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\C6_bmi - C6_bmi.csv")
df
```

Out[2]:

	Gender	Height	Weight	Index
0	Male	174	96	4
1	Male	189	87	2
2	Female	185	110	4
3	Female	195	104	3
4	Male	149	61	3
...	...	...	...	...
495	Female	150	153	5
496	Female	184	121	4
497	Female	141	136	5
498	Male	150	95	5
499	Male	173	131	5

500 rows × 4 columns

In [3]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 500 entries, 0 to 499
Data columns (total 4 columns):
 #   Column  Non-Null Count  Dtype 
---  ------  --------------  ----- 
 0   Gender  500 non-null    object
 1   Height  500 non-null    int64 
 2   Weight  500 non-null    int64 
 3   Index   500 non-null    int64 
dtypes: int64(3), object(1)
memory usage: 15.8+ KB
```

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

Out[5]:
```
5    198
4    130
2     69
3     68
1     22
0     13
Name: Index, dtype: int64
```

In [15]:
```
df1=df[['Height','Weight','Index']]
```

In [58]:
```
x=df1[['Height','Weight','Index']]
y=df1['Index']
```

In [59]:
```
g1={'Index':{'5':0,'4':1,'3':2,'2':3,'1':4,'0':5}}
df1=df1.replace(g1)
print(df1)
```

```
     Height  Weight  Index
0       174      96      4
1       189      87      2
2       185     110      4
3       195     104      3
4       149      61      3
..      ...     ...    ...
495     150     153      5
496     184     121      4
497     141     136      5
498     150      95      5
499     173     131      5

[500 rows x 3 columns]
```

In [60]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=45)
```

In [61]:
```
from sklearn.ensemble import RandomForestClassifier
rfc = RandomForestClassifier()
rfc.fit(x_train,y_train)
```

Out[61]:
```
RandomForestClassifier()
```

In [62]:
```
parameters = {'max_depth':[1,2,3,4,5],
    'min_samples_leaf':[5,10,15,20,25],
    'n_estimators':[10,20,30,40,50]}
```

In [63]:

```
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[63]:

```
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 [64]:

```
grid_search.best_score_
```

Out[64]:

```
0.9538797434113919
```

In [65]:

```
rfc_best = grid_search.best_estimator_
```

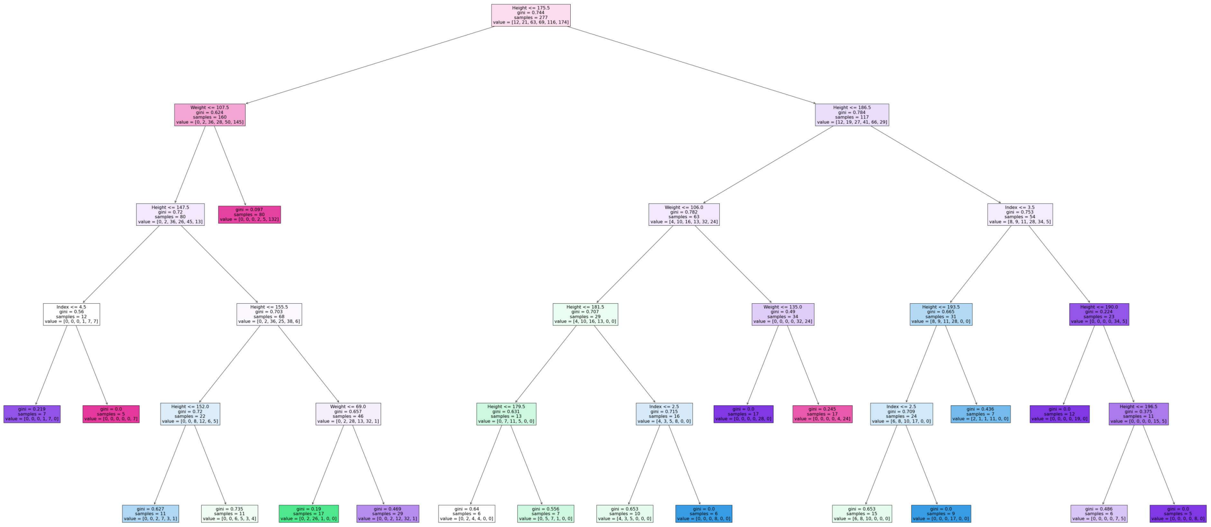
In [69]:

```
from sklearn.tree import plot_tree

plt.figure(figsize=(80,40))
plot_tree(rfc_best.estimators_[4],feature_names=x.columns,filled=True)
```

Out[69]:

```
[Text(1962.0, 1993.2, 'Height <= 175.5\ngini = 0.744\nsamples = 277\nvalue = [1
2, 21, 63, 69, 116, 174]'),
 Text(792.0, 1630.8000000000002, 'Weight <= 107.5\ngini = 0.624\nsamples = 160
\nvalue = [0, 2, 36, 28, 50, 145]'),
 Text(648.0, 1268.4, 'Height <= 147.5\ngini = 0.72\nsamples = 80\nvalue = [0,
2, 36, 26, 45, 13]'),
 Text(288.0, 906.0, 'Index <= 4.5\ngini = 0.56\nsamples = 12\nvalue = [0, 0, 0,
1, 7, 7]'),
 Text(144.0, 543.5999999999999, 'gini = 0.219\nsamples = 7\nvalue = [0, 0, 0,
1, 7, 0]'),
 Text(432.0, 543.5999999999999, 'gini = 0.0\nsamples = 5\nvalue = [0, 0, 0, 0,
0, 7]'),
 Text(1008.0, 906.0, 'Height <= 155.5\ngini = 0.703\nsamples = 68\nvalue = [0,
2, 36, 25, 38, 6]'),
 Text(720.0, 543.5999999999999, 'Height <= 152.0\ngini = 0.72\nsamples = 22\nva
lue = [0, 0, 8, 12, 6, 5]'),
 Text(576.0, 181.19999999999982, 'gini = 0.627\nsamples = 11\nvalue = [0, 0, 2,
7, 3, 1]'),
 Text(864.0, 181.19999999999982, 'gini = 0.735\nsamples = 11\nvalue = [0, 0, 6,
5, 3, 4]'),
 Text(1296.0, 543.5999999999999, 'Weight <= 69.0\ngini = 0.657\nsamples = 46\nv
alue = [0, 2, 28, 13, 32, 1]'),
 Text(1152.0, 181.19999999999982, 'gini = 0.19\nsamples = 17\nvalue = [0, 2, 2
6, 1, 0, 0]'),
 Text(1440.0, 181.19999999999982, 'gini = 0.469\nsamples = 29\nvalue = [0, 0,
2, 12, 32, 1]'),
 Text(936.0, 1268.4, 'gini = 0.097\nsamples = 80\nvalue = [0, 0, 0, 2, 5, 13
2]'),
 Text(3132.0, 1630.8000000000002, 'Height <= 186.5\ngini = 0.784\nsamples = 117
\nvalue = [12, 19, 27, 41, 66, 29]'),
 Text(2520.0, 1268.4, 'Weight <= 106.0\ngini = 0.782\nsamples = 63\nvalue = [4,
10, 16, 13, 32, 24]'),
 Text(2160.0, 906.0, 'Height <= 181.5\ngini = 0.707\nsamples = 29\nvalue = [4,
10, 16, 13, 0, 0]'),
 Text(1872.0, 543.5999999999999, 'Height <= 179.5\ngini = 0.631\nsamples = 13\n
value = [0, 7, 11, 5, 0, 0]'),
 Text(1728.0, 181.19999999999982, 'gini = 0.64\nsamples = 6\nvalue = [0, 2, 4,
4, 0, 0]'),
 Text(2016.0, 181.19999999999982, 'gini = 0.556\nsamples = 7\nvalue = [0, 5, 7,
1, 0, 0]'),
 Text(2448.0, 543.5999999999999, 'Index <= 2.5\ngini = 0.715\nsamples = 16\nval
ue = [4, 3, 5, 8, 0, 0]'),
 Text(2304.0, 181.19999999999982, 'gini = 0.653\nsamples = 10\nvalue = [4, 3,
5, 0, 0, 0]'),
 Text(2592.0, 181.19999999999982, 'gini = 0.0\nsamples = 6\nvalue = [0, 0, 0,
8, 0, 0]'),
 Text(2880.0, 906.0, 'Weight <= 135.0\ngini = 0.49\nsamples = 34\nvalue = [0,
0, 0, 0, 32, 24]'),
 Text(2736.0, 543.5999999999999, 'gini = 0.0\nsamples = 17\nvalue = [0, 0, 0,
0, 28, 0]'),
 Text(3024.0, 543.5999999999999, 'gini = 0.245\nsamples = 17\nvalue = [0, 0, 0,
0, 4, 24]'),
 Text(3744.0, 1268.4, 'Index <= 3.5\ngini = 0.753\nsamples = 54\nvalue = [8, 9,
11, 28, 34, 5]'),
 Text(3456.0, 906.0, 'Height <= 193.5\ngini = 0.665\nsamples = 31\nvalue = [8,
9, 11, 28, 0, 0]'),
 Text(3312.0, 543.5999999999999, 'Index <= 2.5\ngini = 0.709\nsamples = 24\nval
ue = [6, 8, 10, 17, 0, 0]'),
 Text(3168.0, 181.19999999999982, 'gini = 0.653\nsamples = 15\nvalue = [6, 8, 1
0, 0, 0, 0]'),
 Text(3456.0, 181.19999999999982, 'gini = 0.0\nsamples = 9\nvalue = [0, 0, 0, 1
7, 0, 0]'),
 Text(3600.0, 543.5999999999999, 'gini = 0.436\nsamples = 7\nvalue = [2, 1, 1,
11, 0, 0]'),
 Text(4032.0, 906.0, 'Height <= 190.0\ngini = 0.224\nsamples = 23\nvalue = [0,
0, 0, 0, 34, 5]'),
 Text(3888.0, 543.5999999999999, 'gini = 0.0\nsamples = 12\nvalue = [0, 0, 0,
0, 19, 0]'),
 Text(4176.0, 543.5999999999999, 'Height <= 196.5\ngini = 0.375\nsamples = 11\n
value = [0, 0, 0, 0, 15, 5]'),
 Text(4032.0, 181.19999999999982, 'gini = 0.486\nsamples = 6\nvalue = [0, 0, 0,
0, 7, 5]'),
 Text(4320.0, 181.19999999999982, 'gini = 0.0\nsamples = 5\nvalue = [0, 0, 0,
0, 8, 0]')]
```



In [ ]: