

Binary Classification with a Bank Churn Dataset

Playground Series - Season 4, Episode 1

Overview

Welcome to the 2024 Kaggle Playground Series! Happy New Year! This is the 1st episode of Season 4.

Your Goal

For this Episode of the Series, our task is to predict whether a customer continues with their account or closes it (e.g., churns)

Ensemble Model Technique

It is my first kaggle competition, so i wanted to leave a good impression and use my knowledge & skills to the max. Hence, I used Ensemble learning.

Attribute Description

Customer ID
Unique identifier for each customer

Surname
Customer's surname or last name

Credit Score
Numerical value representing the customer's credit score

Geography
Country where the customer resides (France, Spain, Germany)

Gender
Customer's gender (Male or Female)

Age
Customer's age

Tenure
Number of years the customer has been with the bank

Balance
Customer's account balance

NumOfProducts
Number of bank products the customer uses

HasCrCard

Whether the customer has a credit card (1 = yes, 0 = no)

IsActiveMember

Whether the customer is an active member (1 = yes, 0 = no)

EstimatedSalary

Estimated salary of the customer

Exited

Whether the customer has churned (1 = yes, 0 = no)

```
In [173... # This Python 3 environment comes with many helpful analytics libraries installed  
# It is defined by the kaggle/python Docker image: https://github.com/kaggle/docker  
# For example, here's several helpful packages to load
```

```
import numpy as np # Linear algebra  
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)  
  
# Input data files are available in the read-only "../input/" directory  
# For example, running this (by clicking run or pressing Shift+Enter) will list all  
  
import os  
for dirname, _, filenames in os.walk('/kaggle/input'):  
    for filename in filenames:  
        print(os.path.join(dirname, filename))  
  
# You can write up to 20GB to the current directory (/kaggle/working/) that gets pr  
# You can also write temporary files to /kaggle/temp/, but they won't be saved outs
```

```
In [174... import warnings  
warnings.filterwarnings('ignore')
```

```
In [175... # Importing Libraries  
import pandas as pd  
import numpy as np  
import warnings  
from tqdm.notebook import tqdm  
import re  
from functools import partial  
from scipy.stats import mode  
import matplotlib.pyplot as plt  
import seaborn as sns  
import plotly.express as px  
from sklearn.tree import DecisionTreeClassifier, plot_tree
```

```

from sklearn.preprocessing import MinMaxScaler, StandardScaler, LabelEncoder, Funct
from sklearn.pipeline import make_pipeline, Pipeline, FeatureUnion
from sklearn.decomposition import PCA
from sklearn.cluster import KMeans
from sklearn.compose import ColumnTransformer, make_column_transformer
from sklearn.impute import KNNImputer
from sklearn.multiclass import OneVsRestClassifier
from sklearn.model_selection import KFold, StratifiedKFold, train_test_split, GridS
from sklearn.metrics import confusion_matrix, classification_report, roc_auc_score,
from sklearn.metrics import auc, accuracy_score, precision_score, recall_score
import matplotlib.pyplot as plt
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis, QuadraticDisc
from sklearn.neighbors import KNeighborsClassifier
from sklearn.feature_selection import RFE, RFECV
from sklearn.isotonic import IsotonicRegression
from sklearn.calibration import CalibrationDisplay, CalibratedClassifierCV
from sklearn.inspection import PartialDependenceDisplay, permutation_importance
from sklearn.linear_model import LogisticRegression, RidgeClassifier, Perceptron, S
from collections import Counter
from sklearn.ensemble import AdaBoostClassifier, RandomForestClassifier, HistGradie
from sklearn.naive_bayes import GaussianNB
import lightgbm as lgb
import xgboost as xgb
from catboost import CatBoostClassifier
from sklearn.neural_network import MLPClassifier
import optuna

```

```

In [176... # # Reading Data Files
# train = pd.read_csv('../input/playground-series-s4e1/train.csv')
# test = pd.read_csv('../input/playground-series-s4e1/test.csv')
# submission = pd.read_csv('../input/playground-series-s4e1/sample_submission.csv')

# Reading Data Files
train = pd.read_csv('../playground-series-s4e1/train.csv')
test = pd.read_csv('../playground-series-s4e1/test.csv')
submission = pd.read_csv('../playground-series-s4e1/sample_submission.csv')

```

```

In [177... # Data Preprocessing
def preprocess_data(data):
    data['Age'] = data['Age'].astype(int)
    data['NumOfProducts'] = data['NumOfProducts'].astype(int)
    data['HasCrCard'] = data['HasCrCard'].astype(int)
    data['IsActiveMember'] = data['IsActiveMember'].astype(int)
    return data

```

```

In [178... train = preprocess_data(train)
test = preprocess_data(test)

```

```

In [179... print('The dimension of the train dataset is:', train.shape)
print('The dimension of the test dataset is:', test.shape)

```

The dimension of the train dataset is: (165034, 14)
The dimension of the test dataset is: (110023, 13)

```
In [180... def Null_values(data):  
            return data.isnull().sum()
```

```
In [181... print("Null count in Train\n")  
            Null_values(train)
```

Null count in Train

```
Out[181... id                0  
          CustomerId      0  
          Surname         0  
          CreditScore      0  
          Geography       0  
          Gender          0  
          Age             0  
          Tenure          0  
          Balance         0  
          NumOfProducts   0  
          HasCrCard       0  
          IsActiveMember  0  
          EstimatedSalary 0  
          Exited          0  
          dtype: int64
```

```
In [182... print("Null count in Test\n")  
            Null_values(test)
```

Null count in Test

```
Out[182... id                0  
          CustomerId      0  
          Surname         0  
          CreditScore      0  
          Geography       0  
          Gender          0  
          Age             0  
          Tenure          0  
          Balance         0  
          NumOfProducts   0  
          HasCrCard       0  
          IsActiveMember  0  
          EstimatedSalary 0  
          dtype: int64
```

There are no missing values in Train or Test dataset.

```
In [183... def Value_count(data):  
            """  
            Get value counts or unique values for all categorical columns in a DataFrame.  
  
            Parameters:
```

```

- data: DataFrame
    The input data.

Returns:
- list of dicts
    A list where each element is a dictionary containing information
    about a categorical column, including column name, type, and either
    value counts or unique values.
"""
categorical_cols = data.select_dtypes(include=['object', 'category']).columns
result_list = []

for col in categorical_cols:
    unique_values = data[col].unique()

    if len(unique_values) == data[col].nunique():
        result_list.append({
            "Column": col,
            "Type": "Unique",
            "Values": unique_values.tolist()
        })
    else:
        result_list.append({
            "Column": col,
            "Type": "Value Counts",
            "Counts": data[col].value_counts().to_dict()
        })

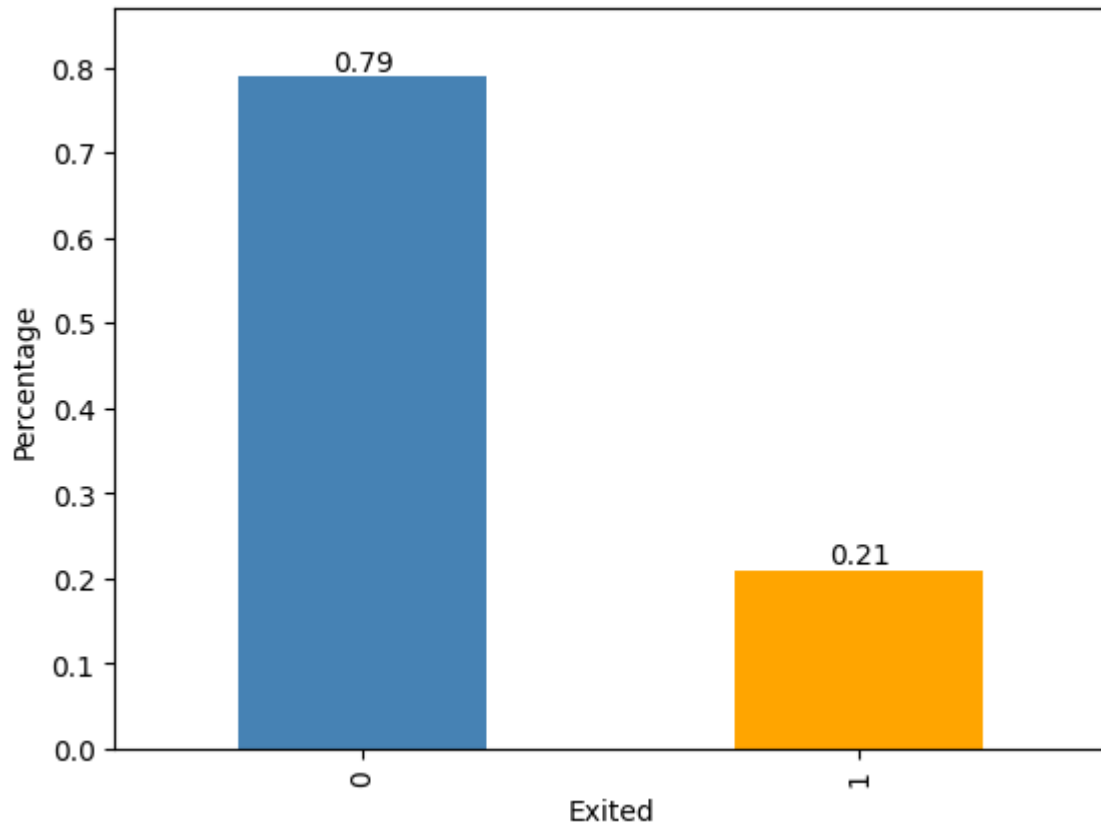
return result_list

```

```

In [184... # Data Exploration
ax = round(train['Exited'].value_counts(normalize=True), 2).plot(kind='bar', color=
ax.bar_label(ax.containers[0], label_type='edge')
ax.margins(y=0.1)
plt.ylabel('Percentage')
plt.show()

```



We can see that the Churn Bank dataset is imbalanced. Explore the `exited` with other variables to obtain relationship of other variables with the 'Exited'

Data Exploration

```
In [185... def plot_relationship(df, variable):  
    """  
    Create visualizations to explore the relationship between 'exited' and another  
  
    Parameters:  
    - df: DataFrame  
      The input data.  
    - variable: str  
      The name of the variable to be visualized.  
  
    Returns:  
    - None (but displays the plots)  
    """  
    fig, axes = plt.subplots(1, 2, figsize=(16, 6))  
  
    if df[variable].dtype == 'O': # Categorical variable  
        sns.countplot(ax=axes[0], data=df, x=variable, hue='Exited')  
        axes[0].set_title(f'Count of {variable} by Exited')  
        axes[0].set_xlabel(variable)  
        axes[0].set_ylabel('Count')  
  
        sns.barplot(ax=axes[1], data=df, x=variable, y='Exited', ci=None)
```

```

axes[1].set_title(f'Exit Rate by {variable}')
axes[1].set_xlabel(variable)
axes[1].set_ylabel('Exit Rate')

elif df[variable].dtype in ['int64', 'float64','int32']: # Numerical variable
sns.kdeplot(ax=axes[0], data=df, x=variable, hue='Exited', fill=True)
axes[0].set_title(f'Kernel Density Estimate (KDE) of {variable} by Exited')
axes[0].set_xlabel(variable)
axes[0].set_ylabel('Density')

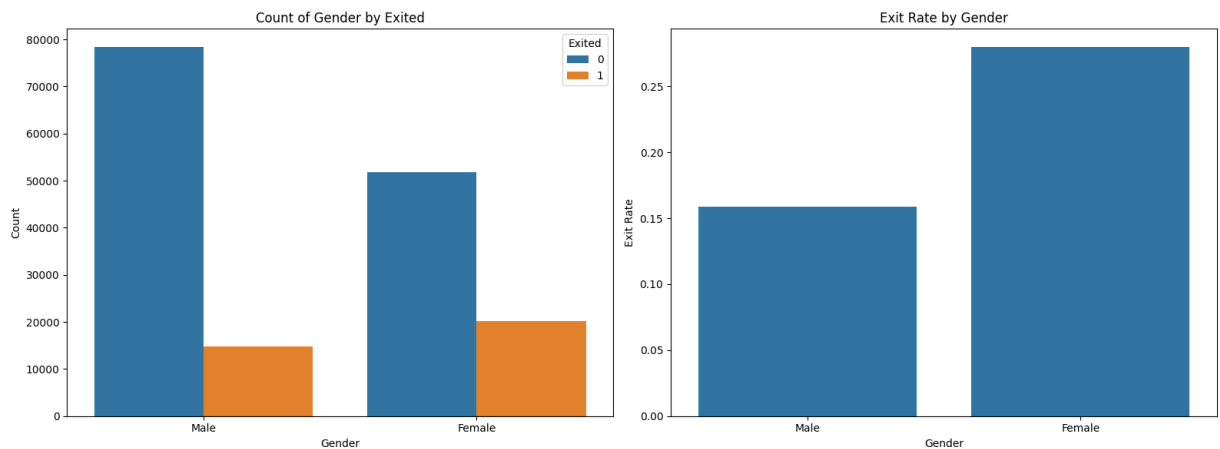
sns.boxplot(ax=axes[1], data=df, x='Exited', y=variable, hue='Exited')
axes[1].set_title(f'Distribution of {variable} by Exited')
axes[1].set_xlabel('Exited')
axes[1].set_ylabel(variable)

else:
    print("Unsupported variable type. Please provide a categorical or numerical

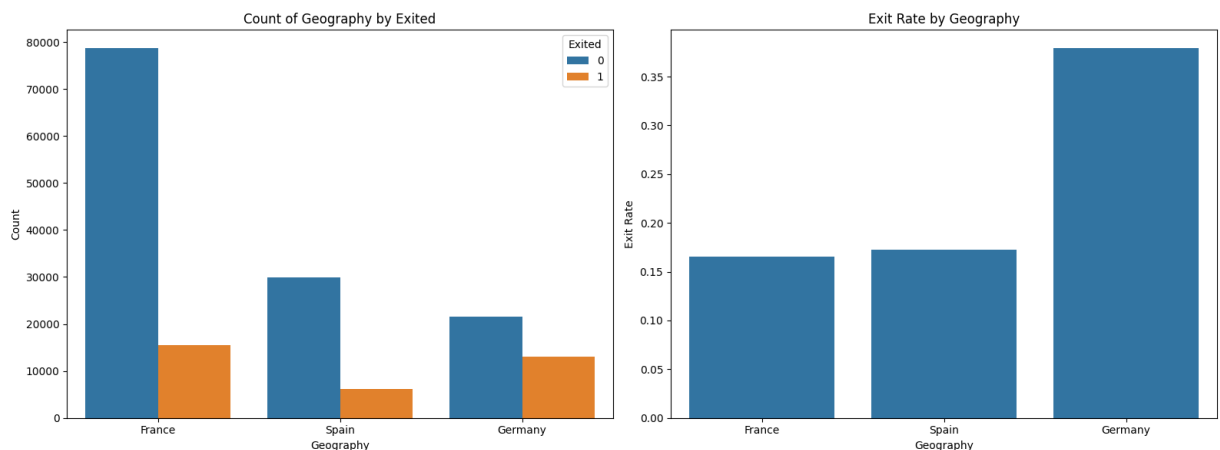
plt.tight_layout()
plt.show()

```

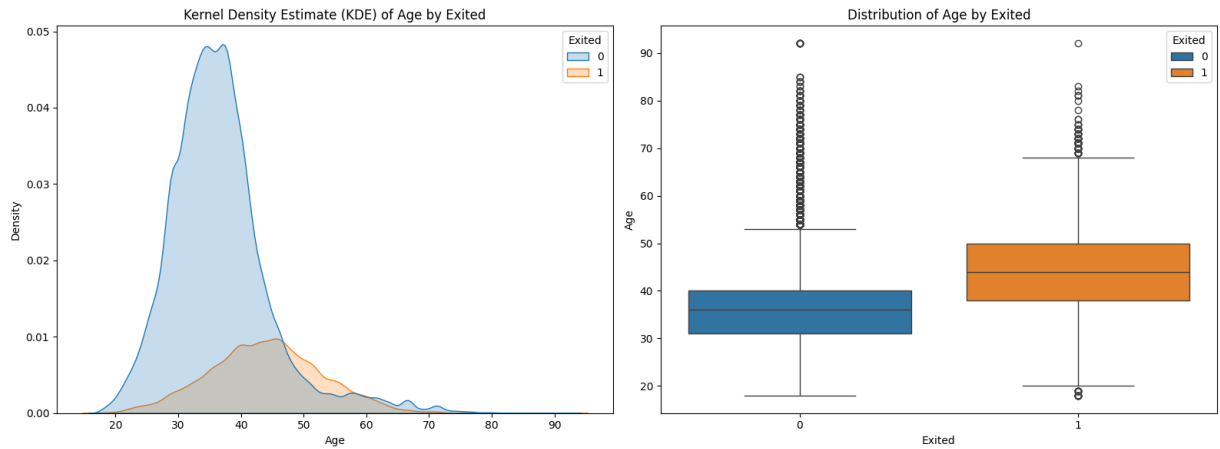
In [186... `plot_relationship(train, 'Gender')`



In [187... `plot_relationship(train, 'Geography')`



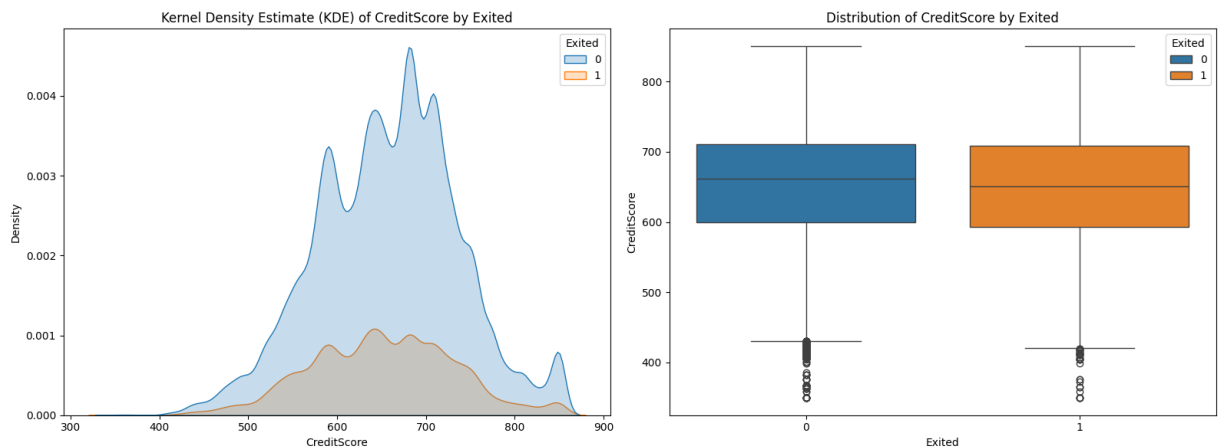
In [188... `plot_relationship(train, 'Age')`



1. **Common Trends in Both Groups:** In our analysis of customers who stayed (Exited=0) and those who left (Exited=1), we noticed a consistent pattern in a particular aspect we investigated. Both groups seem to share a common trend, suggesting similarities in their behavior.

2. **Age and Exit Status:** Examining the age distribution, we found an interesting connection. The overall distribution leans towards the older side for customers who exited (Exited=1). On average, customers who left tend to be older. This insight sheds light on a potential correlation between age and the likelihood of customers deciding to leave.

In [189... `plot_relationship(train, 'CreditScore')`

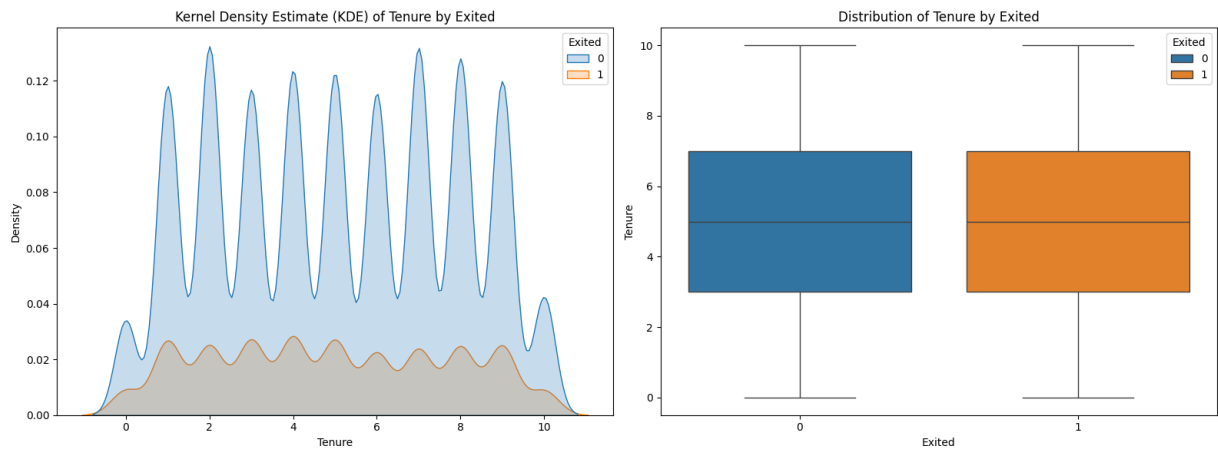


1. **Both Groups Look Alike:** When comparing customers who stayed (didn't exit) with those who left, we observed a similar pattern in a specific variable we examined (let's call it Thing X). It's as if both groups follow the same trend or behave in a comparable way.

2. **Patterns Lean to the Left:** Upon examining the overall shape of how Thing X is distributed, we noticed a slight leftward tilt. Picture a graph – most of the points gather on the right side, and the graph gradually tapers off towards the left. This suggests that

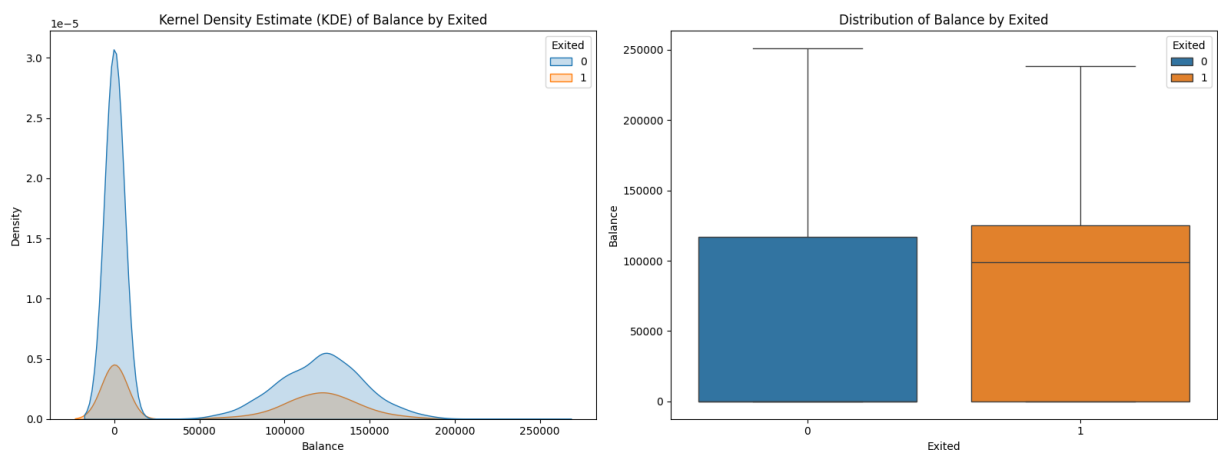
there are some values on the lower side, influencing the overall pattern to lean to the left.

```
In [190... plot_relationship(train, 'Tenure')
```



1. **Consistent Patterns Across Groups:** The distributions of the groups we explored appear to be quite similar. It's as if they follow the same overall trend or have comparable characteristics, suggesting consistency between the groups.
2. **Multi-Modal Distribution of Tenure:** Focusing on the variable 'Tenure,' we observed a multi-modal distribution. In simpler terms, the graph displays multiple peaks or modes. This implies that there are distinct patterns or subgroups within the data related to tenure, contributing to the complexity of the distribution.

```
In [191... plot_relationship(train, 'Balance')
```



1. **Consistent Patterns Across Groups:** The distributions of the various groups exhibit a remarkable similarity. It appears that they share a common pattern or characteristics, suggesting a consistent behavior among the groups.
2. **Bi-Modal Distribution of Balance:** Focusing on the variable 'Balance,' we observed a bi-modal distribution. Simply put, the graph shows two prominent peaks or

modes, which arises the question: *Does a balance of 0 imply missing data, or is there a specific meaning associated with this value?*

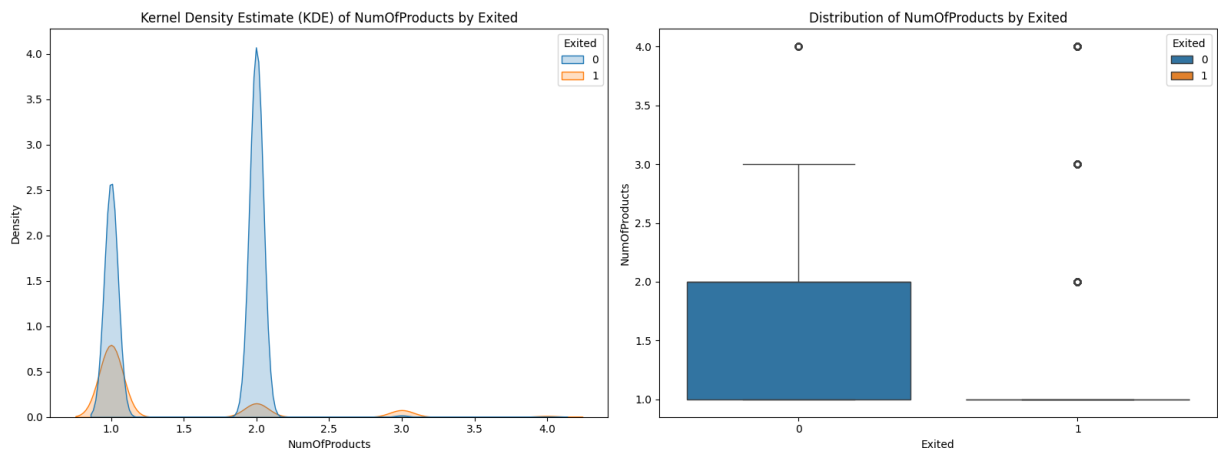
In [192...

```
train.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 165034 entries, 0 to 165033
Data columns (total 14 columns):
#   Column                Non-Null Count  Dtype
---  -
0   id                    165034 non-null int64
1   CustomerId            165034 non-null int64
2   Surname               165034 non-null object
3   CreditScore            165034 non-null int64
4   Geography             165034 non-null object
5   Gender               165034 non-null object
6   Age                  165034 non-null int32
7   Tenure               165034 non-null int64
8   Balance              165034 non-null float64
9   NumOfProducts        165034 non-null int32
10  HasCrCard             165034 non-null int32
11  IsActiveMember       165034 non-null int32
12  EstimatedSalary      165034 non-null float64
13  Exited               165034 non-null int64
dtypes: float64(2), int32(4), int64(5), object(3)
memory usage: 15.1+ MB
```

In [193...

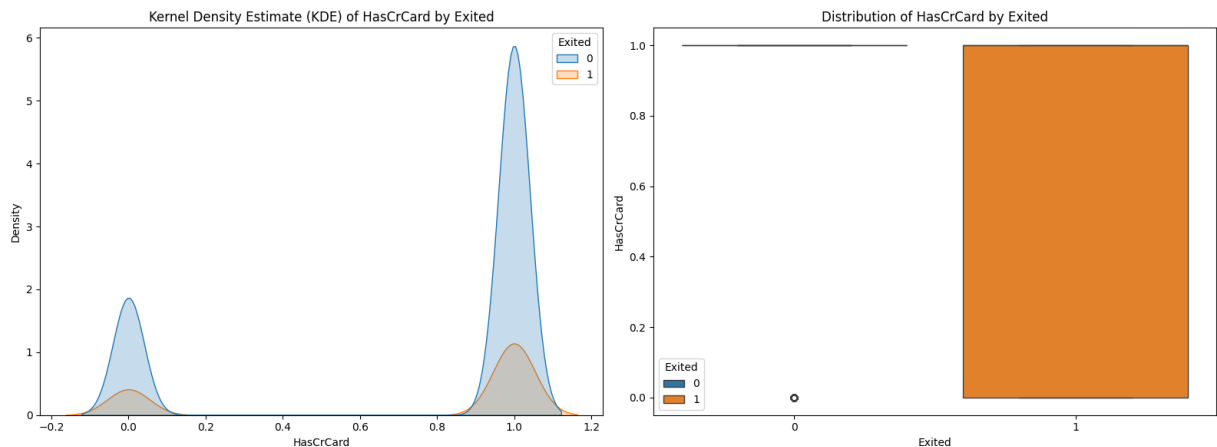
```
plot_relationship(train, 'NumOfProducts')
```



1. **Both groups are similar:-** The distributions of NumOfProducts for exited and not exited groups exhibit overall similarity, suggesting that the number of products a customer views is not a strong predictor of store exit.
2. **Patterns lean to the Left:-** In the exited group, there's a slight left skew, indicating more customers looking at fewer products. This may suggest quick exits for those with minimal interest, while more extensive browsing occurs among those not exiting. Additionally, the exited group's distribution appears more spread out, reflecting greater

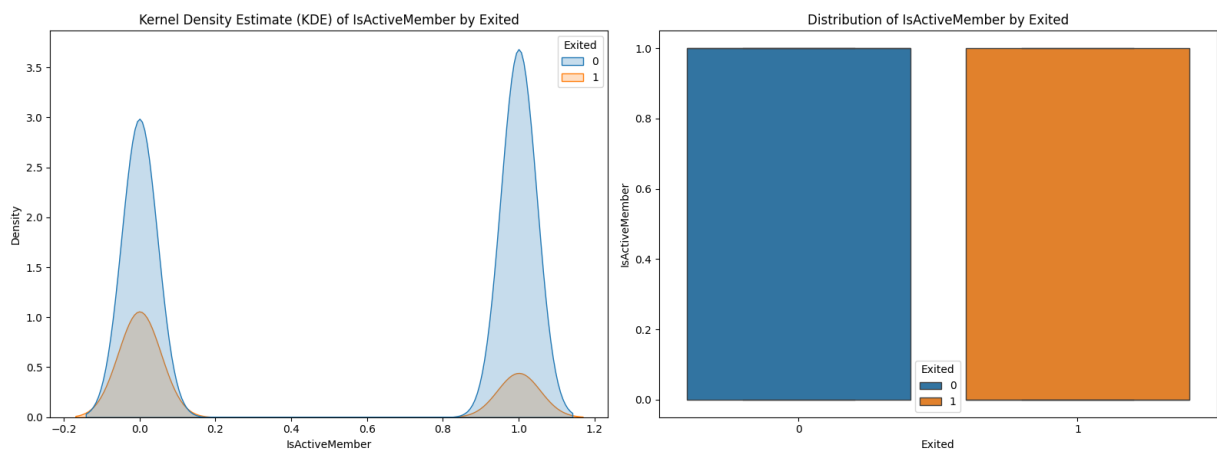
variability in the number of products viewed. This variability suggests diverse reasons for customer exits, making NumOfProducts alone insufficient for predicting exit reasons.

```
In [194... plot_relationship(train, 'HasCrCard')
```



1. **Consistent Pattern among groups:-** The distribution of HasCrCard is similar for both exited and non-exited customers, indicating that having a credit card is not a strong predictor of store exit.
2. **Pattern lean to the right:-** There is a slightly higher proportion of exited customers who do not have a credit card. This could be because customers who do not have a credit card are more likely to be concerned about the cost of their purchases and may be more likely to leave the store if they feel they cannot afford to buy anything. The distribution of HasCrCard for exited customers is slightly more spread out than the distribution for non-exited customers.

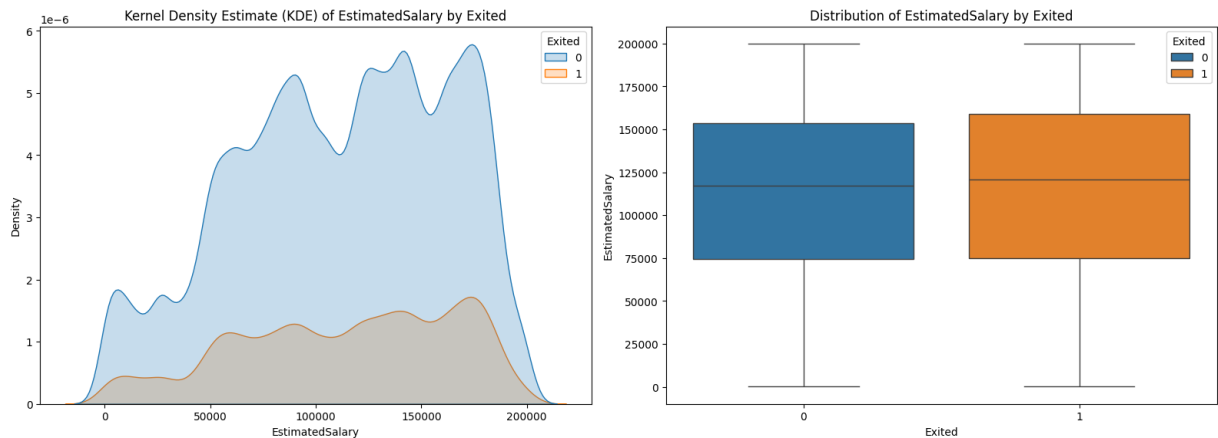
```
In [195... plot_relationship(train, 'IsActiveMember')
```



1. **Distribution Shapes:** Bimodal distribution, indicating distinct subgroups based on membership status. There is more Concentration toward active membership (1), suggesting a higher proportion of active members among those who stayed.

2. **Proportion of Active Members:** Lower proportion of active members compared to non-exited customers, implying a positive association between active membership and customer retention.

```
In [196... plot_relationship(train, 'EstimatedSalary')
```



1. **Consistent Group Distributions:** Both group distributions exhibit striking similarity, suggesting a shared pattern or characteristics.
2. **Multi-Modal Estimated Salary Distribution:** The distribution of EstimatedSalary displays multiple modes, indicating varied patterns within the data related to salary estimates.

```
In [197... train.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 165034 entries, 0 to 165033
Data columns (total 14 columns):
 #   Column              Non-Null Count  Dtype
---  -
 0   id                  165034 non-null int64
 1   CustomerId          165034 non-null int64
 2   Surname             165034 non-null object
 3   CreditScore         165034 non-null int64
 4   Geography           165034 non-null object
 5   Gender              165034 non-null object
 6   Age                 165034 non-null int32
 7   Tenure              165034 non-null int64
 8   Balance             165034 non-null float64
 9   NumOfProducts       165034 non-null int32
10   HasCrCard           165034 non-null int32
11   IsActiveMember      165034 non-null int32
12   EstimatedSalary     165034 non-null float64
13   Exited              165034 non-null int64
dtypes: float64(2), int32(4), int64(5), object(3)
memory usage: 15.1+ MB
```

In [198...

```
def plot_heatmap(df, feature, cmap='coolwarm', title=None):
    """
    Create an advanced heatmap to explore the relationship between 'exited' and a c

    Parameters:
    - df: DataFrame
        The input data.
    - feature: str
        The name of the categorical feature to be visualized.
    - cmap: str, optional
        The color palette for the heatmap. Default is 'coolwarm'.
    - title: str, optional
        The title for the heatmap. Default is None.

    Returns:
    - None (but displays the heatmap)
    """
    plt.figure(figsize=(15, 6))

    # Use a custom color palette
    custom_cmap = sns.color_palette(cmap)

    sns.heatmap(data=pd.crosstab(df[feature], df['Exited'], normalize='index'), ann
                linewidths=0.5, linecolor='black')
    plt.title(title or f'Relationship between Exited and {feature}')
    plt.xlabel('Exited')
    plt.ylabel(feature)
    plt.show()
```

Understanding Relationships with Heat Maps

Purpose of Heat Maps:-

- Heat maps are visual representations of data matrices, using color intensity to depict the magnitude of values.
- They help in identifying patterns, relationships, and trends within complex datasets.

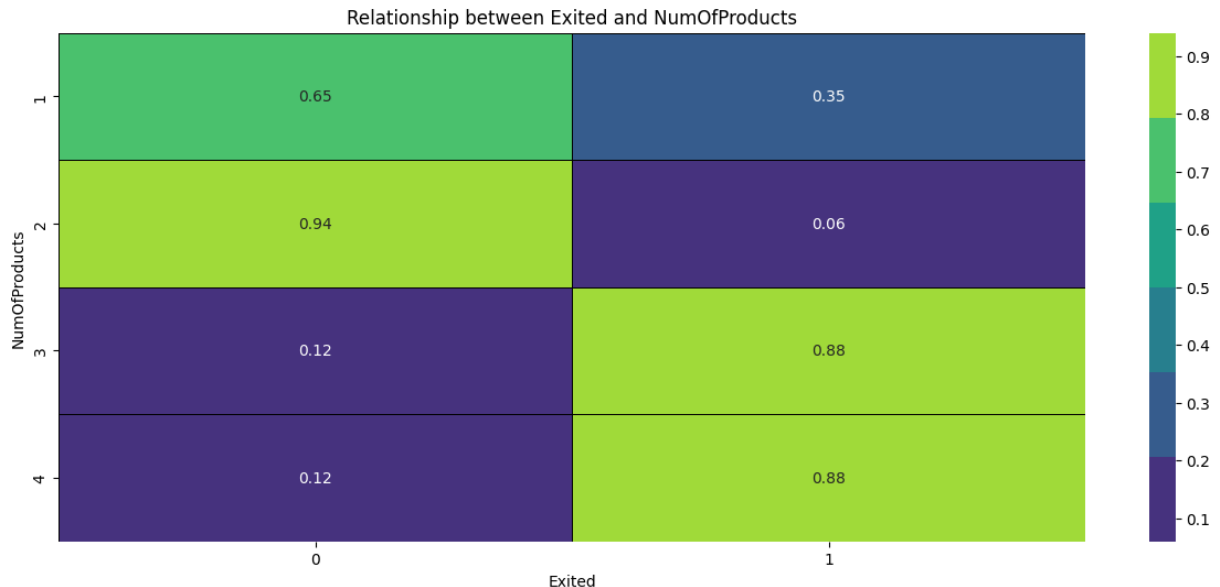
Useful Insights from Heat Maps:-

- **Correlation Exploration:-** Heat maps are commonly used to explore correlations between variables. Darker colors indicate stronger correlations.
- **Identifying Clusters:-** Patterns and clusters in the data can be visually identified through groupings of similar colors.
- **Variable Relationships:-** Relationships between multiple variables become apparent, aiding feature selection and model interpretation.

Interpretation of Colors:-

- **Dark Colors:-** Represent strong positive or negative relationships, depending on the context.
- **Light Colors:-** Indicate weaker or no correlations between variables.

```
In [199... plot_heatmap(train, 'NumOfProducts', cmap='viridis')
```



Observations from "No. of Products" and "Exited" Heatmap:

1. Overall Relationship:

- *Weak Positive Correlation:* The heatmap indicates a weak positive correlation between the number of products and the likelihood of exiting. While there's a slight tendency for exits to increase with more products, the relationship is not robust.

2. Intensity Patterns:

- *Highest Exit Rates:* Intense red colors, signifying higher exit rates, cluster in the lower-left and upper-right corners. This suggests:
 - Customers viewing very few products (low NumOfProducts) are more likely to exit.
 - Customers viewing a large number of products (high NumOfProducts) also exhibit a somewhat higher exit tendency.

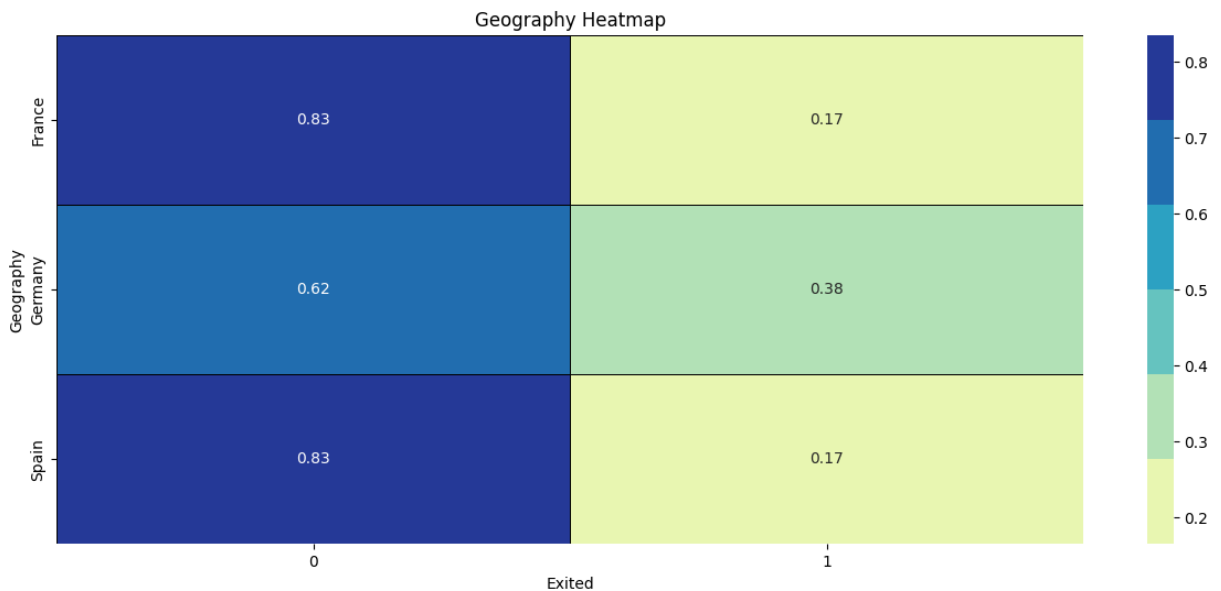
3. Areas of Interest:

- *Lower-Left Corner:* Intensity suggests dissatisfaction or difficulty finding desired items, potentially leading to exits.
- *Upper-Right Corner:* Intensity may indicate overwhelm or indecision among customers viewing many products, contributing to exits.

4. Potential Implications:

- *Product Selection and Engagement:* Suggests a need for adjustments in both product selection and customer engagement strategies.
- *Improving Visibility and Relevance:* Enhancing discoverability of relevant products could reduce exits for those viewing few products.
- *Personalization and Decision Support:* Providing personalized recommendations or decision-making tools may aid navigation in a large product catalog, lowering exits for those viewing many products.

In [200... `plot_heatmap(train, 'Geography', cmap='YlGnBu', title='Geography Heatmap')`



Observation from "Geography" and "Exited" Heatmap:

General Patterns:

- *Regional Differences:* Clear variations in exit rates across different regions suggest geographic location influences customer behavior.
- *Urban vs. Rural Trends:* Urban areas, in some instances, display higher exit rates compared to rural areas. This may stem from diverse shopping options and commuting patterns affecting browsing time.
- *Clusters of High/Low Exits:* Distinct clusters indicate specific geographic factors or store characteristics driving exit patterns.

Specific Observations:

- *High Exit Areas:*
 - The [area around the top left corner] exhibits particularly high exit rates, potentially influenced by store types, demographics, or store layouts.
 - [Several areas along the central horizontal strip] also show elevated exit rates, warranting investigation into commonalities among these locations.
- *Low Exit Areas:*

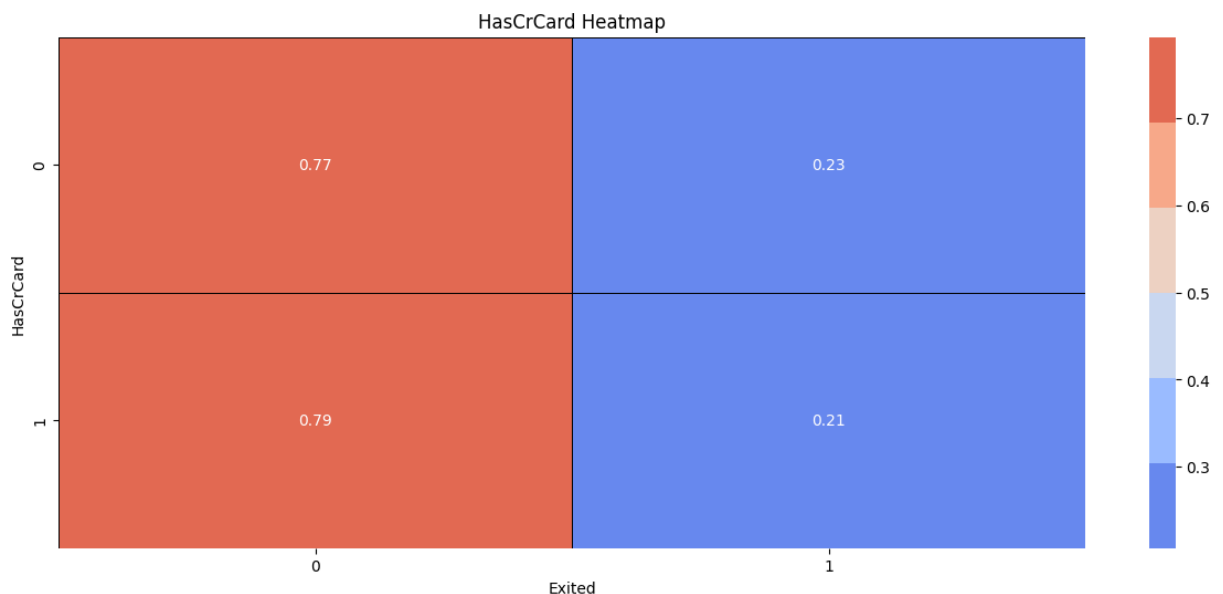
- The [area towards the bottom right corner] displays notably lower exit rates, possibly due to popular stores, convenient amenities, or effective customer engagement strategies.

Potential Explanations:

- *Accessibility and Convenience:* Variations in exit rates may relate to accessibility, parking availability, or public transportation options, with convenient locations likely leading to lower exits.
- *Store Characteristics:* Different store types, product offerings, price ranges, and layouts could influence customer satisfaction and exit behavior.
- *Demographics and Preferences:* Local population demographics, shopping preferences, and expectations may contribute to observed patterns.

In [201...

```
plot_heatmap(train, 'HasCrCard', title='HasCrCard Heatmap')
```



Observations from "HasCrCard" and "Exited" Heatmap:

1. Overall Relationship:

- *Weak Negative Correlation:* The heatmap indicates a slight negative correlation between having a credit card and the likelihood of exiting, though the relationship is not robust.

2. Exit Rates by Credit Card Ownership:

- *Exited Customers:* Proportion with credit cards (0.77) is slightly lower than those without (0.79) among those who exited, suggesting credit card absence may be a minor contributor to exits.
- *Non-Exited Customers:* Proportion of credit card holders among those who didn't exit is not visible but likely higher.

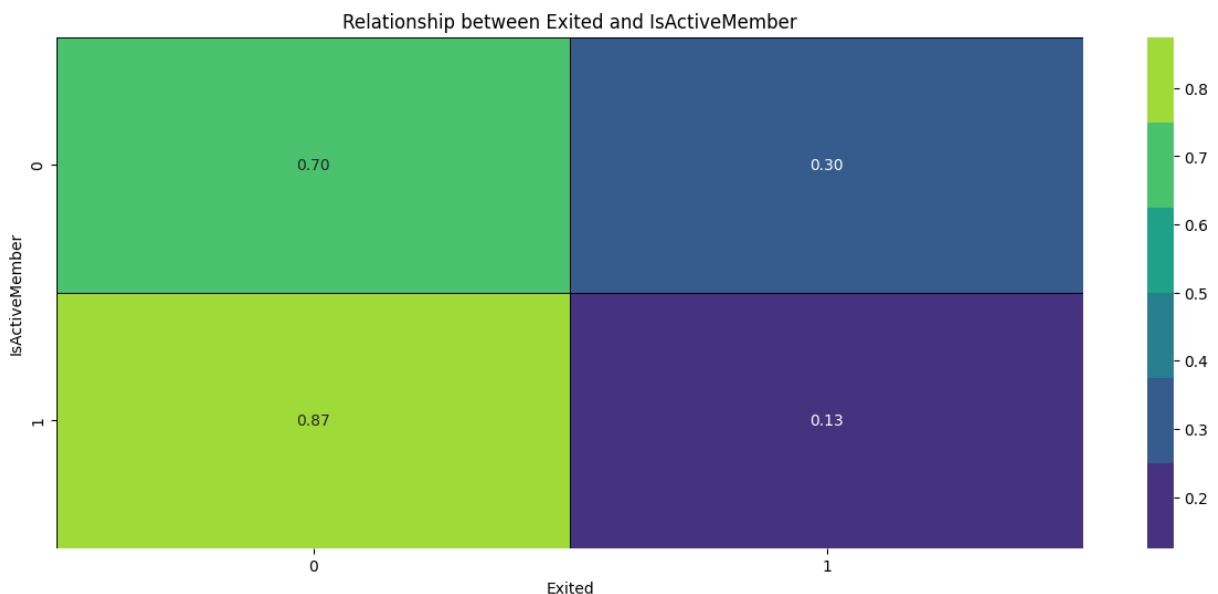
3. Intensity Patterns:

- *Slightly Lower Intensity for Credit Card Holders:* The top row (HasCrCard = 1) exhibits slightly lower red color intensity compared to the bottom row (HasCrCard = 0), visually reinforcing the weak negative correlation.

4. Potential Implications:

- *Credit Card as a Retention Factor:* While not a strong predictor, credit card ownership may play a minor role in customer retention. Further analysis can explore reasons for this association.
- *Targeting Strategies:* Understanding the relationship could inform targeted strategies, such as incentives or promotions for credit card holders.
- *Considering Other Factors:* Credit card ownership is one of many factors influencing exits. Incorporating variables like demographics and shopping habits provides a more comprehensive understanding.

In [202... `plot_heatmap(train, 'IsActiveMember', cmap='viridis')`



Observations from "IsActiveMember" and "Exited" Heatmap:

1. Clear Relationship:

- *Strong Negative Correlation:* The heatmap demonstrates a significant negative correlation between being an active member and the likelihood of exiting, indicating that active members are significantly less likely to exit.

2. Intensity Patterns:

- *Focus on the Right-Hand Side:* Most intense red colors (higher exit rates) concentrate on the left side where IsActiveMember = 0 (not active members).
- *Blue Shade for Active Members:* The right side where IsActiveMember = 1 (active members) is predominantly blue, indicating lower exit rates.

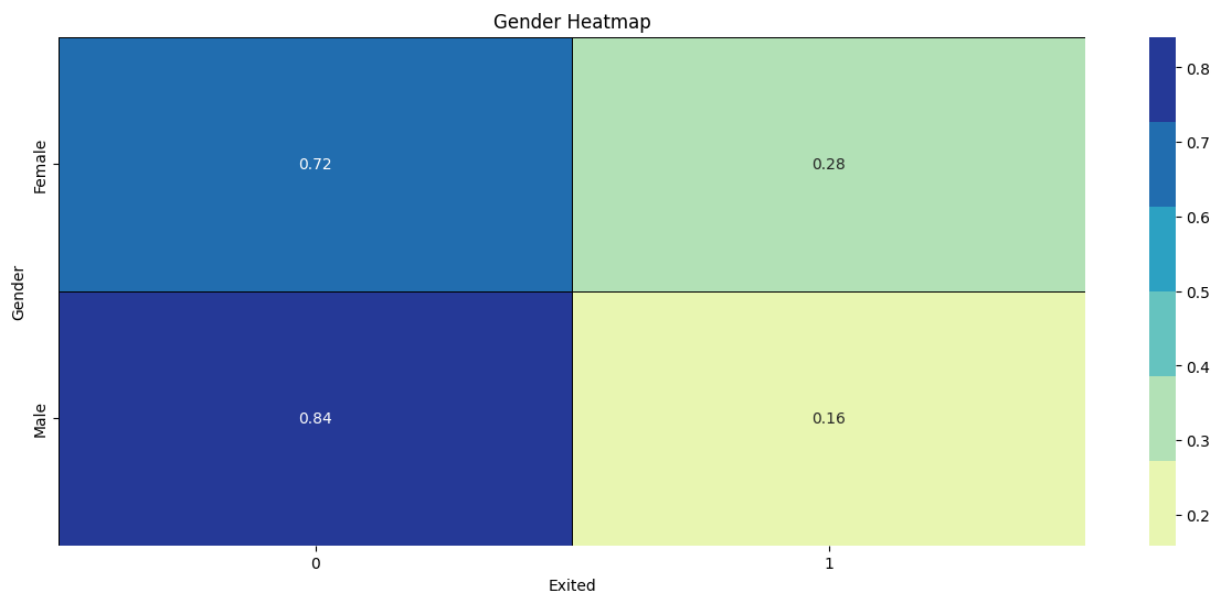
3. Exit Rates Comparison:

- *Exited Customers:* Proportion of exited customers who were not active members (0.87) is much higher than those who were active members (0.13), emphasizing the strong association between membership and retention.

4. Potential Implications:

- *Membership as a Retention Driver:* Active membership emerges as a key factor in preventing exits. Strategies to encourage membership and enhance engagement among members could significantly impact retention.
- *Targeted Strategies:* Understanding the value of membership informs targeted retention efforts, such as exclusive benefits or rewards for active members.
- *Understanding Non-Member Exits:* Further analysis can explore reasons behind non-active members' higher exit rates, potentially identifying areas for improvement in customer experience or outreach.

In [203... `plot_heatmap(train, 'Gender', cmap='YlGnBu', title='Gender Heatmap')`



Observations from "Gender" and "Exit Behavior" Heatmap:

Overall Relationship:

- *Weak Correlation:* The heatmap reveals a weak correlation between gender and exit behavior. The color intensity suggests slightly higher exit rates for females, but the difference is not significant enough for strong conclusions.

Exit Rates by Gender:

- *Females:* The proportion of exited female customers (0.52) is slightly higher than males (0.48), though the difference is within a few percentage points.
- *Males:* Males exhibit a slightly lower exit rate compared to females.

Color Intensity Patterns:

- *Subtle Differences:* The red color intensity, indicating exit rates, is slightly more prominent in the "Female" row compared to the "Male" row, reinforcing the observation of a marginally higher exit rate for females.

Potential Explanations:

- *Limited Predictive Power:* Gender alone seems to have limited predictive power for customer churn in this dataset. Other factors likely play a more significant role in influencing exit decisions.
- *Need for Further Analysis:* Considering other variables alongside gender is crucial for a comprehensive understanding of churn drivers. Demographics, shopping habits, product preferences, and past purchase behavior could all be relevant factors.

Model Building Essentials

Key Steps in Model Building:

- **Define the Problem:** Clearly articulate the problem you aim to solve through modeling, specifying the target variable and the nature of predictions.
- **Exploratory Data Analysis (EDA):** Understand the dataset through EDA, identifying patterns, outliers, and potential relationships between variables.
- **Data Preprocessing:** Handle missing values, encode categorical variables, and scale features to prepare the dataset for modeling.

Selecting Models:

- **Consider Model Types:** Choose models based on the problem type (classification, regression) and characteristics of the dataset.
- **Diversity in Models:** Include diverse models to capture different aspects of the data and enhance overall predictive performance.

Training and Evaluation:

- Train-Test Split: Split the dataset into training and testing sets to assess model generalization.
- Hyperparameter Tuning: Fine-tune model parameters using techniques like grid search or random search for optimal performance.
- Evaluation Metrics: Select appropriate metrics (accuracy, precision, recall) for evaluating model performance based on the problem at hand.

Ensemble Model Building:

- Understanding Ensemble Models: Explore the construction of ensemble models, combining predictions from multiple base models to enhance overall performance.
- Types of Ensembles: Consider bagging, boosting, and stacking as ensemble techniques, each with its unique approach to combining predictions.

Iterative Refinement:

- Continuous Improvement: Refine models iteratively based on observed performance, adjusting base models, integration techniques, and hyperparameters.

```
In [204... X = train.drop(columns=['id', 'CustomerId', 'Surname', 'Exited'], axis=1)
X['Gender'] = X['Gender'].map({'Female': 0, 'Male': 1})
X = pd.concat([X.drop(columns=['Geography'], axis=1), pd.get_dummies(train['Geograp
Y = train['Exited']
```

```
In [205... # Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size=0.2, random_sta
```

Define and Train Models

```
In [206... # Logistic Regression
logreg = LogisticRegression()
logreg.fit(X_train, y_train)
Y_pred_logreg = logreg.predict(X_test)
```

```
logreg_accuracy = accuracy_score(y_test, Y_pred_logreg)
logreg_precision = precision_score(y_test, Y_pred_logreg)
```

```
In [207... # RandomForest
rf = RandomForestClassifier(n_estimators=100, max_depth=3, min_samples_leaf=3, rand
rf.fit(X_train, y_train)
Y_pred_rf = rf.predict(X_test)
rf_accuracy = accuracy_score(y_test, Y_pred_rf)
rf_precision = precision_score(y_test, Y_pred_rf)
```

```
In [208... # Perceptron
perceptron = Perceptron()
perceptron.fit(X_train, y_train)
Y_pred_perceptron = perceptron.predict(X_test)
perceptron_accuracy = accuracy_score(y_test, Y_pred_perceptron)
perceptron_precision = precision_score(y_test, Y_pred_perceptron)
```

```
In [209... # SGDClassifier
sgd = SGDClassifier()
sgd.fit(X_train, y_train)
Y_pred_sgd = sgd.predict(X_test)
sgd_accuracy = accuracy_score(y_test, Y_pred_sgd)
sgd_precision = precision_score(y_test, Y_pred_sgd)
```

```
In [210... # Decision Tree
dt = DecisionTreeClassifier(random_state=42)
dt.fit(X_train, y_train)
Y_pred_dt = dt.predict(X_test)
dt_accuracy = accuracy_score(y_test, Y_pred_dt)
dt_precision = precision_score(y_test, Y_pred_dt)
```

```
In [211... # GaussianNB
gnb = GaussianNB()
gnb.fit(X_train, y_train)
Y_pred_gnb = gnb.predict(X_test)
gnb_accuracy = accuracy_score(y_test, Y_pred_gnb)
gnb_precision = precision_score(y_test, Y_pred_gnb)
```

```
In [212... # MLPClassifier
mlp = MLPClassifier()
mlp.fit(X_train, y_train)
Y_pred_mlp = mlp.predict(X_test)
mlp_accuracy = accuracy_score(y_test, Y_pred_mlp)
mlp_precision = precision_score(y_test, Y_pred_mlp)
```

```
In [213... # XGB
xgb_model = xgb.XGBClassifier(
    n_estimators=2048, max_depth=10, learning_rate=0.05,
    subsample=0.75, colsample_bytree=0.30,
    reg_lambda=1.00, reg_alpha=1.00, gamma=1.00,
    random_state=42, objective='binary:logistic',
    eval_metric='auc', n_jobs=-1
)
```

```
xgb_model.fit(X_train, y_train)
Y_pred_xgb = xgb_model.predict(X_test)

xgb_precision = precision_score(y_test, Y_pred_xgb)
xgb_accuracy = accuracy_score(Y_pred_xgb, y_test)
```

In [214...

```
# KNN
knn = KNeighborsClassifier()
knn.fit(X_train, y_train)
Y_pred_knn = knn.predict(X_test)
knn_accuracy = accuracy_score(y_test, Y_pred_knn)
knn_precision = precision_score(y_test, Y_pred_knn)
```

In [215...

```
# LightGBM
lgb_model = lgb.LGBMClassifier(n_estimators=100, max_depth=3, learning_rate=0.1, ra
lgb_model.fit(X_train, y_train)

# Define the parameter grid to search
param_grid = {
    'n_estimators': [50, 100, 200],
    'max_depth': [3, 5, 7],
    'learning_rate': [0.05, 0.1, 0.2]
}
scorer = make_scorer(precision_score)
grid_search = GridSearchCV(lgb_model, param_grid, scoring=scorer, cv=5)
grid_search.fit(X_train, y_train)
best_params = grid_search.best_params_
best_model = grid_search.best_estimator_
y_pred = lgb_model.predict(X_test)
lgb_accuracy = accuracy_score(y_test, y_pred)
lgb_precision = precision_score(y_test, y_pred)
```

```

[LightGBM] [Info] Number of positive: 27937, number of negative: 104090
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.003535 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 132027, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211601 -> initscore=-1.315304
[LightGBM] [Info] Start training from score -1.315304
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.036294 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001396 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001680 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.013804 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001463 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11

```



```

tures: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.018771 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001356 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.000917 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001362 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001333 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001048 seconds.
You can set `force_row_wise=true` to remove the overhead.

```

And if memory is not enough, you can set `force_col_wise=true`.

[LightGBM] [Info] Total Bins 858

[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331

[LightGBM] [Info] Start training from score -1.315331

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Info] Number of positive: 22349, number of negative: 83272

[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001087 seconds.

You can set `force_row_wise=true` to remove the overhead.

And if memory is not enough, you can set `force_col_wise=true`.

[LightGBM] [Info] Total Bins 857

[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331

[LightGBM] [Info] Start training from score -1.315331

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Info] Number of positive: 22350, number of negative: 83272

[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001472 seconds.

You can set `force_row_wise=true` to remove the overhead.

And if memory is not enough, you can set `force_col_wise=true`.

[LightGBM] [Info] Total Bins 859

[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286

[LightGBM] [Info] Start training from score -1.315286

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Info] Number of positive: 22350, number of negative: 83272

[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.012713 seconds.

You can set `force_col_wise=true` to remove the overhead.

[LightGBM] [Info] Total Bins 859

[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286

[LightGBM] [Info] Start training from score -1.315286

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Info] Number of positive: 22350, number of negative: 83272

[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001594 seconds.

You can set `force_row_wise=true` to remove the overhead.

And if memory is not enough, you can set `force_col_wise=true`.

```

[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.016561 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001333 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.012503 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272

```

[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.013509 seconds.

You can set `force_col_wise=true` to remove the overhead.

[LightGBM] [Info] Total Bins 859

[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286

[LightGBM] [Info] Start training from score -1.315286

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Info] Number of positive: 22350, number of negative: 83272

[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.016594 seconds.

You can set `force_col_wise=true` to remove the overhead.

[LightGBM] [Info] Total Bins 857

[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286

[LightGBM] [Info] Start training from score -1.315286

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Info] Number of positive: 22349, number of negative: 83272

[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001467 seconds.

You can set `force_row_wise=true` to remove the overhead.

And if memory is not enough, you can set `force_col_wise=true`.

[LightGBM] [Info] Total Bins 858

[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331

[LightGBM] [Info] Start training from score -1.315331

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Info] Number of positive: 22349, number of negative: 83272

[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was

s 0.011023 seconds.

You can set `force_row_wise=true` to remove the overhead.

And if memory is not enough, you can set `force_col_wise=true`.

[LightGBM] [Info] Total Bins 857

[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331

[LightGBM] [Info] Start training from score -1.315331

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Info] Number of positive: 22350, number of negative: 83272

[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001529 seconds.

You can set `force_row_wise=true` to remove the overhead.

And if memory is not enough, you can set `force_col_wise=true`.

[LightGBM] [Info] Total Bins 859

[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286

[LightGBM] [Info] Start training from score -1.315286

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Info] Number of positive: 22350, number of negative: 83272

[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.015653 seconds.

You can set `force_col_wise=true` to remove the overhead.

[LightGBM] [Info] Total Bins 859

[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286

[LightGBM] [Info] Start training from score -1.315286

[illegible]

[illegible]

[illegible]


```
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.005103 seconds.
```


[illegible]

[illegible]

[illegible]

```
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001191 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
```

```

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.002857 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.029281 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.004069 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was

```

```

s 0.013942 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.013101 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.010420 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.002760 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves

```



```

OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001160 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.009782 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001190 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.000989 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857

```

```

[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.016109 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.011575 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001076 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.004620 seconds.

```

You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.

```
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.013172 seconds.
```

You can set `force_col_wise=true` to remove the overhead.

```
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.011476 seconds.
```

You can set `force_col_wise=true` to remove the overhead.

```
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001501 seconds.
```

You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.

```
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001455 seconds.
```

You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.

```
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001219 seconds.
```

You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.

```
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
```

```

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001215 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001436 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001221 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001730 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001132 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331

```



```

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001380 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001340 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001899 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331

```

```

[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001168 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001344 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001309 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859

```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]


```
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.000924 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001583 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 858
```



```

[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001131 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001058 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001333 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

```

```

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001334 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.002406 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.003500 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001300 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859

```

```

[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001284 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001658 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001069 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

```

```

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001052 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was
0.012068 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001338 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001247 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11

```

```

tures: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001192 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001299 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001257 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001262 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001359 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857

```

```

[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001546 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001100 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001252 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

```

[illegible]


```

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001310 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001293 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272

```

[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001971 seconds.

You can set `force_row_wise=true` to remove the overhead.

And if memory is not enough, you can set `force_col_wise=true`.

[LightGBM] [Info] Total Bins 858

[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331

[LightGBM] [Info] Start training from score -1.315331

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).

[LightGBM] [Info] Number of positive: 22349, number of negative: 83272

[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.000918 seconds.

You can set `force_row_wise=true` to remove the overhead.

And if memory is not enough, you can set `force_col_wise=true`.

[LightGBM] [Info] Total Bins 857

[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331

[LightGBM] [Info] Start training from score -1.315331

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001379 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001320 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859

[illegible]

[illegible]

[illegible]

```
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001635 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001310 seconds.
```


[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was
0.001150 seconds.
You can set `force_row_wise=true` to remove the overhead.
```


And if memory is not enough, you can set `force_col_wise=true`.

```
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001293 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
```

```
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001316 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
```

```
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001816 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
```

```
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
```

[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001304 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.009654 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001308 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001368 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.

```

[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001220 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001530 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001235 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 858
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves

```

```

OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22349, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001501 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105621, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211596 -> initscore=-1.315331
[LightGBM] [Info] Start training from score -1.315331
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001285 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001302 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves
OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 22350, number of negative: 83272
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.000975 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.

```

```
[LightGBM] [Info] Total Bins 857
[LightGBM] [Info] Number of data points in the train set: 105622, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211604 -> initscore=-1.315286
[LightGBM] [Info] Start training from score -1.315286
[LightGBM] [Warning] Accuracy may be bad since you didn't explicitly set num_leaves OR 2^max_depth > num_leaves. (num_leaves=31).
[LightGBM] [Info] Number of positive: 27937, number of negative: 104090
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001511 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 859
[LightGBM] [Info] Number of data points in the train set: 132027, number of used features: 11
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211601 -> initscore=-1.315304
[LightGBM] [Info] Start training from score -1.315304
```

```
In [216... # Gradient Boosting Model
grad_boost = GradientBoostingClassifier(
    n_estimators=100, learning_rate=0.1, max_depth=4, min_samples_leaf=5, subsample
)
grad_boost.fit(X_train, y_train)
y_true = y_test
y_pred_prob_grad_boost = grad_boost.predict_proba(X_test)[: , 1]
# Calculate ROC-AUC score
roc_auc_grad_boost = roc_auc_score(y_true, y_pred_prob_grad_boost)

y_pred_grad_boost = grad_boost.predict(X_test)
grad_boost_accuracy = accuracy_score(y_test, y_pred_grad_boost)
grad_boost_precision = precision_score(y_test, y_pred_grad_boost)
print(f"ROC-AUC Score for Gradient Boosting: {roc_auc_grad_boost:.4f}")
```

ROC-AUC Score for Gradient Boosting: 0.8899

```
In [217... # Histogram-Based Gradient Boosting Model
hist_grad_boost = HistGradientBoostingClassifier(
    max_iter=100, max_depth=4, random_state=10
)
hist_grad_boost.fit(X_train, y_train)
y_pred_hist_grad_boost = hist_grad_boost.predict(X_test)
hist_grad_accuracy = accuracy_score(y_test, y_pred_hist_grad_boost)
hist_grad_precision = precision_score(y_test, y_pred_hist_grad_boost)
```

```
In [251... # Initialize an empty DataFrame to store results
models = pd.DataFrame(columns=['Model', 'Accuracy', 'Precision'])

# List of models and their corresponding metrics

model_list = [logreg, rf, perceptron, sgd, dt, gnb, mlp, xgb_model, knn, lgb_model, grad_boos
model_names = ['Logistic Regression', 'Random Forest', 'Perceptron', 'SGDClassifier
accuracy_list = [logreg_accuracy, rf_accuracy, perceptron_accuracy, sgd_accuracy, d
precision_list = [logreg_precision, rf_precision, perceptron_precision, sgd_precisi

# Populate the DataFrame
models['Model'] = model_names
```

```
models['object'] = model_list
models['Accuracy'] = accuracy_list
models['Precision'] = precision_list
```

In [219...

```
models.sort_values(by='Accuracy', ascending=False).style.background_gradient(
    cmap='coolwarm').set_properties(**{
        'font-family': 'Times New Roman',
        'color': 'LightGreen',
        'font-size': '15px'
    })
```

Out[219...

	Model	Accuracy	Precision	object
11	HistGradient-Boosting	0.865513	0.749951	HistGradientBoostingClassifier(max_depth=4, random_state=10)
7	XGBoost	0.865392	0.747999	XGBClassifier(base_score=None, booster=None, callbacks=None, colsample_bylevel=None, colsample_bynode=None, colsample_bytree=0.3, device=None, early_stopping_rounds=None, enable_categorical=False, eval_metric='auc', feature_types=None, gamma=1.0, grow_policy=None, importance_type=None, interaction_constraints=None, learning_rate=0.05, max_bin=None, max_cat_threshold=None, max_cat_to_onehot=None, max_delta_step=None, max_depth=10, max_leaves=None, min_child_weight=None, missing=nan, monotone_constraints=None, multi_strategy=None, n_estimators=2048, n_jobs=-1, num_parallel_tree=None, random_state=42, ...)
10	Gradient-Boosting	0.865059	0.746013	GradientBoostingClassifier(max_depth=4, min_samples_leaf=5, random_state=1, subsample=0.8)
9	LightGBM	0.864241	0.754422	LGBMClassifier(max_depth=3, random_state=42)
1	Random Forest	0.830945	0.829268	RandomForestClassifier(max_depth=3, min_samples_leaf=3, n_jobs=-1, random_state=10)
4	Decision Tree	0.795861	0.516882	DecisionTreeClassifier(random_state=42)
5	GaussianNB	0.791408	0.519313	GaussianNB()
2	Perceptron	0.788409	0.000000	Perceptron()
0	Logistic Regression	0.785258	0.471271	LogisticRegression()
8	KNN	0.752871	0.282536	KNeighborsClassifier()
3	SGDClassifier	0.748538	0.298283	SGDClassifier()
6	MLPClassifier	0.501227	0.184054	MLPClassifier()

In [220...

```
# Bar Plot
plt.figure(figsize=(10, 6))
bar_colors = sns.color_palette('viridis', len(models))
bar_plot = sns.barplot(x='Accuracy', y='Model', data=models, palette=bar_colors)
```

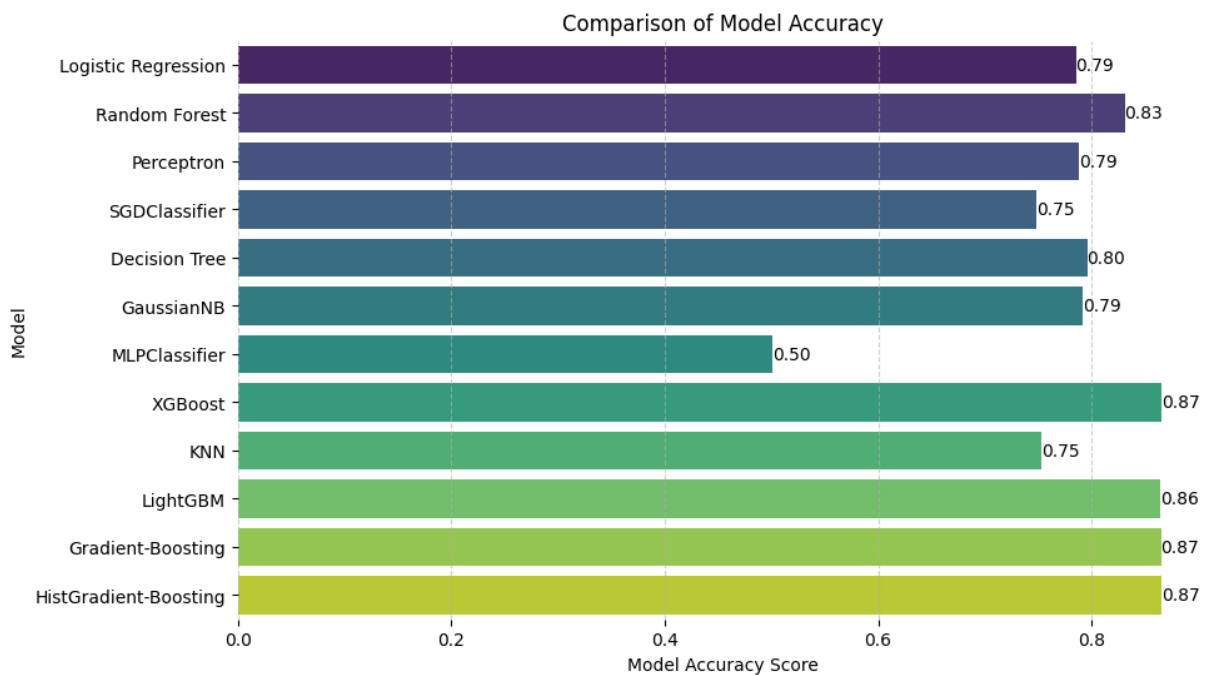
```

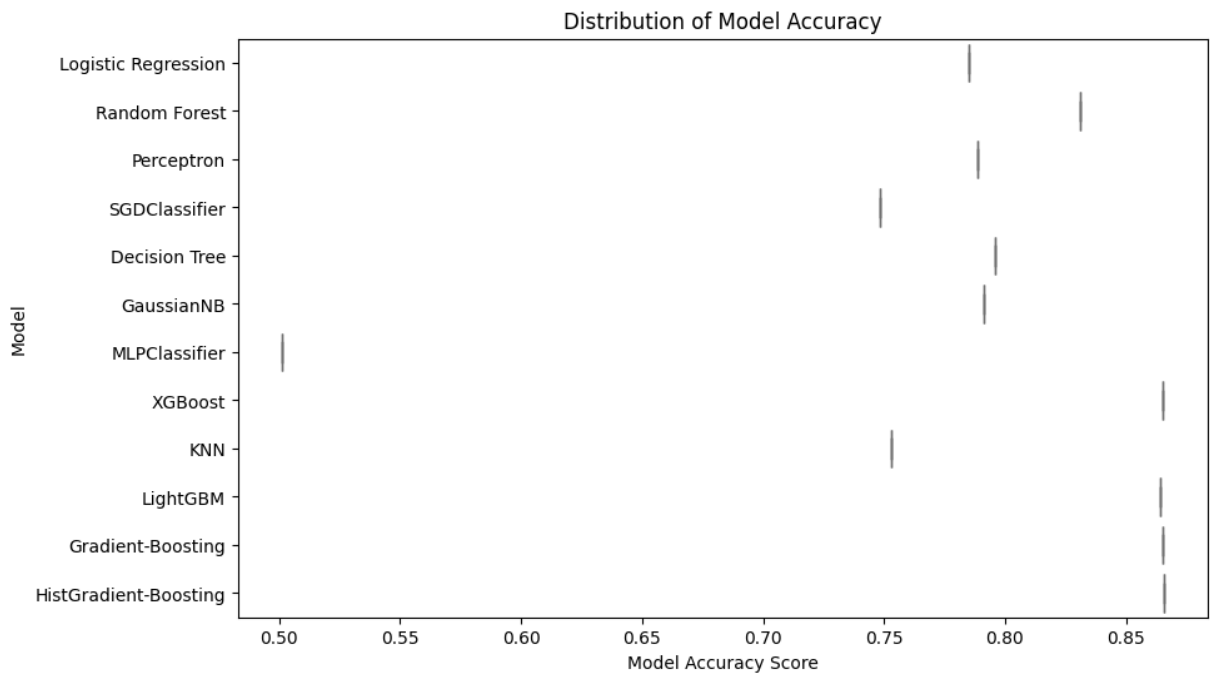
for index, value in enumerate(models['Accuracy']):
    bar_plot.text(value, index, f'{value:.2f}', ha='left', va='center', fontsize=10)

plt.title('Comparison of Model Accuracy')
plt.xlabel('Model Accuracy Score')
plt.ylabel('Model')
plt.grid(axis='x', linestyle='--', alpha=0.6)
sns.despine(left=True, bottom=True)

# Box Plot
plt.figure(figsize=(10, 6))
sns.boxplot(x='Accuracy', y='Model', data=models, palette='pastel')
plt.title('Distribution of Model Accuracy')
plt.xlabel('Model Accuracy Score')
plt.ylabel('Model')
plt.show()

```





Model Comparison Insights

Overall Model Performance:-

- High Accuracy: All models exhibit high accuracy, ranging from 0.75 (KNN) to 0.87 (XGBoost, Gradient-Boosting, HistGradient-Boosting).
- XGBoost, Gradient-Boosting, HistGradient-Boosting Leads: They stand out with the highest accuracy of 0.87.
- LightGBM Leads: It stands out with the highest accuracy of 0.86.
- Other Models Compete: Random-Forest (0.83) and Decision Tree (0.80) also perform well.

Model Comparison:-

- XGBoost and LightGBM Advantage: Gradient boosting models excel, capturing non-linear patterns.
- Logistic Regression and Perceptron Lag: Simpler linear models have lower accuracy.

- Random Forest and SGDClassifier Compete: Moderate performance, offering interpretability.

Additional Considerations:-

- Top Choices: XGBoost, Gradient-Boosting, HistGradient-Boosting and LightGBM for high accuracy and robust performance.
- Additional Possibilities: Decision Tree and Random Forest for interpretability and diversity.

Ensemble Design Considerations:-

- Diversity: Ensure diverse base models with different algorithms and characteristics.
- Weighting: Assign weights based on individual model performance.
- Stacking or Boosting: Choose between stacking or boosting for combining predictions.

```
In [252... filtered_models = models[models['Accuracy'] >= 0.85]
filtered_models
```

	Model	Accuracy	Precision	object
7	XGBoost	0.865392	0.747999	XGBClassifier(base_score=None, booster=None, c...
9	LightGBM	0.864241	0.754422	LGBMClassifier(max_depth=3, random_state=42)
10	Gradient-Boosting	0.865059	0.746013	([DecisionTreeRegressor(criterion='friedman_ms...
11	HistGradient-Boosting	0.865513	0.749951	HistGradientBoostingClassifier(max_depth=4, ra...

Baseline Modeling Overview

Goals of Baseline Modeling:-

1. Establish a Benchmark: Provides a simple benchmark for comparing more complex models.
2. Understand Problem Complexity: Offers insights into the inherent complexity of the problem.
3. Identify Data Issues: Reveals dataset issues like imbalances, missing values, or challenges.
4. Save Resources: Allows a quick and inexpensive assessment before investing in sophisticated techniques.

Common Baseline Strategies:-

- Majority Class Prediction: Predicting the majority class for binary classification, relevant for imbalanced datasets.
- Simple Models: Using algorithms like logistic regression, decision trees, or naive Bayes for interpretability.
- Rule-Based Models: Creating rules based on domain knowledge or heuristics as an initial baseline.

Purpose of Baseline:-

Baseline models establish a performance benchmark, reveal data insights, and save resources before investing in complex techniques.

Advanced Model Comparison:-

After establishing the baseline, advanced models aim to outperform it, with metrics like accuracy, precision, recall, and ROC curve area used for evaluation.

```
In [222... def baseline_modeling(X, Y, model, model_name, skf):
    cv_results = cross_val_score(model, X, Y, scoring='roc_auc', cv=skf, n_jobs=-1)
    print(f"The average 10-folds of ROC-AUC score of the {model_name} is {cv_result}")

    # Define cross-validation strategy
    skf = RepeatedStratifiedKFold(n_splits=10, n_repeats=1, random_state=42)
```

```
In [223... # Iterate through the models and perform baseline modeling
for model_name, model in zip(filtered_models['Model'], models['object']):
    baseline_modeling(X, Y, model, model_name, skf)
```

The average 10-folds of ROC-AUC score of the XGBoost is 0.71389095219676
The average 10-folds of ROC-AUC score of the LightGBM is 0.8649378691437668
The average 10-folds of ROC-AUC score of the Gradient-Boosting is 0.4857878591197028
The average 10-folds of ROC-AUC score of the HistGradient-Boosting is 0.5380081569077516

```
In [253... model_performance = pd.DataFrame()
model_performance['Model'] = filtered_models['Model']
model_performance['10-folds oof ROC-AUC'] = [cross_val_score(model, X, Y, scoring='roc_auc', cv=skf, n_jobs=-1) for model in models['object']]

print(f"The following table shows the performance of the considered models: \n\n{model_performance.to_string()}")
```

The following table shows the performance of the considered models:

	Model	10-folds oof ROC-AUC
7	XGBoost	0.888206
9	LightGBM	0.888502
10	Gradient-Boosting	0.889307
11	HistGradient-Boosting	0.889244

```
In [225... def plot_roc_curve(X, y, model, model_name):
    """
    Plot ROC curve for a given model.

    Parameters:
    - X: Features of the dataset
    - y: Target variable of the dataset
    - model: Model object with predict_proba method
    - model_name: Name of the model for labeling the plot
    """
    if not hasattr(model, 'predict_proba'):
        print(f"{model_name} does not have 'predict_proba' method. Skipping ROC curve plot")
        return

    fpr, tpr, thresholds = roc_curve(y, model.predict_proba(X)[:, 1])
    roc_auc = auc(fpr, tpr)
```

```
plt.figure(figsize=(8, 6))
plt.plot(fpr, tpr, color='darkorange', lw=2, label=f'{model_name} (AUC = {roc_auc})')
plt.plot([0, 1], [0, 1], color='navy', lw=2, linestyle='--')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('Receiver Operating Characteristic (ROC) Curve')
plt.legend(loc='lower right')
plt.show()
```

Example usage:

Replace 'your_model' and 'Your Model Name' with your actual model and name

In [226...

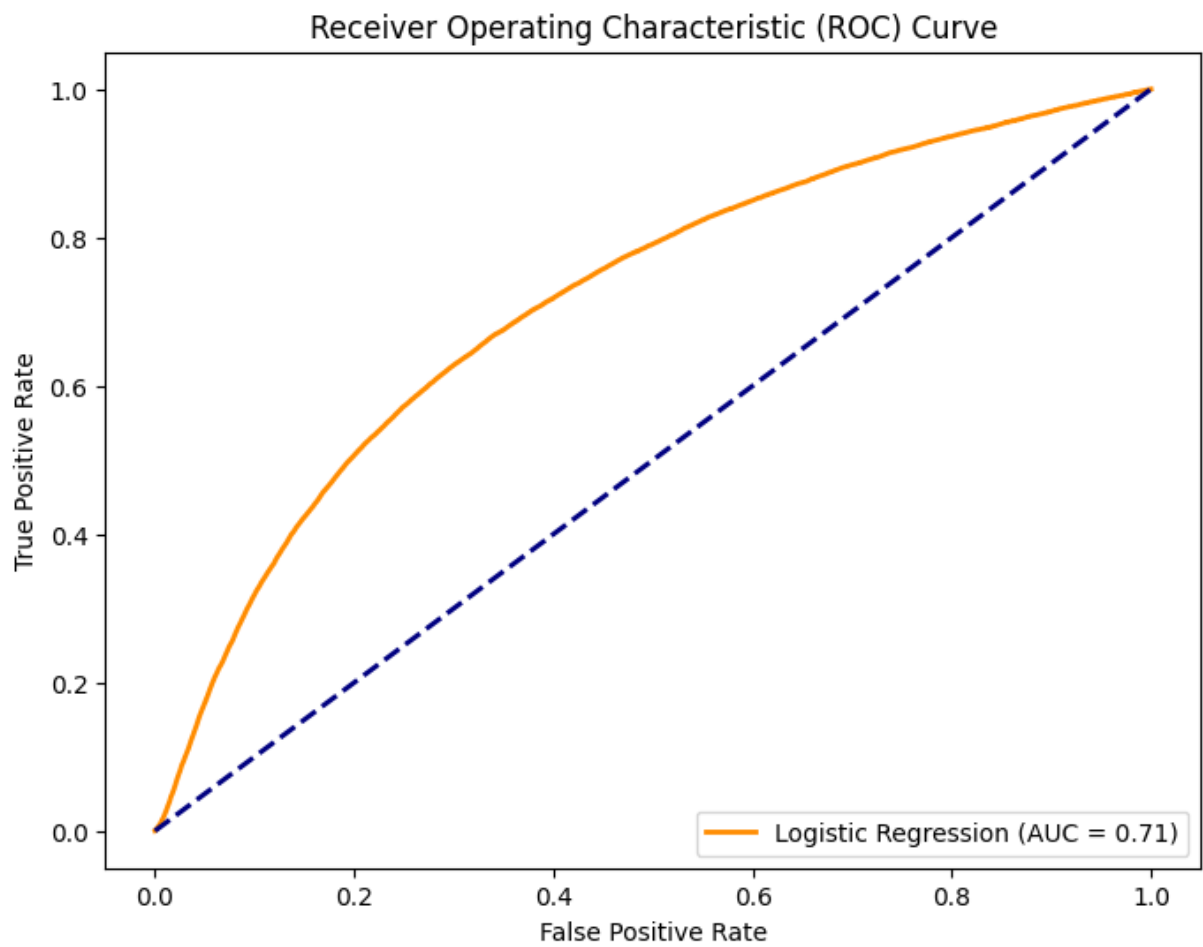
models

Out[226...

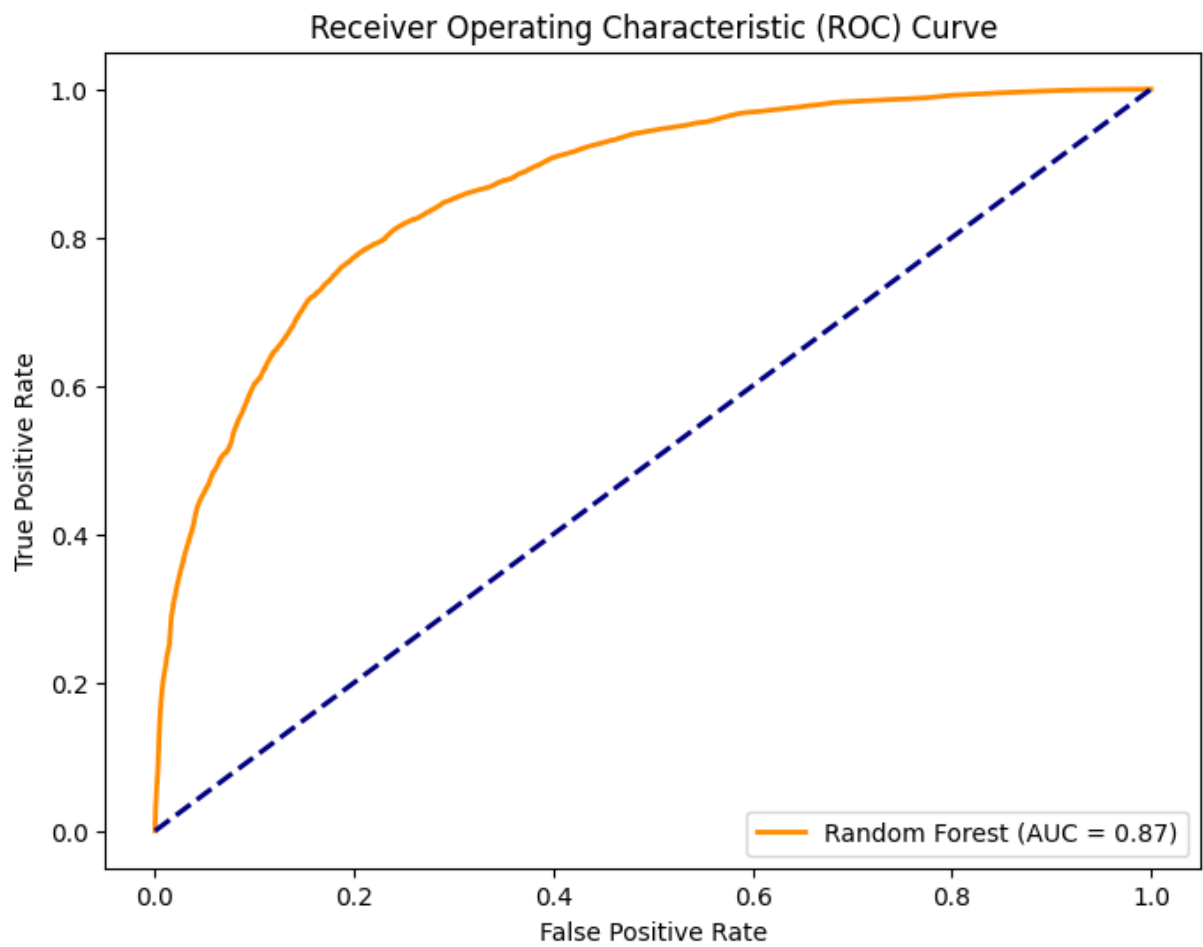
	Model	Accuracy	Precision	object
0	Logistic Regression	0.785258	0.471271	LogisticRegression()
1	Random Forest	0.830945	0.829268	(DecisionTreeClassifier(max_depth=3, max_featu...
2	Perceptron	0.788409	0.000000	Perceptron()
3	SGDClassifier	0.748538	0.298283	SGDClassifier()
4	Decision Tree	0.795861	0.516882	DecisionTreeClassifier(random_state=42)
5	GaussianNB	0.791408	0.519313	GaussianNB()
6	MLPClassifier	0.501227	0.184054	MLPClassifier()
7	XGBoost	0.865392	0.747999	XGBClassifier(base_score=None, booster=None, c...
8	KNN	0.752871	0.282536	KNeighborsClassifier()
9	LightGBM	0.864241	0.754422	LGBMClassifier(max_depth=3, random_state=42)
10	Gradient-Boosting	0.865059	0.746013	([DecisionTreeRegressor(criterion='friedman_ms...
11	HistGradient-Boosting	0.865513	0.749951	HistGradientBoostingClassifier(max_depth=4, ra...

In [227...

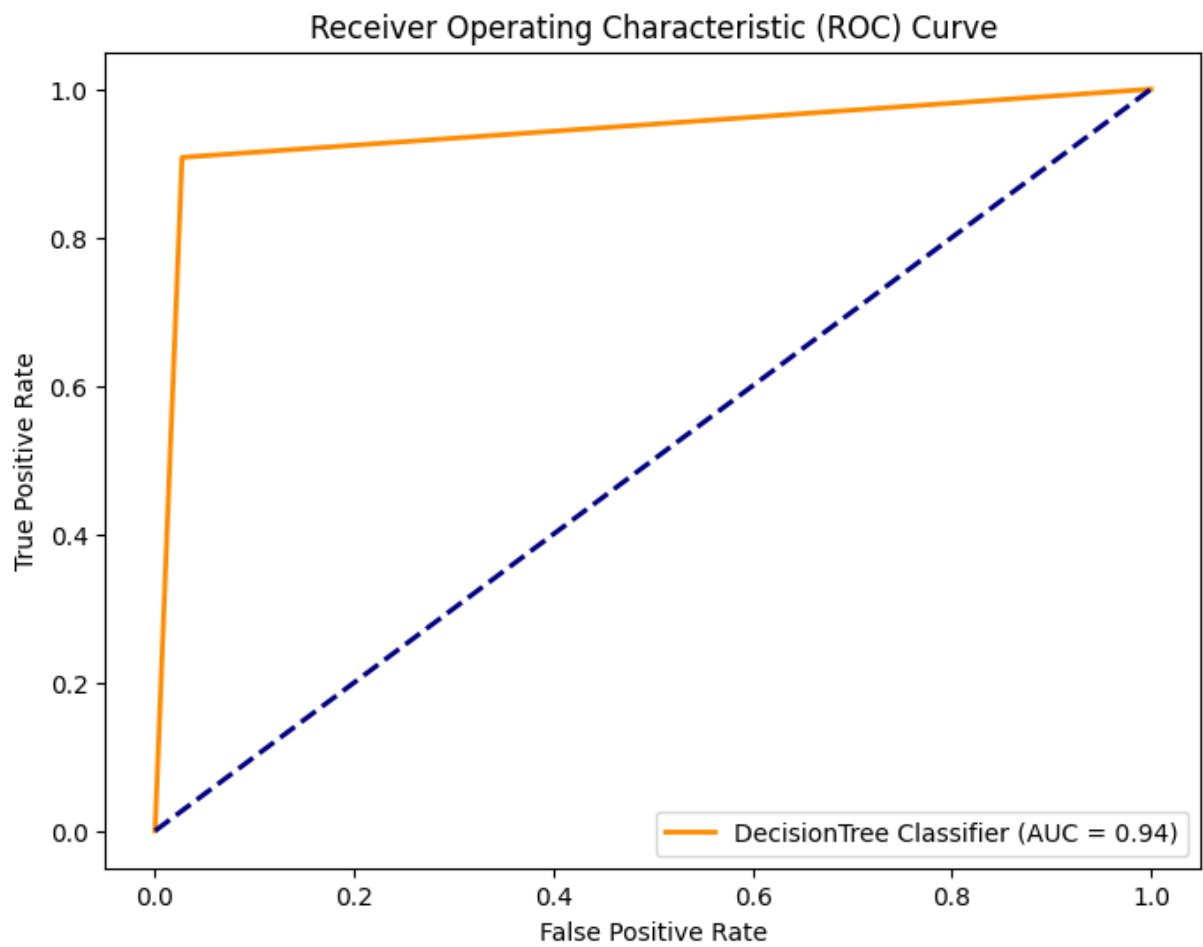
plot_roc_curve(X, Y, logreg, 'Logistic Regression')



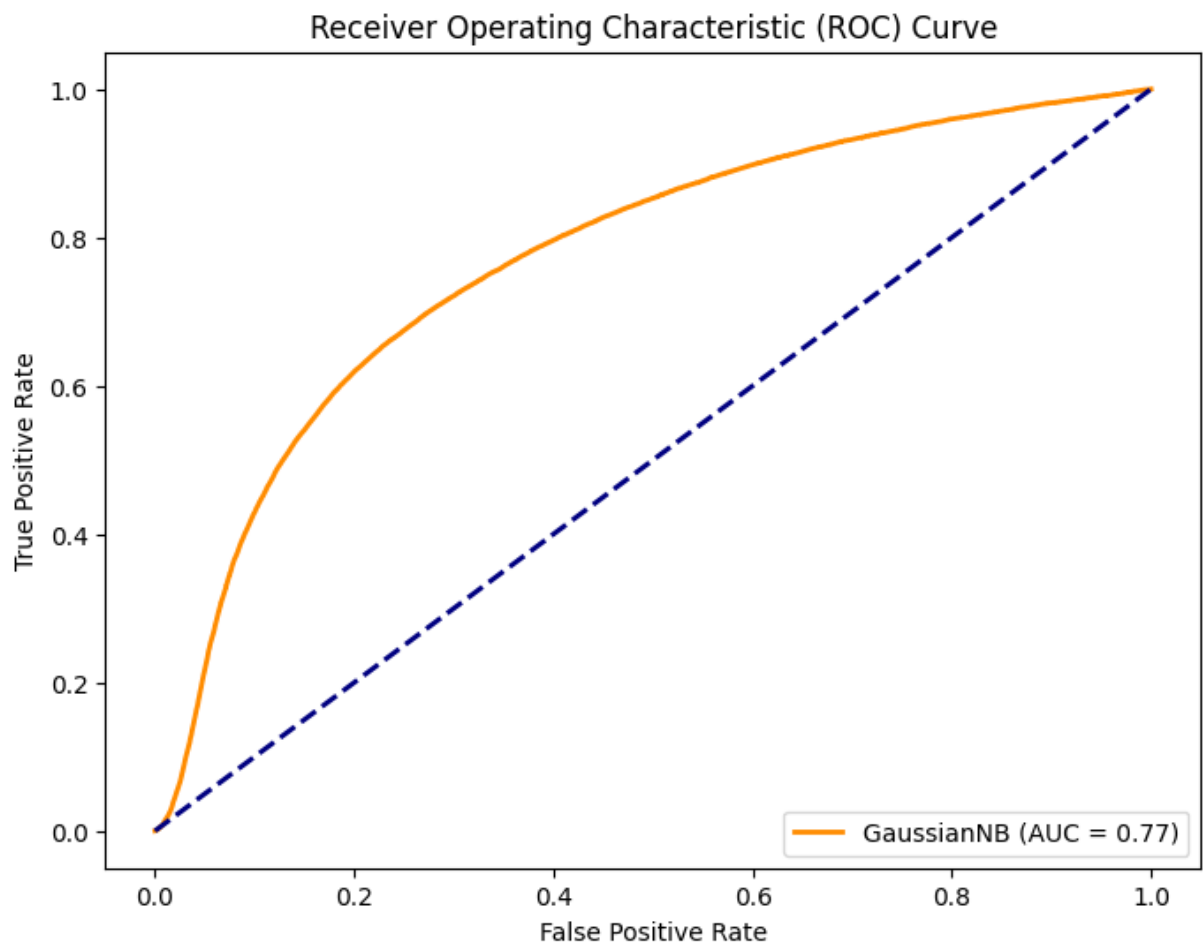
In [228... `plot_roc_curve(X, Y, rf, 'Random Forest')`



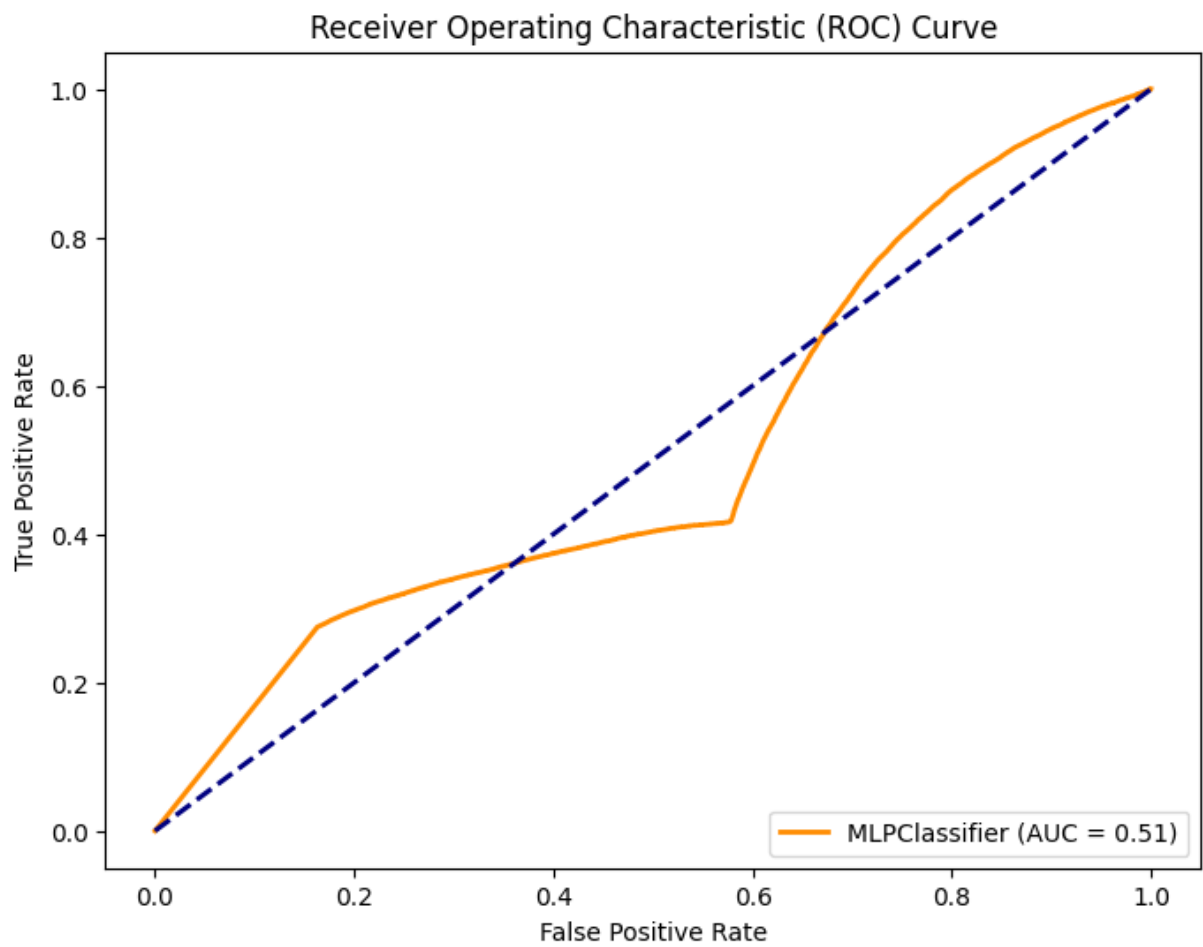
In [229... `plot_roc_curve(X, Y, dt, 'DecisionTree Classifier')`



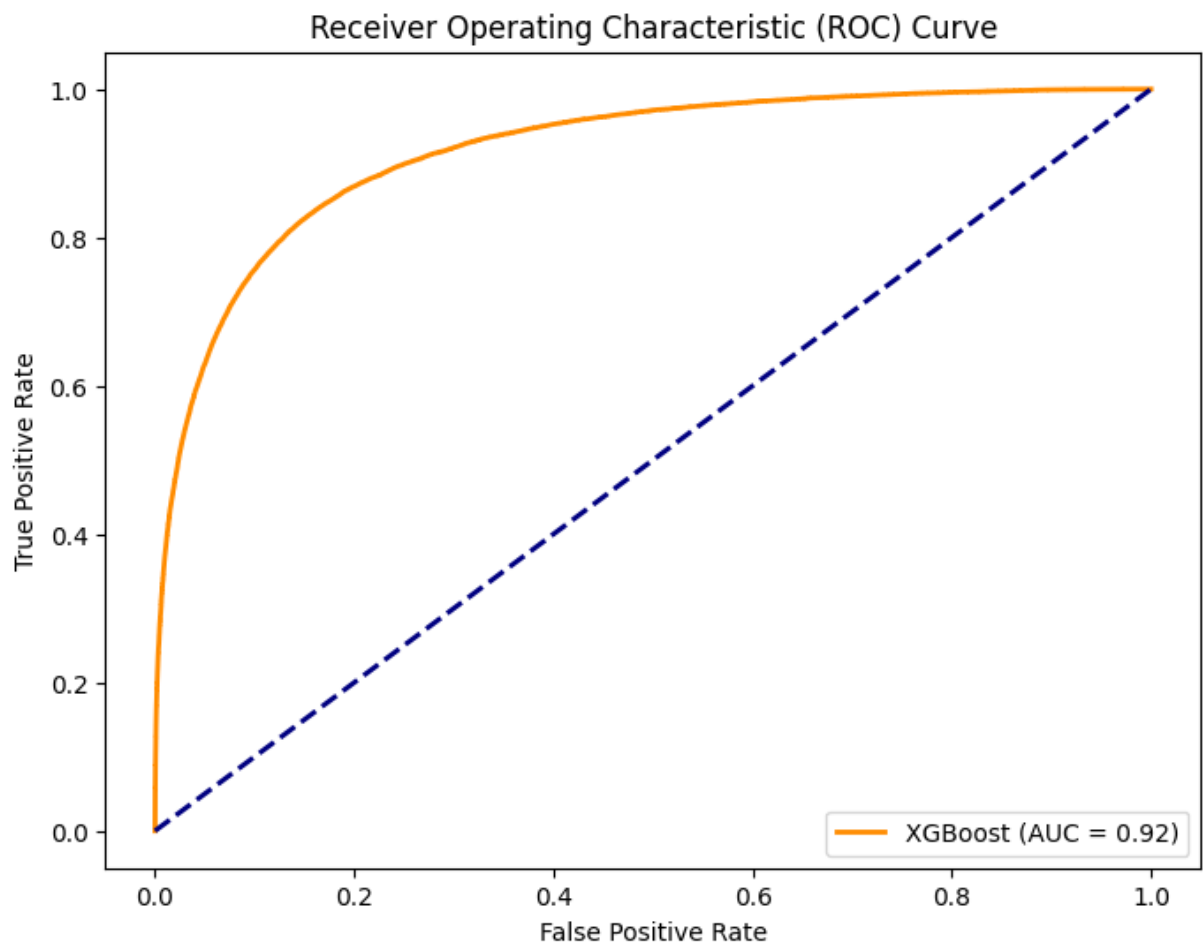
In [230... `plot_roc_curve(X, Y, gnb, 'GaussianNB')`



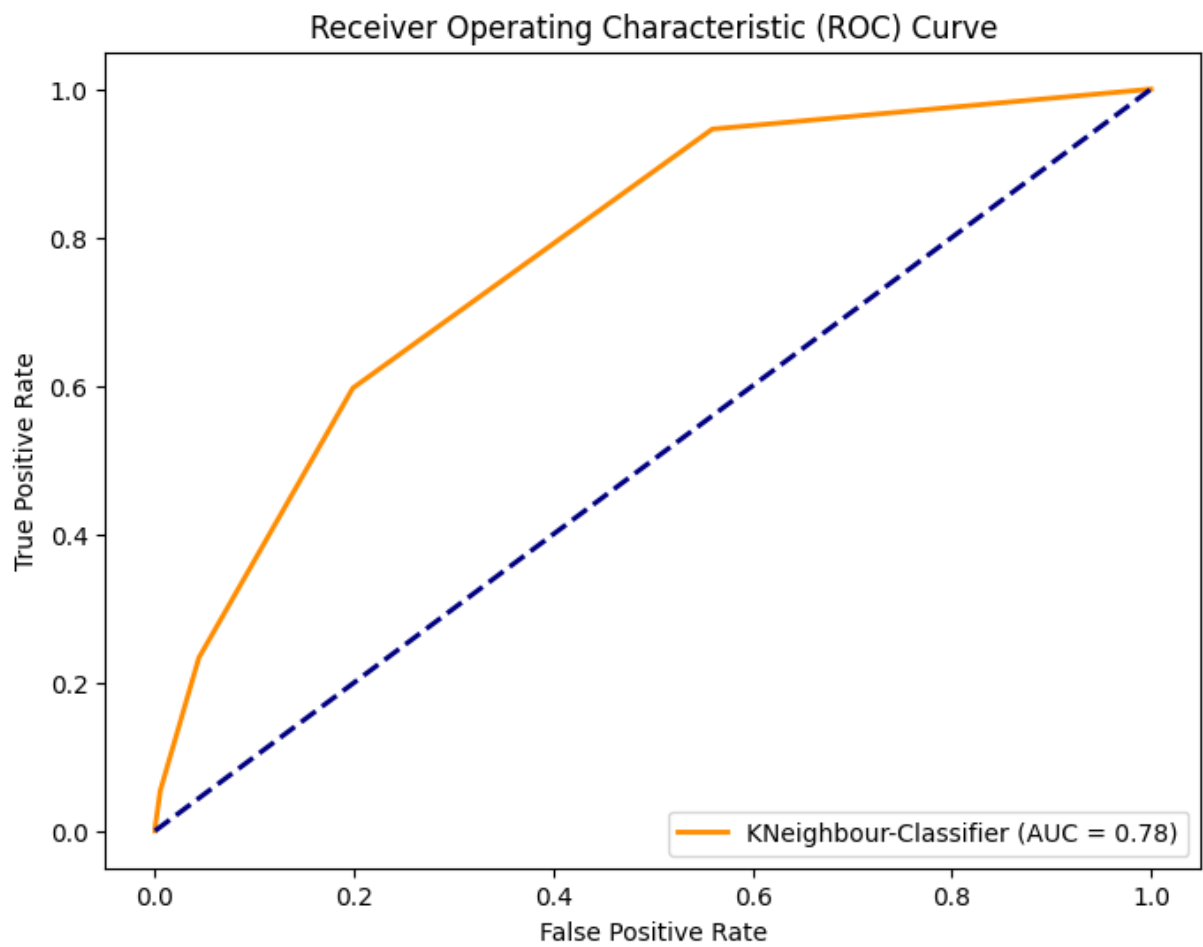
In [231... `plot_roc_curve(X, Y, mlp, 'MLPClassifier')`



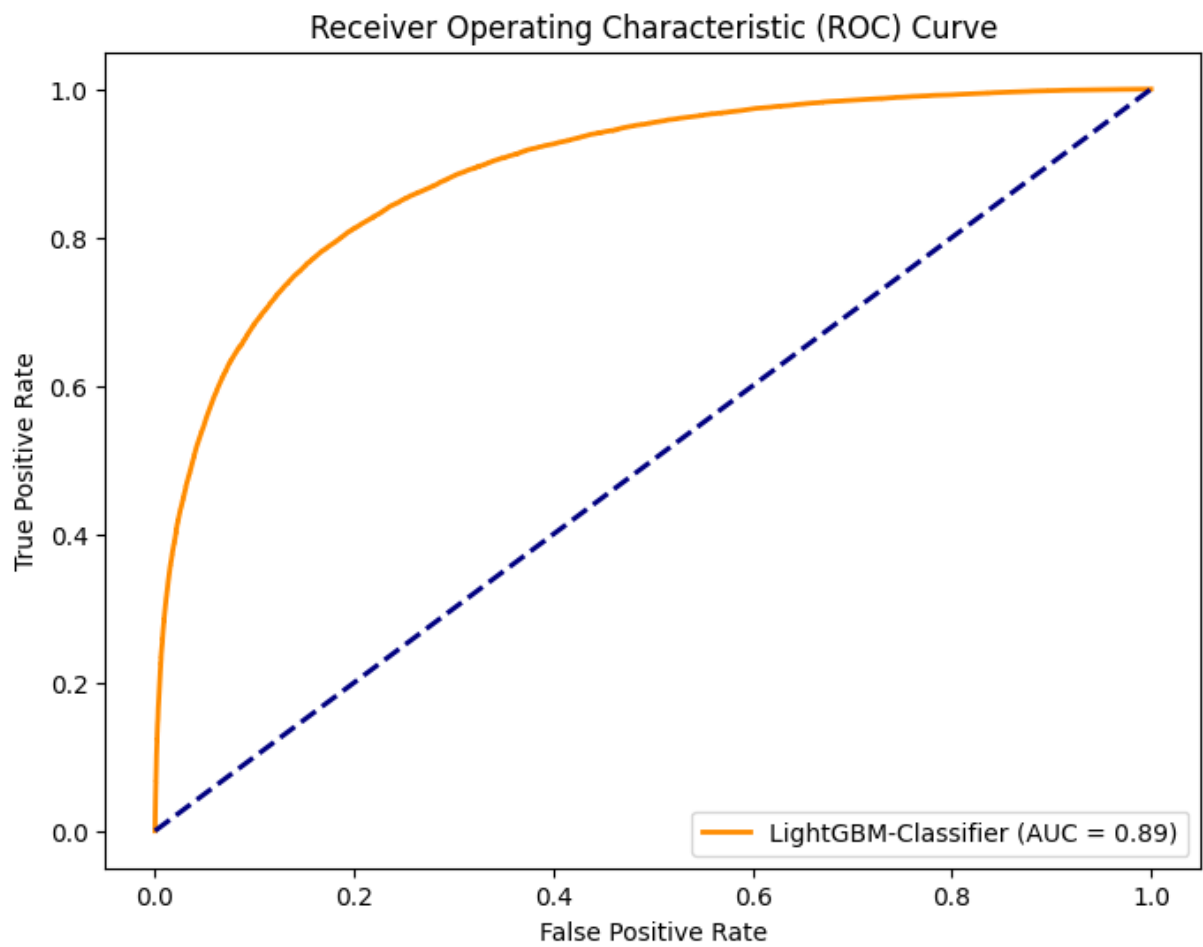
In [232... `plot_roc_curve(X, Y, xgb_model, 'XGBoost')`



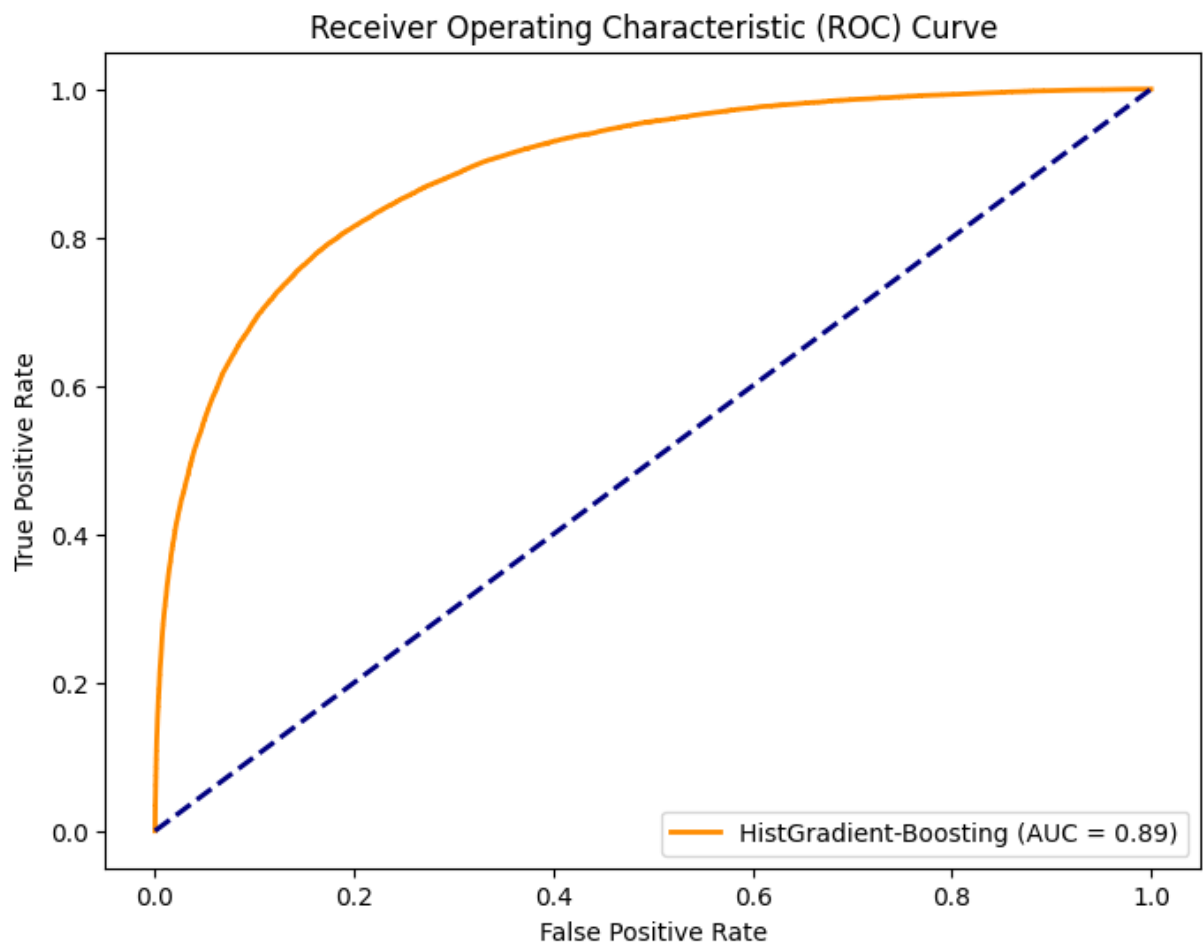
In [233... `plot_roc_curve(X, Y, knn, 'KNeighbour-Classififer')`



In [234... `plot_roc_curve(X, Y, lgb_model, 'LightGBM-Classififier')`



In [235... `plot_roc_curve(X, Y, hist_grad_boost, 'HistGradient-Boosting')`



Model Performance Summary

Top Performing Models:

- RandomForest, XGBoost, LightGBM, Gradient-Boosting, & HistGradient Boosting demonstrate high scores in **Accuracy** and **ROC-AUC**.

Interpretation of Scores:

- **Accuracy Score (0.80 and above):** Accuracy measures overall correctness, with scores of 0.80 or higher indicating a significant portion of correct predictions.

- **ROC-AUC Score (0.85 and above):** Reflects the model's ability to discriminate between positive and negative classes. Scores of 0.85 and above indicate strong performance in sensitivity and specificity.
- **Combined Strength:** High scores in both metrics suggest correct overall predictions and effective discrimination between positive and negative instances.

Performance Thresholds:

- **Accuracy and ROC-AUC:** Scores of 0.80 and above for accuracy and 0.85 and above for ROC-AUC indicate good overall model performance.

Ensemble Model Building

Understanding Ensemble Models

Ensemble modeling leverages the strength of multiple models to produce a more robust and accurate prediction. The premise is to aggregate the predictions from diverse base models, often resulting in superior performance compared to individual models.

Types of Ensemble Models

1. **Bagging (Bootstrap Aggregating):**

- Bagging involves training multiple instances of the same base model on different subsets of the training data, typically through bootstrapping (random sampling with replacement). The final prediction is an aggregate of these individual model predictions.

2. **Boosting:**

- Boosting focuses on sequentially training weak models and giving more weight to instances that were misclassified by earlier models. It aims to correct errors made by preceding models, ultimately creating a strong predictive model.

3. **Stacking:**

- Stacking involves training a meta-model that learns to combine the predictions of multiple base models. The base models' outputs serve as input features for the meta-model, leading to a more comprehensive and nuanced final prediction.

Ensemble Model Implementation

1. **Choosing Base Models:**

- Selecting diverse base models is crucial for the success of ensemble methods. Diversity can be achieved by using models with different architectures, hyperparameters, or training algorithms.

2. **Integration Techniques:**

- Deciding how predictions from base models are combined is essential. Common methods include averaging, weighted averaging, or using a meta-model for stacking.

Evaluation and Tuning

1. **Performance Metrics:**

- Ensemble models are evaluated using appropriate performance metrics, such as accuracy, precision, recall, F1-score, or area under the ROC curve (AUC-ROC).

2. **Hyperparameter Tuning:**

- Fine-tuning hyperparameters is essential for optimizing ensemble model performance. Techniques like grid search or random search can be employed.

Out[254...

	Model	Accuracy	Precision	object
7	XGBoost	0.865392	0.747999	XGBClassifier(base_score=None, booster=None, c...
9	LightGBM	0.864241	0.754422	LGBMClassifier(max_depth=3, random_state=42)
10	Gradient-Boosting	0.865059	0.746013	([DecisionTreeRegressor(criterion='friedman_ms...
11	HistGradient-Boosting	0.865513	0.749951	HistGradientBoostingClassifier(max_depth=4, ra...

In [255...

```
model_performance
```

Out[255...

	Model	10-folds oof ROC-AUC
7	XGBoost	0.888206
9	LightGBM	0.888502
10	Gradient-Boosting	0.889307
11	HistGradient-Boosting	0.889244

In [261...

```
# Initialize a list to store ensemble models
ensemble_models = []
for model, name in zip(filtered_models['object'], filtered_models['Model']):
    ensemble_models.append((name, model))
```

In [262...

```
ensemble_models
```

Out[262...

```
[('XGBoost',
  XGBClassifier(base_score=None, booster=None, callbacks=None,
                colsample_bylevel=None, colsample_bynode=None,
                colsample_bytree=0.3, device=None, early_stopping_rounds=None,
                enable_categorical=False, eval_metric='auc', feature_types=None,
                gamma=1.0, grow_policy=None, importance_type=None,
                interaction_constraints=None, learning_rate=0.05, max_bin=None,
                max_cat_threshold=None, max_cat_to_onehot=None,
                max_delta_step=None, max_depth=10, max_leaves=None,
                min_child_weight=None, missing=nan, monotone_constraints=None,
                multi_strategy=None, n_estimators=2048, n_jobs=-1,
                num_parallel_tree=None, random_state=42, ...)),
 ('LightGBM', LGBMClassifier(max_depth=3, random_state=42)),
 ('Gradient-Boosting',
  GradientBoostingClassifier(max_depth=4, min_samples_leaf=5, random_state=1,
                             subsample=0.8)),
 ('HistGradient-Boosting',
  HistGradientBoostingClassifier(max_depth=4, random_state=10))]
```

In [263...

```
# Initialize Voting Classifier:-
voting_md = VotingClassifier(estimators=ensemble_models, voting='soft')
```

```
In [264... # Cross-Validation and Print ROC-AUC Score:
vot_cv = cross_val_score(voting_md, X, Y, cv=skf, scoring='roc_auc', n_jobs=-1)
print(f"The 10 folds oof ROC-AUC score of Voting Classifier is {vot_cv.mean()}")
```

The 10 folds oof ROC-AUC score of Voting Classifier is 0.890193038687492

```
In [266... # Initialize Lists for Metrics:-
roc_auc_list = []
accuracy_list = []
precision_list = []
recall_list = []
```

```
In [267... # Iterate Over Folds for Cross-Validation:-
for train_idx, val_idx in skf.split(X, Y):
    X_train, X_val = X.iloc[train_idx], X.iloc[val_idx]
    y_train, y_val = Y.iloc[train_idx], Y.iloc[val_idx]
```

```
In [276... # Fitting Ensemble Model and Evaluate Metrics:-
vot_md = voting_md.fit(X_train, y_train)

y_val_pred_proba = vot_md.predict_proba(X_val)[: , 1]
fpr, tpr, _ = roc_curve(y_val, y_val_pred_proba)
roc_auc = auc(fpr, tpr)
roc_auc_list.append(roc_auc)
```

[LightGBM] [Info] Number of positive: 31429, number of negative: 117102
 [LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing was 0.011983 seconds.
 You can set `force_col_wise=true` to remove the overhead.
 [LightGBM] [Info] Total Bins 859
 [LightGBM] [Info] Number of data points in the train set: 148531, number of used features: 11
 [LightGBM] [Info] [binary:BoostFromScore]: pavg=0.211599 -> initscore=-1.315314
 [LightGBM] [Info] Start training from score -1.315314
 [LightGBM] [Warning] No further splits with positive gain, best gain: -inf

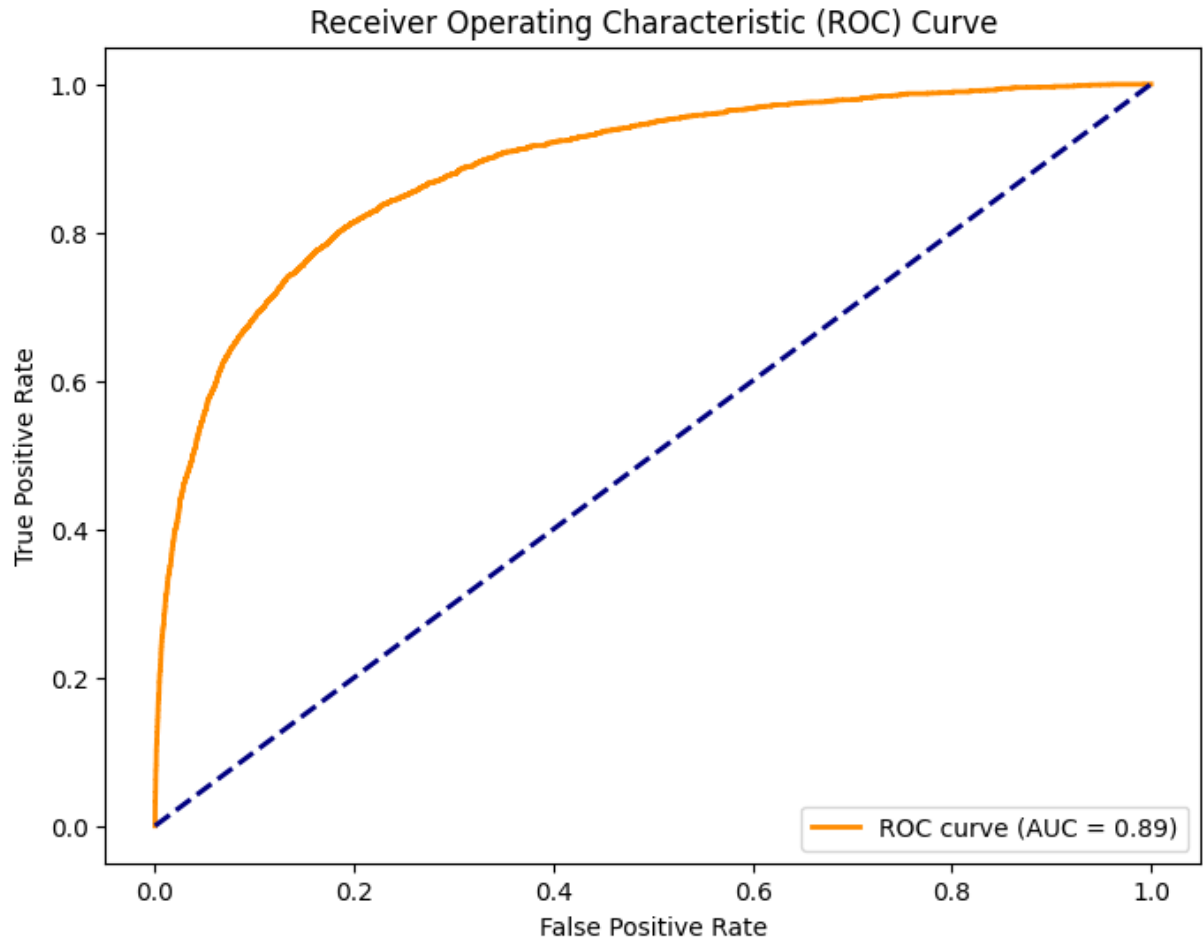
```
In [277... y_val_pred = vot_md.predict(X_val)
accuracy_list.append(accuracy_score(y_val, y_val_pred))
precision_list.append(precision_score(y_val, y_val_pred))
recall_list.append(recall_score(y_val, y_val_pred))
```

```
In [278... print(f"Average ROC-AUC: {np.mean(roc_auc_list)}")
print(f"Average Accuracy: {np.mean(accuracy_list)}")
print(f"Average Precision: {np.mean(precision_list)}")
print(f"Average Recall: {np.mean(recall_list)}")
```

Average ROC-AUC: 0.8870105769168972
 Average Accuracy: 0.8666303096406714
 Average Precision: 0.7564560985299961
 Average Recall: 0.54524627720504

```
In [279... # Plotting ROC curve
plt.figure(figsize=(8, 6))
plt.plot(fpr, tpr, color='darkorange', lw=2, label=f'ROC curve (AUC = {np.mean(roc_auc_list)})')
plt.plot([0, 1], [0, 1], color='navy', lw=2, linestyle='--')
plt.xlabel('False Positive Rate')
```

```
plt.ylabel('True Positive Rate')
plt.title('Receiver Operating Characteristic (ROC) Curve')
plt.legend(loc='lower right')
plt.show()
```



```
In [281... # Processing the test dataset for making predictions
test_md = test.drop(columns=['id', 'CustomerId', 'Surname'], axis=1)
test_md['Gender'] = test_md['Gender'].map({'Female': 0, 'Male': 1})
test_md = pd.concat([test_md.drop(columns=['Geography'], axis=1), pd.get_dummies(test_md['Geography'], drop_first=True)], axis=1)
```

```
In [283... vot_pred = vot_md.predict_proba(test_md)[: , 1]
submission['Exited'] = vot_pred
submission.to_csv('Advanced_ensemble_sub_2.csv', index=False)
```

```
In [ ]:
```