LAB6 : Predictive Analytics for Hospitals

Dinesh Kumar K

225229108

Step 1 : Import Dataset

In [1]:
```
import pandas as pd
```

In [2]:
```
d=pd.read_csv('diabetes.csv')
d
```

Out[2]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFuncti
0	6	148	72	35	0	33.6	0.6
1	1	85	66	29	0	26.6	0.3
2	8	183	64	0	0	23.3	0.6
3	1	89	66	23	94	28.1	0.1
4	0	137	40	35	168	43.1	2.2
...	...	...	...	...	...	...	
763	10	101	76	48	180	32.9	0.1
764	2	122	70	27	0	36.8	0.3
765	5	121	72	23	112	26.2	0.2
766	1	126	60	0	0	30.1	0.3
767	1	93	70	31	0	30.4	0.3

768 rows × 9 columns

In [3]:
```
d.head
```

Out[3]:
```
<bound method NDFrame.head of      Pregnancies  Glucose  BloodPressure  Skin
Thickness  Insulin   BMI  \
0              6      148             72             35        0  33.6
1              1       85             66             29        0  26.6
2              8      183             64              0        0  23.3
3              1       89             66             23       94  28.1
4              0      137             40             35      168  43.1
..           ...      ...            ...            ...      ...   ...
763           10      101             76             48      180  32.9
764            2      122             70             27        0  36.8
765            5      121             72             23      112  26.2
766            1      126             60              0        0  30.1
767            1       93             70             31        0  30.4

     DiabetesPedigreeFunction  Age  Outcome
0                       0.627   50        1
1                       0.351   31        0
2                       0.672   32        1
3                       0.167   21        0
4                       2.288   33        1
..                        ...  ...      ...
763                     0.171   63        0
764                     0.340   27        0
765                     0.245   30        0
766                     0.349   47        1
767                     0.315   23        0

[768 rows x 9 columns]>
```

In [4]:
```
d.shape
```

Out[4]:
```
(768, 9)
```

In [5]:
```
d.columns
```

Out[5]:
```
Index(['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insuli
n',
       'BMI', 'DiabetesPedigreeFunction', 'Age', 'Outcome'],
      dtype='object')
```

In [6]:
```
d.dtypes
```

Out[6]:
```
Pregnancies                 int64
Glucose                     int64
BloodPressure               int64
SkinThickness               int64
Insulin                     int64
BMI                       float64
DiabetesPedigreeFunction  float64
Age                         int64
Outcome                     int64
dtype: object
```

In [7]:
```
d.info
```

Out[7]:
```
<bound method DataFrame.info of      Pregnancies  Glucose  BloodPressure  Sk
inThickness  Insulin   BMI  \
0              6      148             72             35        0  33.6
1              1       85             66             29        0  26.6
2              8      183             64              0        0  23.3
3              1       89             66             23       94  28.1
4              0      137             40             35      168  43.1
..           ...      ...            ...            ...      ...   ...
763           10      101             76             48      180  32.9
764            2      122             70             27        0  36.8
765            5      121             72             23      112  26.2
766            1      126             60              0        0  30.1
767            1       93             70             31        0  30.4

     DiabetesPedigreeFunction  Age  Outcome
0                       0.627   50        1
1                       0.351   31        0
2                       0.672   32        1
3                       0.167   21        0
4                       2.288   33        1
..                        ...  ...      ...
763                     0.171   63        0
764                     0.340   27        0
765                     0.245   30        0
766                     0.349   47        1
767                     0.315   23        0

[768 rows x 9 columns]>
```

In [8]:
```
d.Pregnancies.value_counts()
```

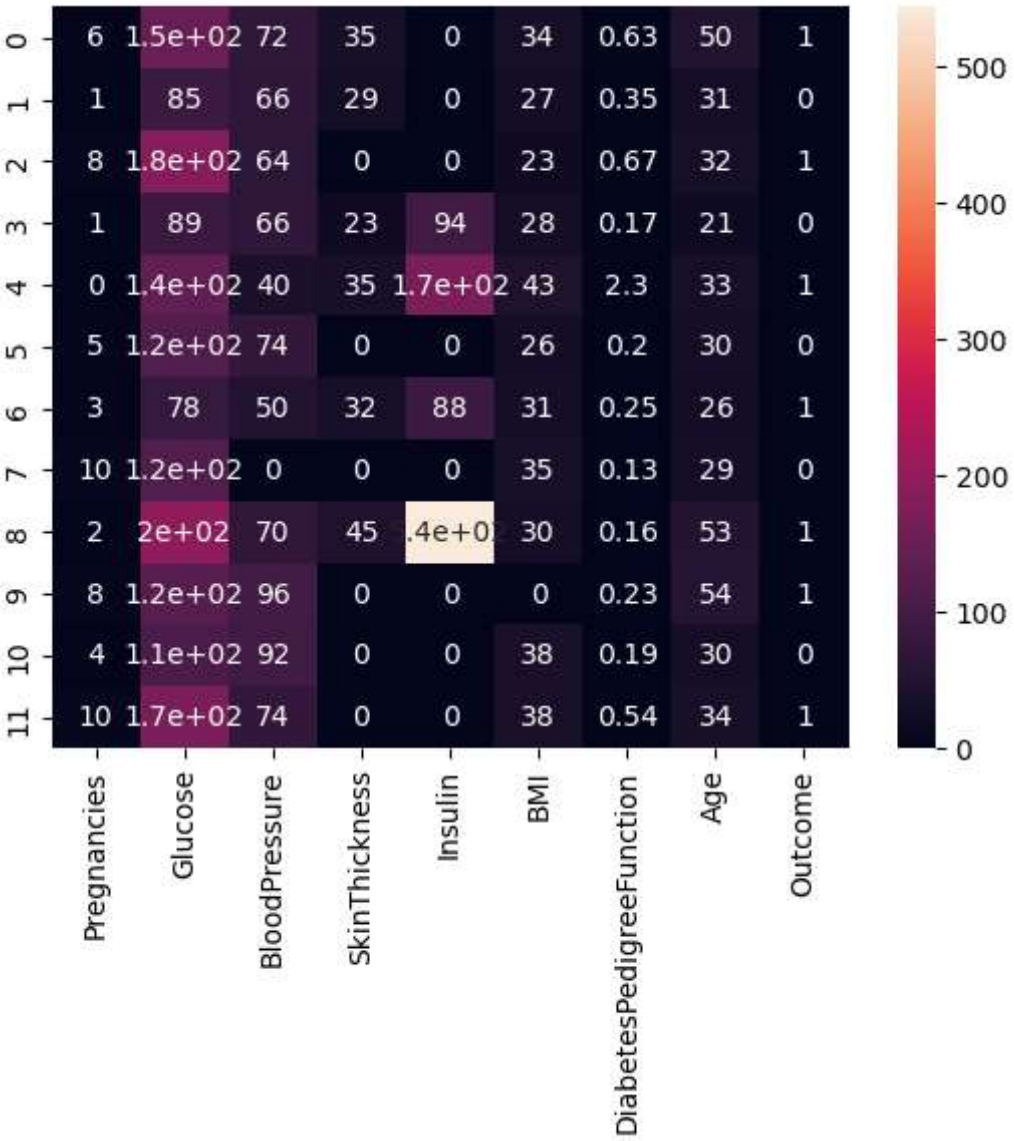
Out[8]:
```
1     135
0     111
2     103
3      75
4      68
5      57
6      50
7      45
8      38
9      28
10     24
11     11
13     10
12      9
14      2
15      1
17      1
Name: Pregnancies, dtype: int64
```

Step 2 : Identify relationships between feature

In [9]:
```
import seaborn as sn
import numpy as np
import matplotlib.pylab as plt
```

In [10]:
```
sn.heatmap(d.head(12), annot=True)
```

Out[10]: <AxesSubplot:>



step 3 : Prediction using one feature

In [11]:
```
X = d[['Age']]
y = d[['Outcome']]
```

In [12]:
```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
```

In [13]:
```
X_train,X_test,y_train,y_test = train_test_split(X,y,test_size=.25,random_sta
LOR=LogisticRegression()
LOR.fit(X_train,y_train)
```

```
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
```

Out[13]:
```
LogisticRegression()
```

In [14]:
```
X_train
```

Out[14]:

	Age
357	44
73	23
352	46
497	25
145	21
...	...
71	26
106	27
270	38
435	29
102	21

576 rows × 1 columns

In [15]:
```
y_pred=LOR.predict(X_test)
y_pred
```

Out[15]:
```
array([0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0,
       0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 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, 1, 0, 1, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0,
       0, 1, 0, 1, 1, 0, 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, 1, 0, 0, 0], dtype=int64)
```

In [16]:
```
print("Coef_ ",LOR.coef_)
print("intercept_",LOR.intercept_)
```

```
Coef_  [[0.05221912]]
intercept_ [-2.39506398]
```

In [17]:
```
LOR.predict([[60]])
```

```
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning:
X does not have valid feature names, but LogisticRegression was fitted with
feature names
  warnings.warn(
```

Out[17]:
```
array([1], dtype=int64)
```

In [18]:
```
lrf=LOR.coef_ * 60 + LOR.intercept_
from scipy.special import expit
dk = expit(lrf)
dk
```

Out[18]:
```
array([[0.67657656]])
```

In [19]:
```
if dk > 0.5:
    print('Yes, he will become diabetic')
else:
    print('No, he will not be diabetic')
```

```
Yes, he will become diabetic
```

Step 4 : Prediction using many features

In [20]:
```
X1=d[['Glucose','BMI','Age']]
```

In [21]:
```
X1_train,X1_test,y1_train,y1_test = train_test_split(X1,y,random_state=42,tes
```

In [22]:
```
from sklearn import linear_model
LOR1 = LogisticRegression()
LOR1.fit(X1_train,y1_train)
LOR1.predict(X1_test)
```

```
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
```

Out[22]:
```
array([0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0,
       0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0,
       0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1,
       0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0,
       0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0,
       0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1,
       0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
       0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0,
       0, 0, 0, 0, 1, 1, 0, 1, 1], dtype=int64)
```

In [23]:
```
print("coef_ : ",LOR1.coef_)
print("intercept_: ",LOR1.intercept_)
```

```
coef_ :  [[0.03292234 0.09635698 0.04398021]]
intercept_:  [-9.39683405]
```

In [24]:
```
lrf1=LOR1.coef_ * 150 * 30 * 40+ LOR1.intercept_
from scipy.special import expit
expit(lrf1)
```

Out[24]:
```
array([[1., 1., 1.]])
```

In [25]:
```
LOR1.predict([[150,30,40]])
```

```
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning:
X does not have valid feature names, but LogisticRegression was fitted with
feature names
  warnings.warn(
```

Out[25]:
```
array([1], dtype=int64)
```

In [26]:
```
LOR1.predict_proba([[150,30,40]])
```

```
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning:
X does not have valid feature names, but LogisticRegression was fitted with
feature names
  warnings.warn(
```

Out[26]:
```
array([[0.45228691, 0.54771309]])
```

Step5 : Build LOR model with all features

In [27]:
```
X2=d.drop(['Outcome'],axis=1)
X2_train,X2_test,y2_train,y2_test = train_test_split(X2,y,test_size=.25,rando
LOR2=LogisticRegression()
LOR2.fit(X2_train,y2_train)
LOR2.predict(X2_test)
```

```
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\linear_model\_logistic.p
y:814: ConvergenceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html (https://scik
it-learn.org/stable/modules/preprocessing.html)
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regre
ssion (https://scikit-learn.org/stable/modules/linear_model.html#logistic-re
gression)
  n_iter_i = _check_optimize_result(
```

Out[27]:
```
array([0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0,
       1, 0, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0,
       0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1,
       0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0,
       0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1,
       0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1,
       0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
       0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0,
       0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0], dtype=int64)
```

In [28]:
```
y2_pred = LOR2.predict(X2_test)
y2_pred
```

Out[28]:
```
array([0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0,
       1, 0, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0,
       0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1,
       0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0,
       0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1,
       0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1,
       0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
       0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0,
       0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0], dtype=int64)
```

In [29]:
```
from sklearn.metrics import roc_auc_score
lor_auc = roc_auc_score(y2_test,y2_pred)
print("Auc: ",lor_auc)
```

```
Auc:  0.7122658183103571
```

Step 6 : forward selection procedure

In [30]:
```
def get_auc(var,tar,d):
    fx = d[var]
    fy = d[tar]
    LOR4=LogisticRegression()
    LOR4.fit(fx,fy)
    pred=LOR4.predict_proba(fx)[:,1]
    auc_val = roc_auc_score(y,pred)
    return auc_val
get_auc(['Glucose',"BMI"],['Outcome'],d)
```

```
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
```

Out[30]:
```
0.8109328358208956
```

In [31]:
```
get_auc(['Pregnancies','BloodPressure','SkinThickness'],['Outcome'],d)
```

```
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
```

Out[31]:
```
0.6444962686567164
```

In [32]:
```
def best_next(current,cand,tar,d):
    best_auc=-1
    best_var=None
    for i in cand:
        auc_v = get_auc(current+[i],tar,d)
        if auc_v>=best_auc:
            best_auc=auc_v
            best_var=i
    return best_var
```

In [33]:
```
current=['Insulin','BMI','DiabetesPedigreeFunction','Age']
cand=['Pregnancies','Glucose','BloodPressure','SkinThickness']
tar=['Outcome']
next_var = best_next(current,cand,tar,d)
next_var
```

```
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
```

Out[33]:
```
'Glucose'
```

In [34]:
```
tar =['Outcome']
current=[]
cand=['Pregnancies','Glucose','BloodPressure','SkinThickness','Insulin','BMI'
max_num = 7
num_it = min(max_num,len(cand))
for i in range(0,num_it):
    next_var = best_next(current,cand,tar,d)
    current += [next_var]
    cand.remove(next_var)
    print("variable add in step "+str(i+1)+' is '+ next_var +" .")
```

```
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:99
3: DataConversionWarning: A column-vector y was passed when a 1d array wa
s expected. Please change the shape of y to (n_samples, ), for example us
ing ravel().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:99
3: DataConversionWarning: A column-vector y was passed when a 1d array wa
s expected. Please change the shape of y to (n_samples, ), for example us
ing ravel().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:99
3: DataConversionWarning: A column-vector y was passed when a 1d array wa
s expected. Please change the shape of y to (n_samples, ), for example us
ing ravel().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:99
3: DataConversionWarning: A column-vector y was passed when a 1d array wa
s expected. Please change the shape of y to (n_samples, ), for example us
ing ravel().
```

In [35]:
```
print(current)
```

```
['Glucose', 'BMI', 'Pregnancies', 'DiabetesPedigreeFunction', 'BloodPressur
e', 'Age', 'SkinThickness']
```

Step 7 : Plot line graph of AUC values and select cut-off

In [36]:
```
X2_train,X2_test,y2_train,y2_test = train_test_split(X2,y,stratify=y,test_siz
```

In [37]:
```
prediction=LOR2.predict_proba(X2_test)
```

In [38]:
```
train = pd.concat([X2_train,y2_train],axis =1)
test = pd.concat([X2_test,y2_test],axis =1)
def auc_train_test (variables,target, train, test):
    X_train = train[variables]
    X_test = test[variables]
    y_train =train[target]
    y_test = test[target]
    Lor=LogisticRegression()
    Lor.fit(X_train,y_train)
    prediction_train = Lor.predict_proba(X_train)[:,1]
    prediction_test = Lor.predict_proba(X_test)[:,1]
    auc_train = roc_auc_score(y_train, prediction_train)
    auc_test = roc_auc_score(y_train,prediction_test)
    return (auc_train,auc_test)
auc_values_train=[]
auc_values_test=[]
variable_evaluate=[]
for v in X2.columns:
    variable_evaluate.append(v)
    auc_train,auc_test = auc_train_test(variable_evaluate,['Outcome'],train,t
    auc_values_train.append(auc_train)
    auc_values_test.append(auc_test)
```

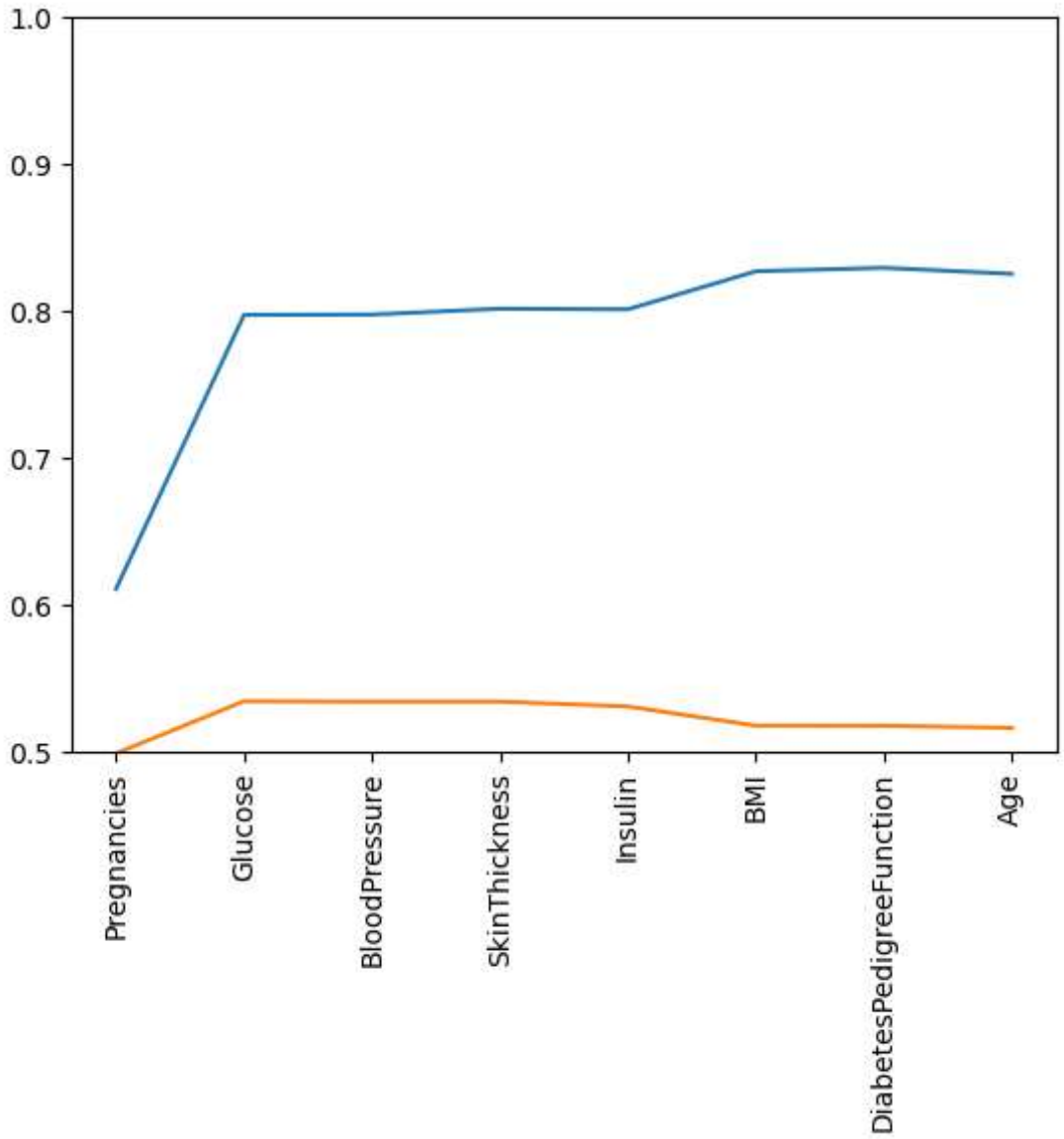
```
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\utils\validation.py:993:
DataConversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples, ), for example using ravel
().
  y = column_or_1d(y, warn=True)
C:\Users\sweth\anaconda3\lib\site-packages\sklearn\linear_model\_logistic.p
y:814: ConvergenceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html (https://scik
it-learn.org/stable/modules/preprocessing.html)
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regre
ssion (https://scikit-learn.org/stable/modules/linear_model.html#logistic-re
gression)
  n_iter_i = _check_optimize_result(
```

In [39]:
```
import numpy as np
x =np.array(range(0,len(auc_values_train)))

my_train = np.array(auc_values_train)
my_test = np.array(auc_values_test)

plt.xticks(x,X2.columns,rotation=90)
plt.plot(x,my_train)
plt.plot(x,my_test)
plt.ylim(0.5,1)
plt.show()
```

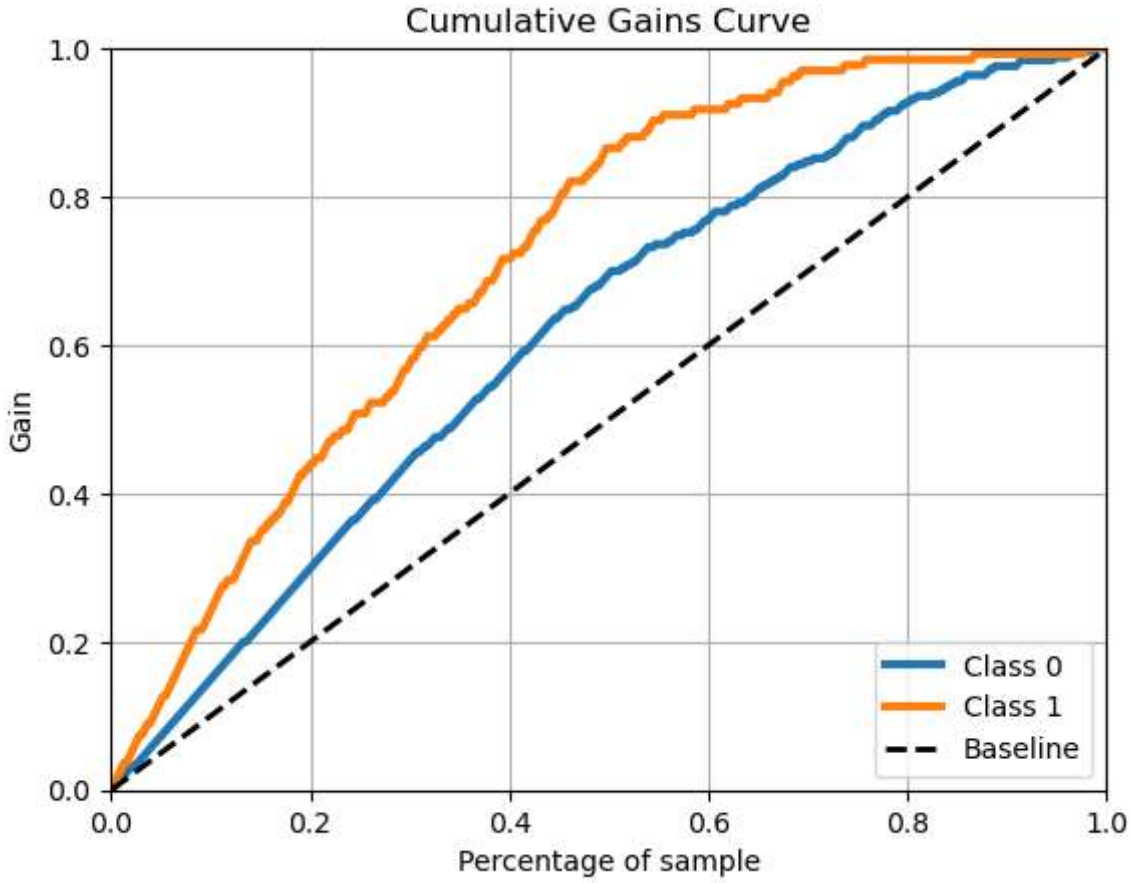


step 8 Draw cumulative gain chart and lift chart

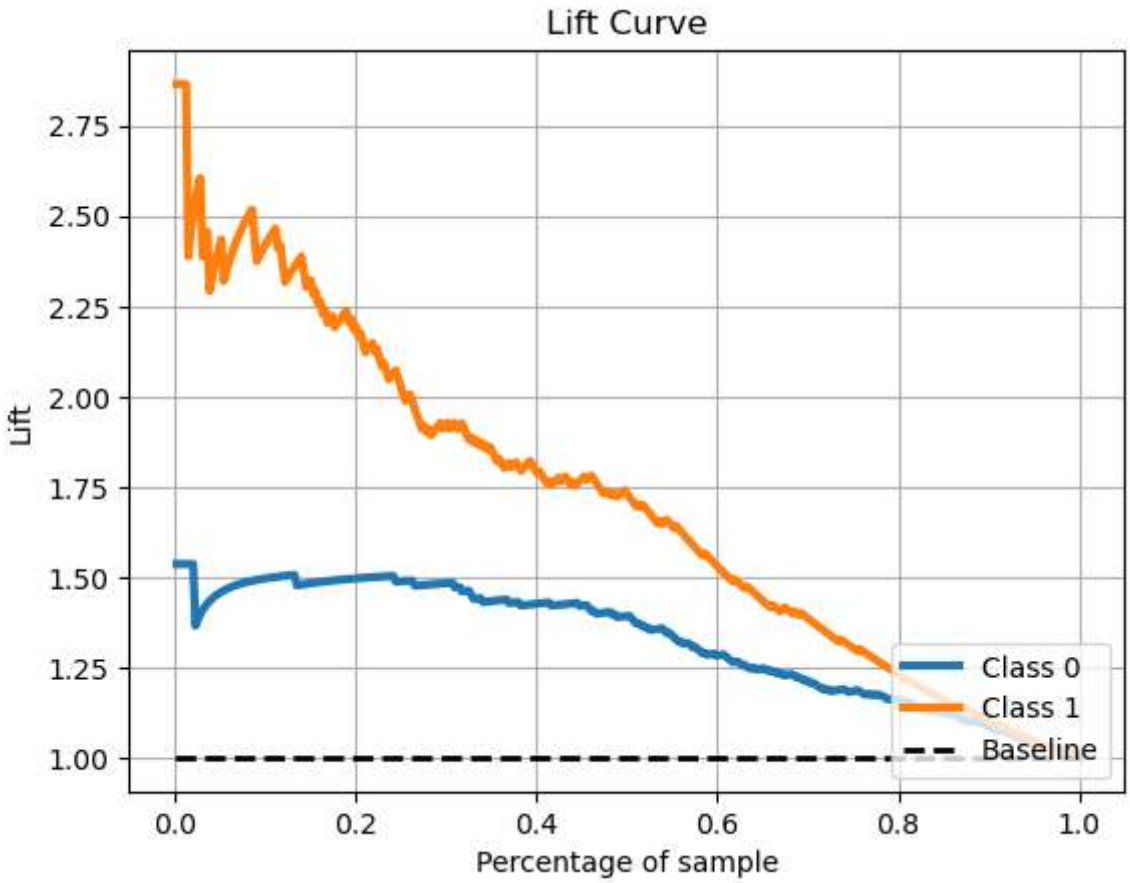
In [40]:
```
!pip install scikit-plot
from scikitplot.estimators import plot_feature_importances
from scikitplot.metrics import plot_confusion_matrix, plot_roc
```

```
Requirement already satisfied: scikit-plot in c:\users\sweth\anaconda3\lib\s
ite-packages (0.3.7)
Requirement already satisfied: scipy>=0.9 in c:\users\sweth\anaconda3\lib\si
te-packages (from scikit-plot) (1.9.1)
Requirement already satisfied: joblib>=0.10 in c:\users\sweth\anaconda3\lib
\site-packages (from scikit-plot) (1.1.0)
Requirement already satisfied: matplotlib>=1.4.0 in c:\users\sweth\anaconda3
\lib\site-packages (from scikit-plot) (3.5.2)
Requirement already satisfied: scikit-learn>=0.18 in c:\users\sweth\anaconda
3\lib\site-packages (from scikit-plot) (1.0.2)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\sweth\anaconda3
\lib\site-packages (from matplotlib>=1.4.0->scikit-plot) (1.4.2)
Requirement already satisfied: numpy>=1.17 in c:\users\sweth\anaconda3\lib\s
ite-packages (from matplotlib>=1.4.0->scikit-plot) (1.21.5)
Requirement already satisfied: pyparsing>=2.2.1 in c:\users\sweth\anaconda3
\lib\site-packages (from matplotlib>=1.4.0->scikit-plot) (3.0.9)
Requirement already satisfied: cycler>=0.10 in c:\users\sweth\anaconda3\lib
\site-packages (from matplotlib>=1.4.0->scikit-plot) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\sweth\anaconda3
\lib\site-packages (from matplotlib>=1.4.0->scikit-plot) (4.25.0)
Requirement already satisfied: packaging>=20.0 in c:\users\sweth\anaconda3\l
ib\site-packages (from matplotlib>=1.4.0->scikit-plot) (21.3)
Requirement already satisfied: pillow>=6.2.0 in c:\users\sweth\anaconda3\lib
\site-packages (from matplotlib>=1.4.0->scikit-plot) (9.2.0)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\sweth\anacon
da3\lib\site-packages (from matplotlib>=1.4.0->scikit-plot) (2.8.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in c:\users\sweth\anacon
da3\lib\site-packages (from scikit-learn>=0.18->scikit-plot) (2.2.0)
Requirement already satisfied: six>=1.5 in c:\users\sweth\anaconda3\lib\site
-packages (from python-dateutil>=2.7->matplotlib>=1.4.0->scikit-plot) (1.16.
0)
```

In [41]:
```
import scikitplot as skplt
skplt.metrics.plot_cumulative_gain(y2_test,prediction)
plt.show()
plt.figure(figsize=(7,7))
skplt.metrics.plot_lift_curve(y2_test, prediction)
plt.show()
```




```
<Figure size 700x700 with 0 Axes>
```


Lift Curve
Lift
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
0.0
0.2
0.4
0.6
0.8
1.0
Percentage of sample
Class 0
Class 1
Baseline

In [ ]:

In [ ]: