In [81]:
```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

In [255]:
```python
df=pd.read_csv(r"C:\USERS\user\Downloads\C9_Data - C9_Data.csv")
```

Out[255]:

	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 [256]:

Out[256]: Index(['row_id', 'user_id', 'timestamp', 'gate_id'], dtype='object')

In [257]: `df=df.head(15)`

Out[257]:

	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
5	5	18	2022-07-29 09:10:34	10
6	6	18	2022-07-29 09:32:47	11
7	7	18	2022-07-29 09:33:12	4
8	8	18	2022-07-29 09:33:13	4
9	9	1	2022-07-29 09:33:16	7
10	10	18	2022-07-29 09:33:23	9
11	11	18	2022-07-29 09:33:23	9
12	12	18	2022-07-29 09:33:41	5
13	13	18	2022-07-29 09:33:42	5
14	14	18	2022-07-29 09:34:04	10

In [270]: `a=df[['row_id', 'user_id','gate_id']]`

Out[270]:

	row_id	user_id	gate_id
0	0	18	7
1	1	18	9
2	2	18	9
3	3	18	5
4	4	18	5
5	5	18	10
6	6	18	11
7	7	18	4
8	8	18	4
9	9	1	7
10	10	18	9
11	11	18	9
12	12	18	5
13	13	18	5
14	14	18	10

In [271]:

Out[271]:
```
18    14
1      1
Name: user_id, dtype: int64
```

In [272]: `x=a.drop('user_id',axis=1)`

In [273]:
```python
g1={"user_id":{'18':1,'1':2}}
a=a.replace(g1)
```

	row_id	user_id	gate_id
0	0	18	7
1	1	18	9
2	2	18	9
3	3	18	5
4	4	18	5
5	5	18	10
6	6	18	11
7	7	18	4
8	8	18	4
9	9	1	7
10	10	18	9
11	11	18	9
12	12	18	5
13	13	18	5
14	14	18	10

In [274]:
```python
from sklearn.model_selection import train_test_split
```

In [275]:
```python
from sklearn.ensemble import RandomForestClassifier

rfc=RandomForestClassifier()
```

Out[275]: RandomForestClassifier()

In [276]:
```python
parameters={'max_depth':[1,2,3,4,5],
            'min_samples_leaf':[5,10,15,20,25],
```

In [277]:
```python
from sklearn.model_selection import GridSearchCV

grid_search=GridSearchCV(estimator=rfc,param_grid=parameters,cv=2,scoring="acc
```

```
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[277]:
```
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')
```
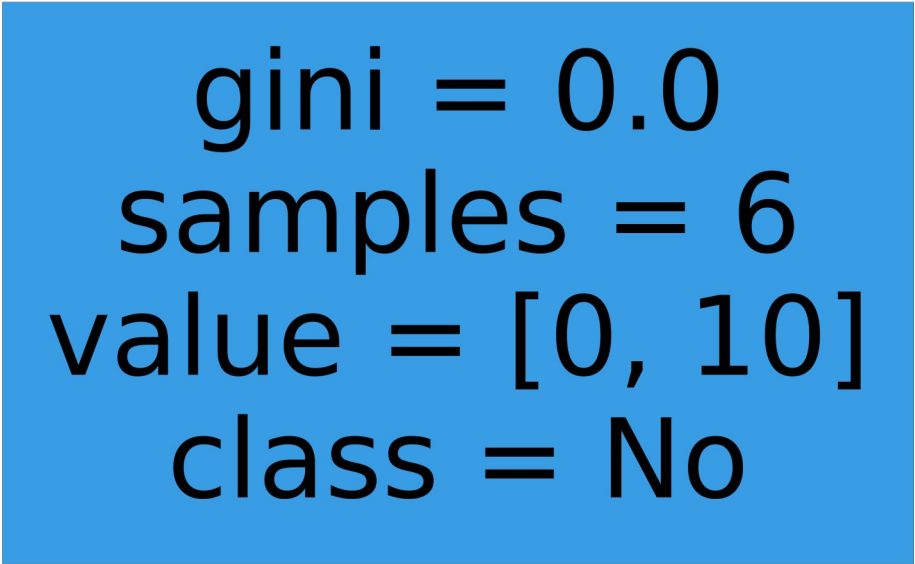
In [278]:

Out[278]: 0.9

In [279]:

In [280]:
```python
from sklearn.tree import plot_tree

plt.figure(figsize=(80,40))
```

Out[280]: [Text(2232.0, 1087.2, 'gini = 0.0\nsamples = 6\nvalue = [0, 10]\nclass = No
')]

In [ ]: