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

In [2]:
```
df=pd.read_csv(r"C1_ionosphere (1).csv")
df
```

Out[2]:

	1	0	0.99539	-0.05889	0.85243	0.02306	0.83398	-0.37708	1.1	0.03760	...	-0.511
0	1	0	1.00000	-0.18829	0.93035	-0.36156	-0.10868	-0.93597	1.00000	-0.04549	...	-0.265
1	1	0	1.00000	-0.03365	1.00000	0.00485	1.00000	-0.12062	0.88965	0.01198	...	-0.402
2	1	0	1.00000	-0.45161	1.00000	1.00000	0.71216	-1.00000	0.00000	0.00000	...	0.906
3	1	0	1.00000	-0.02401	0.94140	0.06531	0.92106	-0.23255	0.77152	-0.16399	...	-0.651
4	1	0	0.02337	-0.00592	-0.09924	-0.11949	-0.00763	-0.11824	0.14706	0.06637	...	-0.015
...	...	...	...	...	...	...	...	...	...	...	...	
345	1	0	0.83508	0.08298	0.73739	-0.14706	0.84349	-0.05567	0.90441	-0.04622	...	-0.042
346	1	0	0.95113	0.00419	0.95183	-0.02723	0.93438	-0.01920	0.94590	0.01606	...	0.013
347	1	0	0.94701	-0.00034	0.93207	-0.03227	0.95177	-0.03431	0.95584	0.02446	...	0.031
348	1	0	0.90608	-0.01657	0.98122	-0.01989	0.95691	-0.03646	0.85746	0.00110	...	-0.020
349	1	0	0.84710	0.13533	0.73638	-0.06151	0.87873	0.08260	0.88928	-0.09139	...	-0.151

350 rows × 35 columns

In [3]:
```
df['g'].value_counts()
```

Out[3]:
```
g    224
b    126
Name: g, dtype: int64
```

In [4]:
```
x=df.drop('g',axis=1)
y=df['g']
```

In [5]:
```
g1={"g":{'g':1,'b':2}}
df=df.replace(g1)
print(df)
```

```
..       ...  ...       ...      ...       ...      ...       ...      ...
345 -0.04622  ...  -0.04202  0.83479   0.00123  1.00000   0.12815  0.86660
346  0.01606  ...   0.01361  0.93522   0.04925  0.93159   0.08168  0.94066
347  0.02446  ...   0.03193  0.92489   0.02542  0.92120   0.02242  0.92459
348  0.00110  ...  -0.02099  0.89147  -0.07760  0.82983  -0.17238  0.96022
349 -0.09139  ...  -0.15114  0.81147  -0.04822  0.78207  -0.00703  0.75747

     -0.54487  0.18641  -0.45300  g
0    -0.06288 -0.13738  -0.02447  2
1    -0.24180  0.56045  -0.38238  1
2     1.00000 -0.32382   1.00000  2
3    -0.59573 -0.04608  -0.65697  1
4     0.00000 -0.00039   0.12011  2
..        ...      ...       ... ..
345  -0.10714  0.90546  -0.04307  1
346  -0.00035  0.91483   0.04712  1
347   0.00442  0.92697  -0.00577  1
348  -0.03757  0.87403  -0.16243  1
349  -0.06678  0.85764  -0.06151  1
```

In [6]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,train_size=0.70)
```

In [7]:
```
from sklearn.ensemble import RandomForestClassifier
rfc=RandomForestClassifier()
rfc.fit(x_train,y_train)
```

Out[7]:
```
RandomForestClassifier()
```

In [8]:
```
parameters={'max_depth':[1,2,3,4,5],
            'min_samples_leaf':[5,10,15,20,25],
            'n_estimators':[10,20,30,40,50]}
```

In [9]:
```
from sklearn.model_selection import GridSearchCV
grid_search=GridSearchCV(estimator=rfc,param_grid=parameters,cv=2,scoring="accu
grid_search.fit(x_train,y_train)
```

Out[9]:
```
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')
```

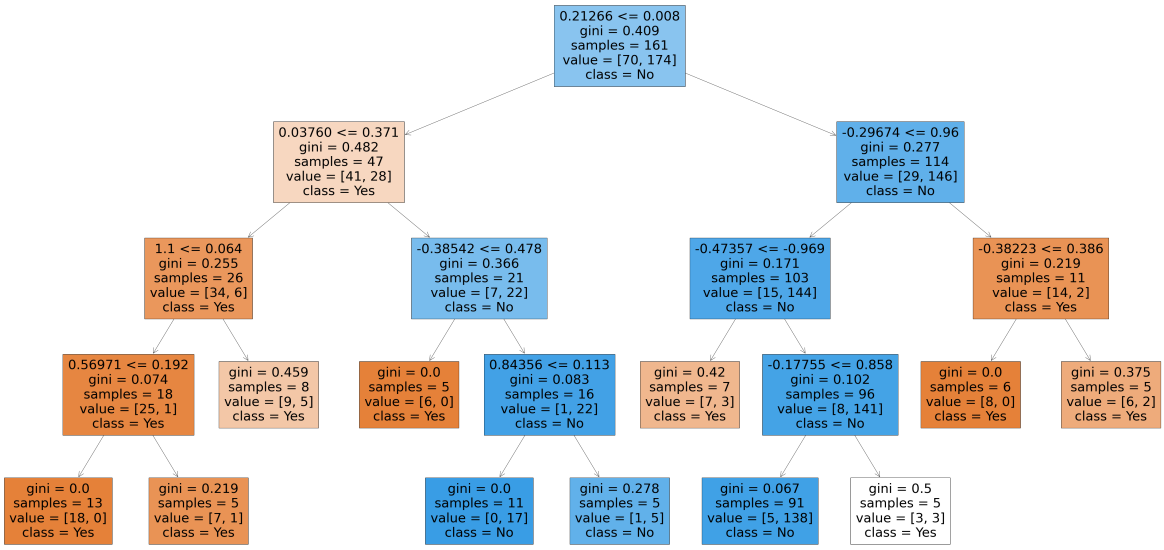
In [10]:
```
grid_search.best_score_
```

Out[10]:
```
0.930327868852459
```

In [11]:
```
rfc_best=grid_search.best_estimator_
```

In [12]:
```
from sklearn.tree import plot_tree
plt.figure(figsize=(80,40))
plot_tree(rfc_best.estimators_[5],feature_names=x.columns,class_names=['Yes','N
```

Out[12]:
```
[Text(2363.2941176470586, 1956.96, '0.21266 <= 0.008\ngini = 0.409\nsamples =
161\nvalue = [70, 174]\nclass = No'),
 Text(1312.941176470588, 1522.0800000000002, '0.03760 <= 0.371\ngini = 0.482
\nsamples = 47\nvalue = [41, 28]\nclass = Yes'),
 Text(787.7647058823529, 1087.2, '1.1 <= 0.064\ngini = 0.255\nsamples = 26\nv
alue = [34, 6]\nclass = Yes'),
 Text(525.1764705882352, 652.3200000000002, '0.56971 <= 0.192\ngini = 0.074\n
samples = 18\nvalue = [25, 1]\nclass = Yes'),
 Text(262.5882352941176, 217.44000000000005, 'gini = 0.0\nsamples = 13\nvalue
= [18, 0]\nclass = Yes'),
 Text(787.7647058823529, 217.44000000000005, 'gini = 0.219\nsamples = 5\nvalu
e = [7, 1]\nclass = Yes'),
 Text(1050.3529411764705, 652.3200000000002, 'gini = 0.459\nsamples = 8\nvalu
e = [9, 5]\nclass = Yes'),
 Text(1838.1176470588234, 1087.2, '-0.38542 <= 0.478\ngini = 0.366\nsamples =
21\nvalue = [7, 22]\nclass = No'),
 Text(1575.5294117647059, 652.3200000000002, 'gini = 0.0\nsamples = 5\nvalue
= [6, 0]\nclass = Yes'),
 Text(2100.705882352941, 652.3200000000002, '0.84356 <= 0.113\ngini = 0.083\n
samples = 16\nvalue = [1, 22]\nclass = No'),
 Text(1838.1176470588234, 217.44000000000005, 'gini = 0.0\nsamples = 11\nvalu
e = [0, 17]\nclass = No'),
 Text(2363.2941176470586, 217.44000000000005, 'gini = 0.278\nsamples = 5\nval
ue = [1, 5]\nclass = No'),
 Text(3413.6470588235293, 1522.0800000000002, '-0.29674 <= 0.96\ngini = 0.277
\nsamples = 114\nvalue = [29, 146]\nclass = No'),
 Text(2888.4705882352937, 1087.2, '-0.47357 <= -0.969\ngini = 0.171\nsamples
= 103\nvalue = [15, 144]\nclass = No'),
 Text(2625.882352941176, 652.3200000000002, 'gini = 0.42\nsamples = 7\nvalue
= [7, 3]\nclass = Yes'),
 Text(3151.0588235294117, 652.3200000000002, '-0.17755 <= 0.858\ngini = 0.102
\nsamples = 96\nvalue = [8, 141]\nclass = No'),
 Text(2888.4705882352937, 217.44000000000005, 'gini = 0.067\nsamples = 91\nva
lue = [5, 138]\nclass = No'),
 Text(3413.6470588235293, 217.44000000000005, 'gini = 0.5\nsamples = 5\nvalue
= [3, 3]\nclass = Yes'),
 Text(3938.8235294117644, 1087.2, '-0.38223 <= 0.386\ngini = 0.219\nsamples =
11\nvalue = [14, 2]\nclass = Yes'),
 Text(3676.235294117647, 652.3200000000002, 'gini = 0.0\nsamples = 6\nvalue =
[8, 0]\nclass = Yes'),
 Text(4201.411764705882, 652.3200000000002, 'gini = 0.375\nsamples = 5\nvalue
= [6, 2]\nclass = Yes')]
```


0.21266 <= 0.008
gini = 0.409
samples = 161
value = [70, 174]
class = No
0.03760 <= 0.371
gini = 0.482
samples = 47
value = [41, 28]
class = Yes
-0.29674 <= 0.96
gini = 0.277
samples = 114
value = [29, 146]
class = No
1.1 <= 0.064
gini = 0.255
samples = 26
value = [34, 6]
class = Yes
-0.38542 <= 0.478
gini = 0.366
samples = 21
value = [7, 22]
class = No
-0.47357 <= -0.969
gini = 0.171
samples = 103
value = [15, 144]
class = No
-0.38223 <= 0.386
gini = 0.219
samples = 11
value = [14, 2]
class = Yes
0.56971 <= 0.192
gini = 0.074
samples = 18
value = [25, 1]
class = Yes
gini = 0.459
samples = 8
value = [9, 5]
class = Yes
gini = 0.0
samples = 5
value = [6, 0]
class = Yes
0.84356 <= 0.113
gini = 0.083
samples = 16
value = [1, 22]
class = No
gini = 0.42
samples = 7
value = [7, 3]
class = Yes
-0.17755 <= 0.858
gini = 0.102
samples = 96
value = [8, 141]
class = No
gini = 0.0
samples = 6
value = [8, 0]
class = Yes
gini = 0.375
samples = 5
value = [6, 2]
class = Yes
gini = 0.0
samples = 13
value = [18, 0]
class = Yes
gini = 0.219
samples = 5
value = [7, 1]
class = Yes
gini = 0.0
samples = 11
value = [0, 17]
class = No
gini = 0.278
samples = 5
value = [1, 5]
class = No
gini = 0.067
samples = 91
value = [5, 138]
class = No
gini = 0.5
samples = 5
value = [3, 3]
class = Yes

In [ ]: