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"1_ionosphere.csv")
df
```

Out[2]:

-0.05889	0.85243	0.02306	0.83398	-0.37708	1.1	0.03760	...	-0.51171	0.41078	-0.46168
-0.18829	0.93035	-0.36156	-0.10868	-0.93597	1.00000	-0.04549	...	-0.26569	-0.20468	-0.18401
-0.03365	1.00000	0.00485	1.00000	-0.12062	0.88965	0.01198	...	-0.40220	0.58984	-0.22145
-0.45161	1.00000	1.00000	0.71216	-1.00000	0.00000	0.00000	...	0.90695	0.51613	1.00000
-0.02401	0.94140	0.06531	0.92106	-0.23255	0.77152	-0.16399	...	-0.65158	0.13290	-0.53206
-0.00592	-0.09924	-0.11949	-0.00763	-0.11824	0.14706	0.06637	...	-0.01535	-0.03240	0.09223
...	...	...	...	...	...	...	...	...	...	..
0.08298	0.73739	-0.14706	0.84349	-0.05567	0.90441	-0.04622	...	-0.04202	0.83479	0.00123
0.00419	0.95183	-0.02723	0.93438	-0.01920	0.94590	0.01606	...	0.01361	0.93522	0.04925
-0.00034	0.93207	-0.03227	0.95177	-0.03431	0.95584	0.02446	...	0.03193	0.92489	0.02542
-0.01657	0.98122	-0.01989	0.95691	-0.03646	0.85746	0.00110	...	-0.02099	0.89147	-0.07760
0.13533	0.73638	-0.06151	0.87873	0.08260	0.88928	-0.09139	...	-0.15114	0.81147	-0.04822

ıns

In [4]:
```
df['g'].value_counts()
```

Out[4]:
```
g    224
b    126
Name: g, dtype: int64
```

In [5]:
```
x=df.drop('g',axis=1)
y=df['g']
```

In [6]:
```
g1={"g":{'g':1,'b':2}}
df=df.replace(g1)
print(df)
```

```
     1  0  0.99539  -0.05889  0.85243  0.02306  0.83398  -0.37708      1.1
\
0    1  0  1.00000  -0.18829  0.93035 -0.36156 -0.10868  -0.93597  1.00000
1    1  0  1.00000  -0.03365  1.00000  0.00485  1.00000  -0.12062  0.88965
2    1  0  1.00000  -0.45161  1.00000  1.00000  0.71216  -1.00000  0.00000
3    1  0  1.00000  -0.02401  0.94140  0.06531  0.92106  -0.23255  0.77152
4    1  0  0.02337  -0.00592 -0.09924 -0.11949 -0.00763  -0.11824  0.14706
..  .. ..      ...       ...      ...      ...      ...       ...      ...
345  1  0  0.83508   0.08298  0.73739 -0.14706  0.84349  -0.05567  0.90441
346  1  0  0.95113   0.00419  0.95183 -0.02723  0.93438  -0.01920  0.94590
347  1  0  0.94701  -0.00034  0.93207 -0.03227  0.95177  -0.03431  0.95584
348  1  0  0.90608  -0.01657  0.98122 -0.01989  0.95691  -0.03646  0.85746
349  1  0  0.84710   0.13533  0.73638 -0.06151  0.87873   0.08260  0.88928

     0.03760  ...  -0.51171  0.41078  -0.46168  0.21266  -0.34090  0.42267
\
0   -0.04549  ...  -0.26569 -0.20468  -0.18401 -0.19040  -0.11593 -0.16626
1    0.01198  ...  -0.40220  0.58984  -0.22145  0.43100  -0.17365  0.60436
2    0.00000  ...   0.90695  0.51613   1.00000  1.00000  -0.20099  0.25682
```

In [38]:
```
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 [39]:
```
from sklearn.ensemble import RandomForestClassifier
rfc=RandomForestClassifier()
rfc.fit(x_train,y_train)
```

Out[39]:
```
RandomForestClassifier()
```

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

In [51]:
```
from sklearn.model_selection import GridSearchCV
grid_search=GridSearchCV(estimator=rfc,param_grid=parameters,cv=2,scoring="acc
grid_search.fit(x_train,y_train)
```

Out[51]:
```
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')
```

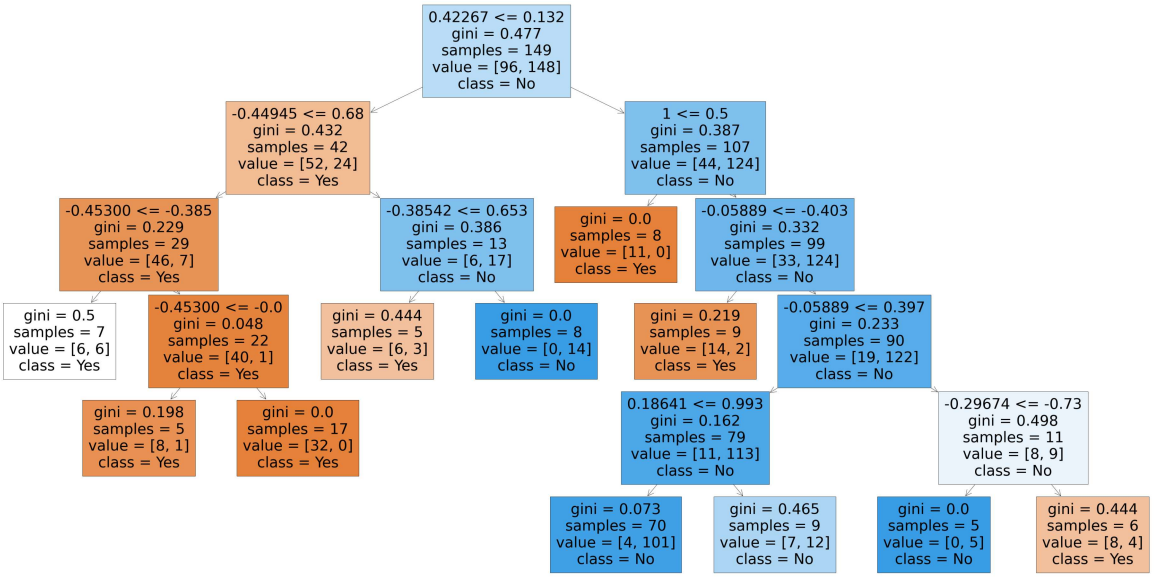
In [52]:
```
grid_search.best_score_
```

Out[52]:
```
0.930327868852459
```

In [53]:
```
rfc_best=grid_search.best_estimator_
```

In [54]:
```
from sklearn.tree import plot_tree
plt.figure(figsize=(80,40))
plot_tree(rfc_best.estimators_[5],feature_names=x.columns,class_names=['Yes','
```

Out[54]:
```
[Text(1934.4, 1993.2, '0.42267 <= 0.132\ngini = 0.477\nsamples = 149\nvalue =
[96, 148]\nclass = No'),
 Text(1190.4, 1630.8000000000002, '-0.44945 <= 0.68\ngini = 0.432\nsamples =
42\nvalue = [52, 24]\nclass = Yes'),
 Text(595.2, 1268.4, '-0.45300 <= -0.385\ngini = 0.229\nsamples = 29\nvalue =
[46, 7]\nclass = Yes'),
 Text(297.6, 906.0, 'gini = 0.5\nsamples = 7\nvalue = [6, 6]\nclass = Yes'),
 Text(892.8000000000001, 906.0, '-0.45300 <= -0.0\ngini = 0.048\nsamples = 22
\nvalue = [40, 1]\nclass = Yes'),
 Text(595.2, 543.5999999999999, 'gini = 0.198\nsamples = 5\nvalue = [8, 1]\nc
lass = Yes'),
 Text(1190.4, 543.5999999999999, 'gini = 0.0\nsamples = 17\nvalue = [32, 0]\n
class = Yes'),
 Text(1785.6000000000001, 1268.4, '-0.38542 <= 0.653\ngini = 0.386\nsamples =
13\nvalue = [6, 17]\nclass = No'),
 Text(1488.0, 906.0, 'gini = 0.444\nsamples = 5\nvalue = [6, 3]\nclass = Ye
s'),
 Text(2083.2000000000003, 906.0, 'gini = 0.0\nsamples = 8\nvalue = [0, 14]\nc
lass = No'),
 Text(2678.4, 1630.8000000000002, '1 <= 0.5\ngini = 0.387\nsamples = 107\nval
ue = [44, 124]\nclass = No'),
 Text(2380.8, 1268.4, 'gini = 0.0\nsamples = 8\nvalue = [11, 0]\nclass = Ye
s'),
 Text(2976.0, 1268.4, '-0.05889 <= -0.403\ngini = 0.332\nsamples = 99\nvalue
= [33, 124]\nclass = No'),
 Text(2678.4, 906.0, 'gini = 0.219\nsamples = 9\nvalue = [14, 2]\nclass = Ye
s'),
 Text(3273.6000000000004, 906.0, '-0.05889 <= 0.397\ngini = 0.233\nsamples =
90\nvalue = [19, 122]\nclass = No'),
 Text(2678.4, 543.5999999999999, '0.18641 <= 0.993\ngini = 0.162\nsamples = 7
9\nvalue = [11, 113]\nclass = No'),
 Text(2380.8, 181.19999999999982, 'gini = 0.073\nsamples = 70\nvalue = [4, 10
1]\nclass = No'),
 Text(2976.0, 181.19999999999982, 'gini = 0.465\nsamples = 9\nvalue = [7, 12]
\nclass = No'),
 Text(3868.8, 543.5999999999999, '-0.29674 <= -0.73\ngini = 0.498\nsamples =
11\nvalue = [8, 9]\nclass = No'),
 Text(3571.2000000000003, 181.19999999999982, 'gini = 0.0\nsamples = 5\nvalue
= [0, 5]\nclass = No'),
 Text(4166.400000000001, 181.19999999999982, 'gini = 0.444\nsamples = 6\nvalu
e = [8, 4]\nclass = Yes')]
```


0.42267 <= 0.132
gini = 0.477
samples = 149
value = [96, 148]
class = No
-0.44945 <= 0.68
gini = 0.432
samples = 42
value = [52, 24]
class = Yes
1 <= 0.5
gini = 0.387
samples = 107
value = [44, 124]
class = No
-0.45300 <= -0.385
gini = 0.229
samples = 29
value = [46, 7]
class = Yes
-0.38542 <= 0.653
gini = 0.386
samples = 13
value = [6, 17]
class = No
gini = 0.0
samples = 8
value = [11, 0]
class = Yes
-0.05889 <= -0.403
gini = 0.332
samples = 99
value = [33, 124]
class = No
gini = 0.5
samples = 7
value = [6, 6]
class = Yes
-0.45300 <= -0.0
gini = 0.048
samples = 22
value = [40, 1]
class = Yes
gini = 0.444
samples = 5
value = [6, 3]
class = Yes
gini = 0.0
samples = 8
value = [0, 14]
class = No
gini = 0.219
samples = 9
value = [14, 2]
class = Yes
-0.05889 <= 0.397
gini = 0.233
samples = 90
value = [19, 122]
class = No
gini = 0.198
samples = 5
value = [8, 1]
class = Yes
gini = 0.0
samples = 17
value = [32, 0]
class = Yes
0.18641 <= 0.993
gini = 0.162
samples = 79
value = [11, 113]
class = No
-0.29674 <= -0.73
gini = 0.498
samples = 11
value = [8, 9]
class = No
gini = 0.073
samples = 70
value = [4, 101]
class = No
gini = 0.465
samples = 9
value = [7, 12]
class = No
gini = 0.0
samples = 5
value = [0, 5]
class = No
gini = 0.444
samples = 6
value = [8, 4]
class = Yes

In [ ]: