

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [2]: df=pd.read_csv(r"C:\Users\user\Downloads\data.csv")
df
```

Out[2]:

	row_id	user_id	timestamp	gate_id
0	0	18	2022-07-29 09:08:54	7
1	1	18	2022-07-29 09:09:54	9
2	2	18	2022-07-29 09:09:54	9
3	3	18	2022-07-29 09:10:06	5
4	4	18	2022-07-29 09:10:08	5
...
37513	37513	6	2022-12-31 20:38:56	11
37514	37514	6	2022-12-31 20:39:22	6
37515	37515	6	2022-12-31 20:39:23	6
37516	37516	6	2022-12-31 20:39:31	9
37517	37517	6	2022-12-31 20:39:31	9

37518 rows × 4 columns

```
In [3]: df.columns
```

Out[3]: Index(['row_id', 'user_id', 'timestamp', 'gate_id'], dtype='object')

```
In [4]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 37518 entries, 0 to 37517
Data columns (total 4 columns):
 #   Column      Non-Null Count  Dtype  
 --- 
 0   row_id      37518 non-null   int64  
 1   user_id     37518 non-null   int64  
 2   timestamp   37518 non-null   object  
 3   gate_id     37518 non-null   int64  
dtypes: int64(3), object(1)
memory usage: 1.1+ MB
```

```
In [5]: df['gate_id'].value_counts()
```

Out[5]:

4	8170
3	5351

```

10    4767
5     4619
11   4090
9    3390
7    3026
6    1800
13   1201
12    698
15    298
-1     48
8     48
1      5
16     4
0      2
14     1
Name: gate_id, dtype: int64

```

In [6]:

```
x=df[['row_id', 'user_id']]
y=df['gate_id']
```

In [7]:

```
g1={"gate_id":{'4':1,'3':2,'10':3,'5':4,'11':5,'9':6,'7':7,'6':8,'13':9,'12':10,'15':11
df=df.replace(g1)
print(df)
```

	row_id	user_id	timestamp	gate_id
0	0	18	2022-07-29 09:08:54	7
1	1	18	2022-07-29 09:09:54	9
2	2	18	2022-07-29 09:09:54	9
3	3	18	2022-07-29 09:10:06	5
4	4	18	2022-07-29 09:10:08	5
...
37513	37513	6	2022-12-31 20:38:56	11
37514	37514	6	2022-12-31 20:39:22	6
37515	37515	6	2022-12-31 20:39:23	6
37516	37516	6	2022-12-31 20:39:31	9
37517	37517	6	2022-12-31 20:39:31	9

[37518 rows x 4 columns]

In [8]:

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.70)
```

In [9]:

```
from sklearn.ensemble import RandomForestClassifier
rfc=RandomForestClassifier()
rfc.fit(x_train,y_train)
```

Out[9]: RandomForestClassifier()

In [10]:

```
parameters= {
    "max_depth":[1,2,3,4,5],
    "min_samples_leaf":[5,10,15,20,25],
    'n_estimators':[10,20,30,40,50]
}
```

In [11]:

```
from sklearn.model_selection import GridSearchCV
grid_search=GridSearchCV(estimator=rfc,param_grid=parameters,cv=2,scoring="accuracy")
```

```
grid_search.fit(x_train,y_train)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\model_selection\_split.py:666: UserWarning: The least populated class in y has only 1 members, which is less than n_splits=2.
    warnings.warn(("The least populated class in y has only %d"

```

```
Out[11]: GridSearchCV(cv=2, estimator=RandomForestClassifier(),
                      param_grid={'max_depth': [1, 2, 3, 4, 5],
                                   'min_samples_leaf': [5, 10, 15, 20, 25],
                                   'n_estimators': [10, 20, 30, 40, 50]},
                      scoring='accuracy')
```

```
In [12]: grid_search.best_score_
```

```
Out[12]: 0.221412596693094
```

```
In [13]: rfc_best=grid_search.best_estimator_
```

```
In [14]: from sklearn.tree import plot_tree
plt.figure(figsize=(80,40))
plot_tree(rfc_best.estimators_[5],feature_names=x.columns,class_names=['a','b','c','d'],
```

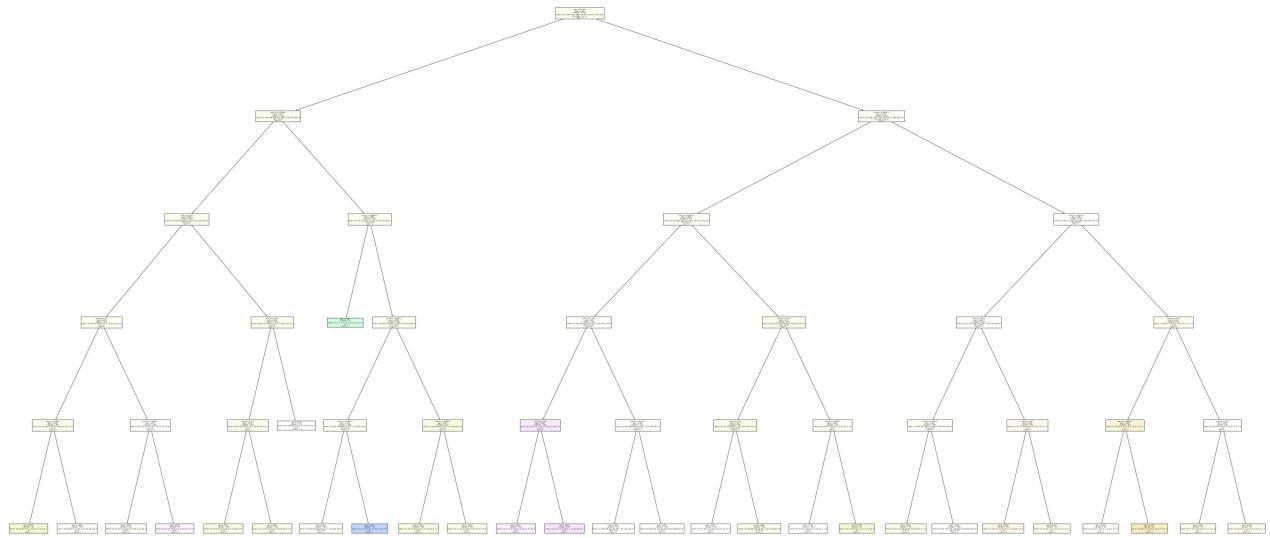
```
Out[14]: [Text(2028.1153846153845, 1993.2, 'user_id <= 16.0\nngini = 0.869\nsamples = 7123\nvalue = [11, 1709, 2441, 1368, 538, 947, 15, 975, 1439, 1182\nn165, 389, 1, 74, 1]\nclass = c'),
Text(965.7692307692307, 1630.8000000000002, 'row_id <= 17593.5\nngini = 0.857\nsamples = 2359\nvalue = [0, 349, 864, 618, 135, 303, 4, 326, 513, 455, 28\nn83, 1, 13, 0]\nclass = c'),
Text(643.8461538461538, 1268.4, 'user_id <= 8.0\nngini = 0.851\nsamples = 1055\nvalue = [0, 179, 420, 301, 63, 121, 0, 107, 195, 203, 24\nn50, 0, 0, 0]\nclass = c'),
Text(343.38461538461536, 906.0, 'user_id <= 4.5\nngini = 0.878\nsamples = 474\nvalue = [0, 98, 150, 92, 50, 48, 0, 70, 87, 76, 18, 45\nn0, 0, 0]\nclass = c'),
Text(171.69230769230768, 543.5999999999999, 'row_id <= 4462.0\nngini = 0.849\nsamples = 265\nvalue = [0, 45, 100, 67, 19, 26, 0, 49, 51, 49, 0, 0, 0\nn0, 0]\nclass = c'),
Text(85.84615384615384, 181.1999999999982, 'gini = 0.798\nsamples = 62\nvalue = [0, 1
2, 34, 16, 2, 7, 0, 8, 5, 11, 0, 0, 0\nn0]\nclass = c'),
Text(257.53846153846155, 181.1999999999982, 'gini = 0.856\nsamples = 203\nvalue = [0,
33, 66, 51, 17, 19, 0, 41, 46, 38, 0, 0, 0\nn0, 0]\nclass = c'),
Text(515.0769230769231, 543.5999999999999, 'row_id <= 10286.5\nngini = 0.887\nsamples = 209\nvalue = [0, 53, 50, 25, 31, 22, 0, 21, 36, 27, 18, 45\nn0, 0, 0]\nclass = b'),
Text(429.23076923076917, 181.19999999999982, 'gini = 0.878\nsamples = 133\nvalue = [0,
40, 30, 15, 23, 15, 0, 12, 16, 9, 13, 18, 0\nn0, 0]\nclass = b'),
Text(600.9230769230769, 181.19999999999982, 'gini = 0.875\nsamples = 76\nvalue = [0, 1
3, 20, 10, 8, 7, 0, 9, 20, 18, 5, 27, 0\nn0, 0]\nclass = 1'),
Text(944.3076923076923, 906.0, 'user_id <= 13.0\nngini = 0.817\nsamples = 581\nvalue = [0,
81, 270, 209, 13, 73, 0, 37, 108, 127, 6, 5\nn0, 0, 0]\nclass = c'),
Text(858.4615384615383, 543.5999999999999, 'user_id <= 11.5\nngini = 0.824\nsamples = 40
3\nvalue = [0, 80, 182, 121, 6, 41, 0, 24, 83, 91, 6, 5, 0\nn0, 0]\nclass = c'),
Text(772.6153846153845, 181.19999999999982, 'gini = 0.813\nsamples = 209\nvalue = [0,
4
9, 102, 62, 3, 29, 0, 0, 41, 35, 6, 4, 0\nn0, 0]\nclass = c'),
Text(944.3076923076923, 181.19999999999982, 'gini = 0.826\nsamples = 194\nvalue = [0,
3
1, 80, 59, 3, 12, 0, 24, 42, 56, 0, 1, 0\nn0, 0]\nclass = c'),
Text(1030.1538461538462, 543.5999999999999, 'gini = 0.778\nsamples = 178\nvalue = [0,
1, 88, 88, 7, 32, 0, 13, 25, 36, 0, 0, 0\nn0, 0]\nclass = c'),
Text(1287.6923076923076, 1268.4, 'row_id <= 18041.0\nngini = 0.859\nsamples = 1304\nvalue = [0, 170, 444, 317, 72, 182, 4, 219, 318, 252, 4\nn33, 1, 13, 0]\nclass = c'),
Text(1201.8461538461538, 906.0, 'gini = 0.768\nsamples = 22\nvalue = [0, 2, 7, 2, 0,
1
3, 0, 3, 3, 0, 0, 0\nn0]\nclass = f'),
Text(1373.5384615384614, 906.0, 'row_id <= 33253.5\nngini = 0.859\nsamples = 1282\nvalue = [0, 168, 437, 315, 72, 169, 4, 216, 315, 249, 4\nn33, 1, 13, 0]\nclass = c'),
```

```

Text(1201.8461538461538, 543.5999999999999, 'row_id <= 32441.5\ngini = 0.859\nsamples = 1025\nvalue = [0, 140, 328, 260, 53, 137, 0, 179, 267, 201, 4\n23, 0, 6, 0]\nnclass = c'),
Text(1116.0, 181.1999999999982, 'gini = 0.86\nsamples = 986\nvalue = [0, 139, 319, 252, 53, 132, 0, 175, 242, 194, 4\n23, 0, 6, 0]\nnclass = c'),
Text(1287.6923076923076, 181.1999999999982, 'gini = 0.753\nsamples = 39\nvalue = [0, 1, 9, 8, 0, 5, 0, 4, 25, 7, 0, 0, 0\n0]\nnclass = i'),
Text(1545.230769230769, 543.5999999999999, 'row_id <= 35805.5\ngini = 0.853\nsamples = 257\nvalue = [0, 28, 109, 55, 19, 32, 4, 37, 48, 48, 0, 10\n1, 7, 0]\nnclass = c'),
Text(1459.3846153846152, 181.1999999999982, 'gini = 0.832\nsamples = 148\nvalue = [0, 16, 70, 40, 15, 14, 4, 11, 25, 23, 0, 2, 1\n4, 0]\nnclass = c'),
Text(1631.0769230769229, 181.1999999999982, 'gini = 0.862\nsamples = 109\nvalue = [0, 12, 39, 15, 4, 18, 0, 26, 23, 25, 0, 8, 0\n3, 0]\nnclass = c'),
Text(3090.461538461538, 1630.8000000000002, 'row_id <= 19091.0\ngini = 0.871\nsamples = 4764\nvalue = [11, 1360, 1577, 750, 403, 644, 11, 649, 926, 727\n137, 306, 0, 61, 1]\nnclass = c'),
Text(2403.6923076923076, 1268.4, 'row_id <= 12802.5\ngini = 0.878\nsamples = 2444\nvalue = [6, 621, 804, 402, 225, 324, 2, 327, 417, 370, 112\n247, 0, 22, 0]\nnclass = c'),
Text(2060.3076923076924, 906.0, 'user_id <= 19.5\ngini = 0.88\nsamples = 1622\nvalue = [6, 429, 490, 270, 137, 225, 1, 211, 276, 217, 86\n184, 0, 12, 0]\nnclass = c'),
Text(1888.6153846153845, 543.5999999999999, 'user_id <= 18.5\ngini = 0.849\nsamples = 305\nvalue = [0, 34, 81, 37, 7, 24, 0, 32, 28, 31, 69, 127\n0, 0, 0]\nnclass = l'),
Text(1802.7692307692307, 181.1999999999982, 'gini = 0.842\nsamples = 159\nvalue = [0, 3, 47, 37, 4, 8, 0, 31, 16, 11, 31, 64, 0\n0, 0]\nnclass = l'),
Text(1974.4615384615383, 181.1999999999982, 'gini = 0.825\nsamples = 146\nvalue = [0, 31, 34, 0, 3, 16, 0, 1, 12, 20, 38, 63, 0\n0, 0]\nnclass = l'),
Text(2232.0, 543.5999999999999, 'row_id <= 6721.0\ngini = 0.868\nsamples = 1317\nvalue = [6, 395, 409, 233, 130, 201, 1, 179, 248, 186, 17\n57, 0, 12, 0]\nnclass = c'),
Text(2146.153846153846, 181.1999999999982, 'gini = 0.865\nsamples = 715\nvalue = [1, 245, 228, 121, 84, 79, 1, 98, 118, 118, 12\n38, 0, 9, 0]\nnclass = b'),
Text(2317.846153846154, 181.1999999999982, 'gini = 0.867\nsamples = 602\nvalue = [5, 150, 181, 112, 46, 122, 0, 81, 130, 68, 5, 19\n0, 3, 0]\nnclass = c'),
Text(2747.076923076923, 906.0, 'user_id <= 51.5\ngini = 0.87\nsamples = 822\nvalue = [0, 192, 314, 132, 88, 99, 1, 116, 141, 153, 26\n63, 0, 10, 0]\nnclass = c'),
Text(2575.3846153846152, 543.5999999999999, 'user_id <= 21.0\ngini = 0.877\nsamples = 643\nvalue = [0, 115, 224, 119, 56, 79, 1, 107, 118, 119, 23\n62, 0, 10, 0]\nnclass = c'),
Text(2489.5384615384614, 181.1999999999982, 'gini = 0.857\nsamples = 114\nvalue = [0, 37, 38, 11, 0, 15, 0, 6, 8, 27, 23, 33, 0\n0, 0]\nnclass = c'),
Text(2661.230769230769, 181.1999999999982, 'gini = 0.869\nsamples = 529\nvalue = [0, 78, 186, 108, 56, 64, 1, 101, 110, 92, 0, 29\n0, 10, 0]\nnclass = c'),
Text(2918.7692307692305, 543.5999999999999, 'row_id <= 17926.5\ngini = 0.809\nsamples = 179\nvalue = [0, 77, 90, 13, 32, 20, 0, 9, 23, 34, 3, 1, 0\n0, 0]\nnclass = c'),
Text(2832.9230769230767, 181.1999999999982, 'gini = 0.797\nsamples = 154\nvalue = [0, 75, 74, 4, 32, 16, 0, 9, 20, 25, 2, 1, 0\n0, 0]\nnclass = b'),
Text(3004.6153846153843, 181.1999999999982, 'gini = 0.769\nsamples = 25\nvalue = [0, 2, 16, 9, 0, 4, 0, 0, 3, 9, 1, 0, 0, 0\n0]\nnclass = c'),
Text(3777.230769230769, 1268.4, 'row_id <= 36135.0\ngini = 0.86\nsamples = 2320\nvalue = [5, 739, 773, 348, 178, 320, 9, 322, 509, 357, 25\n59, 0, 39, 1]\nnclass = c'),
Text(3433.8461538461534, 906.0, 'user_id <= 54.5\ngini = 0.861\nsamples = 2140\nvalue = [5, 671, 722, 329, 169, 285, 9, 302, 490, 338, 18\n58, 0, 33, 0]\nnclass = c'),
Text(3262.1538461538457, 543.5999999999999, 'row_id <= 22464.5\ngini = 0.861\nsamples = 1875\nvalue = [2, 550, 632, 303, 135, 239, 9, 289, 440, 299, 16\n51, 0, 26, 0]\nnclass = c'),
Text(3176.307692307692, 181.1999999999982, 'gini = 0.845\nsamples = 379\nvalue = [1, 13, 167, 42, 20, 39, 4, 60, 84, 61, 5, 12\n0, 8, 0]\nnclass = c'),
Text(3347.9999999999995, 181.1999999999982, 'gini = 0.864\nsamples = 1496\nvalue = [1, 437, 465, 261, 115, 200, 5, 229, 356, 238, 11\n39, 0, 18, 0]\nnclass = c'),
Text(3605.5384615384614, 543.5999999999999, 'user_id <= 55.5\ngini = 0.838\nsamples = 265\nvalue = [3, 121, 90, 26, 34, 46, 0, 13, 50, 39, 2, 7, 0\n7, 0]\nnclass = b'),
Text(3519.6923076923076, 181.1999999999982, 'gini = 0.838\nsamples = 231\nvalue = [0, 107, 72, 23, 31, 41, 0, 11, 50, 36, 0, 5, 0\n7, 0]\nnclass = b'),
Text(3691.3846153846152, 181.1999999999982, 'gini = 0.804\nsamples = 34\nvalue = [3, 14, 18, 3, 3, 5, 0, 2, 0, 3, 2, 2, 0, 0\n0]\nnclass = c'),
Text(4120.615384615385, 906.0, 'user_id <= 34.0\ngini = 0.845\nsamples = 180\nvalue = [0, 68, 51, 19, 9, 35, 0, 20, 19, 19, 7, 1, 0\n6, 1]\nnclass = b'),

```

```
Text(3948.9230769230767, 543.5999999999999, 'row_id <= 36654.0\nngini = 0.823\nsamples = 84\nnvalue = [0, 38, 18, 9, 4, 13, 0, 3, 13, 8, 4, 0, 0, 3\nn1]\nclass = b'),  
Text(3863.076923076923, 181.1999999999982, 'gini = 0.834\nsamples = 34\nnvalue = [0, 1  
0, 9, 7, 0, 8, 0, 1, 3, 3, 0, 0, 0, 0, 3\nn0]\nclass = b'),  
Text(4034.7692307692305, 181.1999999999982, 'gini = 0.784\nsamples = 50\nnvalue = [0, 2  
8, 9, 2, 4, 5, 0, 2, 10, 5, 4, 0, 0, 0\nn1]\nclass = b'),  
Text(4292.307692307692, 543.5999999999999, 'user_id <= 41.0\nngini = 0.846\nsamples = 96  
\nnvalue = [0, 30, 33, 10, 5, 22, 0, 17, 6, 11, 3, 1, 0\nn3, 0]\nclass = c'),  
Text(4206.461538461538, 181.1999999999982, 'gini = 0.858\nsamples = 42\nnvalue = [0, 5,  
14, 8, 4, 5, 0, 9, 5, 2, 3, 1, 0, 0\nn0]\nclass = c'),  
Text(4378.153846153846, 181.1999999999982, 'gini = 0.801\nsamples = 54\nnvalue = [0, 2  
5, 19, 2, 1, 17, 0, 8, 1, 9, 0, 0, 0, 0, 3\nn0]\nclass = b')]
```



In []: