Machine-Learning Experiments

#we use random_state = 42 , To avoid from randomnes that can affect our expirements

Model Decision Tree:

```
1. DecisionTreeClassifier(criterion = "log_loss",max_depth=
20,min_samples_split= 5 ,min_samples_leaf= 6 ,random_state = 42)
```

Criterion = “log loss” (default = gini), max_depth = 20 (default = None) min_sample split = 5 (default = 2) , min sample leaf = 6 (default = 1)

In this expirement we got Recall Score: 0.9907407407407407

```
2. DecisionTreeClassifier(criterion="entropy",random_state=42)
```

Criterion = “entropy” (default = gini).

In this expirement we got Recall Score: 0.9814814814814815

```
3.DecisionTreeClassifier(criterion="entropy",splitter="random",max_depth=1
5,max_features=5,random_state=42)
```

Criterion = entropy(default = gini) , spliter = random(default = best) ,max depth = 15 (default = None) , max features = 5 (default = None).

In this expirement we got Recall Score : 0.9166666666666666

```
4.DecisionTreeClassifier(criterion="entropy",splitter="random",max_depth=4
0,max_features=40,max_leaf_nodes=20,min_samples_split=5,random_state=42)
```

Criterion = entropy(default = gini) , spliter = random(default = best), max depth = 40 (default = None) , max features = 40 (default = None) , max leaf node = 20 (default = None) ,min sample split = 5 (default = 1).

In this expirement we got Recall Score: 0.9629629629629629

```
5. DecisionTreeClassifier(criterion =
"log_loss",splitter="random",max_depth=64,max_features=55
,min_samples_leaf=3,min_impurity_decrease=0.1,random_state=42)
```

Criterion = log loss (default = gini) , spliter = random (default = best),max depth = 64 (default = None) , max features =55 (default = None) , min sample leaf = 3 (default = 1) , min impurity decrease = 0.1 (default = 0.0).

In this expirement we got Recall Score: 0.8240740740740741

Conclusion: **expirement 1** give us the best result .

Model Random Forest:

```
1. RandomForestClassifier(n_estimators=1000 ,
criterion="entropy",max_features=None,random_state=42)
```

N_estimators = 1000(default = 100) , critetion = entropy (default = gini) , max features = None (default = sqrt).
In this expirement we got Recall Score: 0.9722222222222222

```
2. RandomForestClassifier(n_estimators=1000 ,
criterion="log_loss",max_features="log2" , min_samples_split= 4 ,
min_samples_leaf= 10 ,min_weight_fraction_leaf=0.2 ,
max_leaf_nodes=100,random_state=42)
```

N_estimators = 1000(default = 100) , critetion = log loss (default = gini) , max features = log2 (default = sqrt) , min sample split = 4 (default = 2) , min sample leaf = 10 (default = 1) , min weight fraction leaf = 0.2 (default = 0.0) , max leaf nodes = 100 (default = none) .
In this expirement we got Recall Score : 1.0

```
3. RandomForestClassifier(n_estimators=2000 , criterion="log_loss",
min_samples_split= 3 ,min_impurity_decrease=0.1 , min_samples_leaf= 5
,min_weight_fraction_leaf=0.5 , max_leaf_nodes=50,random_state=42)
```

N_estimators = 2000(default = 100) , critetion = log loss (default = gini) , min sample split = 3 (default = 2) , min sample leaf = 5 (default = 1),min impurity decrease = 0.1 (default = 0.0) , min weight fraction leaf = 0.5 (default = 0.0) , max leaf nodes = 50 (default = none) .
In this expirement we got Recall Score : 0.9537037037037037

```
4 RandomForestClassifier(n_estimators=2000, min_samples_leaf= 10  ,
max_leaf_nodes=125,random_state=42)
```

N_estimators = 2000(default = 100), min sample leaf = 10 (default = 1), max leaf nodes = 125 (default = None).
In this expirement we got Recall Score: 0.9907407407407407

```
5. RandomForestClassifier(n_estimators=4000,min_weight_fraction_leaf=0.3 ,
max_features="log2"  , min_samples_leaf= 12  ,
max_leaf_nodes=125,random_state=42)
```

N_estimators = 4000(default = 100), min sample leaf = 12 (default = 1), max_feature = log2(default = sqrt) , min weight fraction leaf = 0.3 (default = 0.0) , max leaf nodes = 125 (default = None).

In this expirement we got Recall Score: 0.9907407407407407

Conclusion: **expirement 2** give us the best result .

Model AdaBoost :

1. `AdaBoostClassifier(n_estimators=100, learning_rate=2.0 , algorithm='SAMME',random_state=42)`

n-estimators = 100 (default = 50) , learning_rate = 2.0 (default = 1.0) algorithm = 'SAMME' (default = 'SAMME.R')
in this expirement we got Recall Score: 0.9351851851851852

2. `AdaBoostClassifier(n_estimators=250, learning_rate=0.5 , algorithm='SAMME',random_state=42)`

n-estimators = 250 (default = 50) , learning_rate = 0.5 (default = 1.0) algorithm = 'SAMME' (default = 'SAMME.R')
in this expirement we got Recall Score: 0.9722222222222222

3. `AdaBoostClassifier(n_estimators=75 , learning_rate=0.5,random_state=42)`

n-estimators = 75 (default = 50) , learning_rate = 0.5 (default = 1.0)
in this expirement we got Recall Score: 0.9629629629629629

4.`AdaBoostClassifier(n_estimators=52,algorithm='SAMME',learning_rate=0.2,random_state=42)`

n-estimators = 52 (default = 50) , learning_rate = 0.2 (default = 1.0) , algorithm = SAMME (default = SAMME.R).
in this expirement we got Recall Score: 0.9814814814814815

5. `AdaBoostClassifier(n_estimators=52,learning_rate=1.5,random_state=42)`

n-estimators = 52 (default = 50) , learning_rate = 1.5 (default = 1.0).
in this expirement we got Recall Score: 0.9722222222222222

Conclusion: **expirement 4** give us the best result .