

In [ ]: Machine learning projects Steps:

```
1-Import data
2-clean the data
3- Split the Data into Training/Testing Sets
4-Select a model to analyze the data
5-Train Model
6-Make prediction
7-Evaluate and improve
```

In [1]: `import os` #OS module in Python provides functions for interacting with the operating sys  
`os.getcwd()` # method to get current working directory  
`os.chdir(r"C:\Users\HP\Downloads\ML Tut")` # method in Python used to change the current

In [2]: `#reading the data`  
`import pandas as pd`  
`music_data = pd.read_csv('music.csv')`  
`music_data`

Out[2]:

	age	gender	genre
0	20	1	HipHop
1	23	1	HipHop
2	25	1	HipHop
3	26	1	Jazz
4	29	1	Jazz
5	30	1	Jazz
6	31	1	Classical
7	33	1	Classical
8	37	1	Classical
9	20	0	Dance
10	21	0	Dance
11	25	0	Dance
12	26	0	Acoustic
13	27	0	Acoustic
14	30	0	Acoustic
15	31	0	Classical
16	34	0	Classical
17	35	0	Classical

In [3]: `# we will remove the genre col as it is the output col used for prediction after trainin`  
`x = music_data.drop(columns=['genre'])` # input data set  
`x`

Out[3]:

	age	gender
0	20	1
1	23	1

<b>2</b>	25	1
<b>3</b>	26	1
<b>4</b>	29	1
<b>5</b>	30	1
<b>6</b>	31	1
<b>7</b>	33	1
<b>8</b>	37	1
<b>9</b>	20	0
<b>10</b>	21	0
<b>11</b>	25	0
<b>12</b>	26	0
<b>13</b>	27	0
<b>14</b>	30	0
<b>15</b>	31	0
<b>16</b>	34	0
<b>17</b>	35	0

```
In [4]: y = music_data['genre'] # output col
        y
```

```
Out[4]: 0      HipHop
        1      HipHop
        2      HipHop
        3      Jazz
        4      Jazz
        5      Jazz
        6      Classical
        7      Classical
        8      Classical
        9      Dance
       10      Dance
       11      Dance
       12      Acoustic
       13      Acoustic
       14      Acoustic
       15      Classical
       16      Classical
       17      Classical
Name: genre, dtype: object
```

```
In [5]: # Building a model algorithm used decision tree
        from sklearn.tree import DecisionTreeClassifier

        model = DecisionTreeClassifier()
        model.fit(x,y)
        prediction = model.predict([ [21,1],[22,0] ]) #2 dim array as the x (input dataset)
        prediction
```

C:\Users\HP\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning: X does not have valid feature names, but DecisionTreeClassifier was fitted with feature names  
warnings.warn(

```
Out[5]: array(['HipHop', 'Dance'], dtype=object)
```

```
In [14]: # Measuring accuracy of a model  
from sklearn.model_selection import train_test_split  
from sklearn.metrics import accuracy_score # to calculate the accuracy  
x_train, x_test, y_train, y_test = train_test_split(x,y, test_size =0.2)
```

```
In [15]: model.fit(x_train , y_train)  
prediction = model.predict(x_test)  
  
accuracy_score(y_test , prediction)
```

Out[15]: 0.75

```
In [8]: import joblib # for saving and loading models  
joblib.dump(model, 'music-recommender.joblib') # to save model in the directiry  
# joblib.load to load model
```

Out[8]: ['music-recommender.joblib']

In [ ]: