

```
In [18]: necessary_columns = ['Age', 'Pregnancies', 'Glucose', 'Blood_Pressure', 'Skin_Thickness', 'Insulin',
'BMI', 'Diabetes_Pedigree_Function']

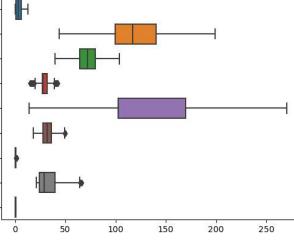
for x in necessary_columns:
    q75,q25 = np.percentile(df.loc[:,x],[75,25])
    intr_qr = q75-q25
    max1 = q75+(1.5*intr_qr)
    min1 = q25-(1.5*intr_qr)
    df.loc[(df[x] < min1) | (df[x] > max1),x] = np.Nan
    df.loc[(df[x] > max1),x] = np.Nan

In [19]: df.isnull().sum()
# Replace Nan values with median of that column and make new column as df_copy[feature_name].
# We saw on df.head() that some features contain 0, it doesn't make sense here and this indicates missing value.
# Give it value 1 is Nan is present else 0
for feature in necessary_columns:
    median_value = df[feature].median()
    df[feature].fillna(median_value, inplace = True)

df.isnull().sum()

Out[19]: Pregnancies      0
Glucose          0
Blood_Pressure   0
Skin_Thickness   0
Insulin          0
BMI              0
Diabetes_Pedigree_Function 0
Age              0
Outcome          0
dtype: int64

In [20]: sns.boxplot(data=df,orient="h")
<AxesSubplot>


```

```
In [21]: ***
for x in necessary_columns:
    q75,q25 = np.percentile(df.loc[:,x],[75,25])
    intr_qr = q75-q25
    max1 = q75+(1.5*intr_qr)
    min1 = q25-(1.5*intr_qr)
    df.loc[(df[x] < min1) | (df[x] > max1),x] = np.Nan
    df.loc[(df[x] > max1),x] = np.Nan

df.isnull().sum()
# Replace Nan values with median of that column and make new column as df_copy[feature_name].
# We saw on df.head() that some features contain 0, it doesn't make sense here and this indicates missing value.
# Below we replace 0 value by Nan.
# Give it value 1 is Nan is present else 0
for feature in necessary_columns:
    median_value = df[feature].median()
    df[feature].fillna(median_value, inplace = True)

df.isnull().sum()
***
```

```
Out[21]: "Infer" in necessary_columns: 0
q75,q25 = np.percentile(df.loc[:,x],[75,25])\n    intr_qr = q75-q25\n    \n    max1 = q75+(1.5*intr_qr)\n    min1 = q25-(1.5*intr_qr)\n    \n    df.loc[(df[x] < min1) | (df[x] > max1),x] = np.Nan\n    df.loc[(df[x] > max1),x] = np.Nan\n\n    df.isnull().sum()\n# Replace Nan Values with median of that column and make new column as df_copy[feature_name].\n# We saw on df.head() that some features contain 0, it doesn't make sense here and this indicates missing value.\n# Below we replace 0 value by Nan.\n# Give it value 1 is Nan is present else 0\nfor feature in necessary_columns:\n    median_value = df[feature].median()\n    df[feature].fillna(median_value, inplace = True)\n\ndf.isnull().sum()
```

```
In [22]: 
```

```
In [23]: 
```

```
In [24]: 
```

```
In [25]: 
```

```
In [26]: 
```

```
In [27]: 
```

```
In [28]: 
```

```
In [29]: 
```

```
In [30]: 
```

```
In [31]: 
```

```
In [32]: 
```

```
In [33]: 
```

```
In [34]: 
```

```
In [35]: 
```

```
In [36]: 
```

```
In [37]: 
```

```
In [38]: 
```

```
In [39]: 
```

```
In [40]: 
```

```
In [41]: 
```

```
In [42]: 
```

```
In [43]: 
```

```
In [44]: 
```

```
In [45]: 
```

```
In [46]: 
```

```
In [47]: 
```

```
In [48]: 
```

```
In [49]: 
```

```
In [50]: 
```

```
In [51]: 
```

```
In [52]: 
```

```
In [53]: 
```

```
In [54]: 
```

```
In [55]: 
```

```
In [56]: 
```

```
In [57]: 
```

```
In [58]: 
```

```
In [59]: 
```

```
In [60]: 
```

```
In [61]: 
```

```
In [62]: 
```

```
In [63]: 
```

```
In [64]: 
```

```
In [65]: 
```

```
In [66]: 
```

```
In [67]: 
```

```
In [68]: 
```

```
In [69]: 
```

```
In [70]: 
```

```
In [71]: 
```

```
In [72]: 
```

```
In [73]: 
```

```
In [74]: 
```

```
In [75]: 
```

```
In [76]: 
```

```
In [77]: 
```

```
In [78]: 
```

```
In [79]: 
```

```
In [80]: 
```

```
In [81]: 
```

```
In [82]: 
```

```
In [83]: 
```

```
In [84]: 
```

```
In [85]: 
```

```
In [86]: 
```

```
In [87]: 
```

```
In [88]: 
```

```
In [89]: 
```

```
In [90]: 
```

```
In [91]: 
```

```
In [92]: 
```

```
In [93]: 
```

```
In [94]: 
```

```
In [95]: 
```

```
In [96]: 
```

```
In [97]: 
```

```
In [98]: 
```

```
In [99]: 
```

```
In [100]: 
```

```
In [101]: 
```

```
In [102]: 
```

```
In [103]: 
```

```
In [104]: 
```

```
In [105]: 
```

```
In [106]: 
```

```
In [107]: 
```

```
In [108]: 
```

```
In [109]: 
```

```
In [110]: 
```

```
In [111]: 
```

```
In [112]: 
```

```
In [113]: 
```

```
In [114]: 
```

```
In [115]: 
```

```
In [116]: 
```

```
In [117]: 
```

```
In [118]: 
```

```
In [119]: 
```

```
In [120]: 
```

```
In [121]: 
```

```
In [122]: 
```

```
In [123]: 
```

```
In [124]: 
```

```
In [125]: 
```

```
In [126]: 
```

```
In [127]: 
```

```
In [128]: 
```

```
In [129]: 
```

```
In [130]: 
```

```
In [131]: 
```

```
In [132]: 
```

```
In [133]: 
```

```
In [134]: 
```

```
In [135]: 
```

```
In [136]: 
```

```
In [137]: 
```

```
In [138]: 
```

```
In [139]: 
```

```
In [140]: 
```

```
In [141]: 
```

```
In [142]: 
```

```
In [143]: 
```

```
In [144]: 
```

```
In [145]: 
```

```
In [146]: 
```

```
In [147]: 
```

```
In [148]: 
```

```
In [149]: 
```

```
In [150]: 
```

```
In [151]: 
```

```
In [152]: 
```

```
In [153]: 
```

```
In [154]: 
```

```
In [155]: 
```

```
In [156]: 
```

```
In [157]: 
```

```
In [158]: 
```

```
In [159]: 
```

```
In [160]: 
```

```
In [161]: 
```

```
In [162]: 
```

```
In [163]: 
```

```
In [164]: 
```

```
In [165]: 
```

```
In [166]: 
```

```
In [167]: 
```

```
In [168]: 
```

```
In [169]: 
```

```
In [170]: 
```

```
In [171]: 
```

```
In [172]: 
```

```
In [173]: 
```

```
In [174]: 
```

```
In [175]: 
```

```
In [176]: 
```

```
In [177]: 
```

```
In [178]: 
```

```
In [179]: 
```

```
In [180]: 
```

```
In [181]: 
```

```
In [182]: 
```

```
In [183]: 
```

```
In [184]: 
```

```
In [185]: 
```

```
In [186]: 
```

```
In [187]: 
```

```
In [188]: 
```

```
In [189]: 
```

```
In [190]: 
```

```
In [191]: 
```

```
In [192]: 
```

```
In [193]: 
```

```
In [194]: 
```

```
In [195]: 
```

```
In [196]: 
```

```
In [197]: 
```

```
In [198]: 
```

```
In [199]: 
```

```
In [200]: 
```

```

Train results: 0.895754723127835
Test results: 0.8958188861038961
Accuracy Score = 0.8958188861038961
[[98  9]
 [ 7 40]]
    precision  recall f1-score support
0       0.93     0.92     0.92    187
1       0.82     0.65     0.65     47

accuracy                           0.89
macro avg       0.87     0.88     0.88    154
weighted avg    0.88     0.88     0.88    154

```

```

In [284]: clf = AdaBoostClassifier()
parameters = {
    "n_estimators": range(1,100)
}
clf = GridSearchCV(ab_clf, parameters, cv=10)
clf.fit(X_train, y_train)

```

```

Out[284]: GridSearchCV
| estimator: AdaBoostClassifier
| | AdaBoostClassifier
| |
| |
| 
```

```

In [285]: # print best parameter after tuning
print(clf.best_params_)
# print how our model looks after hyper-parameter tuning
print(clf.best_estimator_)

['n_estimators': 77]
AdaBoostClassifier(n_estimators=77, random_state=42)

```

## Logitboost

```

In [286]: #pip install logitboost
Collecting logitboost
  Downloading logitboost-0.7-py3-none-any.whl (0.1 kB)
Requirement already satisfied: scikit-learn in c:\users\hp\anaconda3\lib\site-packages (from logitboost) (1.1.2)
Requirement already satisfied: scipy in c:\users\hp\anaconda3\lib\site-packages (from logitboost) (1.7.3)
Requirement already satisfied: numpy in c:\users\hp\anaconda3\lib\site-packages (from logitboost) (1.21.5)
Requirement already satisfied: joblib in c:\users\hp\anaconda3\lib\site-packages (from scikit-learn>logitboost) (1.1.0)
Requirement already satisfied: threadpoolctl>2.0.0 in c:\users\hp\anaconda3\lib\site-packages (from scikit-learn>logitboost) (2.2.0)
Installing collected packages: logitboost
Successfully installed logitboost-0.7

```

```

In [287]: from logitboost import LogitBoost
boost = LogitBoost(n_estimators=16, random_state=0)
boost.fit(X_train, y_train)

print("Train results:", boost.score(X_train, y_train))
print("Test results:", boost.score(X_test, y_test))

y_pred_ann = boost.predict(X_test)
print("Test Accuracy Score = ", metrics.accuracy_score(y_test, y_pred_ann))

# print classification metrics
print(metrics.classification_report(y_test, y_pred_ann, target_names=['Positive', 'Negative']))

# print confusion matrix
print("Confusion matrix:")
print(metrics.confusion_matrix(y_test, y_pred_ann))

Train results: 0.894136878175805
Test results: 0.9025974825974825
Test Accuracy Score = 0.9025974825974826
    precision  recall f1-score support
  Positive       0.95     0.91     0.93    187
  Negative       0.81     0.89     0.85    47

accuracy                           0.89
macro avg       0.88     0.90     0.89    154
weighted avg    0.89     0.90     0.89    154

```

```

Confusion matrix:
[[187  9]
 [ 9 42]]

```

```

In [288]: rate = []
xxx = range(10,18)
for i in xxx:
    boost = LogitBoost(n_estimators=i, random_state=0)
    boost.fit(X_train, y_train)

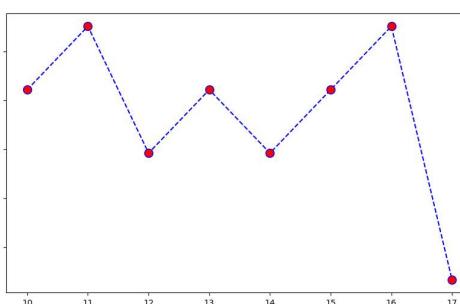
    y_pred = boost.predict(X_test)
    rate.append(metrics.accuracy_score(y_test, y_pred))

plt.figure(figsize=(10,6))
plt.plot(xxx, rate, color='blue', linestyle='dashed', marker='o', markerfacecolor='red', markersize=10)
...
...
```

```

Out[288]: '\n\n'

```



## Gradient Boosting

```

In [251]: from sklearn.ensemble import GradientBoostingClassifier
from sklearn.model_selection import GridSearchCV
from sklearn.metrics import accuracy_score
from sklearn.metrics import recall_score
from sklearn.metrics import f1_score
from sklearn.metrics import make_scorer

```

```

In [252]: clf = GradientBoostingClassifier(n_estimators=55, learning_rate=0.25, max_depth=1, loss='log_loss', random_state=42).fit(X_train, y_train)

y_pred = clf.predict(X_test)
print("Accuracy Score = ", format(metrics.accuracy_score(y_test, y_pred)))
print("Recall Score = ", format(metrics.recall_score(y_test, y_pred)))
print(classification_report(y_test, y_pred))

Accuracy Score = 0.915844155844156
[[181  6]
 [ 7 40]]
    precision  recall f1-score support
0       0.94     0.94     0.94    187
1       0.87     0.85     0.86     47

accuracy                           0.92
macro avg       0.90     0.90     0.90    154
weighted avg    0.92     0.92     0.92    154

```

```

In [253]: # get roc curve points
fpr, tpr, _ = metrics.roc_curve(y_test, y_pred)

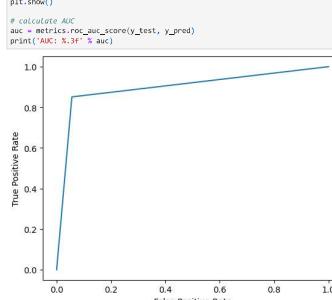
```

```

# create ROC curve
plt.plot(fpr,tpr)
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.show()

# calculate AUC
auc = metrics.roc_auc_score(y_test, y_pred)
print("AUC: %.3f" % auc)

```



```

AUC: 0.916

```

```

In [254]: rate = []
xxx = range(1,100)
for i in xxx:
    clf = GradientBoostingClassifier(n_estimators=i, learning_rate=0.25, max_depth=1, loss='log_loss', random_state=42).fit(X_train, y_train)

    y_pred = clf.predict(X_test)
    rate.append(metrics.accuracy_score(y_test, y_pred))

plt.figure(figsize=(10,6))
plt.plot(xxx, rate, color='blue', linestyle='dashed', marker='o', markerfacecolor='red', markersize=10)

```



































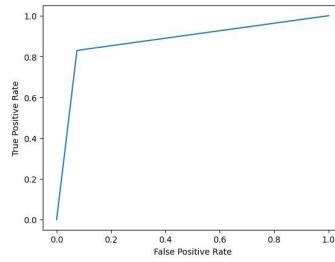
```
{'C': 0.001, 'max_iter': 100, 'penalty': 'l2'}  
Train accuracy: 0.8561038961038961  
Test results: 0.8561038961038961  
Test Accuracy Score = 0.8561038961038961
```

```
precision    recall   f1-score   support  
Positive      0.93      0.93      0.93     107  
Negative      0.83      0.83      0.83      47  
accuracy          0.90      154  
macro avg      0.88      0.88      0.88     154  
weighted avg     0.90      0.90      0.90     154
```

Confusion matrix:

```
[109  10]  
[ 8 39]
```

```
In [32]: # get roc curve points  
fpr, tpr = metrics.roc_curve(y_test, y_pred_ann)  
  
# create ROC curve  
plt.plot(fpr,tpr)  
plt.xlabel('True Positive Rate')  
plt.ylabel('False Positive Rate')  
plt.show()
```



## Support Vector Machine

```
In [33]: from sklearn.model_selection import GridSearchCV  
from sklearn.svm import SVC  
  
# defining parameter range  
param_grid = {'C': [1, 10, 50, 100, 1000, 10000],  
             'gamma': [1, 0.1, 0.01, 0.001, 0.0001, 0.00001, 0.000001],  
             'kernel': ['rbf', 'sigmoid']}
```

```
grid = GridSearchCV(SVC()), param_grid, refit = True, verbose = 3)
```

```
# fitting the model for grid search
```

```
grid.fit(X_train, y_train)
```

```
# print best parameter after tuning
```

```
print(grid.best_params_)
```

```
# print how our model looks after hyper-parameter tuning
```

```
print(grid.best_estimator_)  
  
svc_pred = grid.predict(X_test)  
  
print("Train results:", grid.score(X_train, y_train))  
print("Test results:", grid.score(X_test, y_test))  
print("Test Accuracy score = ", format(metrics.accuracy_score(y_test, svc_pred)))  
  
print(confusion_matrix(y_test, svc_pred))  
print(classification_report(y_test, svc_pred))
```

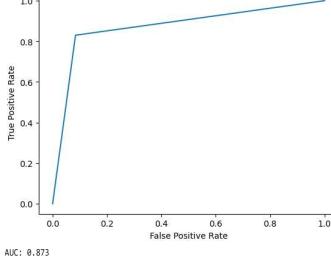




```
[34]: from sklearn import metrics
# get roc curve points
fpr, tpr, _ = metrics.roc_curve(y_test, svc_pred)

# create roc curve
plt.plot(fpr,tpr)
plt.xlabel('True Positive Rate')
plt.xlabel('False Positive Rate')
plt.show()

# calculate AUC
auc = metrics.roc_auc_score(y_test, svc_pred)
print("AUC: %.3f" % auc)
```



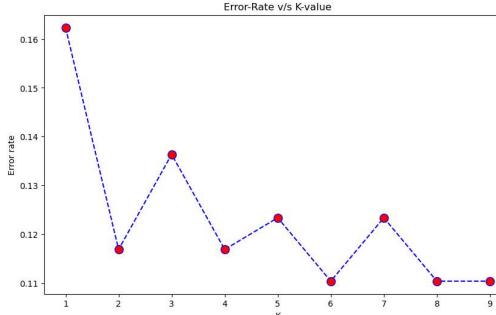
KNN

```

from sklearn.neighbors import KNeighborsClassifier
error_rate = []
for n_neighbors in range(1,10):
    knn = KNeighborsClassifier(n_neighbors = n_neighbors)
    knn.fit(X_train,y_train.ravel())
    pred_i = knn.predict(X_test)
    error_rate.append(np.mean(pred_i != y_test))

plt.figure(figsize=(10,6))
plt.plot(range(1,10), error_rate, color='blue', linestyle='dashed', marker='o', markerfacecolor='red', markersize=10)
plt.title('Error rate vs k-Value')
plt.xlabel('k')
plt.ylabel('Error rate')
Text(0, 6.5, 'Error rate')


```



```
[152]: knn = KNeighborsClassifier(n_neighbors = 8)
knn.fit(X_train,y_train.ravel())
y_pred = knn.predict(X_test)

print("Train results:", knn.score(X_train, y_train))
print("Test results:", knn.score(X_test, y_test))
print("Test Accuracy Score = ", format(metrics.accuracy_score(y_test, y_pred)))

print(confusion_matrix(y_test, y_pred))
print(accuracy_score(y_test, y_pred))
```

```

from sklearn import metrics
# get roc curve points
fpr, tpr, _ = metrics.roc_curve(y_test, y_pred)

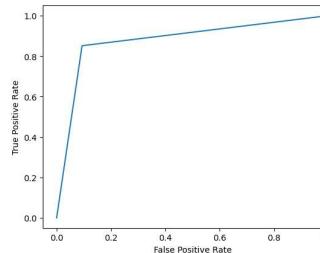
# create ROC plot
plt.plot(fpr, tpr)
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('ROC Curve')
plt.show()

# calculate AUC
auc = metrics.roc_auc_score(y_test, y_pred)
print("AUC: %.3f" % auc)

Train results: 0.877851268564405
Test results: 0.8896103896103896
Confusion Matrix: 0.8896103896103896
[19 18]
[[ 7 48]
 [48 154]]
precision recall f1-score support
0 0.93 0.91 0.92 157
1 0.88 0.85 0.86 47

accuracy: 0.905
macro avg 0.87 0.88 0.88 154
weighted avg 0.89 0.89 0.89 154

```



## Decision Tree

```
In [37]: from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import GridSearchCV
classifier = DecisionTreeClassifier()

from scipy.stats import randint
from sklearn.model_selection import RandomizedSearchCV

parameters = {"max_depth": [1, 2, 3, 5, 10, 15, 20, None],
              "max_features": randint(1, 15),
              "min_samples_leaf": randint(1, 10),
              "criterion": ["gini", "entropy"]}

tree_cv = RandomizedSearchCV(classifier, parameters, cv=10)
tree_cv.fit(X_train, y_train)
tree_cv.best_estimator_
```

```
Out[37]: DecisionTreeClassifier
DecisionTreeClassifier(criterion='entropy', max_depth=20, max_features=7,
                      min_samples_leaf=3)
```

```
In [15]: from sklearn import tree
classifier = DecisionTreeClassifier(criterion='entropy', max_depth=20, max_features=6, min_samples_leaf=4)
```

```
classifier.fit(X_train, y_train)
y_pred = classifier.predict(X_test)
cm = confusion_matrix(y_test, y_pred)
print(cm)
print(accuracy_score(y_test, y_pred))
print(classification_report(y_test, y_pred))

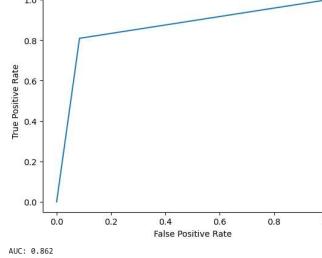
from sklearn import metrics
# get roc curve points
fpr, tpr, _ = metrics.roc_curve(y_test, y_pred)

# create ROC curve
plt.plot(fpr, tpr)
plt.xlabel('True Positive Rate')
plt.ylabel('False Positive Rate')
plt.show()

# calculate AUC
auc = metrics.roc_auc_score(y_test, y_pred)
print("AUC: %.3f" % auc)
```

```
[15]: [98 9]
[9 38]
0.8831168831168831
```

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0            | 0.92      | 0.92   | 0.92     | 107     |
| 1            | 0.81      | 0.81   | 0.81     | 47      |
| accuracy     | 0.86      | 0.86   | 0.86     | 154     |
| macro avg    | 0.88      | 0.88   | 0.88     | 154     |
| weighted avg | 0.88      | 0.88   | 0.88     | 154     |



## Random Forest

```
In [39]: ...
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import GridSearchCV
classifier = RandomForestClassifier()

parameters = {"max_depth": [1, 2, 3, 5, 10, 15, 20, None],
              "n_estimators": [10, 20, 30, 40, 50],
              "min_samples_leaf": randint(1, 10),
              "criterion": ["gini", "entropy"]}
```

```
grid_search = GridSearchCV(classifier=classifier, param_grid=parameters, cv=10, verbose=3)
grid_search.fit(X_train, y_train) # fitting the model for grid search
print(grid_search.best_params_) # print best parameter after tuning
```

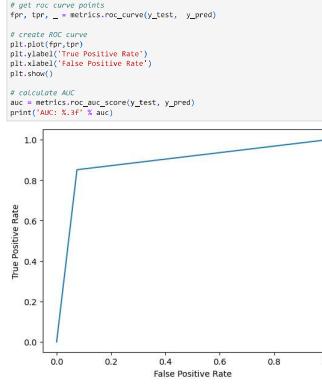
```
# print how our model looks after hyper-parameter tuning
print(grid_search.best_estimator_)
***
```

```
Out[39]: RandomForestClassifier(n_estimators=10, n_jobs=-1, oob_score=False, random_state=None, verbose=0)
In [40]: ...
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import GridSearchCV
parameters = {"max_depth": [1, 2, 3, 5, 10, 15, 20, None],
              "n_estimators": [10, 20, 30, 40, 50],
              "min_samples_leaf": range(1, 10),
              "max_features": range(1, 8),
              "min_samples_leaf": range(1, 10),
              "criterion": ["gini", "entropy"]}
```

```
grid_search = GridSearchCV(classifier=classifier, param_grid=parameters, cv=10, verbose=3)
grid_search.fit(X_train, y_train) # fitting the model for grid search
print(grid_search.best_params_) # print best parameter after tuning
print(grid_search.best_estimator_)
***
```

```
Out[40]: RandomForestClassifier(n_estimators=10, n_jobs=-1, oob_score=False, random_state=None, verbose=0)
```

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0            | 0.93      | 0.94   | 0.94     | 107     |
| 1            | 0.87      | 0.83   | 0.85     | 47      |
| accuracy     | 0.90      | 0.91   | 0.91     | 154     |
| macro avg    | 0.90      | 0.89   | 0.89     | 154     |
| weighted avg | 0.91      | 0.91   | 0.91     | 154     |



```
In [ ]:
```

## Neural Network

```
In [42]: import tensorflow as tf
from tensorflow import keras

class myCallback(tf.keras.callbacks.Callback):
    def on_epoch_end(self, epoch, logs={}):
        if(logs.get('accuracy')>0.99):
            print("\nReached 99% accuracy so cancelling training!")
            self.model.stop_training = True

In [43]: callbacks = myCallback()

In [44]: model = tf.keras.models.Sequential([tf.keras.layers.Flatten(input_shape=(8,)),
                                         tf.keras.layers.Dense(20, activation=tf.nn.relu),
                                         tf.keras.layers.Dense(11, activation=tf.nn.relu),
                                         tf.keras.layers.Dense(11, activation=tf.nn.relu),
                                         #tf.keras.layers.Dropout(0.5),
                                         tf.keras.layers.Dense(8, activation=tf.nn.relu),
                                         tf.keras.layers.Dense(1, activation=tf.nn.sigmoid)])

model.compile(optimizer = tf.optimizers.Adam(lr=0.001),
              loss='binary_crossentropy',
              metrics=['accuracy'])

model.fit(x_train, y_train, epochs=40,batch_size=30,callbacks=[callbacks],validation_data=(x_test, y_test))

C:\Users\shpananconda\lib\site-packages\keras\optimizer_v2\adam.py:185: UserWarning: The 'lr' argument is deprecated, use 'learning_rate' instead.
  super(Adam, self).__init__(name=namewargs)

Epoch 1/40
21/21 [=====] - 0s 66ms/step - loss: 3.6313 - accuracy: 0.4919 - val_loss: 1.4720 - val_accuracy: 0.4156
Epoch 2/40
21/21 [=====] - 0s 66ms/step - loss: 1.6697 - accuracy: 0.5683 - val_loss: 0.9270 - val_accuracy: 0.5714
Epoch 3/40
21/21 [=====] - 0s 66ms/step - loss: 0.7611 - accuracy: 0.6629 - val_loss: 0.7329 - val_accuracy: 0.7078
Epoch 4/40
21/21 [=====] - 0s 66ms/step - loss: 0.6889 - accuracy: 0.6922 - val_loss: 0.6515 - val_accuracy: 0.7273
Epoch 5/40
21/21 [=====] - 0s 66ms/step - loss: 0.6380 - accuracy: 0.7231 - val_loss: 0.5868 - val_accuracy: 0.7483
Epoch 6/40
21/21 [=====] - 0s 66ms/step - loss: 0.5831 - accuracy: 0.7598 - val_loss: 0.5395 - val_accuracy: 0.7727
Epoch 7/40
21/21 [=====] - 0s 66ms/step - loss: 0.5616 - accuracy: 0.7671 - val_loss: 0.5526 - val_accuracy: 0.7727
Epoch 8/40
21/21 [=====] - 0s 66ms/step - loss: 0.5516 - accuracy: 0.7708 - val_loss: 0.5211 - val_accuracy: 0.7922
Epoch 9/40
21/21 [=====] - 0s 66ms/step - loss: 0.5331 - accuracy: 0.7573 - val_loss: 0.5036 - val_accuracy: 0.7857
Epoch 10/40
21/21 [=====] - 0s 66ms/step - loss: 0.5313 - accuracy: 0.7598 - val_loss: 0.5176 - val_accuracy: 0.7727
Epoch 11/40
21/21 [=====] - 0s 66ms/step - loss: 0.5174 - accuracy: 0.7622 - val_loss: 0.5016 - val_accuracy: 0.7857
Epoch 12/40
21/21 [=====] - 0s 66ms/step - loss: 0.5182 - accuracy: 0.7769 - val_loss: 0.4843 - val_accuracy: 0.8117
Epoch 13/40
21/21 [=====] - 0s 66ms/step - loss: 0.5075 - accuracy: 0.7687 - val_loss: 0.4798 - val_accuracy: 0.7987
Epoch 14/40
21/21 [=====] - 0s 66ms/step - loss: 0.5100 - accuracy: 0.7655 - val_loss: 0.4818 - val_accuracy: 0.8052
Epoch 15/40
21/21 [=====] - 0s 66ms/step - loss: 0.5040 - accuracy: 0.7785 - val_loss: 0.5077 - val_accuracy: 0.7792
Epoch 16/40
21/21 [=====] - 0s 66ms/step - loss: 0.4974 - accuracy: 0.7834 - val_loss: 0.4668 - val_accuracy: 0.8052
Epoch 17/40
21/21 [=====] - 0s 66ms/step - loss: 0.4825 - accuracy: 0.7964 - val_loss: 0.4083 - val_accuracy: 0.8117
Epoch 18/40
21/21 [=====] - 0s 66ms/step - loss: 0.4800 - accuracy: 0.7997 - val_loss: 0.4684 - val_accuracy: 0.8112
Epoch 19/40
21/21 [=====] - 0s 66ms/step - loss: 0.4798 - accuracy: 0.7997 - val_loss: 0.4588 - val_accuracy: 0.8112
Epoch 20/40
21/21 [=====] - 0s 66ms/step - loss: 0.4762 - accuracy: 0.7997 - val_loss: 0.4517 - val_accuracy: 0.8092
Epoch 21/40
21/21 [=====] - 0s 66ms/step - loss: 0.4768 - accuracy: 0.7948 - val_loss: 0.4579 - val_accuracy: 0.8177
Epoch 22/40
21/21 [=====] - 0s 66ms/step - loss: 0.4931 - accuracy: 0.7858 - val_loss: 0.4654 - val_accuracy: 0.8312
Epoch 23/40
21/21 [=====] - 0s 66ms/step - loss: 0.4697 - accuracy: 0.7949 - val_loss: 0.4563 - val_accuracy: 0.8442
Epoch 24/40
21/21 [=====] - 0s 66ms/step - loss: 0.4708 - accuracy: 0.7997 - val_loss: 0.4588 - val_accuracy: 0.8312
Epoch 25/40
21/21 [=====] - 0s 66ms/step - loss: 0.4784 - accuracy: 0.7948 - val_loss: 0.5014 - val_accuracy: 0.7922
Epoch 26/40
21/21 [=====] - 0s 66ms/step - loss: 0.4823 - accuracy: 0.7752 - val_loss: 0.4572 - val_accuracy: 0.8312
Epoch 27/40
21/21 [=====] - 0s 66ms/step - loss: 0.4615 - accuracy: 0.7932 - val_loss: 0.4628 - val_accuracy: 0.8312
Epoch 28/40
21/21 [=====] - 0s 66ms/step - loss: 0.4529 - accuracy: 0.8094 - val_loss: 0.4487 - val_accuracy: 0.8377
Epoch 29/40
21/21 [=====] - 0s 66ms/step - loss: 0.4575 - accuracy: 0.8111 - val_loss: 0.4479 - val_accuracy: 0.8312
Epoch 30/40
21/21 [=====] - 0s 66ms/step - loss: 0.4557 - accuracy: 0.8086 - val_loss: 0.4519 - val_accuracy: 0.7979
Epoch 31/40
21/21 [=====] - 0s 66ms/step - loss: 0.4475 - accuracy: 0.8241 - val_loss: 0.4499 - val_accuracy: 0.8312
Epoch 32/40
21/21 [=====] - 0s 66ms/step - loss: 0.4492 - accuracy: 0.8046 - val_loss: 0.4483 - val_accuracy: 0.8312
Epoch 33/40
21/21 [=====] - 0s 66ms/step - loss: 0.4492 - accuracy: 0.8127 - val_loss: 0.4471 - val_accuracy: 0.8312
Epoch 34/40
21/21 [=====] - 0s 66ms/step - loss: 0.4465 - accuracy: 0.8168 - val_loss: 0.4410 - val_accuracy: 0.8306
Epoch 35/40
21/21 [=====] - 0s 66ms/step - loss: 0.4459 - accuracy: 0.8376 - val_loss: 0.4431 - val_accuracy: 0.8247
Epoch 36/40
21/21 [=====] - 0s 66ms/step - loss: 0.4458 - accuracy: 0.8298 - val_loss: 0.4386 - val_accuracy: 0.8312
Epoch 37/40
21/21 [=====] - 0s 66ms/step - loss: 0.4385 - accuracy: 0.8225 - val_loss: 0.4389 - val_accuracy: 0.8247
Epoch 38/40
21/21 [=====] - 0s 66ms/step - loss: 0.4374 - accuracy: 0.8274 - val_loss: 0.4493 - val_accuracy: 0.7922
Epoch 39/40
21/21 [=====] - 0s 66ms/step - loss: 0.4644 - accuracy: 0.7908 - val_loss: 0.4512 - val_accuracy: 0.8182
Epoch 40/40
21/21 [=====] - 0s 66ms/step - loss: 0.4353 - accuracy: 0.8225 - val_loss: 0.4275 - val_accuracy: 0.8571
<keras.callbacks.History at 0x2690151f0b>
```

```
Out[44]: In [45]: # Evaluate the model on the test data using 'evaluate'
print("Evaluate on test data")
results = model.evaluate(x_test, y_test, batch_size=30)
print("test loss, test acc:", results)

Evaluate on test data
6/6 [=====] - 0s 13ms/step - loss: 0.4317 - accuracy: 0.8312
test loss: 0.4317 test acc: 0.8312
```

```
In [45]: y_predan = model.predict(x_test)
y_predan = (y_predan > 0.5)
print("Confusion Matrix")
print(confusion_matrix(y_test, y_predan))
print("Accuracy = ", accuracy_score(y_test, y_predan))
print(classification_report(y_test,y_predan))

from sklearn import metrics
# get true positive rate
y_true = y_test
y_pred = y_predan
# Create ROC curve
fpr, tpr, _ = metrics.roc_curve(y_true, y_pred)
# Create AUC curve
auc = metrics.roc_auc_score(y_true, y_pred)
print("AUC: %.2f" % auc)
```

Confusion Matrix

[93 14]

[12 35]

Accuracy = 0.831688311688312

precision recall f1-score support

|   | 0    | 0.89 | 0.87 | 0.88 | 107 |
|---|------|------|------|------|-----|
| 1 | 0.71 | 0.74 | 0.73 | 47   |     |

accuracy

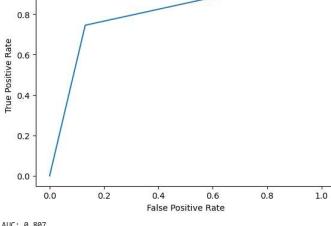
|  | 0.88 | 0.81 | 0.88 | 154 |
|--|------|------|------|-----|
|--|------|------|------|-----|

macro avg

|  | 0.83 | 0.83 | 0.83 | 154 |
|--|------|------|------|-----|
|--|------|------|------|-----|

weighted avg

|  | 0.83 | 0.83 | 0.83 | 154 |
|--|------|------|------|-----|
|--|------|------|------|-----|



## Forward Propagation NN

```
In [48]: import keras
from keras.models import Sequential
from keras.layers import Dense
from keras import optimizers
from keras import metrics

ann_class = Sequential()
ann_class.add(Dense(units=30, kernel_initializer="glorot_uniform", activation='relu', input_dim=8))
ann_class.add(Dense(units=10, kernel_initializer="glorot_uniform", activation='relu'))
ann_class.add(Dense(units=1, kernel_initializer="glorot_uniform", activation='sigmoid'))
ann_class.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
ann_class.compile(loss = 'binary_crossentropy', optimizer = 'adam', metrics = [metrics.accuracy])

ann_class.fit(x_train, y_train,batch_size=40, epochs=40)
```

```

print()
y_y_pred_ANN = ann_clsi.predict(x_test)
y_y_pred_ANN = (y_y_pred_ANN > 0.5)
print(confusion_matrix(y_test, y_y_pred_ANN))
print(accuracy_score(y_test, y_y_pred_ANN))
print(classification_report(y_test, y_y_pred_ANN))

[epoch 1/40]
[16/16] [=====] - 2s 11ms/step - loss: 2.8440 - accuracy: 0.0016
[epoch 2/40] [=====] - 0s 8ms/step - loss: 1.3985 - accuracy: 0.0000e+00
[epoch 3/40] [=====] - 0s 9ms/step - loss: 0.7559 - accuracy: 0.0000e+00
[epoch 4/40] [=====] - 0s 10ms/step - loss: 0.5976 - accuracy: 0.0000e+00
[epoch 5/40] [=====] - 0s 13ms/step - loss: 0.5763 - accuracy: 0.0000e+00
[epoch 6/40] [=====] - 0s 12ms/step - loss: 0.5599 - accuracy: 0.0000e+00
[epoch 7/40] [=====] - 0s 12ms/step - loss: 0.5599 - accuracy: 0.0000e+00
[epoch 8/40] [=====] - 0s 13ms/step - loss: 0.5479 - accuracy: 0.0000e+00
[epoch 9/40] [=====] - 0s 12ms/step - loss: 0.5384 - accuracy: 0.0000e+00
[epoch 10/40] [=====] - 0s 14ms/step - loss: 0.5325 - accuracy: 0.0000e+00
[epoch 11/40] [=====] - 0s 13ms/step - loss: 0.5225 - accuracy: 0.0000e+00
[epoch 12/40] [=====] - 0s 11ms/step - loss: 0.5171 - accuracy: 0.0000e+00
[epoch 13/40] [=====] - 0s 13ms/step - loss: 0.5094 - accuracy: 0.0000e+00
[epoch 14/40] [=====] - 0s 10ms/step - loss: 0.5031 - accuracy: 0.0000e+00
[epoch 15/40] [=====] - 0s 22ms/step - loss: 0.4999 - accuracy: 0.0000e+00
[epoch 16/40] [=====] - 0s 33ms/step - loss: 0.4926 - accuracy: 0.0000e+00
[epoch 17/40] [=====] - 0s 21ms/step - loss: 0.4890 - accuracy: 0.0000e+00
[epoch 18/40] [=====] - 0s 13ms/step - loss: 0.4844 - accuracy: 0.0000e+00
[epoch 19/40] [=====] - 0s 12ms/step - loss: 0.4828 - accuracy: 0.0000e+00
[epoch 20/40] [=====] - 0s 14ms/step - loss: 0.4784 - accuracy: 0.0000e+00
[epoch 21/40] [=====] - 0s 14ms/step - loss: 0.4776 - accuracy: 0.0000e+00
[epoch 22/40] [=====] - 0s 14ms/step - loss: 0.4739 - accuracy: 0.0000e+00
[epoch 23/40] [=====] - 0s 13ms/step - loss: 0.4747 - accuracy: 0.0000e+00
[epoch 24/40] [=====] - 0s 13ms/step - loss: 0.4711 - accuracy: 0.0000e+00
[epoch 25/40] [=====] - 0s 14ms/step - loss: 0.4690 - accuracy: 0.0000e+00
[epoch 26/40] [=====] - 0s 10ms/step - loss: 0.4707 - accuracy: 0.0000e+00
[epoch 27/40] [=====] - 0s 22ms/step - loss: 0.4724 - accuracy: 0.0000e+00
[epoch 28/40] [=====] - 0s 23ms/step - loss: 0.4784 - accuracy: 0.0000e+00
[epoch 29/40] [=====] - 0s 23ms/step - loss: 0.4653 - accuracy: 0.0000e+00
[epoch 30/40] [=====] - 0s 22ms/step - loss: 0.4626 - accuracy: 0.0000e+00
[epoch 31/40] [=====] - 0s 23ms/step - loss: 0.4638 - accuracy: 0.0000e+00
[epoch 32/40] [=====] - 0s 22ms/step - loss: 0.4559 - accuracy: 0.0000e+00
[epoch 33/40] [=====] - 0s 14ms/step - loss: 0.4593 - accuracy: 0.0000e+00
[epoch 34/40] [=====] - 0s 12ms/step - loss: 0.4623 - accuracy: 0.0000e+00
[epoch 35/40] [=====] - 0s 11ms/step - loss: 0.4596 - accuracy: 0.0000e+00
[epoch 36/40] [=====] - 0s 21ms/step - loss: 0.4401 - accuracy: 0.0000e+00
[epoch 37/40] [=====] - 0s 22ms/step - loss: 0.4681 - accuracy: 0.0000e+00
[epoch 38/40] [=====] - 0s 22ms/step - loss: 0.4674 - accuracy: 0.0000e+00
[epoch 39/40] [=====] - 0s 19ms/step - loss: 0.4571 - accuracy: 0.0000e+00
[epoch 40/40] [=====] - 0s 13ms/step - loss: 0.4552 - accuracy: 0.0000e+00
[epoch 41/40] [=====] - 0s 14ms/step - loss: 0.4539 - accuracy: 0.0000e+00

```

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0            | 0.98      | 0.87   | 0.89     | 107     |
| 1            | 0.73      | 0.79   | 0.76     | 47      |
| accuracy     |           |        |          |         |
| macro avg    | 0.81      | 0.83   | 0.82     | 154     |
| weighted avg | 0.85      | 0.84   | 0.85     | 154     |

```

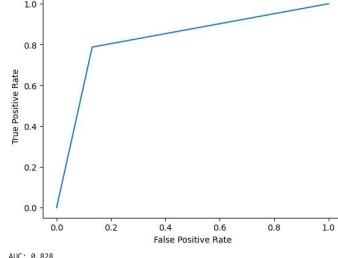
[[93 14]
 [18 37]]
0.8441558441558441

In [51]: #from sklearn import metrics
#fpr, tpr, _ = metrics.roc_curve(y_test, y_y_pred_ANN)

# create ROC curve
plt.plot(fpr,tpr)
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('ROC Curve')

# calculate AUC
auc = metrics.roc_auc_score(y_test, y_y_pred_ANN)
print("AUC: %f" % auc)

```



AUC: 0.828