In [18]:
```
import joblib
from xgboost import XGBClassifier
import pandas as pd
import seaborn as sns
from matplotlib import pyplot as plt
from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassifie
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import cross_validate, GridSearchCV
from sklearn.neighbors import KNeighborsClassifier
from sklearn.svm import SVC
from lightgbm import LGBMClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.preprocessing import StandardScaler
df=sns.load_dataset('titanic')
pd.set_option('display.max_columns',None)
pd.set_option('display.width',500)
```

In [90]:
```
pip install xgboost
```

```
17
   -------- ------------------------------ 21.4/99.8 MB 5.0 MB/s eta 0:00:
16
   -------- ------------------------------ 21.8/99.8 MB 5.0 MB/s eta 0:00:
16
   --------- ------------------------------ 22.5/99.8 MB 5.1 MB/s eta 0:00:
16
   --------- ------------------------------ 22.9/99.8 MB 5.1 MB/s eta 0:00:
16
   --------- ------------------------------ 23.2/99.8 MB 5.1 MB/s eta 0:00:
16
   --------- ------------------------------ 23.2/99.8 MB 5.1 MB/s eta 0:00:
16
   --------- ------------------------------ 23.2/99.8 MB 5.1 MB/s eta 0:00:
16
   --------- ------------------------------ 23.5/99.8 MB 4.8 MB/s eta 0:00:
16
   --------- ------------------------------ 23.6/99.8 MB 4.7 MB/s eta 0:00:
17
   --------- ------------------------------ 23.9/99.8 MB 4.7 MB/s eta 0:00:
17
```

In [17]:
```
pip install lightgbm
```

```
Defaulting to user installation because normal site-packages is not writeable
Collecting lightgbm
  Obtaining dependency information for lightgbm from https://files.pythonhosted.org/packages/e1/4c/4685ccfae9806f561de716e32549190c1f533dde5bcadaf83bdf23972cf0/lightgbm-4.3.0-py3-none-win_amd64.whl.metadata (https://files.pythonhosted.org/packages/e1/4c/4685ccfae9806f561de716e32549190c1f533dde5bcadaf83bdf23972cf0/lightgbm-4.3.0-py3-none-win_amd64.whl.metadata)
  Downloading lightgbm-4.3.0-py3-none-win_amd64.whl.metadata (19 kB)
Requirement already satisfied: numpy in c:\programdata\anaconda3\lib\site-packages (from lightgbm) (1.24.3)
Requirement already satisfied: scipy in c:\programdata\anaconda3\lib\site-packages (from lightgbm) (1.11.1)
Downloading lightgbm-4.3.0-py3-none-win_amd64.whl (1.3 MB)
   ---------------------------------------- 0.0/1.3 MB ? eta -:--:--
   ---------------------------------------- 0.0/1.3 MB ? eta -:--:--
    --------------------------------------- 0.0/1.3 MB 660.6 kB/s eta 0:00:02
   -- ------------------------------------- 0.1/1.3 MB 563.7 kB/s eta 0:00:03
   --- ------------------------------------ 0.1/1.3 MB 798.9 kB/s eta 0:00:02
   ----- ---------------------------------- 0.2/1.3 MB 908.0 kB/s eta 0:00:02
   ------- -------------------------------- 0.3/1.3 MB 1.1 MB/s eta 0:00:02
   --------- ------------------------------ 0.3/1.3 MB 1.1 MB/s eta 0:00:01
   ---------- ----------------------------- 0.4/1.3 MB 1.1 MB/s eta 0:00:01
   ------------ --------------------------- 0.5/1.3 MB 1.2 MB/s eta 0:00:01
   ------------------ --------------------- 0.7/1.3 MB 1.5 MB/s eta 0:00:01
   ------------------------ --------------- 0.8/1.3 MB 1.7 MB/s eta 0:00:01
   ------------------------------- -------- 1.1/1.3 MB 2.0 MB/s eta 0:00:01
   ---------------------------------------  1.3/1.3 MB 2.4 MB/s eta 0:00:01
   ---------------------------------------- 1.3/1.3 MB 2.3 MB/s eta 0:00:00
Installing collected packages: lightgbm
Successfully installed lightgbm-4.3.0
Note: you may need to restart the kernel to use updated packages.
```

In [76]:
```
df.head()
```

Out[76]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	d
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	

In [3]:
```
p=['pclass','age','sibsp','fare']
X = df[p]
y = df['survived']
```

In [6]:
```
X
```

Out[6]:

	pclass	age	sibsp	fare
0	3	22.0	1	7.2500
1	1	38.0	1	71.2833
2	3	26.0	0	7.9250
3	1	35.0	1	53.1000
4	3	35.0	0	8.0500
...	...	...	...	...
886	2	27.0	0	13.0000
887	1	19.0	0	30.0000
888	3	NaN	1	23.4500
889	1	26.0	0	30.0000
890	3	32.0	0	7.7500

891 rows × 4 columns

In [13]:
```
y
```

Out[13]:
```
0      0
1      1
2      1
3      1
4      0
      ..
886    0
887    1
888    0
889    1
890    0
Name: survived, Length: 891, dtype: int64
```

In [14]:
```
y.isnull().values.any()
```

Out[14]:
```
False
```

In [12]:
```
updated_df = X
updated_df['age']=updated_df['age'].fillna(updated_df['age'].mean())
updated_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 4 columns):
 #   Column  Non-Null Count  Dtype  
---  ------  --------------  -----  
 0   pclass  891 non-null    int64  
 1   age     891 non-null    float64
 2   sibsp   891 non-null    int64  
 3   fare    891 non-null    float64
dtypes: float64(2), int64(2)
memory usage: 28.0 KB

C:\Users\User\AppData\Local\Temp\ipykernel_964\925103184.py:2: SettingWithCop
yWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/s
table/user_guide/indexing.html#returning-a-view-versus-a-copy (https://panda
s.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-ver
sus-a-copy)
  updated_df['age']=updated_df['age'].fillna(updated_df['age'].mean())
```

In [15]:
```
def base_models(X,y, scoring="roc_auc"):
    print("Basel Models.....")
    classifiers = [('LR',LogisticRegression()), ('KNN',KNeighborsClassifier())
                   ("RF",RandomForestClassifier()),('Adaboost',AdaBoostClassifi
    for name, classifier in classifiers:
        cv_results = cross_validate(classifier, X,y, cv=3, scoring=scoring)
        print(f"{scoring}: {round(cv_results['test_score'].mean(),4)} ({name})
```

In [19]:
```
base_models(X,y, scoring="accuracy")
```

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

In [ ]: