water-quality-prediction

April 2, 2024

[46]:
```
from google.colab import drive
drive.mount('/content/drive')
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call
drive.mount("/content/drive", force_remount=True).
```

[47]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

[48]:
```
df=pd.read_csv(r'/content/drive/MyDrive/
  ↪Water-Quality-Prediction-using-Machine-Learning-main/
  ↪Water-Quality-Prediction-using-Machine-Learning-main/
  ↪Water-Quality-Prediction-using-Machine-Learning-main/water_potability.csv')
df.head()
```

[48]:

	ph	Hardness	Solids	Chloramines	Sulfate	Conductivity	\
0	NaN	204.890455	20791.318981	7.300212	368.516441	564.308654	
1	3.716080	129.422921	18630.057858	6.635246	NaN	592.885359	
2	8.099124	224.236259	19909.541732	9.275884	NaN	418.606213	
3	8.316766	214.373394	22018.417441	8.059332	356.886136	363.266516	
4	9.092223	181.101509	17978.986339	6.546600	310.135738	398.410813	

	Organic_carbon	Trihalomethanes	Turbidity	Potability
0	10.379783	86.990970	2.963135	0
1	15.180013	56.329076	4.500656	0
2	16.868637	66.420093	3.055934	0
3	18.436524	100.341674	4.628771	0
4	11.558279	31.997993	4.075075	0

1 Exploratory Data Analysis

[49]:
```
df.shape
```

[49]: (3276, 10)

[50]:
```
df.isnull().sum()
```

[50]:
```
ph                491
Hardness            0
Solids              0
Chloramines         0
Sulfate           781
Conductivity        0
Organic_carbon      0
Trihalomethanes   162
Turbidity           0
Potability          0
dtype: int64
```

[51]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3276 entries, 0 to 3275
Data columns (total 10 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   ph               2785 non-null   float64
 1   Hardness         3276 non-null   float64
 2   Solids           3276 non-null   float64
 3   Chloramines      3276 non-null   float64
 4   Sulfate          2495 non-null   float64
 5   Conductivity     3276 non-null   float64
 6   Organic_carbon   3276 non-null   float64
 7   Trihalomethanes  3114 non-null   float64
 8   Turbidity        3276 non-null   float64
 9   Potability       3276 non-null   int64  
dtypes: float64(9), int64(1)
memory usage: 256.1 KB
```

[52]:
```
df.describe()
```

[52]:

	ph	Hardness	Solids	Chloramines	Sulfate	\
count	2785.000000	3276.000000	3276.000000	3276.000000	2495.000000	
mean	7.080795	196.369496	22014.092526	7.122277	333.775777	
std	1.594320	32.879761	8768.570828	1.583085	41.416840	
min	0.000000	47.432000	320.942611	0.352000	129.000000	
25%	6.093092	176.850538	15666.690297	6.127421	307.699498	
50%	7.036752	196.967627	20927.833607	7.130299	333.073546	
75%	8.062066	216.667456	27332.762127	8.114887	359.950170	
max	14.000000	323.124000	61227.196008	13.127000	481.030642	

	Conductivity	Organic_carbon	Trihalomethanes	Turbidity	Potability

```
count   3276.000000     3276.000000      3114.000000  3276.000000  3276.000000
mean     426.205111       14.284970        66.396293     3.966786     0.390110
std       80.824064        3.308162        16.175008     0.780382     0.487849
min      181.483754        2.200000         0.738000     1.450000     0.000000
25%      365.734414       12.065801        55.844536     3.439711     0.000000
50%      421.884968       14.218338        66.622485     3.955028     0.000000
75%      481.792304       16.557652        77.337473     4.500320     1.000000
max      753.342620       28.300000       124.000000     6.739000     1.000000
```

[53]:
```
df['Sulfate'].mean()
```

[53]:
```
333.7757766108135
```

[54]:
```
df.fillna(df.mean(), inplace=True)
df.head()
```

[54]:
```
         ph    Hardness        Solids  Chloramines     Sulfate  Conductivity  \
0  7.080795  204.890455  20791.318981     7.300212  368.516441    564.308654
1  3.716080  129.422921  18630.057858     6.635246  333.775777    592.885359
2  8.099124  224.236259  19909.541732     9.275884  333.775777    418.606213
3  8.316766  214.373394  22018.417441     8.059332  356.886136    363.266516
4  9.092223  181.101509  17978.986339     6.546600  310.135738    398.410813

   Organic_carbon  Trihalomethanes  Turbidity  Potability
0       10.379783        86.990970   2.963135           0
1       15.180013        56.329076   4.500656           0
2       16.868637        66.420093   3.055934           0
3       18.436524       100.341674   4.628771           0
4       11.558279        31.997993   4.075075           0
```

[55]:
```
df.isnull().sum()
```

[55]:
```
ph                0
Hardness          0
Solids            0
Chloramines       0
Sulfate           0
Conductivity      0
Organic_carbon    0
Trihalomethanes   0
Turbidity         0
Potability        0
dtype: int64
```

[56]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 3276 entries, 0 to 3275
Data columns (total 10 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   ph               3276 non-null   float64
 1   Hardness         3276 non-null   float64
 2   Solids           3276 non-null   float64
 3   Chloramines      3276 non-null   float64
 4   Sulfate          3276 non-null   float64
 5   Conductivity     3276 non-null   float64
 6   Organic_carbon   3276 non-null   float64
 7   Trihalomethanes  3276 non-null   float64
 8   Turbidity        3276 non-null   float64
 9   Potability       3276 non-null   int64  
dtypes: float64(9), int64(1)
memory usage: 256.1 KB
```

[57]:
```
df.describe()
```

[57]:

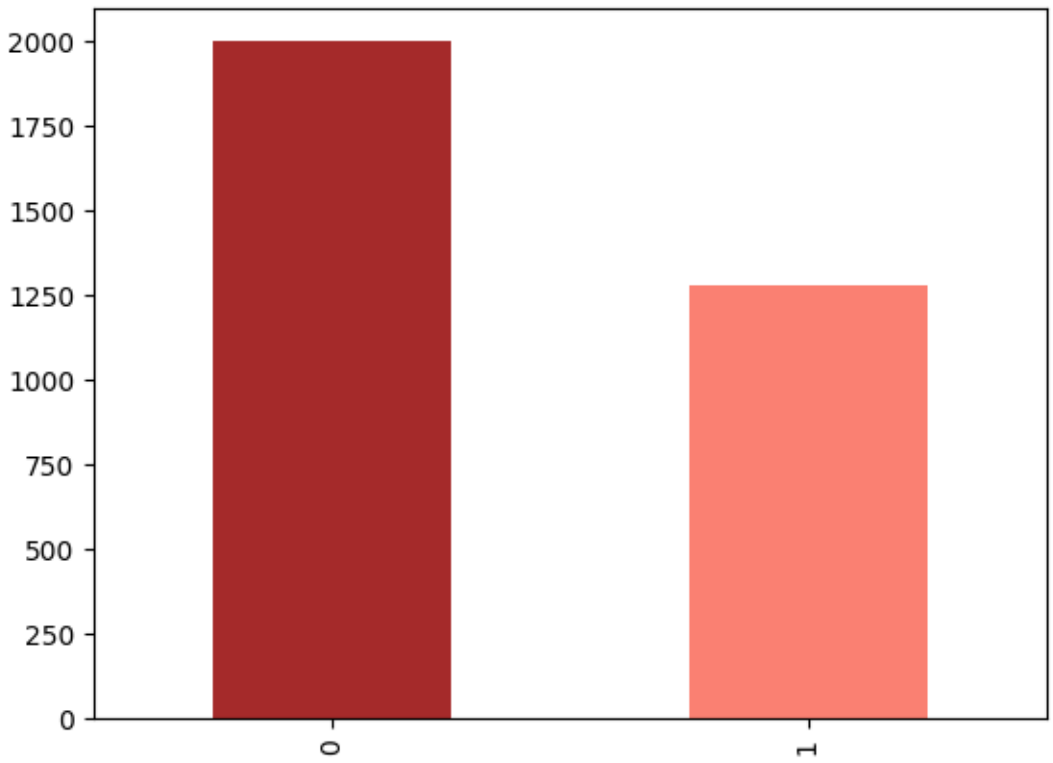
	ph	Hardness	Solids	Chloramines	Sulfate	\
count	3276.000000	3276.000000	3276.000000	3276.000000	3276.000000	
mean	7.080795	196.369496	22014.092526	7.122277	333.775777	
std	1.469956	32.879761	8768.570828	1.583085	36.142612	
min	0.000000	47.432000	320.942611	0.352000	129.000000	
25%	6.277673	176.850538	15666.690297	6.127421	317.094638	
50%	7.080795	196.967627	20927.833607	7.130299	333.775777	
75%	7.870050	216.667456	27332.762127	8.114887	350.385756	
max	14.000000	323.124000	61227.196008	13.127000	481.030642	

	Conductivity	Organic_carbon	Trihalomethanes	Turbidity	Potability
count	3276.000000	3276.000000	3276.000000	3276.000000	3276.000000
mean	426.205111	14.284970	66.396293	3.966786	0.390110
std	80.824064	3.308162	15.769881	0.780382	0.487849
min	181.483754	2.200000	0.738000	1.450000	0.000000
25%	365.734414	12.065801	56.647656	3.439711	0.000000
50%	421.884968	14.218338	66.396293	3.955028	0.000000
75%	481.792304	16.557652	76.666609	4.500320	1.000000
max	753.342620	28.300000	124.000000	6.739000	1.000000

[58]:
```
df.Potability.value_counts()
```

[58]:
```
0    1998
1    1278
Name: Potability, dtype: int64
```

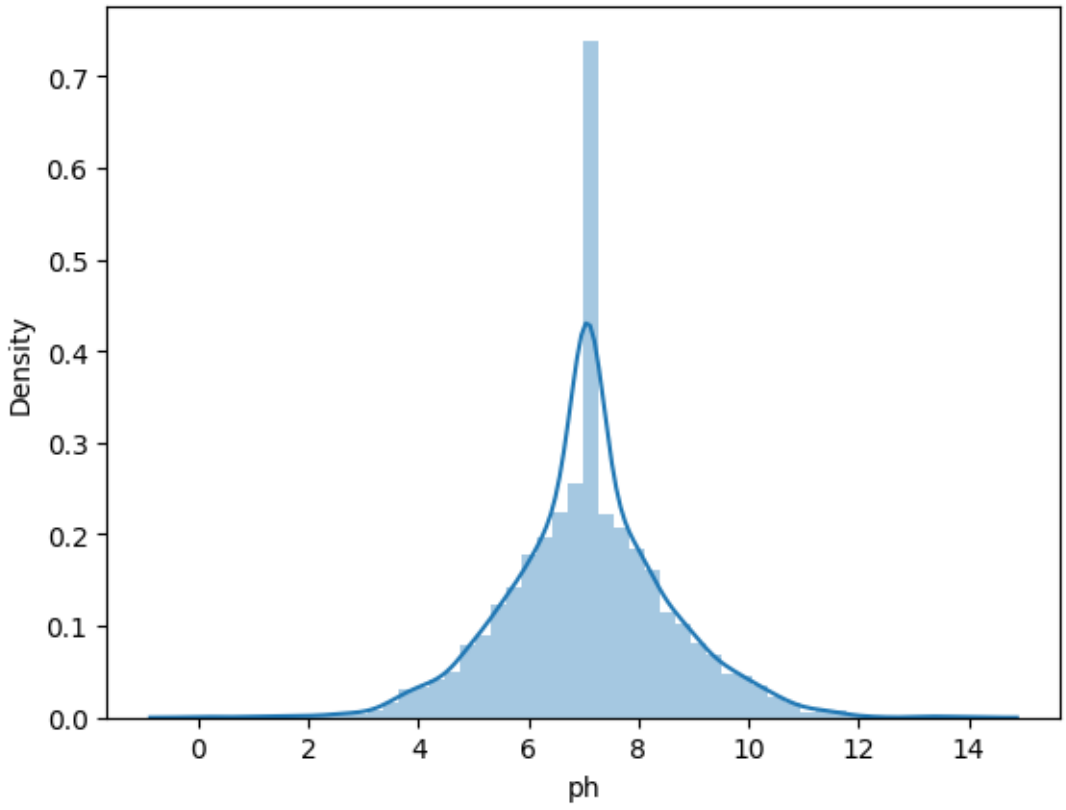
[59]:
```
df.Potability.value_counts().plot(kind="bar", color=["brown", "salmon"])
plt.show()
```



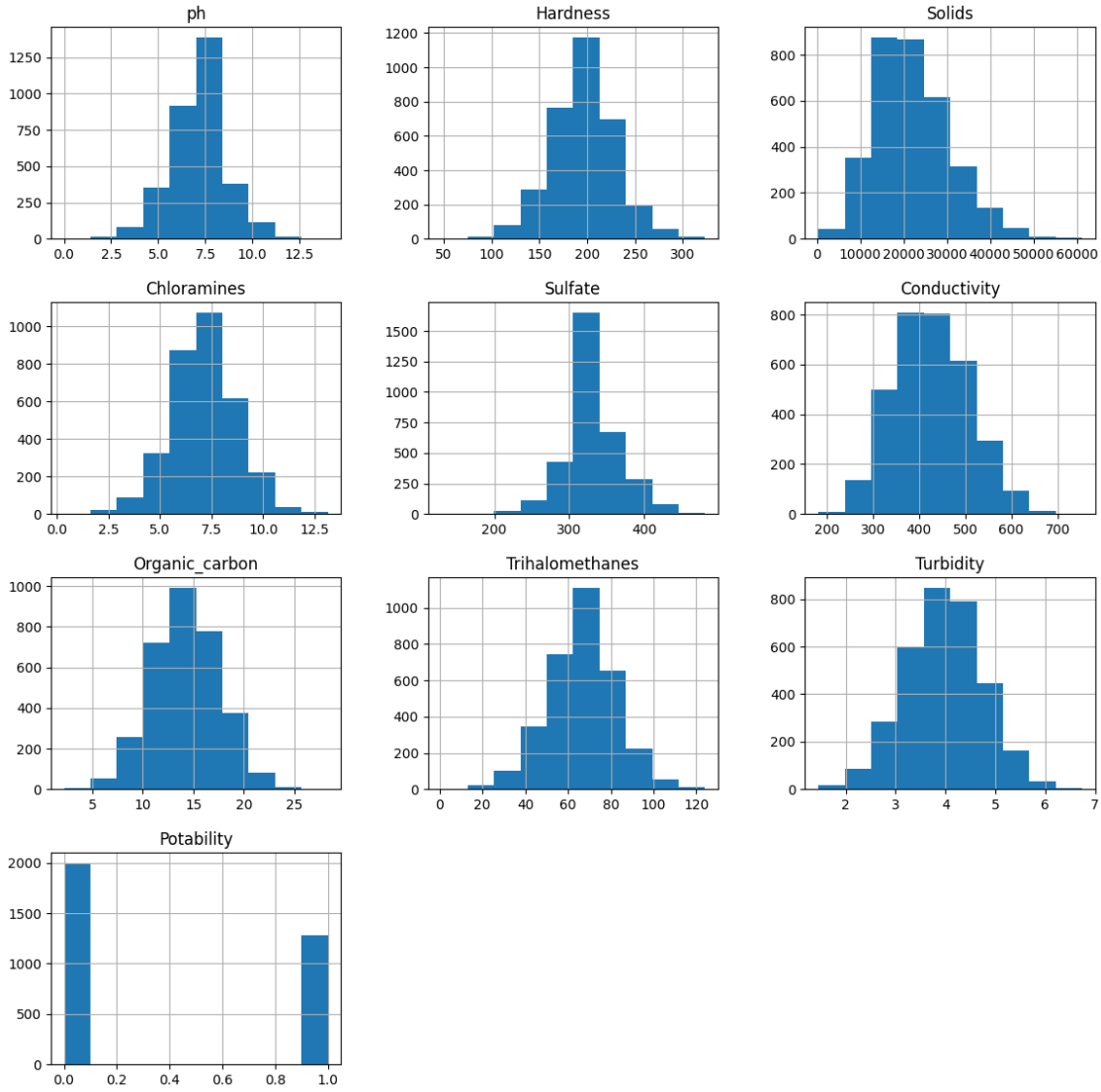
[60]:
```
sns.distplot(df['ph'])
```

```
<ipython-input-60-aa7801fe055a>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df['ph'])
```

[60]: <Axes: xlabel='ph', ylabel='Density'>

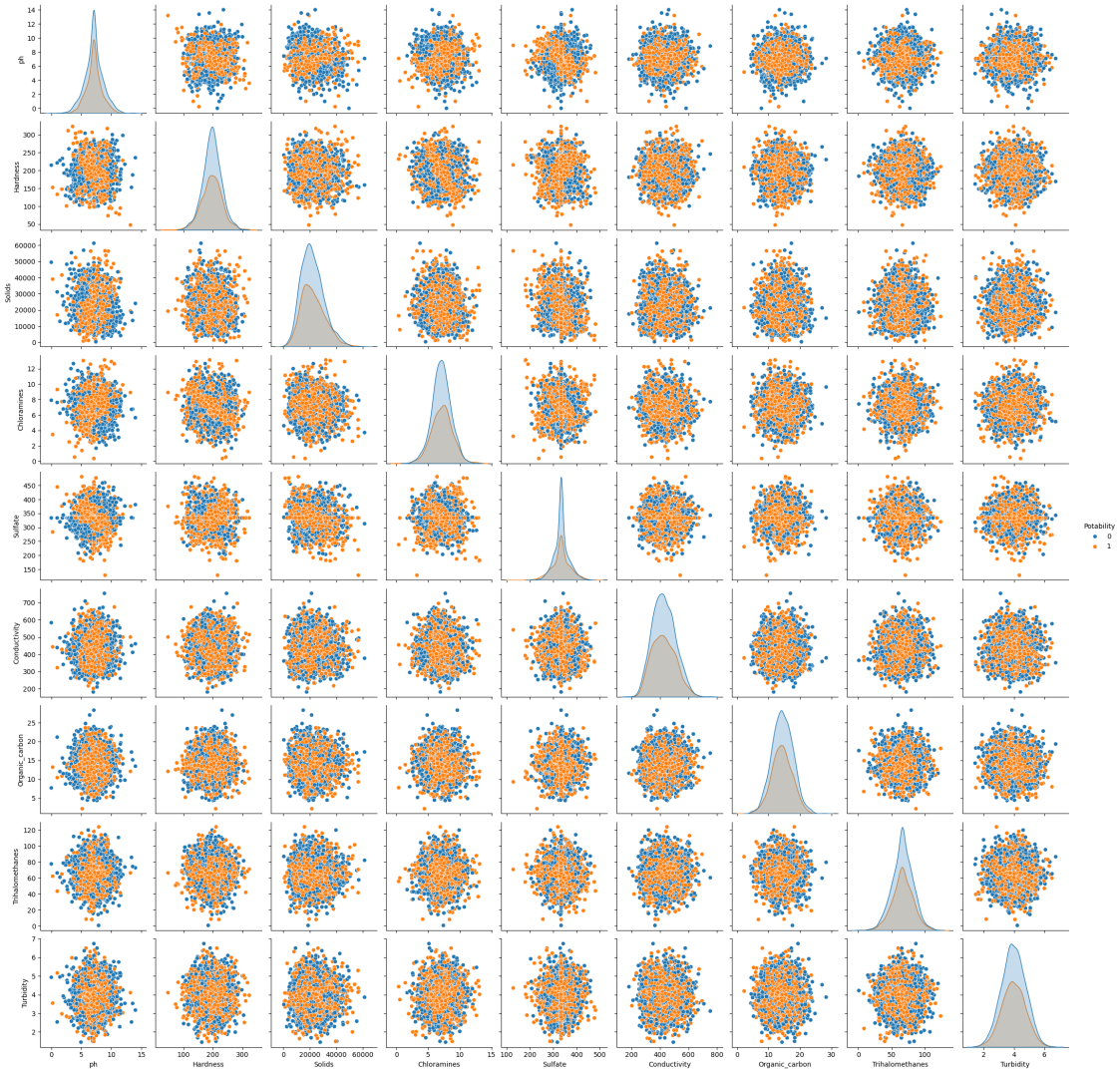


[61]:
```
df.hist(figsize=(14,14))
plt.show()
```



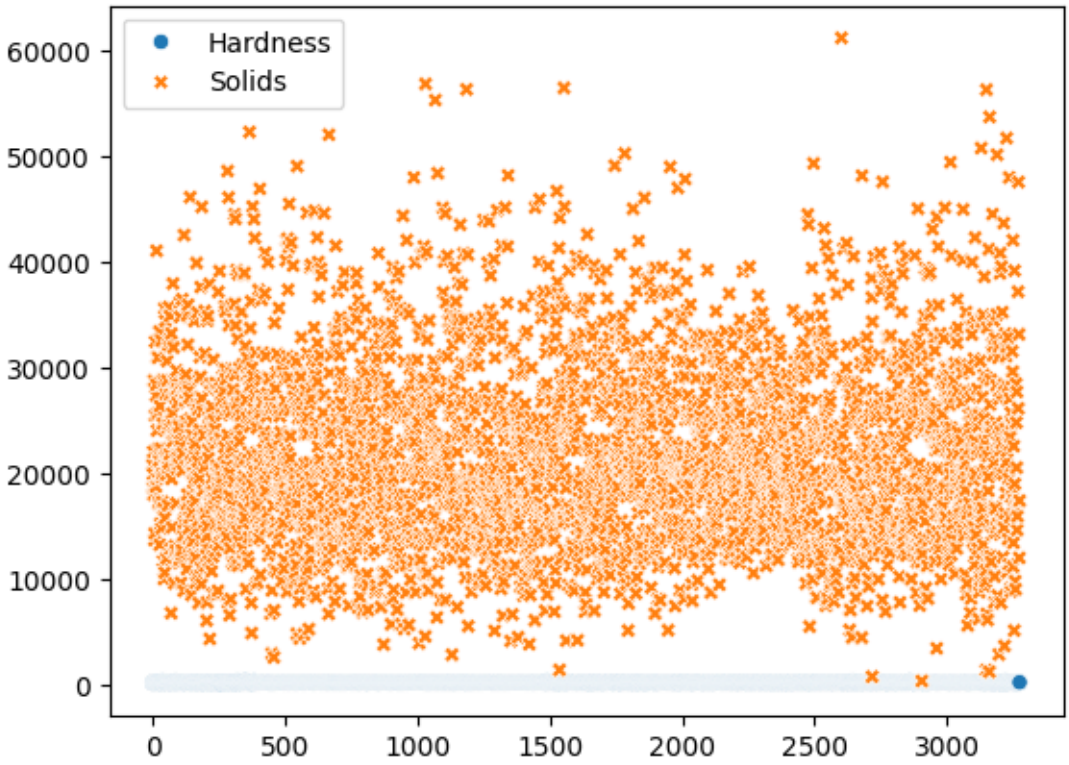
[62]:
```
sns.pairplot(df,hue='Potability')
```

[62]: <seaborn.axisgrid.PairGrid at 0x7de9f97b5360>



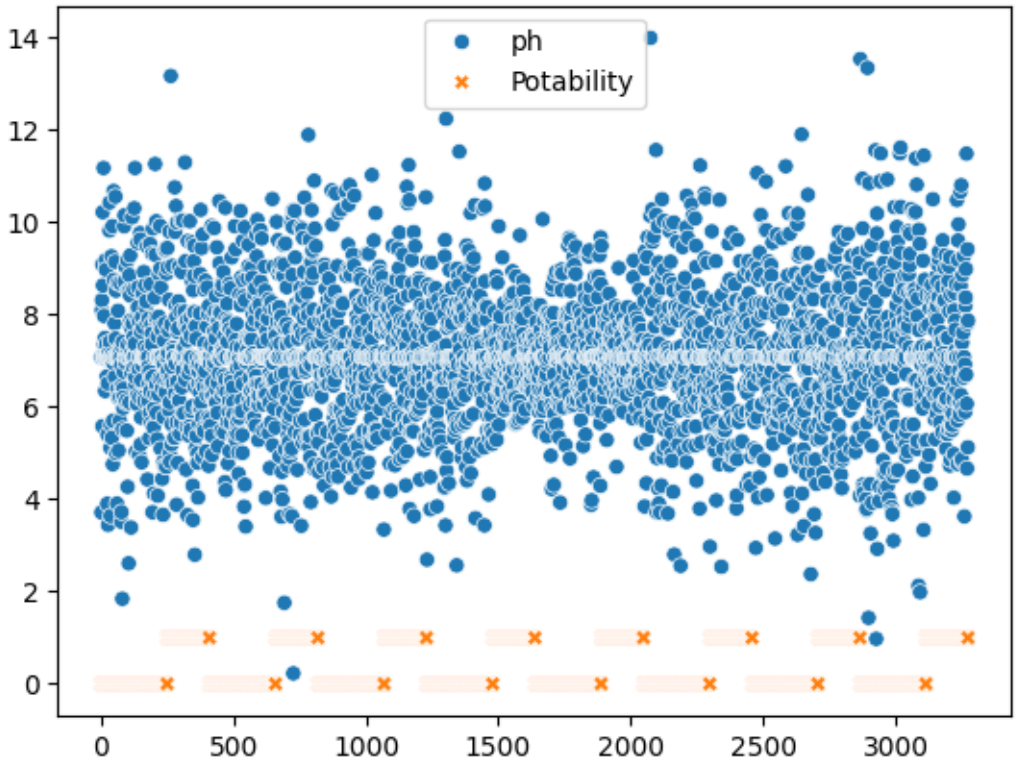
[63]:
```
sns.scatterplot(df[['Hardness','Solids']])
```

[63]: <Axes: >

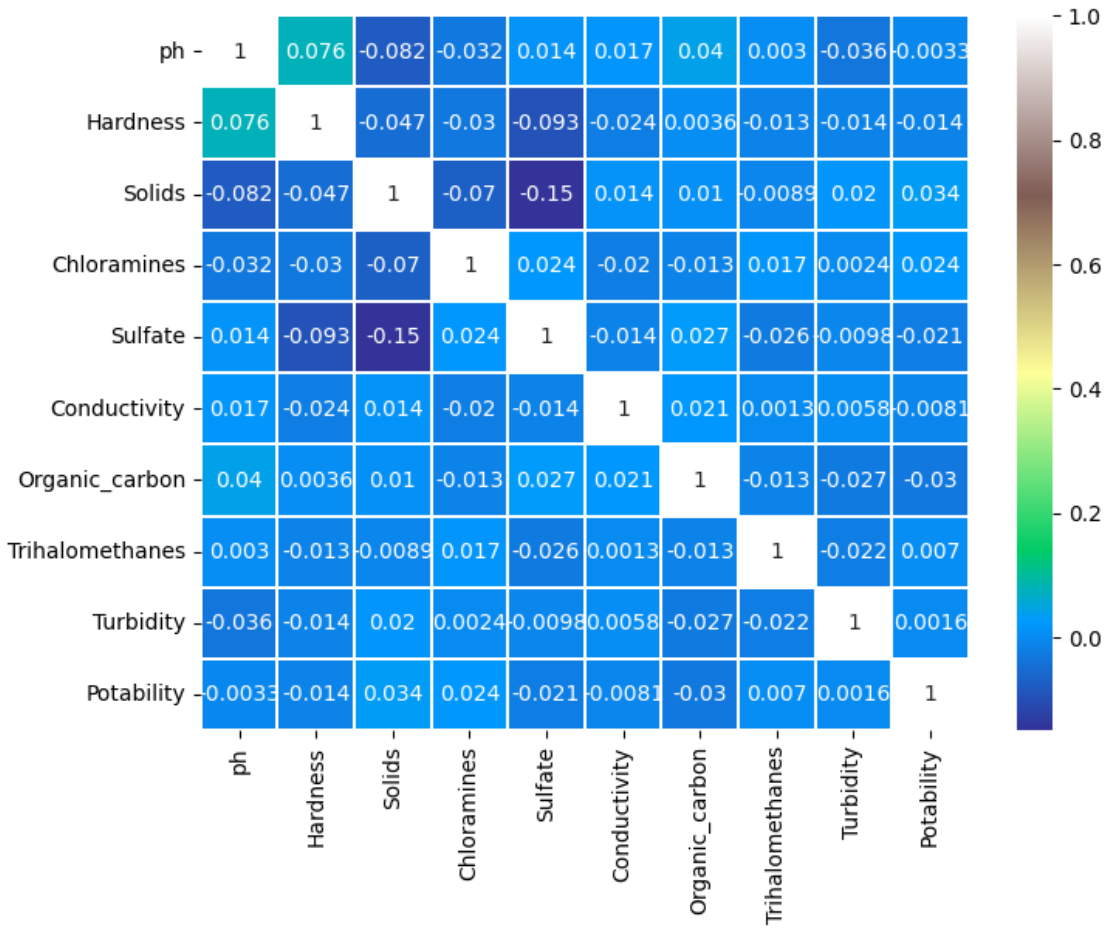


[64]:
```
sns.scatterplot(df[['ph','Potability']])
```

[64]: <Axes: >

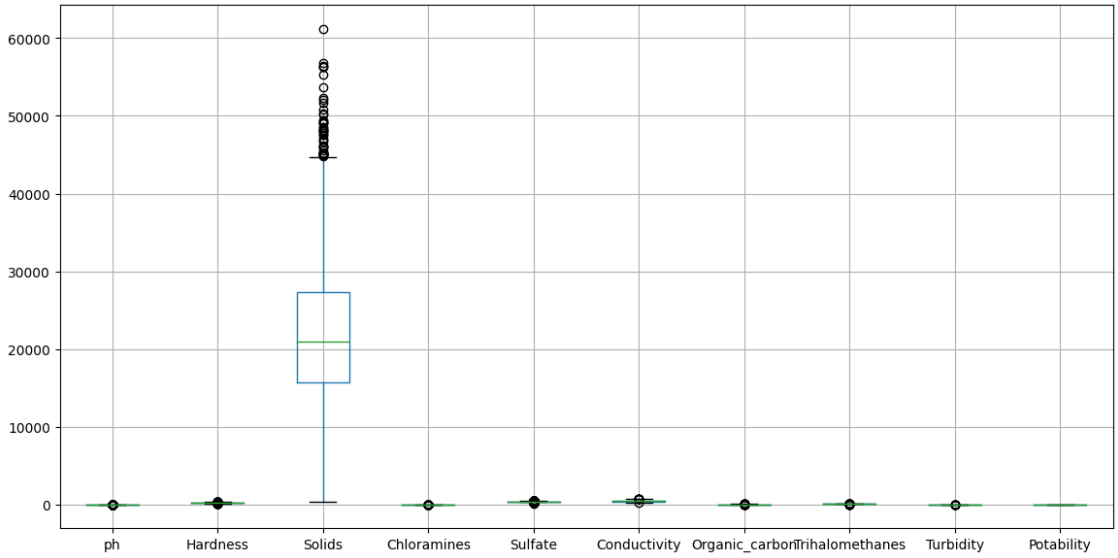


[65]:
```
# create a correlation heatmap
sns.heatmap(df.corr(),annot=True, cmap='terrain', linewidths=0.1)
fig=plt.gcf()
fig.set_size_inches(8,6)
plt.show()
```



[66]:
```
df.boxplot(figsize=(14,7))
```

[66]: <Axes: >



[67]:
```
df['Solids'].describe()
```

[67]:
```
count     3276.000000
mean     22014.092526
std       8768.570828
min        320.942611
25%      15666.690297
50%      20927.833607
75%      27332.762127
max      61227.196008
Name: Solids, dtype: float64
```

[67]:
```

```

2 Partitioning

[68]:
```
X = df.drop('Potability',axis=1)
```

[69]:
```
Y= df['Potability']
```

[70]:
```
from sklearn.model_selection import train_test_split
X_train, X_test, Y_train, Y_test = train_test_split(X,Y, test_size= 0.2,␣
 ↪random_state=101,shuffle=True)
```

[71]:
```
Y_train.value_counts()
```

[71]:
```
0    1596
1    1024
Name: Potability, dtype: int64
```

[72]:
```
Y_test.value_counts()
```

[72]:
```
0    402
1    254
Name: Potability, dtype: int64
```

3 Normalization

[73]:
```
#from sklearn.preprocessing import StandardScaler
#sc=StandardScaler()
```

[74]:
```
#X_train = sc.fit_transform(X_train)
#X_test = sc.transform(X_test)
```

4 Model Building

5 DT

[75]:
```
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score,confusion_matrix,precision_score
dt=DecisionTreeClassifier(criterion= 'gini', min_samples_split= 10, splitter=␣
 ↪'best')
dt.fit(X_train,Y_train)
```

[75]:
```
DecisionTreeClassifier(min_samples_split=10)
```

[76]:
```
prediction=dt.predict(X_test)
accuracy_dt=accuracy_score(Y_test,prediction)*100
accuracy_dt
```

[76]:
```
58.6890243902439
```

[77]:
```
print("Accuracy on training set: {:.3f}".format(dt.score(X_train, Y_train)))
print("Accuracy on test set: {:.3f}".format(dt.score(X_test, Y_test)))
```

```
Accuracy on training set: 0.923
Accuracy on test set: 0.587
```

[77]:
```

```

[78]:
```
accuracy_score(prediction,Y_test)
```

[78]: 0.586890243902439

[79]:
```
print("Feature importances:\n{}".format(dt.feature_importances_))
```

```
Feature importances:
[0.15055508 0.1308659  0.11044326 0.10203988 0.12480112 0.09764863
 0.09708777 0.10508071 0.08147766]
```

[80]:
```
confusion_matrix(prediction,Y_test)
```

[80]:
```
array([[275, 144],
       [127, 110]])
```

6 Prediction on only one set of data

[81]:
```
X_DT=dt.predict([[5.735724, 158.318741,25363.016594,7.728601,377.543291,568.
↪304671,13.626624,75.952337,4.732954]])
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does
not have valid feature names, but DecisionTreeClassifier was fitted with feature
names
  warnings.warn(
```

[82]:
```
X_DT
```

[82]:
```
array([1])
```

[82]:

[82]:

7 KNN

[82]:

[83]:
```
from sklearn.neighbors import KNeighborsClassifier
```

[84]:
```
knn=KNeighborsClassifier(metric='manhattan', n_neighbors=22)
knn.fit(X_train,Y_train)
```

[84]:
```
KNeighborsClassifier(metric='manhattan', n_neighbors=22)
```

[85]:
```
prediction_knn=knn.predict(X_test)
accuracy_knn=accuracy_score(Y_test,prediction_knn)*100
```

```
print('accuracy_score score     :␣
 ↪',accuracy_score(Y_test,prediction_knn)*100,'%')
```

```
accuracy_score score     :  61.737804878048784 %
```

[86]:
```
confusion_matrix(prediction,Y_test)
```

[86]:
```
array([[275, 144],
       [127, 110]])
```

[86]:

[86]:

8 Hyperparameter Tuning / Model Optimization

9 DT HPT

[87]:
```
dt.get_params().keys()
```

[87]:
```
dict_keys(['ccp_alpha', 'class_weight', 'criterion', 'max_depth',
'max_features', 'max_leaf_nodes', 'min_impurity_decrease', 'min_samples_leaf',
'min_samples_split', 'min_weight_fraction_leaf', 'random_state', 'splitter'])
```

[89]:
```
#example of grid searching key hyperparametres for logistic regression
from sklearn.model_selection import RepeatedStratifiedKFold
from sklearn.model_selection import GridSearchCV

# define models and parameters
model = DecisionTreeClassifier()
criterion = ["gini", "entropy"]
splitter = ["best", "random"]
min_samples_split = [2,4,6,8,10]

# define grid search
grid = dict(splitter=splitter, criterion=criterion,␣
 ↪min_samples_split=min_samples_split)
cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)
grid_search_dt = GridSearchCV(estimator=model, param_grid=grid, n_jobs=-1,␣
 ↪cv=cv,
                              scoring='accuracy',error_score=0)
grid_search_dt.fit(X_train, Y_train)

# summarize results
print(f"Best: {grid_search_dt.best_score_:.3f} using {grid_search_dt.
 ↪best_params_}")
```

```
means = grid_search_dt.cv_results_['mean_test_score']
stds = grid_search_dt.cv_results_['std_test_score']
params = grid_search_dt.cv_results_['params']

for mean, stdev, param in zip(means, stds, params):
    print(f"{mean:.3f} ({stdev:.3f}) with: {param}")

print("Training Score:",grid_search_dt.score(X_train, Y_train)*100)
print("Testing Score:", grid_search_dt.score(X_test, Y_test)*100)
```

```
Best: 0.593 using {'criterion': 'entropy', 'min_samples_split': 6, 'splitter':
'random'}
0.577 (0.029) with: {'criterion': 'gini', 'min_samples_split': 2, 'splitter':
'best'}
0.584 (0.034) with: {'criterion': 'gini', 'min_samples_split': 2, 'splitter':
'random'}
0.585 (0.034) with: {'criterion': 'gini', 'min_samples_split': 4, 'splitter':
'best'}
0.576 (0.034) with: {'criterion': 'gini', 'min_samples_split': 4, 'splitter':
'random'}
0.586 (0.029) with: {'criterion': 'gini', 'min_samples_split': 6, 'splitter':
'best'}
0.576 (0.031) with: {'criterion': 'gini', 'min_samples_split': 6, 'splitter':
'random'}
0.588 (0.028) with: {'criterion': 'gini', 'min_samples_split': 8, 'splitter':
'best'}
0.588 (0.027) with: {'criterion': 'gini', 'min_samples_split': 8, 'splitter':
'random'}
0.588 (0.030) with: {'criterion': 'gini', 'min_samples_split': 10, 'splitter':
'best'}
0.575 (0.040) with: {'criterion': 'gini', 'min_samples_split': 10, 'splitter':
'random'}
0.584 (0.028) with: {'criterion': 'entropy', 'min_samples_split': 2, 'splitter':
'best'}
0.574 (0.030) with: {'criterion': 'entropy', 'min_samples_split': 2, 'splitter':
'random'}
0.587 (0.027) with: {'criterion': 'entropy', 'min_samples_split': 4, 'splitter':
'best'}
0.586 (0.036) with: {'criterion': 'entropy', 'min_samples_split': 4, 'splitter':
'random'}
0.587 (0.031) with: {'criterion': 'entropy', 'min_samples_split': 6, 'splitter':
'best'}
0.593 (0.019) with: {'criterion': 'entropy', 'min_samples_split': 6, 'splitter':
'random'}
0.591 (0.028) with: {'criterion': 'entropy', 'min_samples_split': 8, 'splitter':
'best'}
0.582 (0.026) with: {'criterion': 'entropy', 'min_samples_split': 8, 'splitter':
```

```
'random'}
0.586 (0.028) with: {'criterion': 'entropy', 'min_samples_split': 10,
'splitter': 'best'}
0.575 (0.031) with: {'criterion': 'entropy', 'min_samples_split': 10,
'splitter': 'random'}
Training Score: 91.29770992366413
Testing Score: 60.670731707317074
```

[90]:
```
from sklearn.metrics import  make_scorer
from sklearn.model_selection import cross_val_score

def classification_report_with_accuracy_score(Y_test, y_pred2):
    print (classification_report(Y_test, y_pred2)) # print classification report
    return accuracy_score(Y_test, y_pred2) # return accuracy score


nested_score = cross_val_score(grid_search_dt, X=X_train, y=Y_train, cv=cv,
               scoring=make_scorer(classification_report_with_accuracy_score))
print (nested_score)
```

```
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
```

```
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
```

```
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
```

```
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
```

```
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined
```

```
  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
```

```
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
```

```
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
```

```
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
```

```
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
```

```
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(

[nan nan nan nan nan nan nan nan nan nan nan nan nan nan nan nan nan nan
 nan nan nan nan nan nan nan nan nan nan nan nan]

/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-90-5029318a6577>", line 5, in
```

```
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
```

[91]:
```
dt_y_predicted = grid_search_dt.predict(X_test)
dt_y_predicted
```

[91]:
```
array([0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0,
       0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0,
       1, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0,
       0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0,
       1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0,
       0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0,
       0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1,
       1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0,
       0, 1, 1, 1, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0,
       0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1,
       0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1,
       0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0,
       0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0,
       0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0,
       1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0,
       1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0,
       1, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1,
       1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0,
       1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0,
       0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 0,
       0, 0, 0, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0,
       1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 0,
       1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 0,
       0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0,
       0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1,
       1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0,
       1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1,
       1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0])
```

[92]:
```
grid_search_dt.best_params_
```

[92]:
```
{'criterion': 'entropy', 'min_samples_split': 6, 'splitter': 'random'}
```

[93]:
```
dt_grid_score=accuracy_score(Y_test, dt_y_predicted)
dt_grid_score
```

[93]: 0.6067073170731707

[94]:
```
confusion_matrix(Y_test, dt_y_predicted)
```

[94]:
```
array([[276, 126],
       [132, 122]])
```

[ ]:

[ ]:

[ ]:

10 KNN HPT

[96]:
```
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import RepeatedStratifiedKFold
from sklearn.model_selection import GridSearchCV

# define models and parameters
model = KNeighborsClassifier()
n_neighbors = range(1, 31)
weights = ['uniform', 'distance']
metric = ['euclidean', 'manhattan', 'minkowski']

# define grid search
grid = dict(n_neighbors=n_neighbors,weights=weights,metric=metric)
cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=1, random_state=1)
grid_search_knn = GridSearchCV(estimator=model, param_grid=grid, n_jobs=-1,␣
 ↪cv=cv,
                           scoring='accuracy',error_score=0)
grid_search_knn.fit(X_train, Y_train)

# summarize results
print(f"Best: {grid_search_knn.best_score_:.3f} using {grid_search_knn.
 ↪best_params_}")
means = grid_search_knn.cv_results_['mean_test_score']
stds = grid_search_knn.cv_results_['std_test_score']
params = grid_search_knn.cv_results_['params']

for mean, stdev, param in zip(means, stds, params):
    print(f"{mean:.3f} ({stdev:.3f}) with: {param}")
```

```
Best: 0.603 using {'metric': 'manhattan', 'n_neighbors': 22, 'weights':
'uniform'}
0.536 (0.029) with: {'metric': 'euclidean', 'n_neighbors': 1, 'weights':
```

```
'uniform'}
0.536 (0.029) with: {'metric': 'euclidean', 'n_neighbors': 1, 'weights':
'distance'}
0.579 (0.020) with: {'metric': 'euclidean', 'n_neighbors': 2, 'weights':
'uniform'}
0.536 (0.029) with: {'metric': 'euclidean', 'n_neighbors': 2, 'weights':
'distance'}
0.542 (0.023) with: {'metric': 'euclidean', 'n_neighbors': 3, 'weights':
'uniform'}
0.542 (0.024) with: {'metric': 'euclidean', 'n_neighbors': 3, 'weights':
'distance'}
0.574 (0.017) with: {'metric': 'euclidean', 'n_neighbors': 4, 'weights':
'uniform'}
0.542 (0.016) with: {'metric': 'euclidean', 'n_neighbors': 4, 'weights':
'distance'}
0.545 (0.020) with: {'metric': 'euclidean', 'n_neighbors': 5, 'weights':
'uniform'}
0.544 (0.020) with: {'metric': 'euclidean', 'n_neighbors': 5, 'weights':
'distance'}
0.579 (0.020) with: {'metric': 'euclidean', 'n_neighbors': 6, 'weights':
'uniform'}
0.556 (0.025) with: {'metric': 'euclidean', 'n_neighbors': 6, 'weights':
'distance'}
0.561 (0.022) with: {'metric': 'euclidean', 'n_neighbors': 7, 'weights':
'uniform'}
0.564 (0.018) with: {'metric': 'euclidean', 'n_neighbors': 7, 'weights':
'distance'}
0.580 (0.028) with: {'metric': 'euclidean', 'n_neighbors': 8, 'weights':
'uniform'}
0.564 (0.027) with: {'metric': 'euclidean', 'n_neighbors': 8, 'weights':
'distance'}
0.560 (0.024) with: {'metric': 'euclidean', 'n_neighbors': 9, 'weights':
'uniform'}
0.569 (0.026) with: {'metric': 'euclidean', 'n_neighbors': 9, 'weights':
'distance'}
0.585 (0.020) with: {'metric': 'euclidean', 'n_neighbors': 10, 'weights':
'uniform'}
0.566 (0.027) with: {'metric': 'euclidean', 'n_neighbors': 10, 'weights':
'distance'}
0.557 (0.021) with: {'metric': 'euclidean', 'n_neighbors': 11, 'weights':
'uniform'}
0.566 (0.025) with: {'metric': 'euclidean', 'n_neighbors': 11, 'weights':
'distance'}
0.584 (0.015) with: {'metric': 'euclidean', 'n_neighbors': 12, 'weights':
'uniform'}
0.563 (0.025) with: {'metric': 'euclidean', 'n_neighbors': 12, 'weights':
'distance'}
0.565 (0.019) with: {'metric': 'euclidean', 'n_neighbors': 13, 'weights':
```

```
'uniform'}
0.561 (0.027) with: {'metric': 'euclidean', 'n_neighbors': 13, 'weights':
'distance'}
0.588 (0.010) with: {'metric': 'euclidean', 'n_neighbors': 14, 'weights':
'uniform'}
0.569 (0.024) with: {'metric': 'euclidean', 'n_neighbors': 14, 'weights':
'distance'}
0.581 (0.013) with: {'metric': 'euclidean', 'n_neighbors': 15, 'weights':
'uniform'}
0.574 (0.025) with: {'metric': 'euclidean', 'n_neighbors': 15, 'weights':
'distance'}
0.590 (0.011) with: {'metric': 'euclidean', 'n_neighbors': 16, 'weights':
'uniform'}
0.575 (0.024) with: {'metric': 'euclidean', 'n_neighbors': 16, 'weights':
'distance'}
0.578 (0.015) with: {'metric': 'euclidean', 'n_neighbors': 17, 'weights':
'uniform'}
0.572 (0.025) with: {'metric': 'euclidean', 'n_neighbors': 17, 'weights':
'distance'}
0.590 (0.012) with: {'metric': 'euclidean', 'n_neighbors': 18, 'weights':
'uniform'}
0.580 (0.023) with: {'metric': 'euclidean', 'n_neighbors': 18, 'weights':
'distance'}
0.586 (0.012) with: {'metric': 'euclidean', 'n_neighbors': 19, 'weights':
'uniform'}
0.585 (0.025) with: {'metric': 'euclidean', 'n_neighbors': 19, 'weights':
'distance'}
0.595 (0.013) with: {'metric': 'euclidean', 'n_neighbors': 20, 'weights':
'uniform'}
0.581 (0.020) with: {'metric': 'euclidean', 'n_neighbors': 20, 'weights':
'distance'}
0.589 (0.014) with: {'metric': 'euclidean', 'n_neighbors': 21, 'weights':
'uniform'}
0.585 (0.019) with: {'metric': 'euclidean', 'n_neighbors': 21, 'weights':
'distance'}
0.598 (0.015) with: {'metric': 'euclidean', 'n_neighbors': 22, 'weights':
'uniform'}
0.586 (0.021) with: {'metric': 'euclidean', 'n_neighbors': 22, 'weights':
'distance'}
0.592 (0.014) with: {'metric': 'euclidean', 'n_neighbors': 23, 'weights':
'uniform'}
0.587 (0.016) with: {'metric': 'euclidean', 'n_neighbors': 23, 'weights':
'distance'}
0.598 (0.016) with: {'metric': 'euclidean', 'n_neighbors': 24, 'weights':
'uniform'}
0.587 (0.016) with: {'metric': 'euclidean', 'n_neighbors': 24, 'weights':
'distance'}
0.590 (0.016) with: {'metric': 'euclidean', 'n_neighbors': 25, 'weights':
```

```
'uniform'}
0.587 (0.018) with: {'metric': 'euclidean', 'n_neighbors': 25, 'weights':
'distance'}
0.600 (0.015) with: {'metric': 'euclidean', 'n_neighbors': 26, 'weights':
'uniform'}
0.584 (0.018) with: {'metric': 'euclidean', 'n_neighbors': 26, 'weights':
'distance'}
0.590 (0.013) with: {'metric': 'euclidean', 'n_neighbors': 27, 'weights':
'uniform'}
0.588 (0.019) with: {'metric': 'euclidean', 'n_neighbors': 27, 'weights':
'distance'}
0.595 (0.015) with: {'metric': 'euclidean', 'n_neighbors': 28, 'weights':
'uniform'}
0.583 (0.019) with: {'metric': 'euclidean', 'n_neighbors': 28, 'weights':
'distance'}
0.589 (0.018) with: {'metric': 'euclidean', 'n_neighbors': 29, 'weights':
'uniform'}
0.590 (0.015) with: {'metric': 'euclidean', 'n_neighbors': 29, 'weights':
'distance'}
0.590 (0.015) with: {'metric': 'euclidean', 'n_neighbors': 30, 'weights':
'uniform'}
0.588 (0.020) with: {'metric': 'euclidean', 'n_neighbors': 30, 'weights':
'distance'}
0.534 (0.031) with: {'metric': 'manhattan', 'n_neighbors': 1, 'weights':
'uniform'}
0.534 (0.031) with: {'metric': 'manhattan', 'n_neighbors': 1, 'weights':
'distance'}
0.589 (0.017) with: {'metric': 'manhattan', 'n_neighbors': 2, 'weights':
'uniform'}
0.534 (0.031) with: {'metric': 'manhattan', 'n_neighbors': 2, 'weights':
'distance'}
0.550 (0.018) with: {'metric': 'manhattan', 'n_neighbors': 3, 'weights':
'uniform'}
0.544 (0.018) with: {'metric': 'manhattan', 'n_neighbors': 3, 'weights':
'distance'}
0.577 (0.019) with: {'metric': 'manhattan', 'n_neighbors': 4, 'weights':
'uniform'}
0.555 (0.020) with: {'metric': 'manhattan', 'n_neighbors': 4, 'weights':
'distance'}
0.556 (0.020) with: {'metric': 'manhattan', 'n_neighbors': 5, 'weights':
'uniform'}
0.556 (0.016) with: {'metric': 'manhattan', 'n_neighbors': 5, 'weights':
'distance'}
0.585 (0.020) with: {'metric': 'manhattan', 'n_neighbors': 6, 'weights':
'uniform'}
0.564 (0.021) with: {'metric': 'manhattan', 'n_neighbors': 6, 'weights':
'distance'}
0.571 (0.016) with: {'metric': 'manhattan', 'n_neighbors': 7, 'weights':
```

```
'uniform'}
0.571 (0.021) with: {'metric': 'manhattan', 'n_neighbors': 7, 'weights':
'distance'}
0.586 (0.019) with: {'metric': 'manhattan', 'n_neighbors': 8, 'weights':
'uniform'}
0.571 (0.020) with: {'metric': 'manhattan', 'n_neighbors': 8, 'weights':
'distance'}
0.571 (0.021) with: {'metric': 'manhattan', 'n_neighbors': 9, 'weights':
'uniform'}
0.574 (0.025) with: {'metric': 'manhattan', 'n_neighbors': 9, 'weights':
'distance'}
0.581 (0.014) with: {'metric': 'manhattan', 'n_neighbors': 10, 'weights':
'uniform'}
0.571 (0.024) with: {'metric': 'manhattan', 'n_neighbors': 10, 'weights':
'distance'}
0.562 (0.023) with: {'metric': 'manhattan', 'n_neighbors': 11, 'weights':
'uniform'}
0.575 (0.026) with: {'metric': 'manhattan', 'n_neighbors': 11, 'weights':
'distance'}
0.582 (0.019) with: {'metric': 'manhattan', 'n_neighbors': 12, 'weights':
'uniform'}
0.570 (0.031) with: {'metric': 'manhattan', 'n_neighbors': 12, 'weights':
'distance'}
0.572 (0.021) with: {'metric': 'manhattan', 'n_neighbors': 13, 'weights':
'uniform'}
0.571 (0.030) with: {'metric': 'manhattan', 'n_neighbors': 13, 'weights':
'distance'}
0.583 (0.012) with: {'metric': 'manhattan', 'n_neighbors': 14, 'weights':
'uniform'}
0.569 (0.024) with: {'metric': 'manhattan', 'n_neighbors': 14, 'weights':
'distance'}
0.571 (0.017) with: {'metric': 'manhattan', 'n_neighbors': 15, 'weights':
'uniform'}
0.568 (0.023) with: {'metric': 'manhattan', 'n_neighbors': 15, 'weights':
'distance'}
0.585 (0.010) with: {'metric': 'manhattan', 'n_neighbors': 16, 'weights':
'uniform'}
0.574 (0.030) with: {'metric': 'manhattan', 'n_neighbors': 16, 'weights':
'distance'}
0.579 (0.017) with: {'metric': 'manhattan', 'n_neighbors': 17, 'weights':
'uniform'}
0.585 (0.021) with: {'metric': 'manhattan', 'n_neighbors': 17, 'weights':
'distance'}
0.593 (0.011) with: {'metric': 'manhattan', 'n_neighbors': 18, 'weights':
'uniform'}
0.587 (0.021) with: {'metric': 'manhattan', 'n_neighbors': 18, 'weights':
'distance'}
0.590 (0.016) with: {'metric': 'manhattan', 'n_neighbors': 19, 'weights':
```

```
'uniform'}
0.592 (0.022) with: {'metric': 'manhattan', 'n_neighbors': 19, 'weights':
'distance'}
0.597 (0.008) with: {'metric': 'manhattan', 'n_neighbors': 20, 'weights':
'uniform'}
0.589 (0.021) with: {'metric': 'manhattan', 'n_neighbors': 20, 'weights':
'distance'}
0.592 (0.012) with: {'metric': 'manhattan', 'n_neighbors': 21, 'weights':
'uniform'}
0.592 (0.020) with: {'metric': 'manhattan', 'n_neighbors': 21, 'weights':
'distance'}
0.603 (0.013) with: {'metric': 'manhattan', 'n_neighbors': 22, 'weights':
'uniform'}
0.588 (0.022) with: {'metric': 'manhattan', 'n_neighbors': 22, 'weights':
'distance'}
0.595 (0.015) with: {'metric': 'manhattan', 'n_neighbors': 23, 'weights':
'uniform'}
0.592 (0.019) with: {'metric': 'manhattan', 'n_neighbors': 23, 'weights':
'distance'}
0.599 (0.015) with: {'metric': 'manhattan', 'n_neighbors': 24, 'weights':
'uniform'}
0.594 (0.018) with: {'metric': 'manhattan', 'n_neighbors': 24, 'weights':
'distance'}
0.592 (0.013) with: {'metric': 'manhattan', 'n_neighbors': 25, 'weights':
'uniform'}
0.596 (0.019) with: {'metric': 'manhattan', 'n_neighbors': 25, 'weights':
'distance'}
0.595 (0.017) with: {'metric': 'manhattan', 'n_neighbors': 26, 'weights':
'uniform'}
0.593 (0.016) with: {'metric': 'manhattan', 'n_neighbors': 26, 'weights':
'distance'}
0.593 (0.015) with: {'metric': 'manhattan', 'n_neighbors': 27, 'weights':
'uniform'}
0.598 (0.021) with: {'metric': 'manhattan', 'n_neighbors': 27, 'weights':
'distance'}
0.598 (0.016) with: {'metric': 'manhattan', 'n_neighbors': 28, 'weights':
'uniform'}
0.597 (0.017) with: {'metric': 'manhattan', 'n_neighbors': 28, 'weights':
'distance'}
0.592 (0.013) with: {'metric': 'manhattan', 'n_neighbors': 29, 'weights':
'uniform'}
0.599 (0.017) with: {'metric': 'manhattan', 'n_neighbors': 29, 'weights':
'distance'}
0.597 (0.011) with: {'metric': 'manhattan', 'n_neighbors': 30, 'weights':
'uniform'}
0.595 (0.016) with: {'metric': 'manhattan', 'n_neighbors': 30, 'weights':
'distance'}
0.536 (0.029) with: {'metric': 'minkowski', 'n_neighbors': 1, 'weights':
```

```
'uniform'}
0.536 (0.029) with: {'metric': 'minkowski', 'n_neighbors': 1, 'weights':
'distance'}
0.579 (0.020) with: {'metric': 'minkowski', 'n_neighbors': 2, 'weights':
'uniform'}
0.536 (0.029) with: {'metric': 'minkowski', 'n_neighbors': 2, 'weights':
'distance'}
0.542 (0.023) with: {'metric': 'minkowski', 'n_neighbors': 3, 'weights':
'uniform'}
0.542 (0.024) with: {'metric': 'minkowski', 'n_neighbors': 3, 'weights':
'distance'}
0.574 (0.017) with: {'metric': 'minkowski', 'n_neighbors': 4, 'weights':
'uniform'}
0.542 (0.016) with: {'metric': 'minkowski', 'n_neighbors': 4, 'weights':
'distance'}
0.545 (0.020) with: {'metric': 'minkowski', 'n_neighbors': 5, 'weights':
'uniform'}
0.544 (0.020) with: {'metric': 'minkowski', 'n_neighbors': 5, 'weights':
'distance'}
0.579 (0.020) with: {'metric': 'minkowski', 'n_neighbors': 6, 'weights':
'uniform'}
0.556 (0.025) with: {'metric': 'minkowski', 'n_neighbors': 6, 'weights':
'distance'}
0.561 (0.022) with: {'metric': 'minkowski', 'n_neighbors': 7, 'weights':
'uniform'}
0.564 (0.018) with: {'metric': 'minkowski', 'n_neighbors': 7, 'weights':
'distance'}
0.580 (0.028) with: {'metric': 'minkowski', 'n_neighbors': 8, 'weights':
'uniform'}
0.564 (0.027) with: {'metric': 'minkowski', 'n_neighbors': 8, 'weights':
'distance'}
0.560 (0.024) with: {'metric': 'minkowski', 'n_neighbors': 9, 'weights':
'uniform'}
0.569 (0.026) with: {'metric': 'minkowski', 'n_neighbors': 9, 'weights':
'distance'}
0.585 (0.020) with: {'metric': 'minkowski', 'n_neighbors': 10, 'weights':
'uniform'}
0.566 (0.027) with: {'metric': 'minkowski', 'n_neighbors': 10, 'weights':
'distance'}
0.557 (0.021) with: {'metric': 'minkowski', 'n_neighbors': 11, 'weights':
'uniform'}
0.566 (0.025) with: {'metric': 'minkowski', 'n_neighbors': 11, 'weights':
'distance'}
0.584 (0.015) with: {'metric': 'minkowski', 'n_neighbors': 12, 'weights':
'uniform'}
0.563 (0.025) with: {'metric': 'minkowski', 'n_neighbors': 12, 'weights':
'distance'}
0.565 (0.019) with: {'metric': 'minkowski', 'n_neighbors': 13, 'weights':
```

```
'uniform'}
0.561 (0.027) with: {'metric': 'minkowski', 'n_neighbors': 13, 'weights':
'distance'}
0.588 (0.010) with: {'metric': 'minkowski', 'n_neighbors': 14, 'weights':
'uniform'}
0.569 (0.024) with: {'metric': 'minkowski', 'n_neighbors': 14, 'weights':
'distance'}
0.581 (0.013) with: {'metric': 'minkowski', 'n_neighbors': 15, 'weights':
'uniform'}
0.574 (0.025) with: {'metric': 'minkowski', 'n_neighbors': 15, 'weights':
'distance'}
0.590 (0.011) with: {'metric': 'minkowski', 'n_neighbors': 16, 'weights':
'uniform'}
0.575 (0.024) with: {'metric': 'minkowski', 'n_neighbors': 16, 'weights':
'distance'}
0.578 (0.015) with: {'metric': 'minkowski', 'n_neighbors': 17, 'weights':
'uniform'}
0.572 (0.025) with: {'metric': 'minkowski', 'n_neighbors': 17, 'weights':
'distance'}
0.590 (0.012) with: {'metric': 'minkowski', 'n_neighbors': 18, 'weights':
'uniform'}
0.580 (0.023) with: {'metric': 'minkowski', 'n_neighbors': 18, 'weights':
'distance'}
0.586 (0.012) with: {'metric': 'minkowski', 'n_neighbors': 19, 'weights':
'uniform'}
0.585 (0.025) with: {'metric': 'minkowski', 'n_neighbors': 19, 'weights':
'distance'}
0.595 (0.013) with: {'metric': 'minkowski', 'n_neighbors': 20, 'weights':
'uniform'}
0.581 (0.020) with: {'metric': 'minkowski', 'n_neighbors': 20, 'weights':
'distance'}
0.589 (0.014) with: {'metric': 'minkowski', 'n_neighbors': 21, 'weights':
'uniform'}
0.585 (0.019) with: {'metric': 'minkowski', 'n_neighbors': 21, 'weights':
'distance'}
0.598 (0.015) with: {'metric': 'minkowski', 'n_neighbors': 22, 'weights':
'uniform'}
0.586 (0.021) with: {'metric': 'minkowski', 'n_neighbors': 22, 'weights':
'distance'}
0.592 (0.014) with: {'metric': 'minkowski', 'n_neighbors': 23, 'weights':
'uniform'}
0.587 (0.016) with: {'metric': 'minkowski', 'n_neighbors': 23, 'weights':
'distance'}
0.598 (0.016) with: {'metric': 'minkowski', 'n_neighbors': 24, 'weights':
'uniform'}
0.587 (0.016) with: {'metric': 'minkowski', 'n_neighbors': 24, 'weights':
'distance'}
0.590 (0.016) with: {'metric': 'minkowski', 'n_neighbors': 25, 'weights':
```

```
'uniform'}
0.587 (0.018) with: {'metric': 'minkowski', 'n_neighbors': 25, 'weights':
'distance'}
0.600 (0.015) with: {'metric': 'minkowski', 'n_neighbors': 26, 'weights':
'uniform'}
0.584 (0.018) with: {'metric': 'minkowski', 'n_neighbors': 26, 'weights':
'distance'}
0.590 (0.013) with: {'metric': 'minkowski', 'n_neighbors': 27, 'weights':
'uniform'}
0.588 (0.019) with: {'metric': 'minkowski', 'n_neighbors': 27, 'weights':
'distance'}
0.595 (0.015) with: {'metric': 'minkowski', 'n_neighbors': 28, 'weights':
'uniform'}
0.583 (0.019) with: {'metric': 'minkowski', 'n_neighbors': 28, 'weights':
'distance'}
0.589 (0.018) with: {'metric': 'minkowski', 'n_neighbors': 29, 'weights':
'uniform'}
0.590 (0.015) with: {'metric': 'minkowski', 'n_neighbors': 29, 'weights':
'distance'}
0.590 (0.015) with: {'metric': 'minkowski', 'n_neighbors': 30, 'weights':
'uniform'}
0.588 (0.020) with: {'metric': 'minkowski', 'n_neighbors': 30, 'weights':
'distance'}
```

[97]:
```
from sklearn.metrics import  make_scorer
from sklearn.model_selection import cross_val_score

def classification_report_with_accuracy_score(Y_test, y_pred2):
    print (classification_report(Y_test, y_pred2)) # print classification report
    return accuracy_score(Y_test, y_pred2) # return accuracy score


nested_score = cross_val_score(grid_search_knn, X=X_train, y=Y_train, cv=cv,
               scoring=make_scorer(classification_report_with_accuracy_score))
print (nested_score)
```

```
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-97-de1b807e3f87>", line 5, in
```

```
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-97-de1b807e3f87>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-97-de1b807e3f87>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
```

```
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-97-de1b807e3f87>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-97-de1b807e3f87>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-97-de1b807e3f87>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
```

```
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-97-de1b807e3f87>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-97-de1b807e3f87>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-97-de1b807e3f87>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(

[nan nan nan nan nan nan nan nan nan nan]

/usr/local/lib/python3.10/dist-
```

```
packages/sklearn/model_selection/_validation.py:794: UserWarning: Scoring
failed. The score on this train-test partition for these parameters will be set
to nan. Details:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 115, in __call__
    score = scorer._score(cached_call, estimator, *args, **kwargs)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_scorer.py",
line 282, in _score
    return self._sign * self._score_func(y_true, y_pred, **self._kwargs)
  File "<ipython-input-97-de1b807e3f87>", line 5, in
classification_report_with_accuracy_score
    print (classification_report(Y_test, y_pred2)) # print classification report
NameError: name 'classification_report' is not defined

  warnings.warn(
```

[98]:
```
knn_y_predicted = grid_search_knn.predict(X_test)
```

[99]:
```
knn_y_predicted
```

[99]:
```
array([0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1,
       0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0,
       0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0,
       0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0,
```

```
0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0,
0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0])
```

[100]:
```
knn_grid_score=accuracy_score(Y_test, knn_y_predicted)
```

[101]:
```
knn_grid_score
```

[101]:
```
0.6173780487804879
```

[102]:
```
grid_search_knn.best_params_
```

[102]:
```
{'metric': 'manhattan', 'n_neighbors': 22, 'weights': 'uniform'}
```

[103]:
```
confusion_matrix(Y_test, knn_y_predicted)
```

[103]:
```
array([[376,  26],
       [225,  29]])
```

[ ]:
```

```

11 Prediction on only one set of data

[104]:
```
X_KNN=knn.predict([[5.735724, 158.318741,25363.016594,7.728601,377.543291,568.
↪304671,13.626624,75.952337,4.732954]])
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does
not have valid feature names, but KNeighborsClassifier was fitted with feature
names
  warnings.warn(
```

[105]:
```
X_KNN
```

[105]:
```
array([0])
```

[ ]:
```

```