

xgm-and-lgbm

November 25, 2024

```
[1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore")
```

```
[2]: data=pd.read_csv(r"/content/Titanic_train.csv")
data
```

```
[2]: PassengerId  Survived  Pclass  \
0                1         0       3
1                2         1       1
2                3         1       3
3                4         1       1
4                5         0       3
..            ...     ...     ...
886            887         0       2
887            888         1       1
888            889         0       3
889            890         1       1
890            891         0       3
```

```

                                Name    Sex  Age  SibSp  \
0                Braund, Mr. Owen Harris  male  22.0    1
1  Cumings, Mrs. John Bradley (Florence Briggs Th... female  38.0    1
2                Heikkinen, Miss. Laina  female  26.0    0
3  Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0    1
4                Allen, Mr. William Henry  male  35.0    0
..            ...     ...     ...     ...
886                Montvila, Rev. Juozas  male  27.0    0
887                Graham, Miss. Margaret Edith  female  19.0    0
888  Johnston, Miss. Catherine Helen "Carrie"  female   NaN    1
889                Behr, Mr. Karl Howell  male  26.0    0
890                Dooley, Mr. Patrick  male  32.0    0
```

```

Parch  Ticket  Fare Cabin Embarked
```

0	0	A/5 21171	7.2500	NaN	S
1	0	PC 17599	71.2833	C85	C
2	0	STON/O2. 3101282	7.9250	NaN	S
3	0	113803	53.1000	C123	S
4	0	373450	8.0500	NaN	S
..
886	0	211536	13.0000	NaN	S
887	0	112053	30.0000	B42	S
888	2	W./C. 6607	23.4500	NaN	S
889	0	111369	30.0000	C148	C
890	0	370376	7.7500	NaN	Q

[891 rows x 12 columns]

```
[3]: data.isnull().sum()
```

```
[3]: PassengerId      0
      Survived        0
      Pclass          0
      Name            0
      Sex             0
      Age            177
      SibSp           0
      Parch           0
      Ticket          0
      Fare            0
      Cabin          687
      Embarked        2
      dtype: int64
```

```
[4]: train_df = pd.read_csv(r"/content/Titanic_train.csv")
      test_df = pd.read_csv(r"/content/Titanic_test.csv")
      combine = [train_df, test_df]
```

```
[5]: combine
```

```
[5]: [   PassengerId  Survived  Pclass  \
      0           1         0       3
      1           2         1       1
      2           3         1       3
      3           4         1       1
      4           5         0       3
      ..         ...         ...     ...
      886         887         0       2
      887         888         1       1
      888         889         0       3
      889         890         1       1
```

890 891 0 3

		Name	Sex	Age	SibSp	\
0		Braund, Mr. Owen Harris	male	22.0	1	
1	Cummings, Mrs. John Bradley (Florence Briggs Th...		female	38.0	1	
2		Heikkinen, Miss. Laina	female	26.0	0	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)		female	35.0	1	
4		Allen, Mr. William Henry	male	35.0	0	
..			
886		Montvila, Rev. Juozas	male	27.0	0	
887		Graham, Miss. Margaret Edith	female	19.0	0	
888	Johnston, Miss. Catherine Helen "Carrie"		female	NaN	1	
889		Behr, Mr. Karl Howell	male	26.0	0	
890		Dooley, Mr. Patrick	male	32.0	0	

	Parch	Ticket	Fare	Cabin	Embarked
0	0	A/5 21171	7.2500	NaN	S
1	0	PC 17599	71.2833	C85	C
2	0	STON/O2. 3101282	7.9250	NaN	S
3	0	113803	53.1000	C123	S
4	0	373450	8.0500	NaN	S
..	
886	0	211536	13.0000	NaN	S
887	0	112053	30.0000	B42	S
888	2	W./C. 6607	23.4500	NaN	S
889	0	111369	30.0000	C148	C
890	0	370376	7.7500	NaN	Q

[891 rows x 12 columns],

	PassengerId	Pclass	Name	\
0	892	3	Kelly, Mr. James	
1	893	3	Wilkes, Mrs. James (Ellen Needs)	
2	894	2	Myles, Mr. Thomas Francis	
3	895	3	Wirz, Mr. Albert	
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	
..	
413	1305	3	Spector, Mr. Woolf	
414	1306	1	Oliva y Ocana, Dona. Fermina	
415	1307	3	Saether, Mr. Simon Sivertsen	
416	1308	3	Ware, Mr. Frederick	
417	1309	3	Peter, Master. Michael J	

	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	male	34.5	0	0	330911	7.8292	NaN	Q
1	female	47.0	1	0	363272	7.0000	NaN	S
2	male	62.0	0	0	240276	9.6875	NaN	Q
3	male	27.0	0	0	315154	8.6625	NaN	S

4	female	22.0	1	1	3101298	12.2875	NaN	S
..
413	male	NaN	0	0	A.5. 3236	8.0500	NaN	S
414	female	39.0	0	0	PC 17758	108.9000	C105	C
415	male	38.5	0	0	SOTON/O.Q. 3101262	7.2500	NaN	S
416	male	NaN	0	0	359309	8.0500	NaN	S
417	male	NaN	1	1	2668	22.3583	NaN	C

[418 rows x 11 columns]]

```
[6]: train_df.head()
```

```
[6]: PassengerId  Survived  Pclass  \
0             1         0         3
1             2         1         1
2             3         1         3
3             4         1         1
4             5         0         3
```

	Name	Sex	Age	SibSp	\
0	Braund, Mr. Owen Harris	male	22.0	1	
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	
2	Heikkinen, Miss. Laina	female	26.0	0	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	
4	Allen, Mr. William Henry	male	35.0	0	

	Parch	Ticket	Fare	Cabin	Embarked
0	0	A/5 21171	7.2500	NaN	S
1	0	PC 17599	71.2833	C85	C
2	0	STON/O2. 3101282	7.9250	NaN	S
3	0	113803	53.1000	C123	S
4	0	373450	8.0500	NaN	S

```
[7]: train_df.columns
```

```
[7]: Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',
        'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
        dtype='object')
```

```
[8]: train_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
#   Column          Non-Null Count  Dtype
---  -
0   PassengerId     891 non-null   int64
```

```

1  Survived      891 non-null    int64
2  Pclass       891 non-null    int64
3  Name         891 non-null    object
4  Sex          891 non-null    object
5  Age         714 non-null    float64
6  SibSp        891 non-null    int64
7  Parch        891 non-null    int64
8  Ticket       891 non-null    object
9  Fare         891 non-null    float64
10 Cabin        204 non-null    object
11 Embarked     889 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB

```

```
[9]: test_df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 418 entries, 0 to 417
Data columns (total 11 columns):
#   Column          Non-Null Count  Dtype
---  -
0   PassengerId     418 non-null   int64
1   Pclass          418 non-null   int64
2   Name            418 non-null   object
3   Sex             418 non-null   object
4   Age            332 non-null   float64
5   SibSp           418 non-null   int64
6   Parch           418 non-null   int64
7   Ticket          418 non-null   object
8   Fare            417 non-null   float64
9   Cabin           91 non-null    object
10  Embarked        418 non-null   object
dtypes: float64(2), int64(4), object(5)
memory usage: 36.0+ KB

```

```
[10]: train_df.describe()
```

```

[10]:      PassengerId  Survived  Pclass     Age  SibSp  \
count    891.000000    891.000000    891.000000  714.000000  891.000000
mean      446.000000     0.383838     2.308642   29.699118    0.523008
std       257.353842     0.486592     0.836071   14.526497    1.102743
min         1.000000     0.000000     1.000000    0.420000    0.000000
25%       223.500000     0.000000     2.000000   20.125000    0.000000
50%       446.000000     0.000000     3.000000   28.000000    0.000000
75%       668.500000     1.000000     3.000000   38.000000    1.000000
max       891.000000     1.000000     3.000000   80.000000    8.000000

```

	Parch	Fare
count	891.000000	891.000000
mean	0.381594	32.204208
std	0.806057	49.693429
min	0.000000	0.000000
25%	0.000000	7.910400
50%	0.000000	14.454200
75%	0.000000	31.000000
max	6.000000	512.329200

```
[11]: train_df.describe(include=['O'])
```

```
[11]:
```

	Name	Sex	Ticket	Cabin	Embarked
count	891	891	891	204	889
unique	891	2	681	147	3
top	Braund, Mr. Owen Harris	male	347082	B96 B98	S
freq	1	577	7	4	644

```
[12]: test_df
```

```
[12]:
```

	PassengerId	Pclass	Name \
0	892	3	Kelly, Mr. James
1	893	3	Wilkes, Mrs. James (Ellen Needs)
2	894	2	Myles, Mr. Thomas Francis
3	895	3	Wirz, Mr. Albert
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)
..
413	1305	3	Spector, Mr. Woolf
414	1306	1	Oliva y Ocana, Dona. Fermina
415	1307	3	Saether, Mr. Simon Sivertsen
416	1308	3	Ware, Mr. Frederick
417	1309	3	Peter, Master. Michael J

	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	male	34.5	0	0	330911	7.8292	NaN	Q
1	female	47.0	1	0	363272	7.0000	NaN	S
2	male	62.0	0	0	240276	9.6875	NaN	Q
3	male	27.0	0	0	315154	8.6625	NaN	S
4	female	22.0	1	1	3101298	12.2875	NaN	S
..
413	male	NaN	0	0	A.5. 3236	8.0500	NaN	S
414	female	39.0	0	0	PC 17758	108.9000	C105	C
415	male	38.5	0	0	SOTON/O.Q. 3101262	7.2500	NaN	S
416	male	NaN	0	0	359309	8.0500	NaN	S
417	male	NaN	1	1	2668	22.3583	NaN	C

```
[418 rows x 11 columns]
```

```
[13]: survived = train_df[train_df['Survived'] == 1]
not_survived = train_df[train_df['Survived'] == 0]

print ("Survived: %i (%.1f%%)"%(len(survived), float(len(survived))/
↳len(train_df)*100.0))
print ("Not Survived: %i (%.1f%%)"%(len(not_survived), float(len(not_survived))/
↳len(train_df)*100.0))
print ("Total: %i"%len(train_df))
```

```
Survived: 342 (38.4%)
Not Survived: 549 (61.6%)
Total: 891
```

```
[14]: train_df.Pclass.value_counts()
```

```
[14]: Pclass
3      491
1      216
2      184
Name: count, dtype: int64
```

```
[15]: train_df.groupby('Pclass').Survived.value_counts()
```

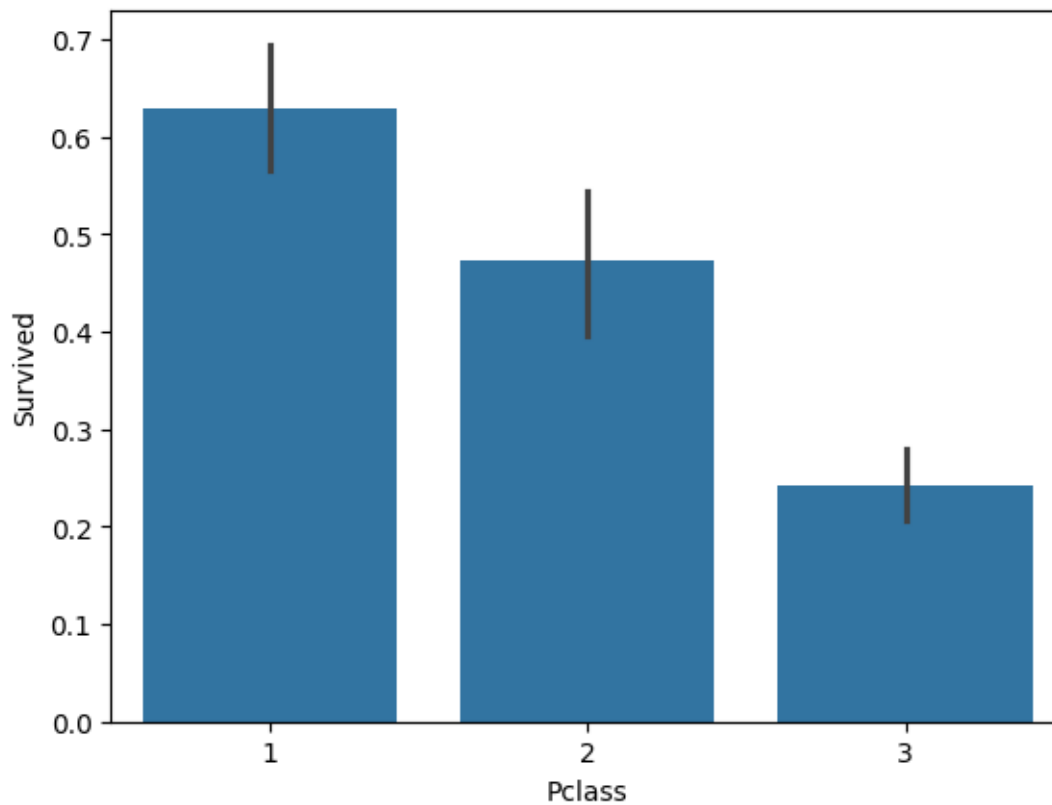
```
[15]: Pclass  Survived
1         1         136
         0          80
2         0          97
         1          87
3         0         372
         1         119
Name: count, dtype: int64
```

```
[16]: train_df[['Pclass', 'Survived']].groupby(['Pclass'], as_index=False).mean()
```

```
[16]:   Pclass  Survived
0      1  0.629630
1      2  0.472826
2      3  0.242363
```

```
[17]: sns.barplot(x='Pclass', y='Survived', data=train_df)
```

```
[17]: <Axes: xlabel='Pclass', ylabel='Survived'>
```



```
[18]: train_df.Sex.value_counts()
```

```
[18]: Sex
male      577
female    314
Name: count, dtype: int64
```

```
[19]: train_df.groupby('Sex').Survived.value_counts()
```

```
[19]: Sex    Survived
female 1         233
       0          81
male   0         468
       1         109
Name: count, dtype: int64
```

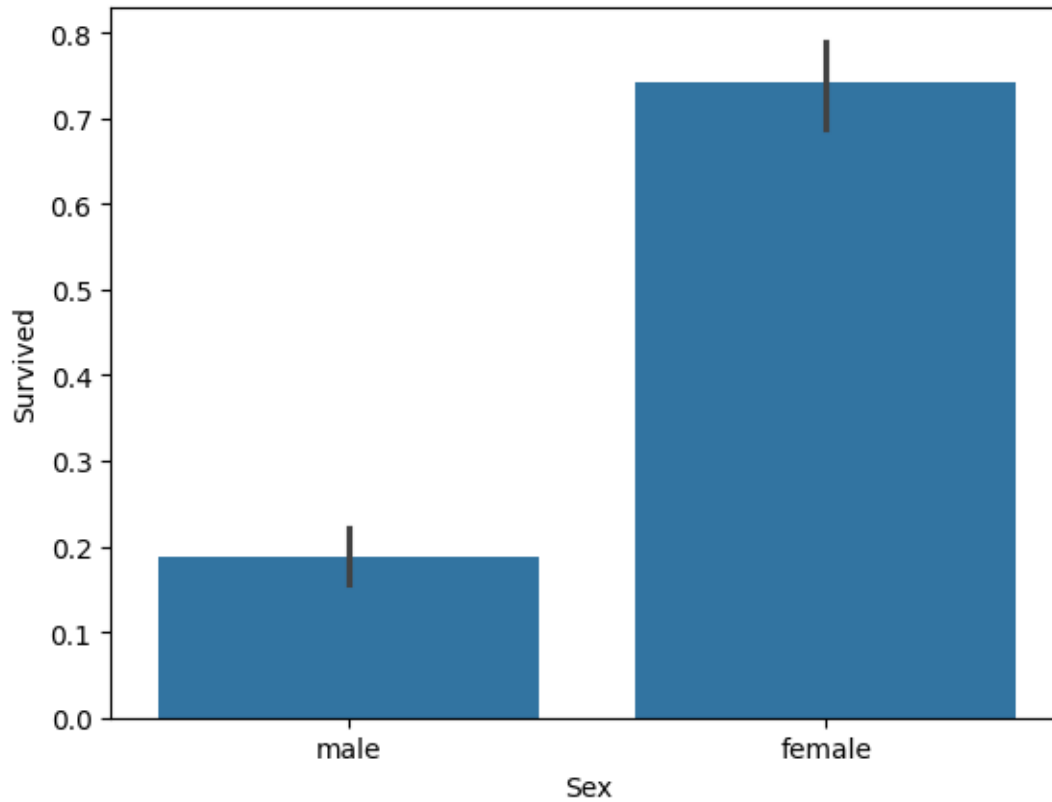
```
[20]: train_df[['Sex', 'Survived']].groupby(['Sex'], as_index=False).mean()
```

```
[20]:      Sex  Survived
0  female  0.742038
1   male   0.188908
```



```
[21]: sns.barplot(x='Sex', y='Survived', data=train_df)
```

```
[21]: <Axes: xlabel='Sex', ylabel='Survived'>
```

