In [1]:
```
import numpy as np
import pandas as pd
from sklearn.ensemble import RandomForestClassifier
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split

data=pd.read_csv('pima.csv')
data.head()
```

Out[1]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunct
0	6	148	72	35	0	33.6	0.
1	1	85	66	29	0	26.6	0.
2	8	183	64	0	0	23.3	0.
3	1	89	66	23	94	28.1	0.
4	0	137	40	35	168	43.1	2.

In [2]:
```
X=data.drop('Outcome',axis=1)
Y=data.Outcome
Y
```

Out[2]:
```
0      1
1      0
2      1
3      0
4      1
      ..
763    0
764    0
765    0
766    1
767    0
Name: Outcome, Length: 768, dtype: int64
```

In [ ]:
```

```

In [3]:
```
xtrain, xtest, ytrain,ytest=train_test_split(X,Y,stratify=Y,test_size=0.20)
```

In [4]:
```
rd=RandomForestClassifier(n_estimators=100)
model1=rd.fit(xtrain,ytrain)
y_pred1=rd.predict(xtest)
```

In [5]:
```
from sklearn.metrics import classification_report, accuracy_score
print('Accuracy : ',accuracy_score(ytest,y_pred1))
```

```
Accuracy :  0.7402597402597403
```

In [6]:
```
from sklearn.tree import DecisionTreeClassifier
dt=DecisionTreeClassifier()
model2=dt.fit(xtrain,ytrain)
y_pred2=dt.predict(xtest)
print('Accuracy : ',accuracy_score(ytest,y_pred2))
```

```
Accuracy :  0.7142857142857143
```

In [ ]:
```

```