

YBI Internship - Telco Customer Churn Analysis

By Hemalatha Kandivan

File loading

```
In [1]: from google.colab import files
        uploaded = files.upload()
```

Choose Files No file chosen

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

Saving synthetic_customer_churn_100k.csv to synthetic_customer_churn_100k.csv

Data Wrangling

```
In [2]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
```

```
In [3]: df = pd.read_csv('synthetic_customer_churn_100k.csv')
        df.head()
```

Out[3]:

	CustomerID	Age	Gender	Tenure	MonthlyCharges	Contract	PaymentMethod	TotalCharges
--	------------	-----	--------	--------	----------------	----------	---------------	--------------

0	1	56	Female	68	147.58	Two year	Bank transfer	1081.83
1	2	69	Male	32	22.54	Month-to-month	Mailed check	726.38
2	3	46	Female	10	52.47	One year	Electronic check	524.70
3	4	32	Male	22	109.67	Month-to-month	Mailed check	2412.76
4	5	60	Female	54	130.98	Month-to-month	Credit card	7072.02



```
In [4]: df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100000 entries, 0 to 99999
Data columns (total 9 columns):
#   Column                Non-Null Count  Dtype
---  -
0   CustomerID            100000 non-null  int64
1   Age                   100000 non-null  int64
2   Gender                 100000 non-null  object
3   Tenure                 100000 non-null  int64
4   MonthlyCharges        100000 non-null  float64
5   Contract               100000 non-null  object
6   PaymentMethod          100000 non-null  object
7   TotalCharges           100000 non-null  float64
8   Churn                  100000 non-null  object
dtypes: float64(2), int64(3), object(4)
memory usage: 6.9+ MB

```

```

In [5]: #check for null values
        df.isnull().sum()

```

```

Out[5]: 0

```

CustomerID	0
Age	0
Gender	0
Tenure	0
MonthlyCharges	0
Contract	0
PaymentMethod	0
TotalCharges	0
Churn	0

dtype: int64

```

In [6]: #drop duplicates if there are any
        df = df.drop_duplicates()
        df.info()

```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100000 entries, 0 to 99999
Data columns (total 9 columns):
#   Column                Non-Null Count  Dtype
---  -
0   CustomerID            100000 non-null  int64
1   Age                   100000 non-null  int64
2   Gender                 100000 non-null  object
3   Tenure                 100000 non-null  int64
4   MonthlyCharges        100000 non-null  float64
5   Contract               100000 non-null  object
6   PaymentMethod          100000 non-null  object
7   TotalCharges           100000 non-null  float64
8   Churn                  100000 non-null  object
dtypes: float64(2), int64(3), object(4)
memory usage: 6.9+ MB

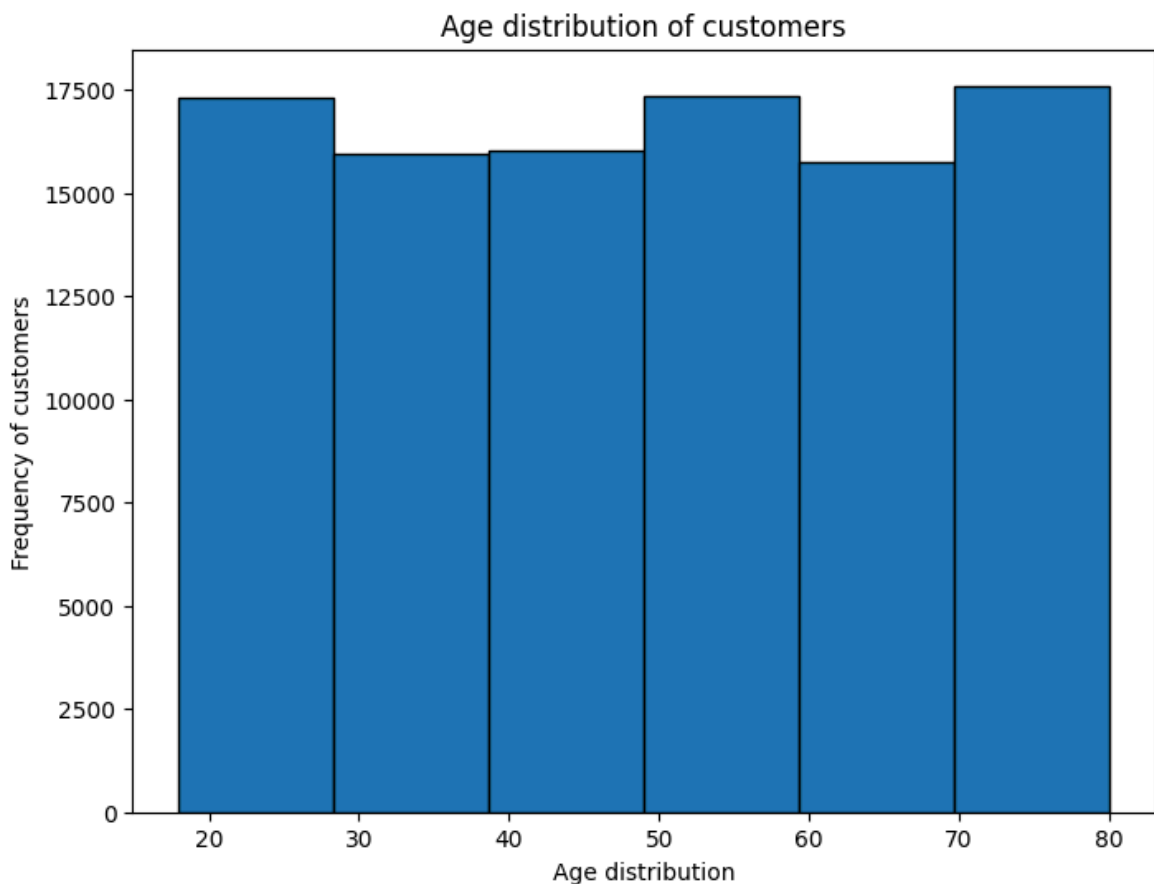
```

In [7]: *#visualize patterns*

```

plt.figure(figsize=(8,6))
plt.hist(df['Age'], bins=6, edgecolor='black')
plt.xlabel('Age distribution')
plt.ylabel('Frequency of customers')
plt.title('Age distribution of customers')
plt.show()

```

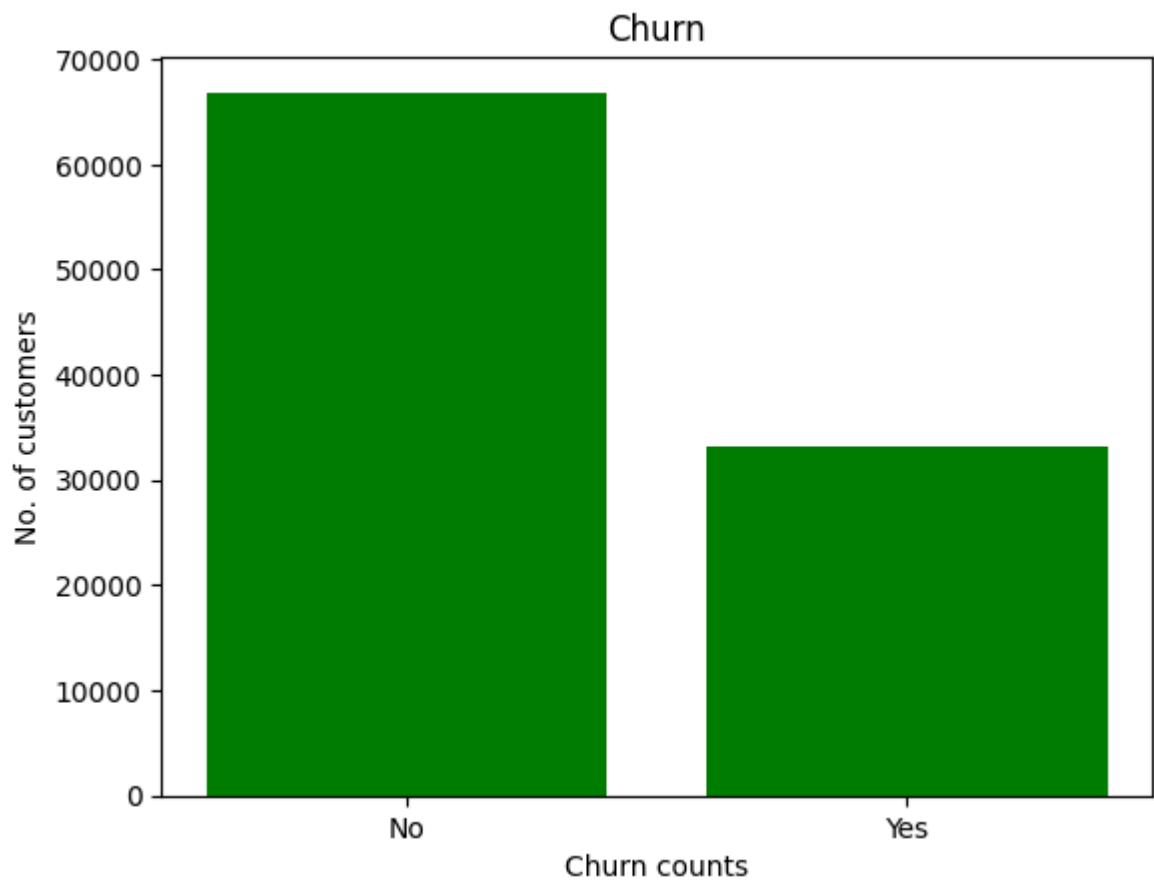


In [8]:

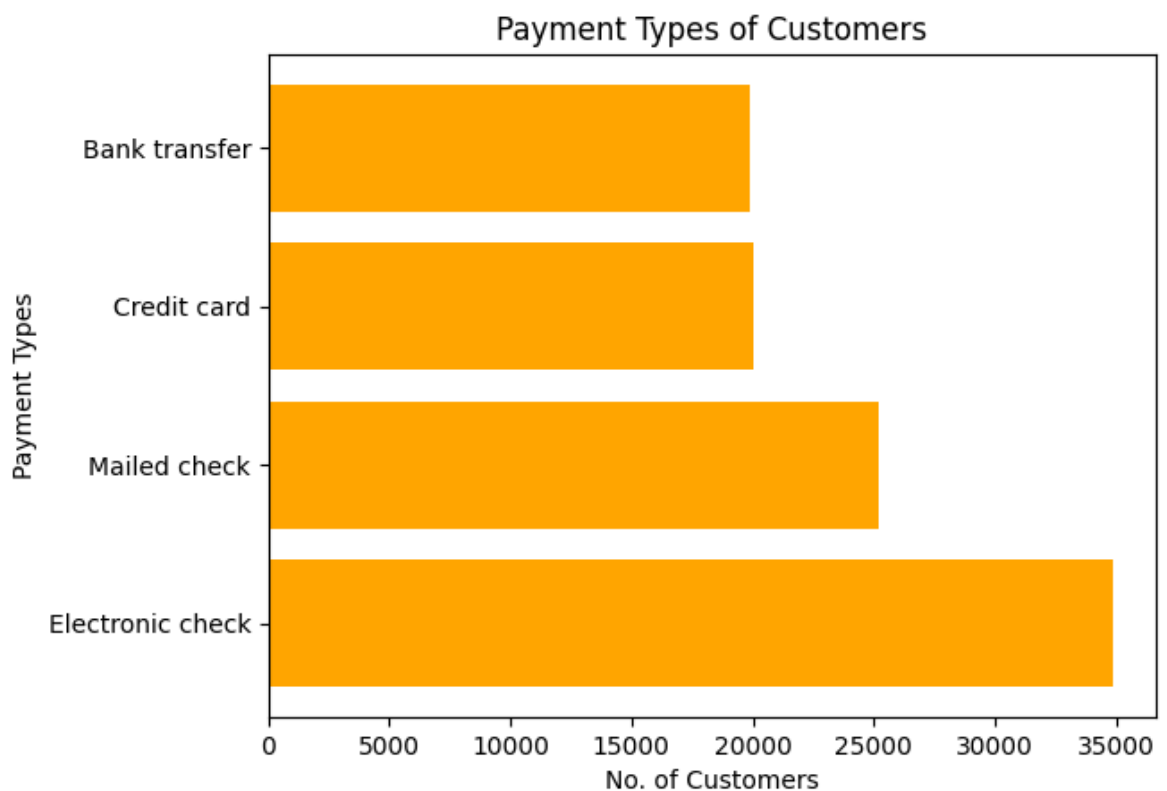
```

churn_counts = df['Churn'].value_counts()
plt.bar(churn_counts.index, churn_counts.values, color='green')
plt.xlabel('Churn counts')
plt.ylabel('No. of customers')
plt.title('Churn')
plt.show()

```



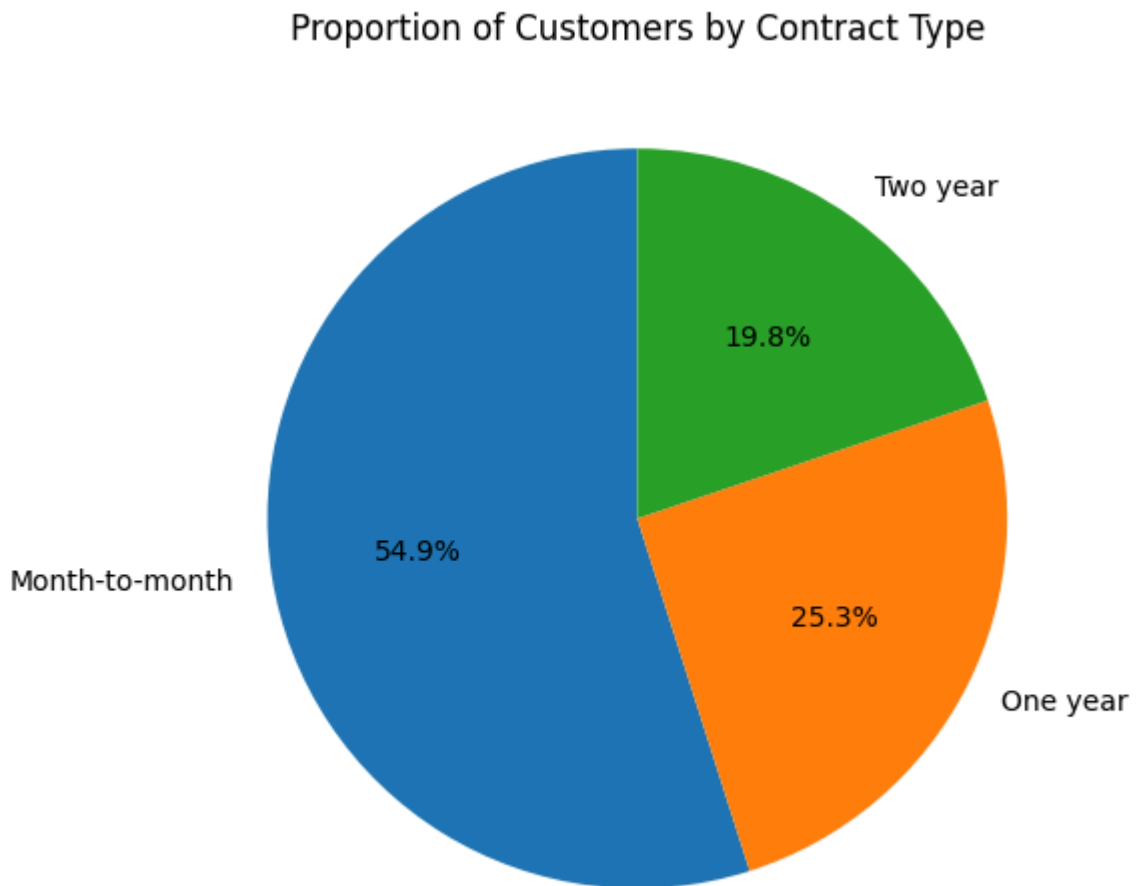
```
In [9]: payment_counts = df['PaymentMethod'].value_counts()
plt.barh(payment_counts.index, payment_counts.values, color='orange')
plt.xlabel('No. of Customers')
plt.ylabel('Payment Types')
plt.title('Payment Types of Customers')
plt.show()
```



```
In [10]: con_counts = df['Contract'].value_counts()

plt.figure(figsize=(6,6))
plt.pie(
    con_counts,
    labels=con_counts.index,
    autopct='%1.1f%%',
    startangle=90,
)

plt.title('Proportion of Customers by Contract Type')
plt.show()
```



Data cleaning

```
In [11]: #convert object data types into numerical form for ML training using label encoder
from sklearn.preprocessing import LabelEncoder

le = LabelEncoder()
df['Gender_encoded'] = le.fit_transform(df['Gender'])
df['Churn_encoded'] = le.fit_transform(df['Churn']) #0 is No and 1 is Yes
df['Contract_encoded'] = le.fit_transform(df['Contract'])
df['PaymentMethod_encoded'] = le.fit_transform(df['PaymentMethod'])
df.head()
```

Out[11]:

	CustomerID	Age	Gender	Tenure	MonthlyCharges	Contract	PaymentMethod	TotalCharges
0	1	56	Female	68	147.58	Two year	Bank transfer	1034.34
1	2	69	Male	32	22.54	Month-to-month	Mailed check	815.27
2	3	46	Female	10	52.47	One year	Electronic check	524.95
3	4	32	Male	22	109.67	Month-to-month	Mailed check	2400.96
4	5	60	Female	54	130.98	Month-to-month	Credit card	7066.48

Data Training

```
In [13]: from sklearn.preprocessing import StandardScaler
from lightgbm import LGBMClassifier
from xgboost import XGBClassifier
from sklearn.linear_model import LogisticRegression
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
#X is scaled after split to ensure no data leakage

X = df.drop(['Gender', 'Contract', 'PaymentMethod', 'Churn', 'Churn_encoded'], axis=1)
y = df['Churn_encoded']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.20, random_state=42)

scaler = StandardScaler()
scaled_xtrain = scaler.fit_transform(X_train)
scaled_xtest = scaler.fit_transform(X_test)

print(scaled_xtrain)
print(scaled_xtest)
```

```
[ 0.87282776 -1.43176326 -0.94232825 ... 0.77816513 -0.82113055
 0.32523686]
[-0.03695691 -1.26672572 -1.56742879 ... 0.77816513 -0.82113055
 0.32523686]
[-0.17513057 0.76873731 1.31765064 ... 2.53039466 -0.82113055
 0.32523686]
...
[ 0.92824964 -0.60657555 -0.26914305 ... 0.77816513 1.71092735
 0.32523686]
[-1.70290376 -1.48677577 0.64446544 ... -0.97406439 0.4448984
 0.32523686]
[-1.18557522 -0.60657555 -1.08658222 ... 0.77816513 0.4448984
 -1.55951132]]
[[ 0.89437411 -0.7718974 -0.1563886 ... -0.98239718 -0.8231433
 -0.61466417]
 [ 1.0490198 1.53857149 -0.34908329 ... 0.7673529 0.44109154
 -0.61466417]
 [-1.0411052 -0.99194206 1.04795321 ... 0.7673529 -0.8231433
 1.26375028]
 ...
 [-0.41746346 -0.88191973 -1.36073042 ... -0.98239718 0.44109154
 -1.55387139]
 [ 0.12166543 -0.22178576 -0.39725696 ... 0.7673529 -0.8231433
 -1.55387139]
 [-0.25987247 0.27331472 -1.36073042 ... 0.7673529 -0.8231433
 0.32454306]]
```

```
In [24]: lgbm = LGBMClassifier(class_weight='balanced', n_estimators=300, learning_rate=0.05,
                             scale_pos_weight = len(y_train[y_train==0]) / len(y_train[y_train==1]))
xgb = XGBClassifier(scale_pos_weight=scale_pos_weight)

lr = LogisticRegression()

rfc = RandomForestClassifier()
```

Data Evaluation

```
In [25]: from sklearn.metrics import classification_report, accuracy_score, roc_auc_score

models = {'LightGBM': lgbm,
          'XGBoost': xgb,
          'Logistic Regression': lr,
          'Random Forest Classifier': rfc
        }

target_names = ['Class 0', 'Class 1']

for name, model in models.items():

    model.fit(scaled_xtrain, y_train)

    preds = model.predict(scaled_xtest)

    probs = model.predict_proba(scaled_xtest)[:,-1]

    evals = classification_report(y_test, preds, target_names=target_names)
```

```

accu_score = accuracy_score(y_test, preds)

roc = roc_auc_score(y_test, preds)

print(f'For {name}:')
print(f'Accuracy score: {accu_score*100:.2f}%')
print(f'Classification report: \n{evals}')
print(f'ROC-AUC score: {roc:.3f}\n')

```

[LightGBM] [Info] Number of positive: 26438, number of negative: 53562
 [LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.002152 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 915

[LightGBM] [Info] Number of data points in the train set: 80000, number of used features: 8

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000

[LightGBM] [Info] Start training from score 0.000000

/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names

warnings.warn(

/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names

warnings.warn(

For LightGBM:
 Accuracy score: 73.77%
 Classification report:

	precision	recall	f1-score	support
Class 0	0.81	0.79	0.80	13294
Class 1	0.60	0.64	0.62	6706
accuracy			0.74	20000
macro avg	0.71	0.71	0.71	20000
weighted avg	0.74	0.74	0.74	20000

ROC-AUC score: 0.714

For XGBoost:
 Accuracy score: 66.05%
 Classification report:

	precision	recall	f1-score	support
Class 0	0.82	0.62	0.71	13294
Class 1	0.50	0.74	0.59	6706
accuracy			0.66	20000
macro avg	0.66	0.68	0.65	20000
weighted avg	0.71	0.66	0.67	20000

ROC-AUC score: 0.679

For Logistic Regression:
 Accuracy score: 72.34%
 Classification report:

	precision	recall	f1-score	support
Class 0	0.75	0.86	0.81	13294
Class 1	0.62	0.44	0.52	6706
accuracy			0.72	20000
macro avg	0.69	0.65	0.66	20000
weighted avg	0.71	0.72	0.71	20000

ROC-AUC score: 0.654

For Random Forest Classifier:
 Accuracy score: 74.49%
 Classification report:

	precision	recall	f1-score	support
Class 0	0.78	0.87	0.82	13294
Class 1	0.66	0.51	0.57	6706
accuracy			0.74	20000
macro avg	0.72	0.69	0.69	20000
weighted avg	0.74	0.74	0.74	20000

ROC-AUC score: 0.686

```
In [38]: #since LGBM showed best ROC-AUC score, we'll fine tune LGBM for better results

from sklearn.model_selection import GridSearchCV
```

```
params = {
    'num_leaves': [31, 63],
    'max_depth': [5, 10],
    'learning_rate': [0.01, 0.05],
    'n_estimators': [200, 400]
}

grid = GridSearchCV(lgbm, params, scoring='roc_auc', cv=5)
grid.fit(scaled_xtrain, y_train)

best_model = grid.best_estimator_
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001682 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001603 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001614 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001581 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001613 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.003028 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001666 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001596 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001595 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001599 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.002373 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001587 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.002934 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001590 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001599 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.002900 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001611 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001616 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(
```

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(
```

[illegible]

[illegible]

[illegible]

```

/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(

```

[illegible]

[illegible]

[illegible]

```

/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(

```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```

/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(

```

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(
```

[illegible]

[illegible]

[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
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
```

```
warnings.warn(
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```

/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but LGBMClassifier was fitted with feature names
  warnings.warn(

```



```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001685 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001631 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001650 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001598 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001610 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001618 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.002945 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.002846 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001624 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001603 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001588 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001621 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001951 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.003072 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001616 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```
[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001644 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
```

```
warnings.warn(
```

```

[LightGBM] [Info] Number of positive: 21151, number of negative: 42849
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001638 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=-0.000000
[LightGBM] [Info] Start training from score -0.000000

/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
  warnings.warn(
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing col-wise multi-threading, the overhead of testing
was 0.005579 seconds.
You can set `force_col_wise=true` to remove the overhead.
[LightGBM] [Info] Total Bins 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000

/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
  warnings.warn(
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.003026 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000

/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
  warnings.warn(
[LightGBM] [Info] Number of positive: 21150, number of negative: 42850
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.001611 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 915
[LightGBM] [Info] Number of data points in the train set: 64000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000

/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
  warnings.warn(

```

```
[LightGBM] [Info] Number of positive: 26438, number of negative: 53562
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.003540 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 915
[LightGBM] [Info] Number of data points in the train set: 80000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
In [41]: print(best_model)
```

```
LGBMClassifier(class_weight='balanced', learning_rate=0.01, max_depth=10,
               n_estimators=400)
```

```
In [42]: #refit tuned model with data
```

```
new_lgbm = LGBMClassifier(class_weight='balanced', learning_rate=0.01,
                          max_depth=10, n_estimators=400)
```

```
In [43]: new_lgbm.fit(scaled_xtrain, y_train)
```

```
y_pred = new_lgbm.predict(scaled_xtest)
```

```
[LightGBM] [Info] Number of positive: 26438, number of negative: 53562
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing
was 0.007544 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 915
[LightGBM] [Info] Number of data points in the train set: 80000, number of used f
eatures: 8
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
warnings.warn(
```

```
In [44]: target_names = ['Class 0', 'Class 1']
```

```
probs = new_lgbm.predict_proba(scaled_xtest)[:,-1]
```

```
evals = classification_report(y_test, y_pred, target_names=target_names)
```

```
accu_score = accuracy_score(y_test, y_pred)
```

```
roc = roc_auc_score(y_test, y_pred)
```

```
print(f'Accuracy score: {accu_score*100:.2f}%')
```

```
print(f'Classification report: \n{evals}')
```

```
print(f'ROC-AUC score: {roc:.3f}\n')
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWar
ning: X does not have valid feature names, but LGBMClassifier was fitted with fea
ture names
warnings.warn(
```

Accuracy score: 74.89%

Classification report:

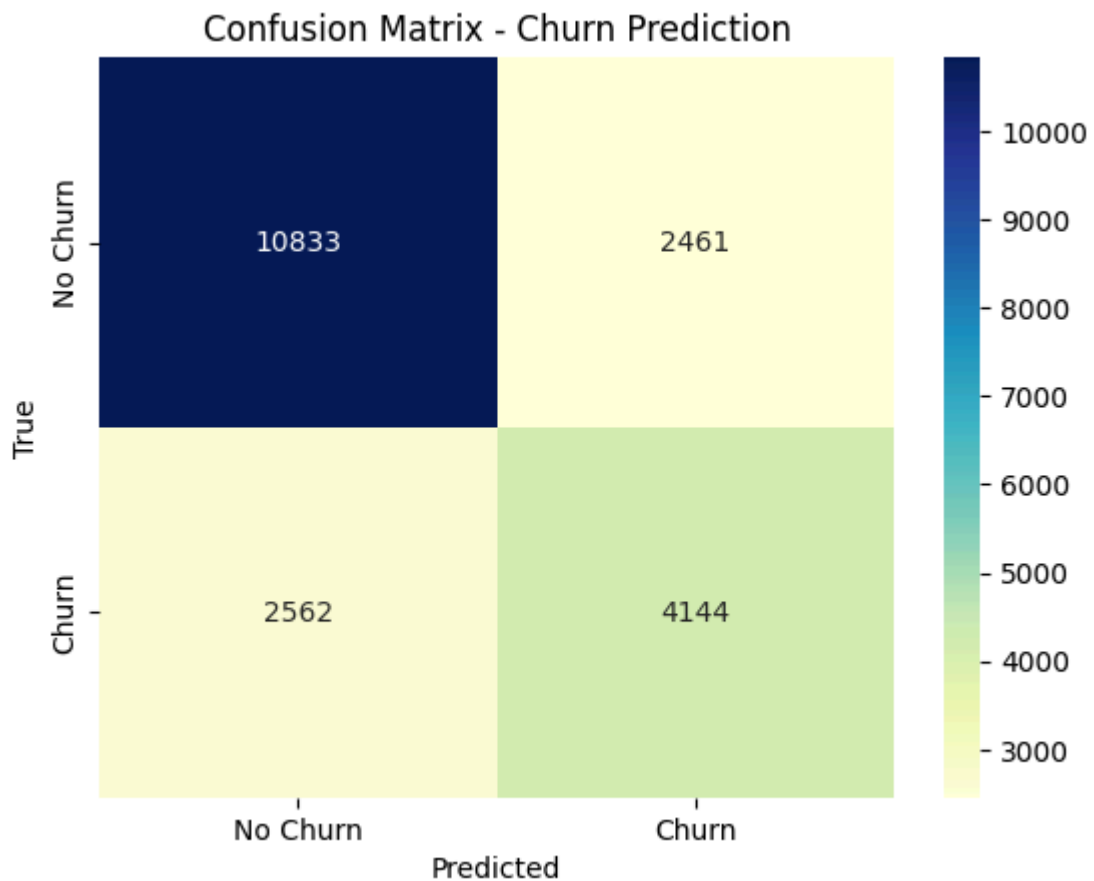
	precision	recall	f1-score	support
Class 0	0.81	0.81	0.81	13294
Class 1	0.63	0.62	0.62	6706
accuracy			0.75	20000
macro avg	0.72	0.72	0.72	20000
weighted avg	0.75	0.75	0.75	20000

ROC-AUC score: 0.716

```
In [52]: from sklearn.metrics import confusion_matrix
import seaborn as sns

cm = confusion_matrix(y_test, y_pred)

sns.heatmap(cm,
            annot=True,
            fmt='d',
            cmap='YlGnBu',
            xticklabels=['No Churn', 'Churn'],
            yticklabels=['No Churn', 'Churn']
            )
plt.xlabel('Predicted')
plt.ylabel('True')
plt.title('Confusion Matrix - Churn Prediction')
plt.show()
```



Dissussion and Conclusion

The accuracy of the fine-tuned LGBM model is 74.89% which is good baseline but other metrics than accuracy are more important to be looked at.

The ROC-AUC metric is used to check how well a model works in binary classification. With the best ROC-AUC score among other models, the LGBM model can distinguish churn vs not churn 71.6% of the time. Of all churners, LGBM model caught 62% of churners correctly with recall of 62% and precision of 63%. The balance between precision and recall, F1 score is 0.62.

This confusion matrix shows that the model correctly predicted 10,833 customers as not churned (True Negative) and 4144 customers as churned (True Positive). False Negative is 2562 which means that 2562 customers churned but model predicts as not churned and False Positive is 2461.

Overall, this project evaluated multiple classification machine learning models for churn prediction using a large synthetic dataset designed to simulate realistic customer behavior patterns. The results demonstrate the model's ability to capture complex nonlinear relationships present in churn behavior while highlighting inherent prediction challenges caused by overlapping customer profiles. The synthetic nature of the dataset intentionally incorporates behavioral variability, making perfect separation unrealistic and reflective of real-world scenarios.

Future work will focus on enhanced feature engineering, recall-focused optimization, advanced ensemble methods, and temporal modeling to further improve churn detection performance.

Source of data: <https://www.kaggle.com/datasets/dhrubangtalukdar/telco-customer-churn-data>

In []: