```
In [ ]:
```

# In [25]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plot
import seaborn as sns
from sklearn.preprocessing import StandardScaler

import warnings
warnings.filterwarnings('ignore')
```

# In [ ]:

# In [26]:

cardioDataset = pd.read\_csv('heart.csv')

### In [27]:

cardioDataset.head(30)

# Out[27]:

	age	sex	ср	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	8.0	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
5	57	1	0	140	192	0	1	148	0	0.4	1	0	1	1
6	56	0	1	140	294	0	0	153	0	1.3	1	0	2	1
7	44	1	1	120	263	0	1	173	0	0.0	2	0	3	1
8	52	1	2	172	199	1	1	162	0	0.5	2	0	3	1
9	57	1	2	150	168	0	1	174	0	1.6	2	0	2	1
10	54	1	0	140	239	0	1	160	0	1.2	2	0	2	1
11	48	0	2	130	275	0	1	139	0	0.2	2	0	2	1
12	49	1	1	130	266	0	1	171	0	0.6	2	0	2	1
13	64	1	3	110	211	0	0	144	1	1.8	1	0	2	1
14	58	0	3	150	283	1	0	162	0	1.0	2	0	2	1
15	50	0	2	120	219	0	1	158	0	1.6	1	0	2	1
16	58	0	2	120	340	0	1	172	0	0.0	2	0	2	1
17	66	0	3	150	226	0	1	114	0	2.6	0	0	2	1
18	43	1	0	150	247	0	1	171	0	1.5	2	0	2	1
19	69	0	3	140	239	0	1	151	0	1.8	2	2	2	1
20	59	1	0	135	234	0	1	161	0	0.5	1	0	3	1

```
fbs
                                                           oldpeak
                          chol
                                   restecg
                                            thalach
                                                                              thal
                                                                                   target
                                                    exang
    42
              0
                     140
                          226
                                 0
                                         1
                                                178
                                                         0
                                                                0.0
                                                                        2 0
                                                                                 2
                                                                                        1
22
          1
23
    61
          1
              2
                     150
                          243
                                 1
                                         1
                                                137
                                                         1
                                                                1.0
                                                                        1
                                                                            0
                                                                                 2
                                                                                        1
                                 0
                                          1
                                                                                        1
24
    40
          1
              3
                     140
                          199
                                                178
                                                         1
                                                                1.4
                                                                        2
                                                                            0
                                                                                 3
25
    71
          0
              1
                     160
                          302
                                 0
                                          1
                                                162
                                                         0
                                                                0.4
                                                                        2
                                                                            2
                                                                                 2
                                                                                        1
              2
                     150
                                                157
                                                                        2 0
                                                                                 2
26
    59
                          212
                                 1
                                                         0
                                                                1.6
    51
              2
                     110 175
                                 0
                                         1
                                                123
                                                         0
                                                                0.6
                                                                        2
                                                                                 2
                                                                                        1
27
          1
                                                                           0
              2
28
    65
          0
                     140
                          417
                                         0
                                                157
                                                                8.0
                                                                        2
                                                                                 2
              2
                     130 197
                                         0
                                                152
                                                                1.2
                                                                        0
                                                                          0
                                                                                 2
29
    53
                                 1
                                                         0
                                                                                        1
          1
```

### In [28]:

cardioDataset.shape

### Out[28]:

(303, 14)

### In [29]:

cardioDataset.describe()

### Out[29]:

	age	sex	ср	trestbps	chol	fbs	restecg	thalach	exang	oldpe
count	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.0000
mean	54.366337	0.683168	0.966997	131.623762	246.264026	0.148515	0.528053	149.646865	0.326733	1.0396
std	9.082101	0.466011	1.032052	17.538143	51.830751	0.356198	0.525860	22.905161	0.469794	1.1610
min	29.000000	0.000000	0.000000	94.000000	126.000000	0.000000	0.000000	71.000000	0.000000	0.0000
25%	47.500000	0.000000	0.000000	120.000000	211.000000	0.000000	0.000000	133.500000	0.000000	0.0000
50%	55.000000	1.000000	1.000000	130.000000	240.000000	0.000000	1.000000	153.000000	0.000000	0.8000
75%	61.000000	1.000000	2.000000	140.000000	274.500000	0.000000	1.000000	166.000000	1.000000	1.6000
max	77.000000	1.000000	3.000000	200.000000	564.000000	1.000000	2.000000	202.000000	1.000000	6.2000
4							1			<b>)</b>

# In [30]:

cardioDataset.isnull().sum()

# Out[30]:

age sex ср trestbps chol fbs restecg thalach exang oldpeak slope са thal target dtype: int64

## In [31]:

cardioDataset.columns = ['Age', 'Sex', 'ChestPainType', 'RestingBloodPressure', 'Cholest

#### In [32]:

cardioDataset.dtypes

### Out[32]:

Age int64 int64 Sex ChestPainType int64 RestingBloodPressure int64 Cholesterol int64 FastingBloodSugar int64 RestingECG int64 MaxHeartRate int64 ExerciseInducedAngina int64 PreviousPeak float64 Slope int64 MajorBloodVessels int64 ThalRate int64 ProbHA int64

dtype: object

### In [33]:

```
#Correlation matrix

plot.figure(figsize = (15, 10))

correlation = cardioDataset.corr()

mask = np.triu(np.ones_like(correlation, dtype=bool))

sns.heatmap(correlation, annot = True, cmap='BuGn', linewidths=1)
plot.show()
```

1.0

- 0.8

- 0.6

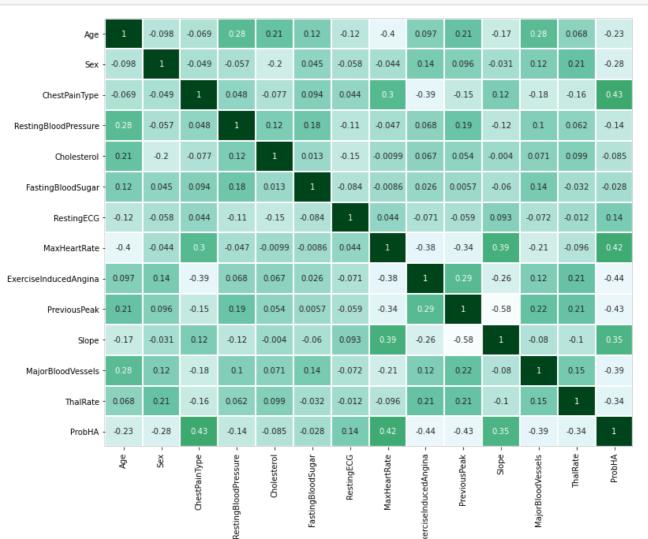
- 04

- 0.2

0.0

- -0.2

-0.4



### In [34]:

```
chart = plot.figure(figsize = (14, 14))
ax = chart.gca()
cardioDataset.hist(ax=ax, color="skyblue")
```

### Out[34]:

125

100

75 50

25

0.00

0.25

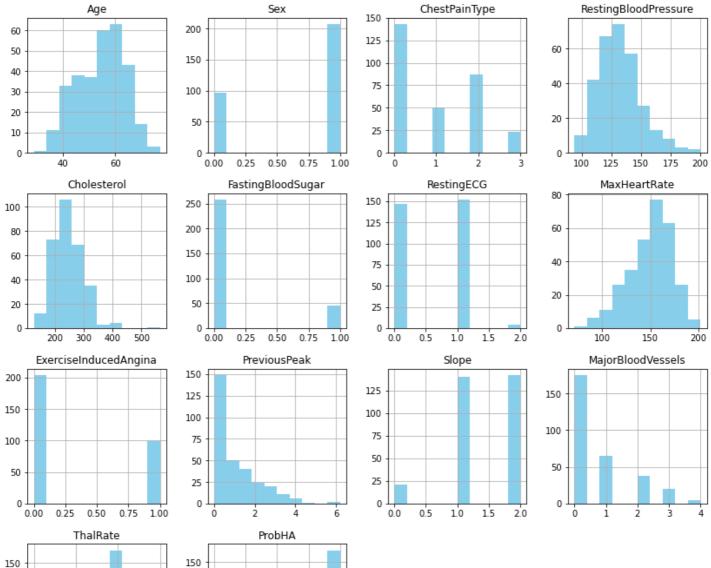
0.50

0.75

125

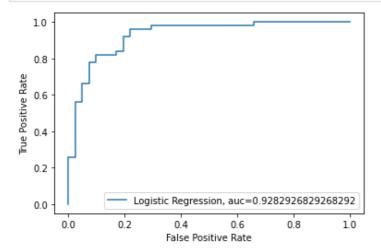
100 75

> 50 25



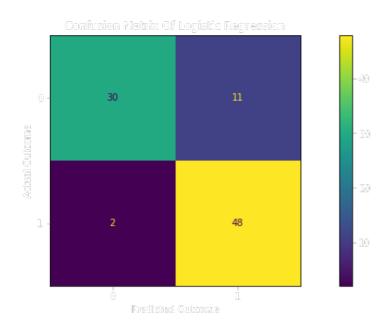
```
In [35]:
cardioDataset.var()
Out[35]:
                           82.484558
Age
                            0.217166
Sex
ChestPainType
                            1.065132
RestingBloodPressure
                          307.586453
                         2686.426748
Cholesterol
FastingBloodSugar
                            0.126877
RestingECG
                            0.276528
                          524.646406
MaxHeartRate
ExerciseInducedAngina
                          0.220707
PreviousPeak
                            1.348095
                            0.379735
Slope
MajorBloodVessels
                            1.045724
ThalRate
                            0.374883
ProbHA
                            0.248836
dtype: float64
In [36]:
cardioDataset['RestingBloodPressure']=np.log(cardioDataset['RestingBloodPressure'])
cardioDataset["Cholesterol"] = np.log(cardioDataset["Cholesterol"])
cardioDataset["MaxHeartRate"] = np.log(cardioDataset["MaxHeartRate"])
np.var(cardioDataset[["RestingBloodPressure", "Cholesterol", "MaxHeartRate"]])
Out[36]:
RestingBloodPressure
                        0.016894
Cholesterol
                        0.041401
MaxHeartRate
                        0.027054
dtype: float64
In [37]:
x = cardioDataset.drop('ProbHA', axis=1)
y = cardioDataset["ProbHA"]
In [38]:
#splitting the dataset
from sklearn.model selection import train test split
x train, x test, y train, y test = train test split(x,y,test size=0.30, random state=43)
In [39]:
#Logistic Regression
accuracies Of Algorithms= {}
from sklearn.linear model import LogisticRegression
from sklearn.metrics import accuracy score, confusion matrix, classification report, f1 s
core, roc auc score, roc curve, plot confusion matrix, precision recall curve, plot precisi
on recall curve
def logisticModel():
    logisticRegression = LogisticRegression(penalty='12')
    logisticRegression.fit(x train, y train)
    y pred = logisticRegression.predict(x test)
    y predProbability = logisticRegression.predict proba(x test)[::,1]
    falsePositiveRate, truePositiveRate, = roc curve(y test, y predProbability)
    auc = roc_auc_score(y_test, y_predProbability)
    plot.plot(falsePositiveRate, truePositiveRate, label="Logistic Regression, auc="+str
(auc))
    plot.legend(loc=4)
    plot.ylabel("True Positive Rate")
```

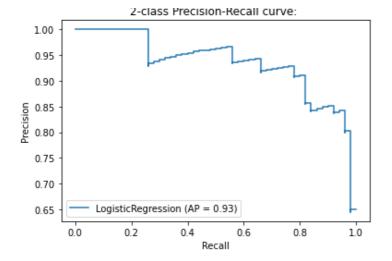
```
plot.xlabel("False Positive Rate")
    plot.show()
    accuracy = accuracy_score(y_test, y_pred)
    accuracies Of Algorithms["Logistic Regression"] = accuracy*100
   print('Accuracy score of Logistic Regression is:' , accuracy score(y test, y pred)*1
00, "%")
   print("Confusion Matrix of Logistic Regression", confusion matrix(y test, y pred))
   print("ClassificationReport", classification report(y test, y pred))
   plottedMatrix = plot confusion matrix(logisticRegression, x test, y test)
   plottedMatrix.ax .set title('Confusion Matrix Of Logistic Regression', color='white')
    plot.xlabel("Predicted Outcome", color='white')
   plot.ylabel("Actual Outcome", color='white')
    plot.gcf().axes[1].tick params(colors='white')
    plot.gcf().axes[0].tick params(colors='white')
    plot.gcf().set size inches(15,5)
    plot.show()
   disp = plot precision recall curve(logisticRegression, x test, y test)
    disp.ax .set title('2-class Precision-Recall curve: ')
logisticModel()
```



Accuracy score of Logistic Regression is: 85.71428571428571 % Confusion Matrix of Logistic Regression [[30 11] [ 2 48]]

Classifica	tionRepo	rt		precision	recall	f1-score	support
	0	0.94	0.73	0.82	41		
	1	0.81	0.96	0.88	50		
accura	су			0.86	91		
macro a	vg	0.88	0.85	0.85	91		
weighted a	vg	0.87	0.86	0.85	91		



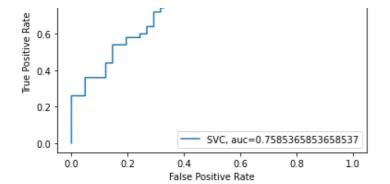


### In [40]:

1.0

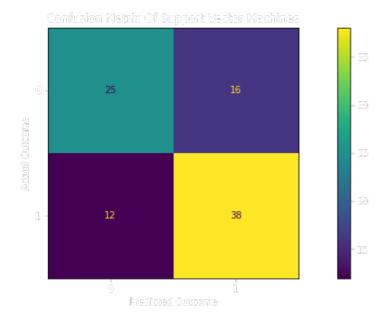
0.8

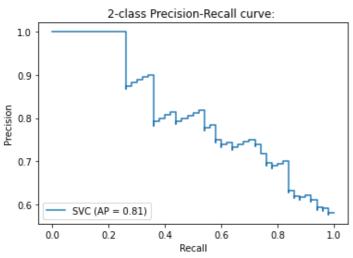
```
#support vector
from sklearn.svm import SVC
def svcClassifier():
    svc = SVC(probability=True)
    svc.fit(x_train, y_train)
    y pred1 = svc.predict(x test)
    y predProbability1 = svc.predict proba(x test)[::,1]
    falsePositiveRate, truePositiveRate, = roc curve(y test, y predProbability1)
    auc = roc auc score(y test, y predProbability1)
   plot.plot(falsePositiveRate, truePositiveRate, label="SVC, auc="+str(auc))
   plot.ylabel("True Positive Rate")
   plot.xlabel("False Positive Rate")
   plot.legend(loc=4)
   plot.show()
    accuracy1 = accuracy score(y test, y pred1)
    accuracies Of Algorithms['supportVectorMachine'] = accuracy1*100
    accuracy_score(y_train, svc.predict(x_train))
   print("Accuracy score of the model is:", accuracy score(y test, y pred1)*100, "%")
   print("Confusion matrix of the model", confusion_matrix(y_test, y_pred1))
    print("Classification Report", classification report(y test, y pred1))
    plottedMatrix = plot confusion matrix(svc, x test, y test)
   plottedMatrix.ax_.set_title('Confusion Matrix Of Support Vector Machines', color='whi
te')
   plot.xlabel("Predicted Outcome", color='white')
   plot.ylabel("Actual Outcome", color='white')
   plot.gcf().axes[1].tick_params(colors='white')
    plot.gcf().axes[0].tick params(colors='white')
    plot.gcf().set size inches(15,5)
    plot.show()
    disp = plot precision recall curve(svc, x test, y test)
    disp.ax_.set_title('2-class Precision-Recall curve: '
svcClassifier()
```



Accuracy score of the model is: 69.23076923076923 % Confusion matrix of the model [[25 16] [12 38]]

Classification R	eport		precision	recall	f1-score	support
0	0.68	0.61	0.64	41		
1	0.70	0.76	0.73	50		
accuracy			0.69	91		
macro avg	0.69	0.68	0.69	91		
weighted avg	0.69	0.69	0.69	91		

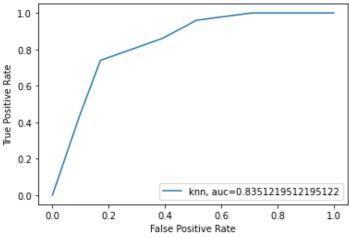




In [41]:

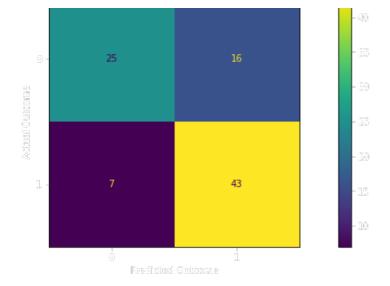
```
#kNearestNeighbours
from sklearn.neighbors import KNeighborsClassifier
def knnClassifier():
```

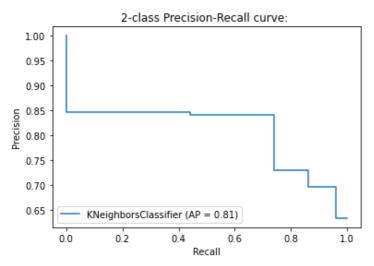
```
knn = KNeighborsClassifier()
   knn.fit(x train, y train)
   y pred2 = knn.predict(x test)
   y predProbability2 = knn.predict proba(x test)[::,1]
   falsePositiveRate, truePositiveRate, = roc curve(y_test, y_predProbability2)
   auc = roc_auc_score(y_test, y_predProbability2)
   plot.plot(falsePositiveRate, truePositiveRate, label="knn, auc="+str(auc))
   plot.legend(loc=4)
   plot.ylabel("True Positive Rate")
   plot.xlabel("False Positive Rate")
   plot.show()
   accuracy2 = accuracy_score(y_test, y_pred2)
   accuracies Of Algorithms['KNeighborsClassifier'] = accuracy2*100
   accuracy_score(y_train, knn.predict(x_train))
   print("Accuracy score of the KNN:", accuracy score(y test, y pred2)*100, "%")
   print("Confusion matrix of the model", confusion_matrix(y_test, y_pred2))
   print("Classification Report", classification report(y test, y pred2))
   plottedMatrix = plot confusion matrix(knn,x_test, y_test)
   plottedMatrix.ax .set title('Confusion Matrix Of knn', color='white')
   plot.xlabel("Predicted Outcome", color='white')
   plot.ylabel("Actual Outcome", color='white')
   plot.gcf().axes[1].tick params(colors='white')
   plot.gcf().axes[0].tick params(colors='white')
   plot.gcf().set size inches(15,5)
   plot.show()
   disp = plot_precision_recall_curve(knn, x_test, y_test)
   disp.ax .set title('2-class Precision-Recall curve: ')
knnClassifier()
```



Accuracy score of the KNN: 74.72527472527473 % Confusion matrix of the model [[25 16] [ 7 43]]

Classific	-	Report		precision	recall	f1-score	support
	0	0.78	0.61	0.68	41		
	1	0.73	0.86	0.79	50		
accui	racy			0.75	91		
macro	avg	0.76	0.73	0.74	91		
weighted	avg	0.75	0.75	0.74	91		



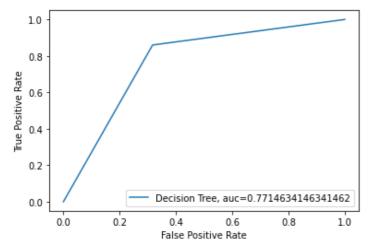


## In [42]:

