```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
import pandas as pd
df_tennis = pd.read_csv("/content/drive/MyDrive/Personal/Studies/MSC Data Science Material/SEM2/ML/Practical/data_set/PlayTennis.csv")
print("\n Given Play Tennis Data Set: \n\n", df_tennis)
```

```
 Given Play Tennis Data Set:

     Outlook Temperature Humidity    Wind PlayTennis
0      Sunny         Hot     High    Weak         No
1      Sunny         Hot     High  Strong         No
2   Overcast         Hot     High    Weak        Yes
3       Rain        Mild     High    Weak        Yes
4       Rain        Cool   Normal    Weak        Yes
5       Rain        Cool   Normal  Strong         No
6   Overcast        Cool   Normal  Strong        Yes
7      Sunny        Mild     High    Weak         No
8      Sunny        Cool   Normal    Weak        Yes
9       Rain        Mild   Normal    Weak        Yes
10     Sunny        Mild   Normal  Strong        Yes
11  Overcast        Mild     High  Strong        Yes
12  Overcast         Hot   Normal    Weak        Yes
13      Rain        Mild     High  Strong         No
```

```
#Function to calculate the entropy of probability of observations

def entropy(probs):
  import math
  return sum( [-prob*math.log(prob,2) for prob in probs])

#Function to calculate the entropy of given dataset
def entropy_of_list(a_list):
  #print("A-list", a_list)
  from collections import Counter
  cnt = Counter(x for x in a_list)   #Counter calculates proportion of class
  num_instances = len(a_list)*1.0    #=14

  print("\n Number of Instances of the Current Sub Class is {0}:".format(num_instances))
  probs = [x / num_instances for x in cnt.values()]    #x means no of YES/NO
  print("\n Classes: ",min(cnt), max(cnt))
  print("\n Probabilities of Class {0} is {1}: ".format(min(cnt),min(probs)))
  print("\n Probabilities of Class {0} is {1}: ".format(max(cnt),max(probs)))
  return entropy (probs)

#The initial entropy of the YES/NO attribute for datasets
print("\n Input Dataset for entropy calculation:\n",df_tennis['PlayTennis'])

total_entropy = entropy_of_list(df_tennis['PlayTennis'])

print("\n Total Entropy of Play Tennis Dataset: ", total_entropy)
```

```
 Input Dataset for entropy calculation:
 0      No
1      No
2     Yes
3     Yes
4     Yes
5      No
6     Yes
7      No
8     Yes
9     Yes
10    Yes
11    Yes
12    Yes
13     No
Name: PlayTennis, dtype: object

 Number of Instances of the Current Sub Class is 14.0:

 Classes:  No Yes

 Probabilities of Class No is 0.35714285714285715:

 Probabilities of Class Yes is 0.6428571428571429:

 Total Entropy of Play Tennis Dataset:  0.9402859586706309
```

✓ 0s completed at 11:45 PM