[2]:
```
import numpy as np
import pandas as pd
from IPython.core.interactiveshell import InteractiveShell

InteractiveShell.ast_node_interactivity = "all"


pd.set_option("display.max_columns", 100)
pd.set_option("display.max_rows", 100)

data = pd.read_csv('DataSetForPhishingVSBenignUrl.csv', header=0)

data.head(10)
```

[2]:

	Querylength	domain_token_count	path_token_count	avgdomaintokenlen
0	0	4	5	5.5
1	0	4	5	5.5
2	0	4	5	5.5
3	0	4	12	5.5
4	0	4	6	5.5
5	0	4	8	5.5
6	0	4	5	5.5
7	0	4	7	5.5
8	0	4	6	5.5
9	0	4	5	5.5

	longdomaintokenlen	avgpathtokenlen	tld	charcompvowels	charcompace
0	14	4.400000	4	8	3
1	14	6.000000	4	12	4
2	14	5.800000	4	12	5
3	14	5.500000	4	32	16
4	14	7.333334	4	18	11
5	14	6.500000	4	22	10
6	14	7.800000	4	17	10
7	14	6.285714	4	16	9
8	14	6.500000	4	16	10
9	14	3.600000	4	7	3

	ldl_url	ldl_domain	ldl_path	ldl_filename	ldl_getArg	dld_url
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0

	dld_domain	dld_path	dld_filename	dld_getArg	urlLen	domainlength

0	0	0	0	0	58	25
1	0	0	0	0	66	25
2	0	0	0	0	65	25
3	0	0	0	0	109	25
4	0	0	0	0	81	25
5	0	0	0	0	91	25
6	0	0	0	0	75	25
7	0	0	0	0	82	25
8	0	0	0	0	76	25
9	0	0	0	0	54	25

	pathLength	subDirLen	fileNameLen	this.fileExtLen	ArgLen	pathurlRatio	\
0	26	26	13	1	2	0.448276	
1	34	34	2	2	2	0.515151	
2	33	33	2	2	2	0.507692	
3	77	77	2	2	2	0.706422	
4	49	49	2	2	2	0.604938	
5	59	59	2	2	2	0.648352	
6	43	43	2	2	2	0.573333	
7	50	50	2	2	2	0.609756	
8	44	44	2	2	2	0.578947	
9	22	22	9	1	2	0.407407	

	ArgUrlRatio	argDomanRatio	domainUrlRatio	pathDomainRatio	argPathRatio	\
0	0.034483	0.08	0.431034	1.04	0.0769231	
1	0.030303	0.08	0.378788	1.36	0.0588235	
2	0.030769	0.08	0.384615	1.32	0.0606061	
3	0.018349	0.08	0.229358	3.08	0.025974	
4	0.024691	0.08	0.308642	1.96	0.0408163	
5	0.021978	0.08	0.274725	2.36	0.0338983	
6	0.026667	0.08	0.333333	1.72	0.0465116	
7	0.024390	0.08	0.304878	2.00	0.04	
8	0.026316	0.08	0.328947	1.76	0.0454545	
9	0.037037	0.08	0.462963	0.88	0.0909091	

	executable	isPortEighty	NumberofDotsinURL	ISIpAddressInDomainName	\
0	0	-1	5	-1	
1	0	-1	4	-1	
2	0	-1	4	-1	
3	0	-1	4	-1	
4	0	-1	4	-1	
5	0	-1	4	-1	
6	0	-1	4	-1	
7	0	-1	4	-1	
8	0	-1	4	-1	
9	0	-1	5	-1	

	CharacterContinuityRate	LongestVariableValue	URL_DigitCount	\
0	0.6	-1	1	
1	0.6	-1	0	
2	0.6	-1	0	

3	0.6	-1	0
4	0.6	-1	0
5	0.6	-1	0
6	0.6	-1	0
7	0.6	-1	8
8	0.6	-1	0
9	0.6	-1	1

	host_DigitCount	Directory_DigitCount	File_name_DigitCount	\
0	0	0	0	
1	0	0	0	
2	0	0	0	
3	0	0	0	
4	0	0	0	
5	0	0	0	
6	0	0	0	
7	0	0	0	
8	0	0	0	
9	0	0	0	

	Extension_DigitCount	Query_DigitCount	URL_Letter_Count	\
0	1	-1	47	
1	0	-1	56	
2	0	-1	55	
3	0	-1	92	
4	0	-1	70	
5	0	-1	78	
6	0	-1	65	
7	8	-1	62	
8	0	-1	65	
9	1	-1	43	

	host_letter_count	Directory_LetterCount	Filename_LetterCount	\
0	22	8	13	
1	22	8	13	
2	22	8	13	
3	22	8	13	
4	22	8	13	
5	22	8	13	
6	22	8	13	
7	22	8	13	
8	22	8	13	
9	22	8	9	

	Extension_LetterCount	Query_LetterCount	LongestPathTokenLength	\
0	0	-1	13	
1	9	-1	13	
2	8	-1	13	
3	45	-1	52	
4	23	-1	24	
5	31	-1	34	

```
6                    18                -1                    18
7                    15                -1                    25
8                    18                -1                    19
9                     0                -1                     9

   Domain_LongestWordLength  Path_LongestWordLength  \
0                        14                      13
1                        14                      13
2                        14                      13
3                        14                      13
4                        14                      13
5                        14                      15
6                        14                      18
7                        14                      13
8                        14                      13
9                        14                       9

   sub-Directory_LongestWordLength  Arguments_LongestWordLength  \
0                                5                           -1
1                                5                           -1
2                                5                           -1
3                               13                           -1
4                               13                           -1
5                               13                           -1
6                                5                           -1
7                               13                           -1
8                               13                           -1
9                                5                           -1

   URL_sensitiveWord  URLQueries_variable  spcharUrl  delimeter_Domain  \
0                  0                    0          3                 0
1                  0                    0          4                 0
2                  0                    0          4                 0
3                  0                    0          4                 0
4                  0                    0          4                 0
5                  0                    0          4                 0
6                  0                    0          4                 0
7                  0                    0          4                 0
8                  0                    0          4                 0
9                  0                    0          3                 0

   delimeter_path  delimeter_Count  NumberRate_URL  NumberRate_Domain  \
0               2               -1        0.017241                0.0
1               1               -1        0.000000                0.0
2               1               -1        0.000000                0.0
3               8               -1        0.000000                0.0
4               2               -1        0.000000                0.0
5               4               -1        0.000000                0.0
6               1               -1        0.000000                0.0
7               3               -1        0.097561                0.0
8               2               -1        0.000000                0.0
```

```
9               2              -1        0.018519                0.0

   NumberRate_DirectoryName  NumberRate_FileName  NumberRate_Extension  \
0                       0.0             0.066667                   1.0
1                       0.0             0.000000                   NaN
2                       0.0             0.000000                   NaN
3                       0.0             0.000000                   NaN
4                       0.0             0.000000                   NaN
5                       0.0             0.000000                   NaN
6                       0.0             0.000000                   NaN
7                       0.0             0.320000                   NaN
8                       0.0             0.000000                   NaN
9                       0.0             0.090909                   1.0

   NumberRate_AfterPath  SymbolCount_URL  SymbolCount_Domain  \
0                  -1.0                8                   3
1                  -1.0                8                   3
2                  -1.0                8                   3
3                  -1.0                8                   3
4                  -1.0                8                   3
5                  -1.0                8                   3
6                  -1.0                8                   3
7                  -1.0                8                   3
8                  -1.0                8                   3
9                  -1.0                8                   3

   SymbolCount_Directoryname  SymbolCount_FileName  SymbolCount_Extension  \
0                          2                     1                      0
1                          3                     0                      0
2                          3                     0                      0
3                          3                     0                      0
4                          3                     0                      0
5                          3                     0                      0
6                          3                     0                      0
7                          3                     0                      0
8                          3                     0                      0
9                          2                     1                      0

   SymbolCount_Afterpath  Entropy_URL  Entropy_Domain  Entropy_DirectoryName  \
0                     -1     0.726298        0.784493               0.894886
1                     -1     0.688635        0.784493               0.814725
2                     -1     0.695049        0.784493               0.814725
3                     -1     0.640130        0.784493               0.814725
4                     -1     0.681307        0.784493               0.814725
5                     -1     0.666676        0.784493               0.814725
6                     -1     0.682440        0.784493               0.814725
7                     -1     0.709396        0.784493               0.814725
8                     -1     0.678242        0.784493               0.814725
9                     -1     0.740950        0.784493               0.894886

   Entropy_Filename  Entropy_Extension  Entropy_Afterpath URL_Type_obf_Type
```

0	0.850608	NaN	-1.0	Defacement
1	0.859793	0.0	-1.0	Defacement
2	0.801880	0.0	-1.0	Defacement
3	0.663210	0.0	-1.0	Defacement
4	0.804526	0.0	-1.0	Defacement
5	0.755658	0.0	-1.0	Defacement
6	0.766719	0.0	-1.0	Defacement
7	0.797498	0.0	-1.0	Defacement
8	0.732258	0.0	-1.0	Defacement
9	0.894886	NaN	-1.0	Defacement

Handle columns with Nulls in the current data set

[3]:
```
data.isnull().sum()
```

[3]:
```
Querylength                  0
domain_token_count           0
path_token_count             0
avgdomaintokenlen            0
longdomaintokenlen           0
avgpathtokenlen            280
tld                          0
charcompvowels               0
charcompace                  0
ldl_url                      0
ldl_domain                   0
ldl_path                     0
ldl_filename                 0
ldl_getArg                   0
dld_url                      0
dld_domain                   0
dld_path                     0
dld_filename                 0
dld_getArg                   0
urlLen                       0
domainlength                 0
pathLength                   0
subDirLen                    0
fileNameLen                  0
this.fileExtLen              0
ArgLen                       0
pathurlRatio                 0
ArgUrlRatio                  0
argDomanRatio                0
domainUrlRatio               0
pathDomainRatio              0
argPathRatio                 0
executable                   0
isPortEighty                 0
NumberofDotsinURL            0
ISIpAddressInDomainName      0
```

```
CharacterContinuityRate                0
LongestVariableValue                   0
URL_DigitCount                         0
host_DigitCount                        0
Directory_DigitCount                   0
File_name_DigitCount                   0
Extension_DigitCount                   0
Query_DigitCount                       0
URL_Letter_Count                       0
host_letter_count                      0
Directory_LetterCount                  0
Filename_LetterCount                   0
Extension_LetterCount                  0
Query_LetterCount                      0
LongestPathTokenLength                 0
Domain_LongestWordLength               0
Path_LongestWordLength                 0
sub-Directory_LongestWordLength        0
Arguments_LongestWordLength            0
URL_sensitiveWord                      0
URLQueries_variable                    0
spcharUrl                              0
delimeter_Domain                       0
delimeter_path                         0
delimeter_Count                        0
NumberRate_URL                         0
NumberRate_Domain                      0
NumberRate_DirectoryName              10
NumberRate_FileName                   10
NumberRate_Extension               10130
NumberRate_AfterPath                   3
SymbolCount_URL                        0
SymbolCount_Domain                     0
SymbolCount_Directoryname              0
SymbolCount_FileName                   0
SymbolCount_Extension                  0
SymbolCount_Afterpath                  0
Entropy_URL                            0
Entropy_Domain                         0
Entropy_DirectoryName               8468
Entropy_Filename                     236
Entropy_Extension                     40
Entropy_Afterpath                      6
URL_Type_obf_Type                      0
dtype: int64
```

[4]:
```
data_clean_na = data.dropna()          #(subset=['Entropy_DirectoryName'])
```

[16]:
```
lst = ['phishing' , 'benign']

data_clean = data_clean_na[data_clean_na.URL_Type_obf_Type.isin(lst)  ]
```

```
data_clean['URL_Type_obf_Type'].unique()
```

[16]:
```
array(['benign', 'phishing'], dtype=object)
```

[17]:
```
data_clean.dtypes
```

[17]:
```
Querylength                    int64
domain_token_count             int64
path_token_count               int64
avgdomaintokenlen            float64
longdomaintokenlen             int64
avgpathtokenlen              float64
tld                            int64
charcompvowels                 int64
charcompace                    int64
ldl_url                        int64
ldl_domain                     int64
ldl_path                       int64
ldl_filename                   int64
ldl_getArg                     int64
dld_url                        int64
dld_domain                     int64
dld_path                       int64
dld_filename                   int64
dld_getArg                     int64
urlLen                         int64
domainlength                   int64
pathLength                     int64
subDirLen                      int64
fileNameLen                    int64
this.fileExtLen                int64
ArgLen                         int64
pathurlRatio                 float64
ArgUrlRatio                  float64
argDomanRatio                float64
domainUrlRatio               float64
pathDomainRatio              float64
argPathRatio                  object
executable                     int64
isPortEighty                   int64
NumberofDotsinURL              int64
ISIpAddressInDomainName        int64
CharacterContinuityRate      float64
LongestVariableValue           int64
URL_DigitCount                 int64
host_DigitCount                int64
Directory_DigitCount           int64
File_name_DigitCount           int64
Extension_DigitCount           int64
Query_DigitCount               int64
URL_Letter_Count               int64
```

```
host_letter_count                    int64
Directory_LetterCount                int64
Filename_LetterCount                 int64
Extension_LetterCount                int64
Query_LetterCount                    int64
LongestPathTokenLength               int64
Domain_LongestWordLength             int64
Path_LongestWordLength               int64
sub-Directory_LongestWordLength      int64
Arguments_LongestWordLength          int64
URL_sensitiveWord                    int64
URLQueries_variable                  int64
spcharUrl                            int64
delimeter_Domain                     int64
delimeter_path                       int64
delimeter_Count                      int64
NumberRate_URL                     float64
NumberRate_Domain                  float64
NumberRate_DirectoryName           float64
NumberRate_FileName                float64
NumberRate_Extension               float64
NumberRate_AfterPath               float64
SymbolCount_URL                      int64
SymbolCount_Domain                   int64
SymbolCount_Directoryname            int64
SymbolCount_FileName                 int64
SymbolCount_Extension                int64
SymbolCount_Afterpath                int64
Entropy_URL                        float64
Entropy_Domain                     float64
Entropy_DirectoryName              float64
Entropy_Filename                   float64
Entropy_Extension                  float64
Entropy_Afterpath                  float64
URL_Type_obf_Type                   object
dtype: object
```

[18]:
```
# split into input and output elements


data_numpy = data_clean.values
X, y = data_numpy[:, :-1], data_numpy[:, -1]


#X = data_clean.drop('URL_Type_obf_Type',axis = 'columns')
#y =  data_clean.URL_Type_obf_Type
```

SMOT Oversampling the data to removed skewed classes

[19]:
```
from sklearn.preprocessing import LabelEncoder
from imblearn.over_sampling import SMOTE
from collections import Counter


y = LabelEncoder().fit_transform(y)

# transform the dataset
oversample = SMOTE()
X, y = oversample.fit_resample(X, y)
# summarize distribution
```

[20]:
```
from matplotlib import pyplot

counter = Counter(y)
for k,v in counter.items():
    per = v / len(y) * 100
    print('Class=%d, n=%d (%.3f%%)' % (k, v, per))
# plot the distribution
pyplot.bar(counter.keys(), counter.values())
pyplot.show()
```

```
Class=0, n=4014 (50.000%)
Class=1, n=4014 (50.000%)
```

[20]:
```
<BarContainer object of 2 artists>
```


Test and Training Split and Scaling the data

[21]:
```
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler


X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=1, stratify=y)


sc = StandardScaler()
sc.fit(X_train)

X_train_std = sc.transform(X_train)
X_test_std = sc.transform(X_test)

print('Labels count in y:', np.bincount(y))
print('Labels count in y_train:', np.bincount(y_train))

print('Labels count in y_test:', np.bincount(y_test))
```

[21]:
```
StandardScaler()
```

```
Labels count in y: [4014 4014]
Labels count in y_train: [3211 3211]
Labels count in y_test: [803 803]
```

TRAIN MODEL & CREATE PREDICTIONS USING SKLEARN - GINI CRITERION

[73]:
```
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import AdaBoostClassifier
from sklearn import metrics


#developing a model with gini , tree and adboost


clf_gini_d1 = DecisionTreeClassifier(criterion = 'gini' , max_depth = 1 , splitter =␣
↪'best')
clf_gini_d3 = DecisionTreeClassifier(criterion = 'gini' , max_depth = 3 , splitter =␣
↪'best')
clf_gini_d6 = DecisionTreeClassifier(criterion = 'gini' , max_depth = 6 , splitter =␣
↪'best')
clf_gini_d9 = DecisionTreeClassifier(criterion = 'gini' , max_depth = 9 , splitter =␣
↪'best')
clf_gini_d12 = DecisionTreeClassifier(criterion = 'gini' , max_depth = 12 , splitter =␣
↪'best')
clf_gini_d15 = DecisionTreeClassifier(criterion = 'gini' , max_depth = 15 , splitter =␣
↪'best')
clf_gini_d18 = DecisionTreeClassifier(criterion = 'gini' , max_depth = 18 , splitter =␣
↪'best')



clf_g_d1 = AdaBoostClassifier(base_estimator = clf_gini_d1 ,n_estimators=50,␣
↪random_state=0)
clf_g_d3 = AdaBoostClassifier(base_estimator = clf_gini_d3 ,n_estimators=50,␣
↪random_state=0)
clf_g_d6 = AdaBoostClassifier(base_estimator = clf_gini_d6 ,n_estimators=50,␣
↪random_state=0)
clf_g_d9 = AdaBoostClassifier(base_estimator = clf_gini_d9 ,n_estimators=50,␣
↪random_state=0)
clf_g_d12 = AdaBoostClassifier(base_estimator = clf_gini_d12 ,n_estimators=50,␣
↪random_state=0)
clf_g_d15 = AdaBoostClassifier(base_estimator = clf_gini_d15 ,n_estimators=50,␣
↪random_state=0)
clf_g_d18 = AdaBoostClassifier(base_estimator = clf_gini_d18 ,n_estimators=50,␣
↪random_state=0)



clf_ada_g_1 = clf_g_d1.fit(X_train_std,y_train)
```

```
clf_ada_g_3 = clf_g_d3.fit(X_train_std,y_train)
clf_ada_g_6 = clf_g_d6.fit(X_train_std,y_train)
clf_ada_g_9 = clf_g_d9.fit(X_train_std,y_train)
clf_ada_g_12 = clf_g_d12.fit(X_train_std,y_train)
clf_ada_g_15 = clf_g_d15.fit(X_train_std,y_train)
clf_ada_g_18 = clf_g_d18.fit(X_train_std,y_train)


y_pred_g_1 = clf_ada_g_1.predict(X_test_std)
y_pred_g_3 = clf_ada_g_3.predict(X_test_std)
y_pred_g_6 = clf_ada_g_6.predict(X_test_std)
y_pred_g_9 = clf_ada_g_9.predict(X_test_std)
y_pred_g_12 = clf_ada_g_12.predict(X_test_std)
y_pred_g_15 = clf_ada_g_15.predict(X_test_std)
y_pred_g_18 = clf_ada_g_18.predict(X_test_std)
```

[74]:
```
print("Accuracy Gini 1:",metrics.accuracy_score(y_test, y_pred_g_1))
print("Accuracy Gini 3:",metrics.accuracy_score(y_test, y_pred_g_3))
print("Accuracy Gini 6:",metrics.accuracy_score(y_test, y_pred_g_6))
print("Accuracy Gini 9:",metrics.accuracy_score(y_test, y_pred_g_9))
print("Accuracy Gini 12:",metrics.accuracy_score(y_test, y_pred_g_12))
print("Accuracy Gini 15:",metrics.accuracy_score(y_test, y_pred_g_15))
print("Accuracy Gini 18:",metrics.accuracy_score(y_test, y_pred_g_18))
```

```
Accuracy Gini 1: 0.9645080946450809
Accuracy Gini 3: 0.975093399750934
Accuracy Gini 6: 0.9782067247820673
Accuracy Gini 9: 0.9788293897882939
Accuracy Gini 12: 0.9825653798256538
Accuracy Gini 15: 0.9813200498132005
Accuracy Gini 18: 0.9838107098381071
```

TRAIN MODEL & CREATE PREDICTIONS USING SKLEARN - ENTROPY

[76]:
```
clf_entropy_d1 = DecisionTreeClassifier(criterion = 'entropy' , max_depth = 1 ,␣
↪splitter = 'best')
clf_entropy_d3 = DecisionTreeClassifier(criterion = 'entropy' , max_depth = 3 ,␣
↪splitter = 'best')
clf_entropy_d6 = DecisionTreeClassifier(criterion = 'entropy' , max_depth = 6 ,␣
↪splitter = 'best')
clf_entropy_d9 = DecisionTreeClassifier(criterion = 'entropy' , max_depth = 9 ,␣
↪splitter = 'best')
clf_entropy_d12 = DecisionTreeClassifier(criterion = 'entropy' , max_depth = 12 ,␣
↪splitter = 'best')
clf_entropy_d15 = DecisionTreeClassifier(criterion = 'entropy' , max_depth = 15 ,␣
↪splitter = 'best')
clf_entropy_d18 = DecisionTreeClassifier(criterion = 'entropy' , max_depth = 18 ,␣
↪splitter = 'best')
```

```
clf_e_d1 = AdaBoostClassifier(base_estimator = clf_entropy_d1 ,n_estimators=50,␣
 ↪random_state=0)
clf_e_d3 = AdaBoostClassifier(base_estimator = clf_entropy_d3 ,n_estimators=50,␣
 ↪random_state=0)
clf_e_d6 = AdaBoostClassifier(base_estimator = clf_entropy_d6 ,n_estimators=50,␣
 ↪random_state=0)
clf_e_d9 = AdaBoostClassifier(base_estimator = clf_entropy_d9 ,n_estimators=50,␣
 ↪random_state=0)
clf_e_d12 = AdaBoostClassifier(base_estimator = clf_entropy_d12 ,n_estimators=50,␣
 ↪random_state=0)
clf_e_d15 = AdaBoostClassifier(base_estimator = clf_entropy_d15 ,n_estimators=50,␣
 ↪random_state=0)
clf_e_d18 = AdaBoostClassifier(base_estimator = clf_entropy_d18 ,n_estimators=50,␣
 ↪random_state=0)



clf_ada_e_1 = clf_e_d1.fit(X_train_std,y_train)
clf_ada_e_3 = clf_e_d3.fit(X_train_std,y_train)
clf_ada_e_6 = clf_e_d6.fit(X_train_std,y_train)
clf_ada_e_9 = clf_e_d9.fit(X_train_std,y_train)
clf_ada_e_12 = clf_e_d12.fit(X_train_std,y_train)
clf_ada_e_15 = clf_e_d15.fit(X_train_std,y_train)
clf_ada_e_18 = clf_e_d18.fit(X_train_std,y_train)



y_pred_e_1 = clf_ada_e_1.predict(X_test_std)
y_pred_e_3 = clf_ada_e_3.predict(X_test_std)
y_pred_e_6 = clf_ada_e_6.predict(X_test_std)
y_pred_e_9 = clf_ada_e_9.predict(X_test_std)
y_pred_e_12 = clf_ada_e_12.predict(X_test_std)
y_pred_e_15 = clf_ada_e_15.predict(X_test_std)
y_pred_e_18 = clf_ada_e_18.predict(X_test_std)
```

[77]:
```
print("Accuracy Entropy 1:",metrics.accuracy_score(y_test, y_pred_e_1))
print("Accuracy Entropy 3:",metrics.accuracy_score(y_test, y_pred_e_3))
print("Accuracy Entropy 6:",metrics.accuracy_score(y_test, y_pred_e_6))
print("Accuracy Entropy 9:",metrics.accuracy_score(y_test, y_pred_e_9))
print("Accuracy Entropy 12:",metrics.accuracy_score(y_test, y_pred_e_12))
print("Accuracy Entropy 15:",metrics.accuracy_score(y_test, y_pred_e_15))
print("Accuracy Entropy 18:",metrics.accuracy_score(y_test, y_pred_e_18))
```

```
Accuracy Entropy 1: 0.9676214196762142
Accuracy Entropy 3: 0.9800747198007472
Accuracy Entropy 6: 0.9794520547945206
Accuracy Entropy 9: 0.9825653798256538
Accuracy Entropy 12: 0.9819427148194272
Accuracy Entropy 15: 0.9694894146948941
Accuracy Entropy 18: 0.9694894146948941
```

ANSWER : The accuracy using Ensemble is much higer at a greater performance as compared

to using Decision Tree independently. The max accuracy reached with Ensemble was around 98% however around 88% with only using decision tree without ensemble.