

## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



Notes



Help

### Jupyter Lab

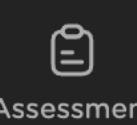
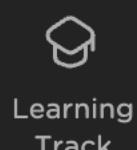
This Lab will get reset on 29th December 2022, 11:45 AM

jupyter

Used 10.4 of 50 hours in Dec, 2022

▶ Start Lab

■ End Lab



File Edit View Run Kernel Tabs Settings Help

project.ipynb

[1]: `import pandas as pd`  
`cep = pd.read_excel("cep1.xlsx")`  
`cep`

[1]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
298	57	0	0	140	241	0	1	123	1	0.2	1	0	3	0
299	45	1	3	110	264	0	1	132	0	1.2	1	0	3	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	3	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	3	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	2	0

303 rows × 14 columns

## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



### Jupyter Lab

This Lab will get reset on 29th December 2022, 11:45 AM

jupyter

Used 10.4 of 50 hours in Dec, 2022

▶ Start Lab

■ End Lab



The screenshot shows the Jupyter Lab interface. On the left is a sidebar with icons for Learning Track, Live Classes, Assessment, and Certificate. The main area has a file browser on the left and a code editor on the right. The code editor shows two cells:

```
[2]: cep.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 303 entries, 0 to 302
Data columns (total 14 columns):
 #   Column   Non-Null Count   Dtype  
--- 
 0   age       303 non-null    int64  
 1   sex       303 non-null    int64  
 2   cp        303 non-null    int64  
 3   trestbps  303 non-null    int64  
 4   chol      303 non-null    int64  
 5   fbs       303 non-null    int64  
 6   restecg   303 non-null    int64  
 7   thalach   303 non-null    int64  
 8   exang     303 non-null    int64  
 9   oldpeak   303 non-null    float64 
 10  slope     303 non-null    int64  
 11  ca        303 non-null    int64  
 12  thal      303 non-null    int64  
 13  target    303 non-null    int64  
dtypes: float64(1), int64(13)
memory usage: 33.3 KB

[3]: cep.describe()
```

The status bar at the bottom indicates "Simple" mode, "Python 3 | Idle", "Mode: Edit", "Ln 1, Col 1", and the file name "project.ipynb".



## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



Notes



Help

### Jupyter Lab

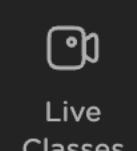
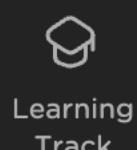
This Lab will get reset on 29th December 2022, 11:45 AM

jupyter

Used 10.4 of 50 hours in Dec, 2022

Start Lab

End Lab



File Edit View Run Kernel Tabs Settings Help

project.ipynb

[3]: cep.describe()

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slop
count	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000
mean	54.366337	0.683168	0.966997	131.623762	246.264026	0.148515	0.528053	149.646865	0.326733	1.039604	1.39934
std	9.082101	0.466011	1.032052	17.538143	51.830751	0.356198	0.525860	22.905161	0.469794	1.161075	0.61622
min	29.000000	0.000000	0.000000	94.000000	126.000000	0.000000	0.000000	71.000000	0.000000	0.000000	0.000000
25%	47.500000	0.000000	0.000000	120.000000	211.000000	0.000000	0.000000	133.500000	0.000000	0.000000	1.00000
50%	55.000000	1.000000	1.000000	130.000000	240.000000	0.000000	1.000000	153.000000	0.000000	0.800000	1.00000
75%	61.000000	1.000000	2.000000	140.000000	274.500000	0.000000	1.000000	166.000000	1.000000	1.600000	2.00000
max	77.000000	1.000000	3.000000	200.000000	564.000000	1.000000	2.000000	202.000000	1.000000	6.200000	2.00000

[4]: import statsmodels.formula.api as smf

[5]: log\_reg = smf.logit("target ~ age + sex + cp + trestbps + chol + fbs + restecg + thalach + exang + oldpeak + slope + slos")

Optimization terminated successfully.  
Current function value: 0.348904  
Iterations 7

[6]: print(log\_reg.summary())

## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



Notes



Help

### Jupyter Lab

This Lab will get reset on 29th December 2022, 11:45 AM

jupyter

Used 10.4 of 50 hours in Dec, 2022

▶ Start Lab

■ End Lab



File Edit View Run Kernel Tabs Settings Help

project.ipynb

[5]: `log_reg = smf.logit("target ~ age + sex + cp + trestbps + chol + fbs + restecg + thalach + exang + oldpeak + slope +`

Optimization terminated successfully.  
Current function value: 0.348904  
Iterations 7

[6]: `print(log_reg.summary())`

Logit Regression Results

Dep. Variable:	target	No. Observations:	303
Model:	Logit	Df Residuals:	289
Method:	MLE	Df Model:	13
Date:	Wed, 14 Dec 2022	Pseudo R-squ.:	0.4937
Time:	11:27:47	Log-Likelihood:	-105.72
converged:	True	LL-Null:	-208.82
Covariance Type:	nonrobust	LLR p-value:	7.262e-37

	coef	std err	z	P> z	[0.025	0.975]
Intercept	3.4505	2.571	1.342	0.180	-1.590	8.490
age	-0.0049	0.023	-0.212	0.832	-0.050	0.041
sex	-1.7582	0.469	-3.751	0.000	-2.677	-0.839
cp	0.8599	0.185	4.638	0.000	0.496	1.223
trestbps	-0.0195	0.010	-1.884	0.060	-0.040	0.001
chol	-0.0046	0.004	-1.224	0.221	-0.012	0.003
fbs	0.0349	0.529	0.066	0.947	-1.003	1.073
restecg	0.1662	0.242	1.220	0.101	0.216	1.140

## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



### Jupyter Lab

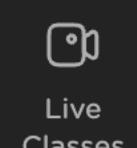
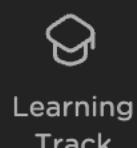
This Lab will get reset on 29th December 2022, 11:45 AM

jupyter

Used 10.4 of 50 hours in Dec, 2022

▶ Start Lab

■ End Lab



File Edit View Run Kernel Tabs Settings Help

project.ipynb

[8]:

```
import numpy as np
odds_ratios = pd.DataFrame(
    {
        "OR": log_reg.params,
        "Lower CI": log_reg.conf_int()[0],
        "Upper CI": log_reg.conf_int()[1],
    }
)
odds_ratios = np.exp(odds_ratios)
print(odds_ratios)
```

	OR	Lower CI	Upper CI
Intercept	31.515277	0.204020	4868.205549
age	0.995104	0.950914	1.041346
sex	0.172358	0.068772	0.431970
cp	2.362808	1.642927	3.398121
trestbps	0.980712	0.961039	1.000787
chol	0.995380	0.988029	1.002787
fbs	1.035503	0.366834	2.923029
restecg	1.594057	0.805483	3.154650
thalach	1.023482	1.002713	1.044682
exang	0.375318	0.168109	0.837932
oldpeak	0.582589	0.383119	0.885912
slope	1.784767	0.899137	3.542726
ca	0.461465	0.317437	0.670842
thal	0.406394	0.230152	0.717595

## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



### Jupyter Lab

This Lab will get reset on 29th December 2022, 11:45 AM

jupyter

Used 10.4 of 50 hours in Dec, 2022

Start Lab

End Lab



File Edit View Run Kernel Tabs Settings Help

project.ipynb

[10]: import seaborn as sns  
sns.pairplot(cep)

[10]: <seaborn.axisgrid.PairGrid at 0x7f109d11c0d0>

Simple 0 \$ 1 Python 3 | Idle Mode: Edit Ln 1, Col 1 project.ipynb



## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



Notes



Help

### Jupyter Lab

This Lab will get reset on 29th December 2022, 11:45 AM

jupyter ●

Used 10.4 of 50 hours in Dec, 2022

▶ Start Lab

■ End Lab



The screenshot shows a Jupyter Notebook interface. On the left, there's a sidebar with icons for Learning Track, Live Classes, Assessment, and Certificate. The main area has a file browser on the left showing files like datasets, boston.csv, cep1.xlsx, project.ipynb, seed.csv, seeds.csv, Untitled.ipynb, Untitled1.ipynb, Untitled2.ipynb, and varietiesne... The notebook itself has two cells. Cell [15] contains:

```
y_pred=logreg.predict(x_test)
y_pred
```

Cell [15] outputs:

```
array([1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0,
       1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1,
       1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1,
       1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1,
       1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1,
       1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1,
       1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1, 0, 1, 1,
```

Cell [16] contains:

```
y_pred=logreg.predict_proba(x_test)
y_pred
```

Cell [16] outputs:

```
array([[1.71096108e-02, 9.82890389e-01],
       [9.32881253e-01, 6.71187470e-02],
       [9.80779586e-01, 1.92204142e-02],
       [9.92705108e-01, 7.29489192e-03],
       [7.90593197e-02, 9.20940680e-01],
       [9.89432909e-01, 1.05670907e-02],
       [6.72549265e-02, 9.32745074e-01],
       [6.38502884e-01, 3.61497116e-01],
       [2.10315881e-02, 9.78968412e-01],
       [4.05232046e-01, 5.94767954e-01],
       [9.12260038e-01, 8.77399618e-02],
       [7.87928986e-01, 2.12071014e-01],
       [9.98360816e-01, 1.63918382e-03],
```

## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



### Jupyter Lab

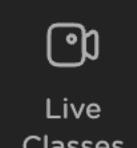
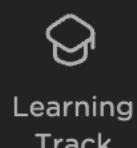
This Lab will get reset on 29th December 2022, 11:45 AM

jupyter

Used 10.4 of 50 hours in Dec, 2022

▶ Start Lab

■ End Lab



The screenshot shows a Jupyter Notebook interface with a sidebar on the left containing icons for Learning Track, Live Classes, Assessment, and Certificate. The main area has tabs for 'File', 'Edit', 'View', 'Run', 'Kernel', 'Tabs', 'Settings', and 'Help'. A file tree on the left shows files like 'datasets', 'boston.csv', 'cep1.xlsx', 'project.ipynb' (which is selected), 'seed.csv', 'seeds.csv', and several unnamed files. The notebook cell [17] contains Python code for calculating confusion matrix and classification report:

```
[17]: from sklearn.metrics import confusion_matrix, classification_report
print(confusion_matrix(y_test, y_pred))
print(classification_report(y_test, y_pred))
```

The output shows the confusion matrix and a detailed classification report:

	precision	recall	f1-score	support
0	0.80	0.72	0.76	72
1	0.77	0.84	0.80	80
accuracy			0.78	152
macro avg	0.79	0.78	0.78	152
weighted avg	0.78	0.78	0.78	152

## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



Notes



Help

### Jupyter Lab

This Lab will get reset on 29th December 2022, 11:45 AM

jupyter

Used 10.4 of 50 hours in Dec, 2022

▶ Start Lab

■ End Lab



Learning Track

Live Classes

Assessment

Certificate

File Edit View Run Kernel Tabs Settings Help

project.ipynb

[19]:

```
from sklearn.metrics import roc_curve
import matplotlib.pyplot as plt
from sklearn.metrics import roc_auc_score
y_pred_proba = logreg.predict_proba(x_test)[:,1]
fpr, tpr, thresh = roc_curve(y_test, y_pred_proba)
auc = roc_auc_score(y_test, y_pred_proba)
plt.plot(fpr,tpr,label="AUC="+str(auc))
plt.ylabel('True Positive Rate')
plt.xlabel('False Positive Rate')
plt.legend(loc=4)
plt.show()
```

The figure shows an ROC curve with the x-axis labeled 'False Positive Rate' and the y-axis labeled 'True Positive Rate'. The curve starts at (0,0) and ends at (1,1). It is a step function with approximately 10 steps. A legend in the bottom right corner indicates 'AUC=0.8803819444444444'.

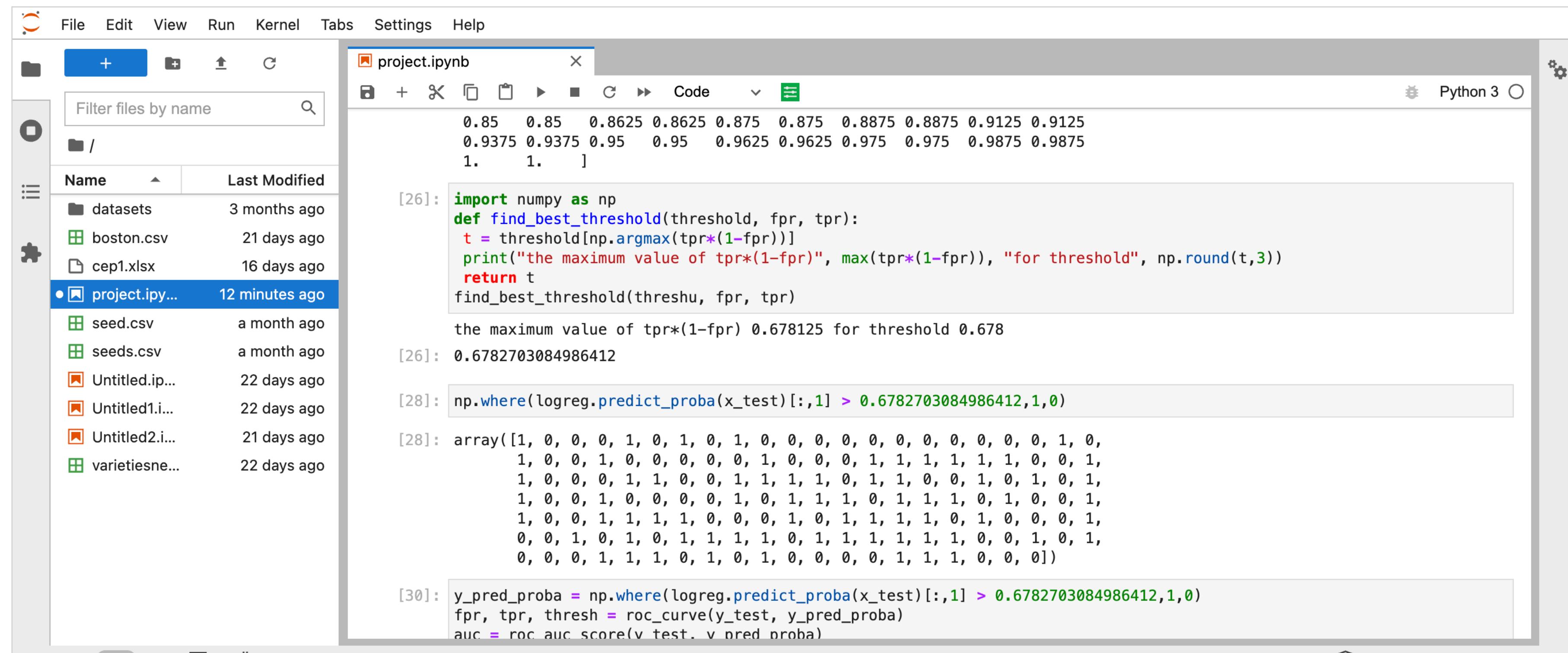
Jupyter Lab

This Lab will get reset on 29th December 2022, 11:45 AM

jupyter

Used 10.4 of 50 hours in Dec. 2022

► Start Lab ■ End Lab ✎



## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



### Jupyter Lab

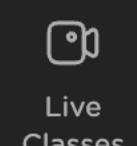
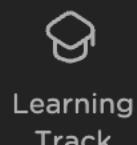
This Lab will get reset on 29th December 2022, 11:45 AM

jupyter

Used 10.4 of 50 hours in Dec, 2022

▶ Start Lab

■ End Lab



File Edit View Run Kernel Tabs Settings Help

project.ipynb

[30]:

```
y_pred_proba = np.where(logreg.predict_proba(x_test)[:,1] > 0.6782703084986412,1,0)
fpr, tpr, thresh = roc_curve(y_test, y_pred_proba)
auc = roc_auc_score(y_test, y_pred_proba)
plt.plot(fpr,tpr,label="AUC="+str(auc))
plt.ylabel('True Positive Rate')
plt.xlabel('False Positive Rate')
plt.legend(loc=4)
plt.show()
```

True Positive Rate

False Positive Rate

AUC=0.8180555555555555



## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



### Jupyter Lab

This Lab will get reset on 29th December 2022, 11:45 AM

jupyter

Used 10.4 of 50 hours in Dec, 2022

▶ Start Lab

■ End Lab



File Edit View Run Kernel Tabs Settings Help

project.ipynb

Code Python 3

True Positive Rate

AUC=0.8180555555555555

False Positive Rate

[31]:

```
from sklearn.metrics import classification_report, confusion_matrix
print(confusion_matrix(y_test, y_pred_proba))
print(classification_report(y_test, y_pred_proba))

[[62 10]
 [18 62]]

precision    recall   f1-score   support
      0       0.78      0.86      0.82       72
      1       0.86      0.78      0.82       80

accuracy          0.82      0.82      0.82      152
macro avg       0.82      0.82      0.82      152
weighted avg    0.82      0.82      0.82      152
```

Simple 0 \$ 1 Python 3 | Idle Mode: Edit Ln 1, Col 1 project.ipynb

## PC AIML - Machine Learning

Class completed: 1 | 89% of Self-Learning Videos Watched | Projects completed: 0/2



This Lab will get reset on 29th December 2022, 11:45 AM

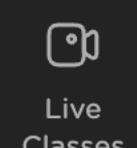
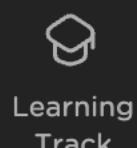
### Jupyter Lab

jupyter

Used 10.4 of 50 hours in Dec, 2022

▶ Start Lab

■ End Lab



File Edit View Run Kernel Tabs Settings Help

project.ipynb x

weighted avg 0.82 0.82 0.82 152

Python 3

[33]: `from sklearn.ensemble import RandomForestClassifier  
classifier_rf = RandomForestClassifier(random_state=42, max_depth=5,n_estimators=100,oob_score=True)  
classifier_rf.fit(x_train, y_train)  
RandomForestClassifier(max_depth=5, oob_score=True, random_state=42)  
classifier_rf.oob_score_`

[33]: `0.8675496688741722`

[34]: `from sklearn.metrics import classification_report  
from sklearn.metrics import confusion_matrix  
print(confusion_matrix(y_test, y_pred))  
print(classification_report(y_test, y_pred))`

[[52 20]  
[13 67]]

	precision	recall	f1-score	support
0	0.80	0.72	0.76	72
1	0.77	0.84	0.80	80
accuracy			0.78	152
macro avg	0.79	0.78	0.78	152
weighted avg	0.78	0.78	0.78	152