```
from sklearn.tree import DecisionTreeClassifier
def decisionClassifier():
   dt = DecisionTreeClassifier()
   dt.fit(x train, y train)
   y pred3 = dt.predict(x test)
   y_predProbability3 = dt.predict_proba(x_test)[::,1]
   falsePositiveRate, truePositiveRate, = roc curve(y test, y predProbability3)
   auc = roc auc score(y test, y predProbability3)
   plot.plot(falsePositiveRate, truePositiveRate, label="Decision Tree, auc="+str(auc))
   plot.legend(loc=4)
   plot.ylabel("True Positive Rate")
   plot.xlabel("False Positive Rate")
   plot.show()
   accuracy3= accuracy_score(y_test, y_pred3)
   accuracies_Of_Algorithms['DecisionTreeClassifier'] = accuracy3*100
   accuracy score(y train, dt.predict(x train))
   print("Accuracy score of the model is:", accuracy score(y test,y pred3)*100, "%")
   print("Confusion matrix of the model", confusion matrix(y test, y pred3))
   print("Classification Report", classification report(y test, y pred3))
   plottedMatrix = plot confusion matrix(dt, x test, y test)
   plottedMatrix.ax .set title('Confusion Matrix Of Decision Tree', color='white')
   plot.xlabel("Predicted Outcome", color='white')
   plot.ylabel("Actual Outcome", color='white')
```

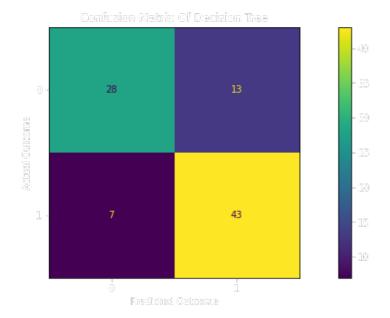
```
plot.gcf().axes[1].tick_params(colors='white')
plot.gcf().axes[0].tick_params(colors='white')
plot.gcf().set_size_inches(15,5)
disp = plot_precision_recall_curve(dt, x_test, y_test)
disp.ax_.set_title('2-class Precision-Recall curve: ')

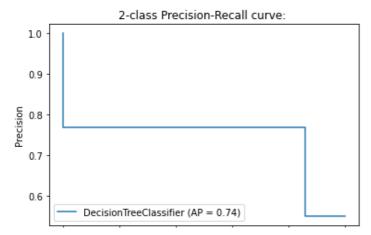
decisionClassifier()
```



Accuracy score of the model is: 78.02197802197803 % Confusion matrix of the model [[28 13] [ 7 43]]

Classification R	eport		precision	recall	f1-score	support
0	0.80	0.68	0.74	41		
1	0.77	0.86	0.81	50		
accuracy			0.78	91		
macro avg	0.78	0.77	0.77	91		
weighted avg	0.78	0.78	0.78	91		



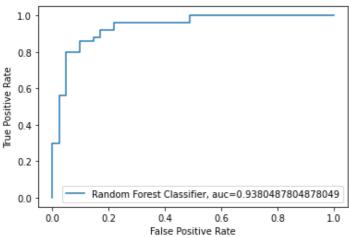


```
0.0 0.2 0.4 0.6 0.8 1.0
```

### In [43]:

```
from sklearn.ensemble import RandomForestClassifier
def randomForest():
   rfc = RandomForestClassifier(criterion = 'gini', max depth = 7, max features = 'sqr
t',
                           min samples leaf = 2, min samples split = 4, n estimators =
180)
   rfc.fit(x train, y train)
    y pred5 = rfc.predict(x test)
   y_predProbability5 = rfc.predict_proba(x_test)[::,1]
    falsePositiveRate, truePositiveRate, = roc curve(y test, y predProbability5)
    auc = roc_auc_score(y_test, y_predProbability5)
   plot.plot(falsePositiveRate, truePositiveRate, label="Random Forest Classifier, auc=
"+str(auc))
   plot.legend(loc=4)
   plot.ylabel("True Positive Rate")
   plot.xlabel("False Positive Rate")
   plot.show()
   accuracy5 = accuracy score(y test, y pred5)
   accuracies Of Algorithms['RandomForestClassifier'] = accuracy5*100
   accuracy score(y train, rfc.predict(x train))
   print("Accuracy score of the model is:", accuracy score(y test, y pred5)*100, "%")
   print("Confusion matrix of the model", confusion_matrix(y_test, y_pred5))
   print("Classification Report", classification_report(y_test, y_pred5))
   plottedMatrix = plot confusion matrix(rfc, x test, y test)
   plottedMatrix.ax_.set_title('Confusion Matrix Of Random Forest', color='white')
    plot.xlabel("Predicted Outcome", color='blue')
    plot.ylabel("Actual Outcome", color='white')
    plot.gcf().axes[1].tick params(colors='white')
   plot.gcf().axes[0].tick params(colors='white')
   plot.gcf().set_size_inches(15,5)
   plot.show()
    disp = plot precision recall curve(rfc, x test, y test)
    disp.ax .set title('2-class Precision-Recall curve: ')
randomForest()
```

sunnort

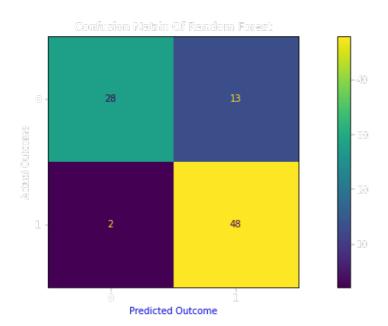


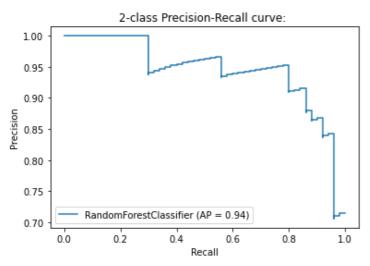
```
Accuracy score of the model is: 83.51648351648352 %

Confusion matrix of the model [[28 13]
  [2 48]]

Classification Report precision recall f1-score
```

O140011110461011 1	CPOT C		PTCCT0T011	100411	11 00010	Dabborc
0	0.93	0.68	0.79	41		
1	0.79	0.96	0.86	50		
accuracy			0.84	91		
macro avg	0.86	0.82	0.83	91		
weighted avg	0.85	0.84	0.83	91		





# In [44]:

```
from sklearn.ensemble import GradientBoostingClassifier

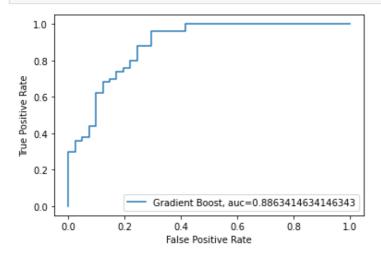
def GradientBoost():
    gbc = GradientBoostingClassifier()

    gbc.fit(x_train, y_train)
    y_pred6 = gbc.predict(x_test)

    y_predProbability6 = gbc.predict_proba(x_test)[::,1]
    falsePositiveRate, truePositiveRate, _ = roc_curve(y_test, y_predProbability6)
    auc = roc_auc_score(y_test, y_predProbability6)

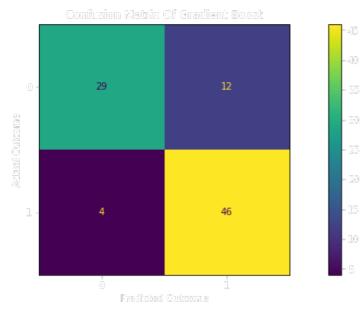
    plot.plot(falsePositiveRate, truePositiveRate, label="Gradient Boost, auc="+str(auc)")
    plot.legend(loc=4)
    plot.ylabel("True Positive Rate")
    plot.xlabel("False Positive Rate")
    plot.show()
```

```
accuracy6=accuracy_score(y_test, y_pred6)
    accuracies_Of_Algorithms['GradientBoosting'] = accuracy6*100
   print("Accuracy score of the model is:", accuracy_score(y_test, y_pred6)*100, "%")
   print("Confusion matrix of the model", confusion_matrix(y test, y pred6))
   print("Classification Report", classification report(y test, y pred6))
   plottedMatrix = plot confusion matrix(gbc, x test, y test)
   plottedMatrix.ax .set title('Confusion Matrix Of Gradient Boost', color='white')
   plot.xlabel("Predicted Outcome", color='white')
   plot.ylabel("Actual Outcome", color='white')
    plot.gcf().axes[1].tick params(colors='white')
    plot.gcf().axes[0].tick params(colors='white')
    plot.gcf().set size inches(15,5)
    plot.show()
   disp = plot precision recall curve(gbc, x test, y test)
    disp.ax .set title('2-class Precision-Recall curve: ')
GradientBoost()
```

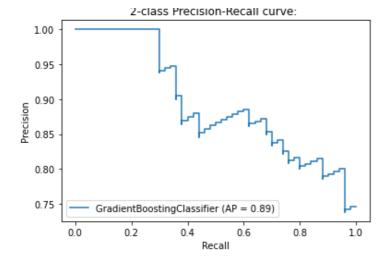


Accuracy score of the model is: 82.41758241758241 % Confusion matrix of the model [[29 12] [ 4 46]]

Classification	Report		precision	recall	f1-score	support
0	0.88	0.71	0.78	41		
1	0.79	0.92	0.85	50		
accuracy			0.82	91		
macro avg	0.84	0.81	0.82	91		
weighted avg	0.83	0.82	0.82	91		

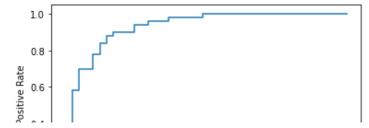


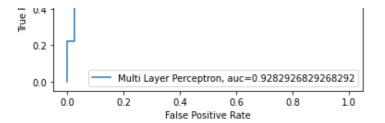
\_ . \_ . . \_ ..



### In [45]:

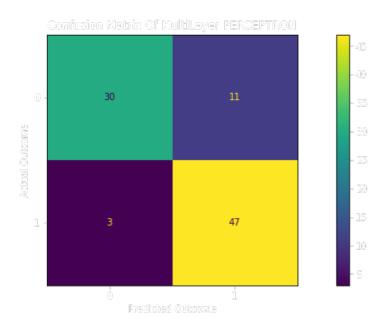
```
from sklearn.neural network import MLPClassifier
def multiLayerClassifier():
    multiLayerPerceptron = MLPClassifier(random state = 10, max iter=300, activation="relu
",
                                     hidden layer sizes=(34,34,34))
    multiLayerPerceptron.fit(x_train, y_train)
    y pred7 = multiLayerPerceptron.predict(x test)
    y predProbability7 = multiLayerPerceptron.predict proba(x test)[::,1]
    falsePositiveRate, truePositiveRate, = roc curve(y test, y predProbability7)
    auc = roc_auc_score(y_test, y_predProbability7)
   plot.plot(falsePositiveRate, truePositiveRate, label="Multi Layer Perceptron, auc="+
str(auc))
    plot.legend(loc=4)
    plot.ylabel("True Positive Rate")
    plot.xlabel("False Positive Rate")
    plot.show()
    accuracy7 = accuracy score(y test, y pred7)
    accuracies Of Algorithms['MLPClassifier'] = accuracy7*100
    print("Confusion Matrix of MLPClassifier", confusion matrix(y test, y pred7))
    print("Classification Report", classification report(y test, y pred7))
    plottedMatrix = plot confusion matrix(multiLayerPerceptron, x test, y test)
    plottedMatrix.ax .set title('Confusion Matrix Of MultiLayer PERCEPTRON', color='white
• )
    plot.xlabel("Predicted Outcome", color='white')
    plot.ylabel("Actual Outcome", color='white')
    plot.gcf().axes[1].tick params(colors='white')
    plot.gcf().axes[0].tick params(colors='white')
    plot.gcf().set_size_inches(15,5)
    plot.show()
    accuracy score(y test, y pred7)
    print("Accuracy score of the model is:", accuracy score(y test, y pred7)*100, "%")
    disp = plot precision recall curve(multiLayerPerceptron, x test, y test)
    disp.ax .set title('2-class Precision-Recall curve: ')
multiLayerClassifier()
```



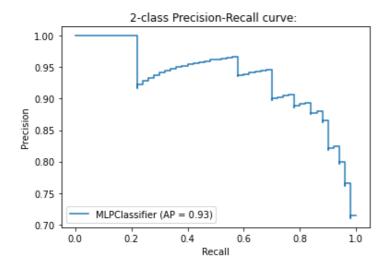


Confusion Matrix of MLPClassifier [[30 11]

Classification F	Report		precision	recall	f1-score	support
0	0.91	0.73	0.81	41		
1	0.81	0.94	0.87	50		
accuracy			0.85	91		
macro avg	0.86	0.84	0.84	91		
weighted avg	0.85	0.85	0.84	91		



Accuracy score of the model is: 84.61538461538461 %



### In [46]:

90

80

Accuracy Of Algorithms