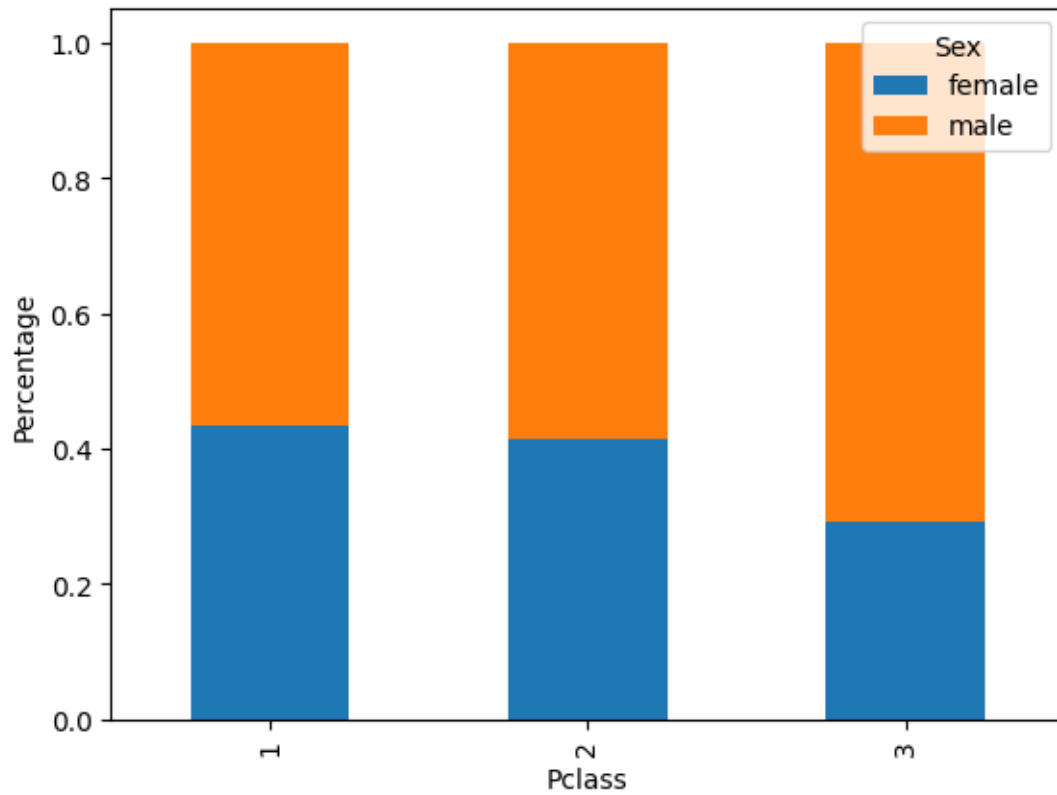


```
[24]: tab = pd.crosstab(train_df['Pclass'], train_df['Sex'])
print (tab)

tab.div(tab.sum(1).astype(float), axis=0).plot(kind="bar", stacked=True)
plt.xlabel('Pclass')
plt.ylabel('Percentage')
```

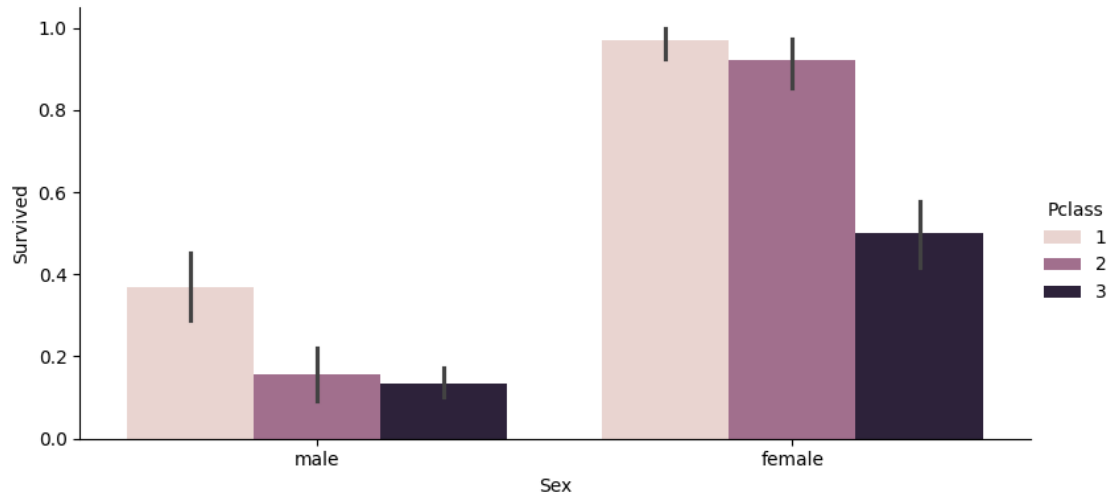
Sex	female	male
Pclass		
1	94	122
2	76	108
3	144	347

```
[24]: Text(0, 0.5, 'Percentage')
```



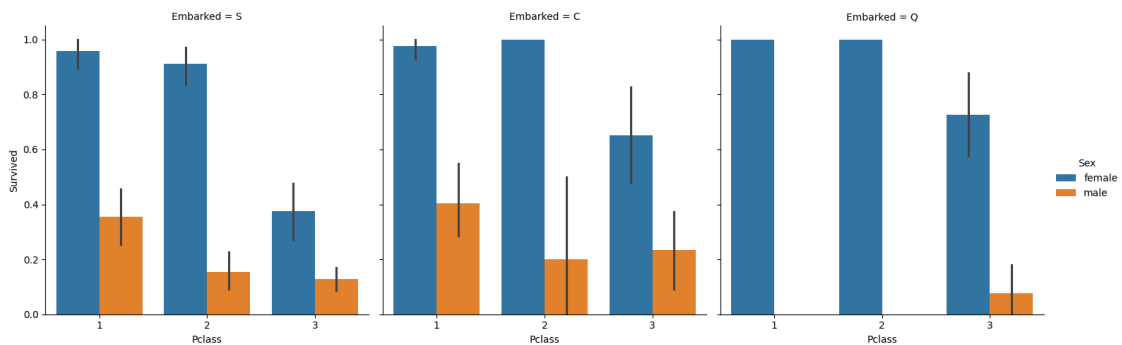
```
[31]: import seaborn as sns
import matplotlib.pyplot as plt

# Use catplot instead of factorplot
sns.catplot(x='Sex', y='Survived', hue='Pclass', kind='bar', data=train_df,
            height=4, aspect=2)
plt.show()
```



```
[32]: import seaborn as sns
import matplotlib.pyplot as plt

# Use catplot instead of factorplot
sns.catplot(x='Pclass', y='Survived', hue='Sex', col='Embarked', kind='bar',
            data=train_df)
# kind='bar' specifies a bar plot, which is the default for factorplot
plt.show()
```



```
[33]: train_df.Embarked.value_counts()
```

```
[33]: Embarked
S      644
C      168
Q       77
Name: count, dtype: int64
```

```
[34]: train_df.groupby('Embarked').Survived.value_counts()
```

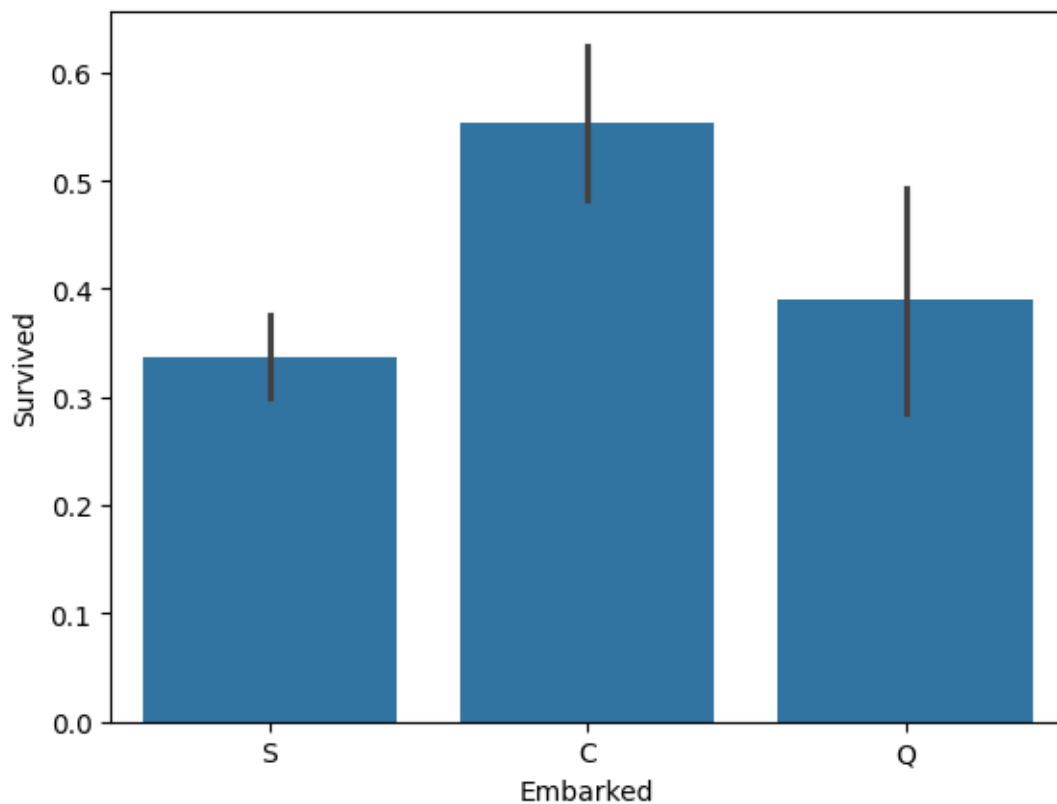
```
[34]: Embarked  Survived
C          1          93
         0          75
Q          0          47
         1          30
S          0         427
         1         217
Name: count, dtype: int64
```

```
[35]: train_df[['Embarked', 'Survived']].groupby(['Embarked'], as_index=False).mean()
```

```
[35]:   Embarked  Survived
0        C  0.553571
1        Q  0.389610
2        S  0.336957
```

```
[36]: sns.barplot(x='Embarked', y='Survived', data=train_df)
```

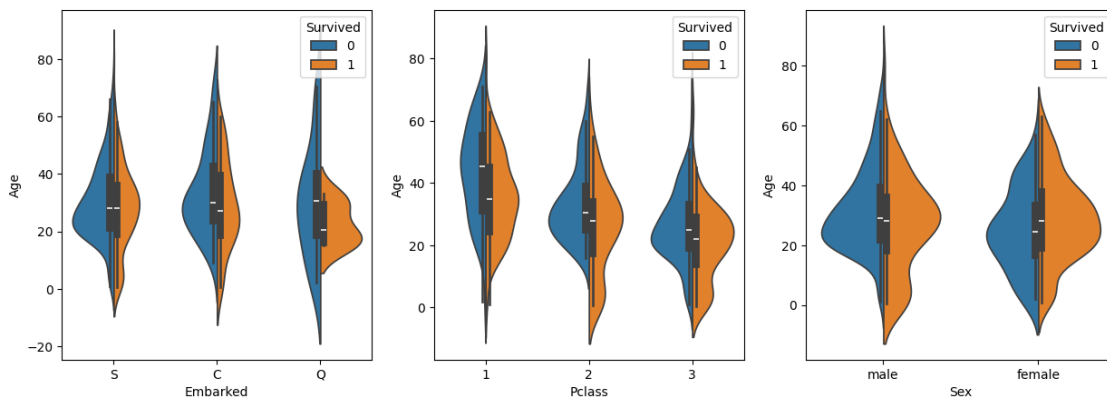
```
[36]: <Axes: xlabel='Embarked', ylabel='Survived'>
```



```
[39]: fig = plt.figure(figsize=(15,5))
ax1 = fig.add_subplot(131)
ax2 = fig.add_subplot(132)
ax3 = fig.add_subplot(133)

sns.violinplot(x="Embarked", y="Age", hue="Survived", data=train_df,
               ↪split=True, ax=ax1)
sns.violinplot(x="Pclass", y="Age", hue="Survived", data=train_df, split=True,
               ↪ax=ax2)
sns.violinplot(x="Sex", y="Age", hue="Survived", data=train_df, split=True,
               ↪ax=ax3)
```

[39]: <Axes: xlabel='Sex', ylabel='Age'>



```
[40]: !pip install seaborn
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np

# Assuming train_df is your DataFrame
plt.figure(figsize=(15,6))

# Select only numeric columns for correlation calculation
numeric_df = train_df.select_dtypes(include=np.number)

# Calculate and plot the correlation matrix
sns.heatmap(numeric_df.drop('PassengerId',axis=1).corr(), vmax=0.6,
            ↪square=True, annot=True)
plt.show()
```

Requirement already satisfied: seaborn in /usr/local/lib/python3.10/dist-packages (0.13.2)

Requirement already satisfied: numpy!=1.24.0,>=1.20 in /usr/local/lib/python3.10/dist-packages (from seaborn) (1.26.4)

Requirement already satisfied: pandas>=1.2 in /usr/local/lib/python3.10/dist-packages (from seaborn) (2.2.2)

Requirement already satisfied: matplotlib!=3.6.1,>=3.4 in /usr/local/lib/python3.10/dist-packages (from seaborn) (3.8.0)

Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.3.1)

Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (0.12.1)

Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (4.55.0)

Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.4.7)

Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (24.2)

Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (11.0.0)

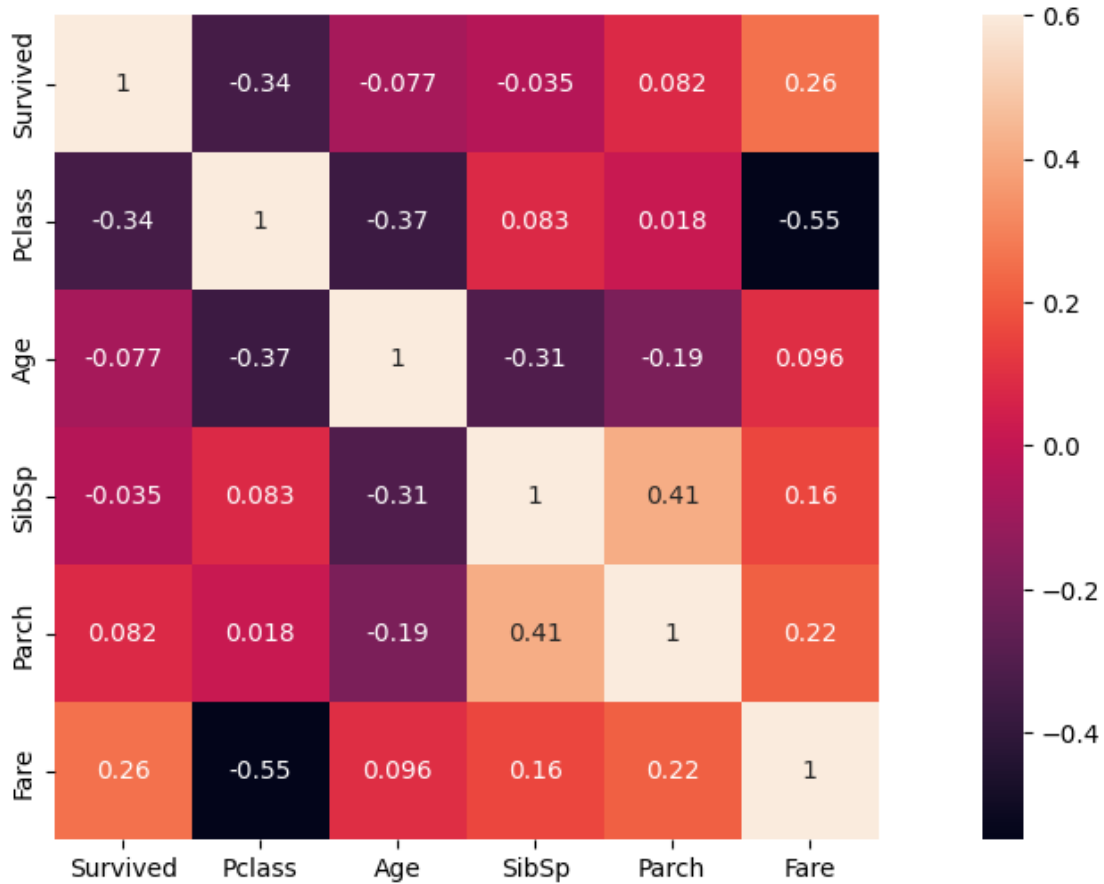
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (3.2.0)

Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (2.8.2)

Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas>=1.2->seaborn) (2024.2)

Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.10/dist-packages (from pandas>=1.2->seaborn) (2024.2)

Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.7->matplotlib!=3.6.1,>=3.4->seaborn) (1.16.0)



```
[41]: train_test_data = [train_df, test_df] # combining train and test dataset
```

```
for dataset in train_test_data:
    dataset['Title'] = dataset.Name.str.extract(' ([A-Za-z]+)\.')
```

```
[42]: train_df.tail()
```

```
[42]:
```

	PassengerId	Survived	Pclass	Name \
886	887	0	2	Montvila, Rev. Juozas
887	888	1	1	Graham, Miss. Margaret Edith
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"
889	890	1	1	Behr, Mr. Karl Howell
890	891	0	3	Dooley, Mr. Patrick

	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	Title
886	male	27.0	0	0	211536	13.00	NaN	S	Rev
887	female	19.0	0	0	112053	30.00	B42	S	Miss
888	female	NaN	1	2	W./C. 6607	23.45	NaN	S	Miss
889	male	26.0	0	0	111369	30.00	C148	C	Mr

890	male	32.0	0	0	370376	7.75	NaN	Q	Mr
-----	------	------	---	---	--------	------	-----	---	----

```
[43]: pd.crosstab(train_df['Title'], train_df['Sex'])
```

```
[43]: Sex      female  male
Title
Capt         0      1
Col           0      2
Countess      1      0
Don           0      1
Dr            1      6
Jonkheer      0      1
Lady          1      0
Major         0      2
Master        0     40
Miss         182      0
Mlle          2      0
Mme           1      0
Mr            0     517
Mrs          125      0
Ms            1      0
Rev           0      6
Sir           0      1
```

```
[44]: for dataset in train_test_data:
        dataset['Title'] = dataset['Title'].replace(['Lady', 'Countess', 'Capt', 'Col', \
        ↪ 'Col', \
            'Don', 'Dr', 'Major', 'Rev', 'Sir', 'Jonkheer', 'Dona'], 'Other')

        dataset['Title'] = dataset['Title'].replace('Mlle', 'Miss')
        dataset['Title'] = dataset['Title'].replace('Ms', 'Miss')
        dataset['Title'] = dataset['Title'].replace('Mme', 'Mrs')

train_df[['Title', 'Survived']].groupby(['Title'], as_index=False).mean()
```

```
[44]: Title  Survived
0  Master  0.575000
1   Miss  0.702703
2    Mr   0.156673
3   Mrs  0.793651
4  Other  0.347826
```

```
[45]: title_mapping = {"Mr": 1, "Miss": 2, "Mrs": 3, "Master": 4, "Other": 5}
for dataset in train_test_data:
    dataset['Title'] = dataset['Title'].map(title_mapping)
    dataset['Title'] = dataset['Title'].fillna(0)
```



```
[46]: train_df
```

```
[46]:
```

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	
..	
886	887	0	2	
887	888	1	1	
888	889	0	3	
889	890	1	1	
890	891	0	3	

	Name	Sex	Age	SibSp	\
0	Braund, Mr. Owen Harris	male	22.0	1	
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	
2	Heikkinen, Miss. Laina	female	26.0	0	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	
4	Allen, Mr. William Henry	male	35.0	0	
..	
886	Montvila, Rev. Juozas	male	27.0	0	
887	Graham, Miss. Margaret Edith	female	19.0	0	
888	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	
889	Behr, Mr. Karl Howell	male	26.0	0	
890	Dooley, Mr. Patrick	male	32.0	0	

	Parch	Ticket	Fare	Cabin	Embarked	Title
0	0	A/5 21171	7.2500	NaN	S	1
1	0	PC 17599	71.2833	C85	C	3
2	0	STON/O2. 3101282	7.9250	NaN	S	2
3	0	113803	53.1000	C123	S	3
4	0	373450	8.0500	NaN	S	1
..
886	0	211536	13.0000	NaN	S	5
887	0	112053	30.0000	B42	S	2
888	2	W./C. 6607	23.4500	NaN	S	2
889	0	111369	30.0000	C148	C	1
890	0	370376	7.7500	NaN	Q	1

```
[891 rows x 13 columns]
```

```
[47]: for dataset in train_test_data:
      dataset['Sex'] = dataset['Sex'].map( {'female': 1, 'male': 0} ).astype(int)
```

```
[48]: train_df
```

```
[48]: PassengerId  Survived  Pclass  \
0            1         0         3
1            2         1         1
2            3         1         3
3            4         1         1
4            5         0         3
..          ...         ...         ...
886          887         0         2
887          888         1         1
888          889         0         3
889          890         1         1
890          891         0         3
```

```

                                Name  Sex  Age  SibSp  \
0                        Braund, Mr. Owen Harris    0  22.0    1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...    1  38.0    1
2                        Heikkinen, Miss. Laina    1  26.0    0
3  Futrelle, Mrs. Jacques Heath (Lily May Peel)    1  35.0    1
4                        Allen, Mr. William Henry    0  35.0    0
..          ...         ...         ...         ...
886                        Montvila, Rev. Juozas    0  27.0    0
887                        Graham, Miss. Margaret Edith    1  19.0    0
888  Johnston, Miss. Catherine Helen "Carrie"    1   NaN    1
889                        Behr, Mr. Karl Howell    0  26.0    0
890                        Dooley, Mr. Patrick    0  32.0    0
```

```

      Parch      Ticket    Fare Cabin Embarked  Title
0         0      A/5 21171    7.2500   NaN      S      1
1         0      PC 17599   71.2833   C85      C      3
2         0  STON/O2. 3101282    7.9250   NaN      S      2
3         0      113803   53.1000  C123      S      3
4         0      373450    8.0500   NaN      S      1
..        ...         ...         ...         ...
886        0      211536   13.0000   NaN      S      5
887        0      112053   30.0000   B42      S      2
888        2      W./C. 6607   23.4500   NaN      S      2
889        0      111369   30.0000  C148      C      1
890        0      370376    7.7500   NaN      Q      1
```

[891 rows x 13 columns]

```
[49]: train_df.Embarked.unique()
```

```
[49]: array(['S', 'C', 'Q', nan], dtype=object)
```

```
[50]: train_df['Embarked'].value_counts()
```

```
[50]: Embarked
S      644
C      168
Q       77
Name: count, dtype: int64
```

```
[51]: for dataset in train_test_data:
      dataset['Embarked'] = dataset['Embarked'].fillna('S')
```

```
[52]: train_df['Embarked'].unique()
```

```
[52]: array(['S', 'C', 'Q'], dtype=object)
```

```
[53]: for dataset in train_test_data:

      dataset['Embarked'] = dataset['Embarked'].map( {'S': 0, 'C': 1, 'Q': 2} ).
      ↪astype(int)
```

```
[54]: train_df
```

```
[54]:      PassengerId  Survived  Pclass  \
0                1         0        3
1                2         1        1
2                3         1        3
3                4         1        1
4                5         0        3
..            ...     ...     ...
886            887         0        2
887            888         1        1
888            889         0        3
889            890         1        1
890            891         0        3
```

```

                                Name  Sex  Age  SibSp  \
0                        Braund, Mr. Owen Harris    0  22.0    1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...    1  38.0    1
2                        Heikkinen, Miss. Laina    1  26.0    0
3  Futrelle, Mrs. Jacques Heath (Lily May Peel)    1  35.0    1
4                        Allen, Mr. William Henry    0  35.0    0
..            ...     ...     ...     ...
886                        Montvila, Rev. Juozas    0  27.0    0
887                        Graham, Miss. Margaret Edith    1  19.0    0
888      Johnston, Miss. Catherine Helen "Carrie"    1   NaN    1
889                        Behr, Mr. Karl Howell    0  26.0    0
890                        Dooley, Mr. Patrick    0  32.0    0
```

```
      Parch      Ticket     Fare Cabin Embarked Title
```

0	0	A/5 21171	7.2500	NaN	0	1
1	0	PC 17599	71.2833	C85	1	3
2	0	STON/O2. 3101282	7.9250	NaN	0	2
3	0	113803	53.1000	C123	0	3
4	0	373450	8.0500	NaN	0	1
..
886	0	211536	13.0000	NaN	0	5
887	0	112053	30.0000	B42	0	2
888	2	W./C. 6607	23.4500	NaN	0	2
889	0	111369	30.0000	C148	1	1
890	0	370376	7.7500	NaN	2	1

[891 rows x 13 columns]

```
[55]: for dataset in train_test_data:
    age_avg = dataset['Age'].mean()
    age_std = dataset['Age'].std()
    age_null_count = dataset['Age'].isnull().sum()

    age_null_random_list = np.random.randint(age_avg - age_std, age_avg +
    age_std, size=age_null_count)
    dataset['Age'][np.isnan(dataset['Age'])] = age_null_random_list
    dataset['Age'] = dataset['Age'].astype(int)

train_df['AgeBand'] = pd.cut(train_df['Age'], 5)

print (train_df[['AgeBand', 'Survived']].groupby(['AgeBand'], as_index=False).
    mean())
```

	AgeBand	Survived
0	(-0.08, 16.0]	0.537037
1	(16.0, 32.0]	0.349558
2	(32.0, 48.0]	0.378486
3	(48.0, 64.0]	0.434783
4	(64.0, 80.0]	0.090909

```
[56]: train_df.head()
```

```
[56]: PassengerId  Survived  Pclass  \
0             1         0         3
1             2         1         1
2             3         1         3
3             4         1         1
4             5         0         3

                                Name  Sex  Age  SibSp  Parch  \
0      Braund, Mr. Owen Harris     0   22      1      0
```

1	Cumings, Mrs. John Bradley (Florence Briggs Th...	1	38	1	0
2	Heikkinen, Miss. Laina	1	26	0	0
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	1	35	1	0
4	Allen, Mr. William Henry	0	35	0	0

	Ticket	Fare	Cabin	Embarked	Title	AgeBand
0	A/5 21171	7.2500	NaN	0	1	(16.0, 32.0]
1	PC 17599	71.2833	C85	1	3	(32.0, 48.0]
2	STON/O2. 3101282	7.9250	NaN	0	2	(16.0, 32.0]
3	113803	53.1000	C123	0	3	(32.0, 48.0]
4	373450	8.0500	NaN	0	1	(32.0, 48.0]

```
[57]: for dataset in train_test_data:
      dataset.loc[ dataset['Age'] <= 16, 'Age'] = 0
      dataset.loc[(dataset['Age'] > 16) & (dataset['Age'] <= 32), 'Age'] = 1
      dataset.loc[(dataset['Age'] > 32) & (dataset['Age'] <= 48), 'Age'] = 2
      dataset.loc[(dataset['Age'] > 48) & (dataset['Age'] <= 64), 'Age'] = 3
      dataset.loc[ dataset['Age'] > 64, 'Age'] = 4
```

```
[58]: train_df.head()
```

```
[58]: PassengerId  Survived  Pclass  \
0             1         0         3
1             2         1         1
2             3         1         3
3             4         1         1
4             5         0         3
```

	Name	Sex	Age	SibSp	Parch	\
0	Braund, Mr. Owen Harris	0	1	1	0	
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	1	2	1	0	
2	Heikkinen, Miss. Laina	1	1	0	0	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	1	2	1	0	
4	Allen, Mr. William Henry	0	2	0	0	

	Ticket	Fare	Cabin	Embarked	Title	AgeBand
0	A/5 21171	7.2500	NaN	0	1	(16.0, 32.0]
1	PC 17599	71.2833	C85	1	3	(32.0, 48.0]
2	STON/O2. 3101282	7.9250	NaN	0	2	(16.0, 32.0]
3	113803	53.1000	C123	0	3	(32.0, 48.0]
4	373450	8.0500	NaN	0	1	(32.0, 48.0]

```
[59]: for dataset in train_test_data:
      dataset['Fare'] = dataset['Fare'].fillna(train_df['Fare'].median())
```

```
[60]: train_df['FareBand'] = pd.qcut(train_df['Fare'], 4)
```

```
print (train_df[['FareBand', 'Survived']].groupby(['FareBand'], as_index=False).
      ↪mean())
```

	FareBand	Survived
0	(-0.001, 7.91]	0.197309
1	(7.91, 14.454]	0.303571
2	(14.454, 31.0]	0.454955
3	(31.0, 512.329]	0.581081

```
[61]: for dataset in train_test_data:
      dataset.loc[ dataset['Fare'] <= 7.91, 'Fare'] = 0
      dataset.loc[(dataset['Fare'] > 7.91) & (dataset['Fare'] <= 14.454), 'Fare'] ↪
      ↪= 1
      dataset.loc[(dataset['Fare'] > 14.454) & (dataset['Fare'] <= 31), 'Fare'] ↪
      ↪= 2
      dataset.loc[ dataset['Fare'] > 31, 'Fare'] = 3
      dataset['Fare'] = dataset['Fare'].astype(int)
```

```
[62]: train_df
```

```
[62]:
```

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	
..	
886	887	0	2	
887	888	1	1	
888	889	0	3	
889	890	1	1	
890	891	0	3	

	Name	Sex	Age	SibSp	\
0	Braund, Mr. Owen Harris	0	1	1	
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	1	2	1	
2	Heikkinen, Miss. Laina	1	1	0	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	1	2	1	
4	Allen, Mr. William Henry	0	2	0	
..	
886	Montvila, Rev. Juozas	0	1	0	
887	Graham, Miss. Margaret Edith	1	1	0	
888	Johnston, Miss. Catherine Helen "Carrie"	1	1	1	
889	Behr, Mr. Karl Howell	0	1	0	
890	Dooley, Mr. Patrick	0	1	0	

	Parch	Ticket	Fare	Cabin	Embarked	Title	AgeBand	\
--	-------	--------	------	-------	----------	-------	---------	---

0	0	A/5 21171	0	NaN	0	1	(16.0, 32.0]
1	0	PC 17599	3	C85	1	3	(32.0, 48.0]
2	0	STON/O2. 3101282	1	NaN	0	2	(16.0, 32.0]
3	0	113803	3	C123	0	3	(32.0, 48.0]
4	0	373450	1	NaN	0	1	(32.0, 48.0]
..
886	0	211536	1	NaN	0	5	(16.0, 32.0]
887	0	112053	2	B42	0	2	(16.0, 32.0]
888	2	W./C. 6607	2	NaN	0	2	(16.0, 32.0]
889	0	111369	2	C148	1	1	(16.0, 32.0]
890	0	370376	0	NaN	2	1	(16.0, 32.0]

FareBand	
0	(-0.001, 7.91]
1	(31.0, 512.329]
2	(7.91, 14.454]
3	(31.0, 512.329]
4	(7.91, 14.454]
..	...
886	(7.91, 14.454]
887	(14.454, 31.0]
888	(14.454, 31.0]
889	(14.454, 31.0]
890	(-0.001, 7.91]

[891 rows x 15 columns]

```
[63]: for dataset in train_test_data:
        dataset['FamilySize'] = dataset['SibSp'] + dataset['Parch'] + 1

        print (train_df[['FamilySize', 'Survived']].groupby(['FamilySize'],
        ↪as_index=False).mean())
```

	FamilySize	Survived
0	1	0.303538
1	2	0.552795
2	3	0.578431
3	4	0.724138
4	5	0.200000
5	6	0.136364
6	7	0.333333
7	8	0.000000
8	11	0.000000

```
[64]: for dataset in train_test_data:
        dataset['IsAlone'] = 0
        dataset.loc[dataset['FamilySize'] == 1, 'IsAlone'] = 1
```

```
print (train_df[['IsAlone', 'Survived']].groupby(['IsAlone'], as_index=False).
      ↪mean())
```

```
      IsAlone  Survived
0         0    0.505650
1         1    0.303538
```

```
[65]: train_df.head()
```

```
[65]:   PassengerId  Survived  Pclass  \
0             1         0        3
1             2         1        1
2             3         1        3
3             4         1        1
4             5         0        3
```

```
      Name Sex Age SibSp Parch  \
0  Braund, Mr. Owen Harris    0   1     1     0
1  Cumings, Mrs. John Bradley (Florence Briggs Th...    1   2     1     0
2  Heikkinen, Miss. Laina    1   1     0     0
3  Futrelle, Mrs. Jacques Heath (Lily May Peel)    1   2     1     0
4  Allen, Mr. William Henry    0   2     0     0
```

```
      Ticket  Fare Cabin Embarked  Title  AgeBand  \
0  A/5 21171    0  NaN      0      1  (16.0, 32.0]
1  PC 17599    3  C85      1      3  (32.0, 48.0]
2  STON/O2. 3101282    1  NaN      0      2  (16.0, 32.0]
3  113803    3  C123      0      3  (32.0, 48.0]
4  373450    1  NaN      0      1  (32.0, 48.0]
```

```
      FareBand  FamilySize  IsAlone
0  (-0.001, 7.91]         2         0
1  (31.0, 512.329]         2         0
2  (7.91, 14.454]         1         1
3  (31.0, 512.329]         2         0
4  (7.91, 14.454]         1         1
```

```
[66]: features_drop = ['Name', 'SibSp', 'Parch', 'Ticket', 'Cabin', 'FamilySize']
train_df = train_df.drop(features_drop, axis=1)
test_df = test_df.drop(features_drop, axis=1)
train_df = train_df.drop(['PassengerId', 'AgeBand', 'FareBand'], axis=1)
```

```
[67]: train_df.head()
```

```
[67]:   Survived  Pclass  Sex  Age  Fare  Embarked  Title  IsAlone
0         0        3    0    1    0         0      1         0
```


1	1	1	1	2	3	1	3	0
2	1	3	1	1	1	0	2	1
3	1	1	1	2	3	0	3	0
4	0	3	0	2	1	0	1	1

```
[68]: test_df.head()
```

```
[68]: PassengerId  Pclass  Sex  Age  Fare  Embarked  Title  IsAlone
0          892      3    0    2    0         2      1      1
1          893      3    1    2    0         0      3      0
2          894      2    0    3    1         2      1      1
3          895      3    0    1    1         0      1      1
4          896      3    1    1    1         0      3      0
```

```
[69]: X_train = train_df.drop('Survived', axis=1)
      y_train = train_df['Survived']
      X_test = test_df.drop("PassengerId", axis=1).copy()

      X_train.shape, y_train.shape, X_test.shape
```

```
[69]: ((891, 7), (891,), (418, 7))
```

```
[70]: from sklearn.linear_model import LogisticRegression
      from sklearn.svm import SVC, LinearSVC
      from sklearn.neighbors import KNeighborsClassifier
      from sklearn.tree import DecisionTreeClassifier
      from sklearn.ensemble import RandomForestClassifier
      from sklearn.naive_bayes import GaussianNB
      from sklearn.linear_model import Perceptron
      from sklearn.linear_model import SGDClassifier
```

```
[71]: clf = DecisionTreeClassifier()
      clf.fit(X_train, y_train)
      y_pred_decision_tree = clf.predict(X_test)
      acc_decision_tree = round(clf.score(X_train, y_train) * 100, 2)
      print (acc_decision_tree)
```

87.65

```
[72]: clf = RandomForestClassifier(n_estimators=100)
      clf.fit(X_train, y_train)
      y_pred_random_forest = clf.predict(X_test)
      acc_random_forest = round(clf.score(X_train, y_train) * 100, 2)
      print (acc_random_forest)
```

87.65

```
[73]: clf = KNeighborsClassifier(n_neighbors = 3)
      clf.fit(X_train, y_train)
      y_pred_knn = clf.predict(X_test)
      acc_knn = round(clf.score(X_train, y_train) * 100, 2)
      print (acc_knn)
```

85.07

```
[74]: pip install xgboost
```

Requirement already satisfied: xgboost in /usr/local/lib/python3.10/dist-packages (2.1.2)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from xgboost) (1.26.4)
Requirement already satisfied: nvidia-nccl-cu12 in /usr/local/lib/python3.10/dist-packages (from xgboost) (2.23.4)
Requirement already satisfied: scipy in /usr/local/lib/python3.10/dist-packages (from xgboost) (1.13.1)

```
[75]: import xgboost as xgb
```

```
[76]: model = xgb.XGBClassifier()
```

```
[77]: clf_xgb=model.fit(X_train, y_train)
```

```
[78]: y_pred_xgb = clf_xgb.predict(X_test)
      acc_xgb = round(clf_xgb.score(X_train, y_train) * 100, 2)
      print (acc_xgb)
```

87.43

```
[79]: y_pred_xgb = clf_xgb.predict(X_test)
      acc_xgb = round(clf_xgb.score(X_train, y_train) * 100, 2)
      print (acc_xgb)
```

87.43

```
[80]: from sklearn.model_selection import GridSearchCV
```

```
# Define hyperparameters grid
param_grid = {
    'max_depth': [3, 6, 9],
    'learning_rate': [0.1, 0.01, 0.05],
    'n_estimators': [100, 200, 300],
    'subsample': [0.8, 0.9, 1.0],
    'colsample_bytree': [0.8, 0.9, 1.0]
}
```

```

# Initialize XGBoost classifier
xgb_model = xgb.XGBClassifier()

# Initialize Grid Search with cross-validation
grid_search = GridSearchCV(estimator=xgb_model, param_grid=param_grid, cv=3,
    ↪scoring='accuracy')

# Perform grid search
grid_search.fit(X_train, y_train)

# Get best parameters and best score
best_params = grid_search.best_params_
best_score = grid_search.best_score_

# Train final model with best parameters
final_model = xgb.XGBClassifier(**best_params)
final_model.fit(X_train, y_train)

```

```

[80]: XGBClassifier(base_score=None, booster=None, callbacks=None,
    colsample_bylevel=None, colsample_bynode=None,
    colsample_bytree=0.9, device=None, early_stopping_rounds=None,
    enable_categorical=False, eval_metric=None, feature_types=None,
    gamma=None, 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=3, max_leaves=None,
    min_child_weight=None, missing=nan, monotone_constraints=None,
    multi_strategy=None, n_estimators=300, n_jobs=None,
    num_parallel_tree=None, random_state=None, ...)

```

```

[81]: print(best_params)
    print(best_score)

```

```

{'colsample_bytree': 0.9, 'learning_rate': 0.05, 'max_depth': 3, 'n_estimators':
300, 'subsample': 0.8}
0.8204264870931538

```

```

[82]: params = {
    'max_depth': 3,
    'learning_rate': 0.05,
    'n_estimators': 300,
    'subsample': 0.8,
    'colsample_bytree': 0.8
}

```

```
model = xgb.XGBClassifier(**params)
clf_xgb=model.fit(X_train, y_train)
y_pred_train_xgb=clf_xgb.predict(X_train)
y_pred_xgb = clf_xgb.predict(X_test)
acc_xgb = round(clf_xgb.score(X_train, y_train) * 100, 2)
print (acc_xgb)
```

85.63

```
[83]: y_train.value_counts()
```

```
[83]: Survived
0      549
1      342
Name: count, dtype: int64
```

```
[84]: from sklearn.metrics import precision_score, recall_score, f1_score

precision=precision_score(y_train,y_pred_train_xgb)
precision
```

```
[84]: 0.8664383561643836
```

```
[85]: recall = recall_score(y_train, y_pred_train_xgb)

# Calculate F1-score
f1 = f1_score(y_train, y_pred_train_xgb)

print("Precision:", precision)
print("Recall:", recall)
print("F1-score:", f1)
```

```
Precision: 0.8664383561643836
Recall: 0.7397660818713451
F1-score: 0.7981072555205048
```

```
[87]: pip install lightgbm
```

```
Requirement already satisfied: lightgbm in /usr/local/lib/python3.10/dist-
packages (4.5.0)
Requirement already satisfied: numpy>=1.17.0 in /usr/local/lib/python3.10/dist-
packages (from lightgbm) (1.26.4)
Requirement already satisfied: scipy in /usr/local/lib/python3.10/dist-packages
(from lightgbm) (1.13.1)
```

```
[88]: import lightgbm as lgb
```


[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
```

```
[91]: y_pred_test = clf_lgb.predict(X_test)
y_pred_train=clf_lgb.predict(X_train)
acc_lgb = round(clf_lgb.score(X_train, y_train) * 100, 2)
print (acc_lgb)
```

86.53

```
[92]: print(clf_lgb.feature_importances_)
```

[365 171 563 541 417 306 195]

```
[93]: X_train.columns
```

```
[93]: Index(['Pclass', 'Sex', 'Age', 'Fare', 'Embarked', 'Title', 'IsAlone'],
dtype='object')
```

```
[94]: from sklearn.model_selection import RandomizedSearchCV
```

```
# Define parameter space
param_grid = {
    'num_leaves': [20, 30, 40, 50],
    'learning_rate': [0.01, 0.05, 0.1],
    'max_depth': [-1, 5, 10, 15],
    'min_child_samples': [20, 30, 40, 50]
}

# Define the objective function
def objective(params):
    model = lgb.LGBMClassifier(**params)
```

```

model.fit(X_train, y_train)
y_pred_train = model.predict(X_train)
accuracy = accuracy_score(y_train, y_pred_train)
return -accuracy # Minimize negative accuracy

# # Split data into training and validation sets
# X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.2,
↳ random_state=42)

# Initialize RandomizedSearchCV
rs = RandomizedSearchCV(estimator=lgb.LGBMClassifier(),
                        param_distributions=param_grid,
                        n_iter=50,
                        scoring='accuracy',
                        cv=3,
                        random_state=42)

# Perform hyperparameter tuning
rs.fit(X_train, y_train)

# Get the best parameters
best_params = rs.best_params_
print("Best parameters:", best_params)

# # Evaluate the best model
# best_model = rs.best_estimator_
# y_pred_test = best_model.predict(X_test)
# test_accuracy = accuracy_score(y_test, y_pred_test)
# print("Test accuracy:", test_accuracy)

```

Streaming output truncated to the last 5000 lines.

```

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

```



```

[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -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: 228, number of negative: 366
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of
testing was 0.000097 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 26
[LightGBM] [Info] Number of data points in the train set: 594, number of used
features: 7
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.383838 -> initscore=-0.473288
[LightGBM] [Info] Start training from score -0.473288
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best 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]

[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
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -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: 228, number of negative: 366
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of
testing was 0.000101 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 26
[LightGBM] [Info] Number of data points in the train set: 594, number of used
features: 7
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.383838 -> initscore=-0.473288
[LightGBM] [Info] Start training from score -0.473288
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best 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]

[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] [Info] Number of positive: 228, number of negative: 366
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of
testing was 0.000096 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 26
[LightGBM] [Info] Number of data points in the train set: 594, number of used
features: 7
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.383838 -> initscore=-0.473288
[LightGBM] [Info] Start training from score -0.473288
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best 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]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

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

```
[LightGBM] [Info] Total Bins 26
```

```
[LightGBM] [Info] Number of data points in the train set: 594, number of used
features: 7
```

```
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.383838 -> initscore=-0.473288
```

```
[LightGBM] [Info] Start training from score -0.473288
```

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

```
[LightGBM] [Warning] No further splits with positive gain, best 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]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

testing was 0.000076 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 26

[LightGBM] [Info] Number of data points in the train set: 594, number of used features: 7

[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.383838 -> initscore=-0.473288

[LightGBM] [Info] Start training from score -0.473288

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[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
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -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: 228, number of negative: 366
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of
testing was 0.000094 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 26
[LightGBM] [Info] Number of data points in the train set: 594, number of used
features: 7
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.383838 -> initscore=-0.473288
[LightGBM] [Info] Start training from score -0.473288
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best 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]

[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] [Info] Number of positive: 228, number of negative: 366
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of
testing was 0.000095 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 26
[LightGBM] [Info] Number of data points in the train set: 594, number of used
features: 7
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.383838 -> initscore=-0.473288
[LightGBM] [Info] Start training from score -0.473288
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best gain: -inf
[LightGBM] [Warning] No further splits with positive gain, best 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]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

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

```
[LightGBM] [Info] Total Bins 26
```

```
[LightGBM] [Info] Number of data points in the train set: 594, number of used
features: 7
```

```
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.383838 -> initscore=-0.473288
```

```
[LightGBM] [Info] Start training from score -0.473288
```

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

```
[LightGBM] [Warning] No further splits with positive gain, best 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]

[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
85.75
```

```
[96]: labels = ['Accuracy']
lgb_scores = [acc_xgb]
xgb_scores = [acc_lgb]

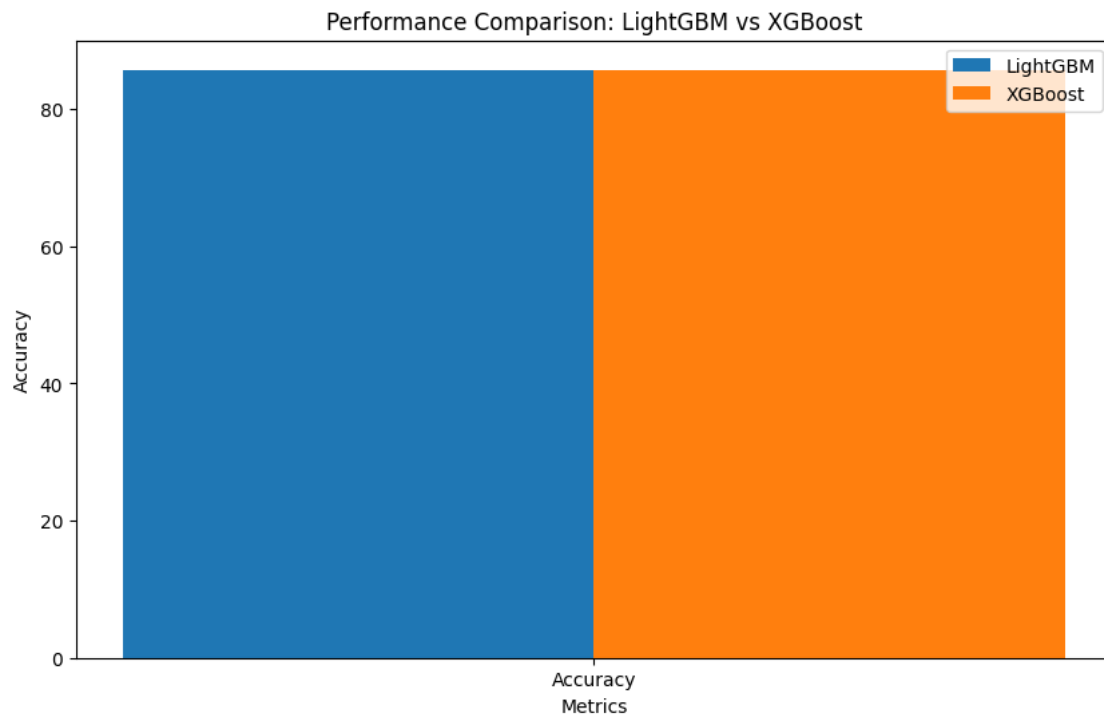
x = range(len(labels))

plt.figure(figsize=(10, 6))

plt.bar(x, lgb_scores, width=0.4, label='LightGBM', align='center')
plt.bar([i + 0.4 for i in x], xgb_scores, width=0.4, label='XGBoost',
        align='center')

plt.xlabel('Metrics')
plt.ylabel('Accuracy')
plt.xticks([i + 0.2 for i in x], labels)
plt.title('Performance Comparison: LightGBM vs XGBoost')
plt.legend()
plt.show()

# Interpretation
print("LightGBM Accuracy:", acc_lgb)
print("XGBoost Accuracy:", acc_xgb)
```



LightGBM Accuracy: 85.75
XGBoost Accuracy: 85.63

[]: