
Accuracy: 0.8842910839456842
Precision: 0.7042394014962593
Recall: 0.7211440245148111
F1 Score: 0.7125914711077466

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.97      0.96      0.96     45852
           1       0.65      0.70      0.68      5317

    accuracy                           0.93     51169
   macro avg       0.81      0.83      0.82     51169
weighted avg       0.93      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43852  2000]
 [ 1578  3739]]
```

```
Churn Samples: 39374, Non-Churn Samples: 157496
Churn Ratio: 20.0% | Max Depth: 30
Validation Metrics:
Accuracy: 0.8814087368777515
Precision: 0.7000333667000334
Recall: 0.7111864406779661
F1 Score: 0.7055658315116865

Test Metrics:
Accuracy: 0.8831058887271004
Precision: 0.7019006335445148
Recall: 0.7167177391896493
F1 Score: 0.7092318059299191

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.96      0.96      0.96     45852
           1       0.65      0.69      0.67      5317

    accuracy                           0.93     51169
   macro avg       0.81      0.83      0.82     51169
weighted avg       0.93      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43873  1979]
 [ 1623  3694]]
```

```
Churn Samples: 39374, Non-Churn Samples: 157496
Churn Ratio: 20.0% | Max Depth: 35
Validation Metrics:
Accuracy: 0.8801219099221131
Precision: 0.6981195433176629
Recall: 0.7047457627118644
F1 Score: 0.701417004048583

Test Metrics:
Accuracy: 0.8820900071111714
Precision: 0.703747870528109
Recall: 0.7032686414708886
F1 Score: 0.7035081743869209

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.96      0.96      0.96     45852
           1       0.65      0.70      0.67      5317

    accuracy                           0.93     51169
   macro avg       0.81      0.83      0.82     51169
weighted avg       0.93      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43849  2003]
 [ 1606  3711]]
```

```
Churn Samples: 39374, Non-Churn Samples: 157496
Churn Ratio: 20.0% | Max Depth: 40
Validation Metrics:
Accuracy: 0.8811716898069759
Precision: 0.7014321819713564
Recall: 0.7055932203389831
F1 Score: 0.7035065483734686

Test Metrics:
Accuracy: 0.8818868307879855
Precision: 0.7030983997276132
Recall: 0.7030983997276132
F1 Score: 0.7030983997276132

Classification Report for April 2024:
              precision    recall  f1-score   support

           0       0.96      0.96      0.96     45852
           1       0.65      0.69      0.67      5317

    accuracy                           0.93     51169
   macro avg       0.81      0.82      0.81     51169
weighted avg       0.93      0.93      0.93     51169

Confusion Matrix for April 2024:
[[43887  1965]
 [ 1654  3663]]
```