



```
In [20]: import numpy as np
import pandas as pd
import csv
import math
```

```
In [16]: data = pd.read_csv("Weather-D.csv")
print(data)
```

	Outlook	Temperature	Humidity	Windy	Play Football
0	Sunny	Hot	High	Weak	No
1	Sunny	Hot	High	Strong	No
2	Overcast	Hot	High	Weak	Yes
3	Rainy	Mild	High	Weak	Yes
4	Rainy	Cool	Normal	Weak	Yes
5	Rainy	Cool	Normal	Strong	No
6	Overcast	Cool	Normal	Strong	Yes
7	Sunny	Mild	High	Weak	No
8	Sunny	Cool	Normal	Weak	Yes
9	Rainy	Mild	Normal	Weak	Yes
10	Sunny	Mild	Normal	Strong	Yes
11	Overcast	Mild	High	Strong	Yes
12	Overcast	Hot	Normal	Weak	Yes
13	Rainy	Mild	High	Strong	No

```
In [12]: Sunny = len(data[data["Outlook"] == "Sunny" ])
Sunny_Y = len(data[(data["Outlook"] == "Sunny") & (data["Play Football"] == "Y
Sunny_N = len(data[(data["Outlook"] == "Sunny") & (data["Play Football"] == "N
print("Sunny: ", Sunny, "Sunny_Y: ", Sunny_Y, "Sunny_N: ", Sunny_N)
```

```
Sunny: 5 Sunny_Y: 2 Sunny_N: 3
```

```
In [11]: Overcast = len(data[data['Outlook'] == 'Overcast'])
Overcast_Y = len(data[(data['Outlook'] == 'Overcast') & (data['Play Football']
Overcast_N = len(data[(data['Outlook'] == 'Overcast') & (data['Play Football']
print("Overcast: ", Overcast, " ", "Overcast_Y: ", Overcast_Y, " ", "Overcast_
```

```
Overcast: 4 Overcast_Y: 4 Overcast_N: 0
```

```
In [18]: Rainy = len(data[data['Outlook'] == 'Rainy'])
Rainy_Y = len(data[(data['Outlook'] == 'Rainy') & (data['Play Football'] == 'Y
Rainy_N = len(data[(data['Outlook'] == 'Rainy') & (data['Play Football'] == 'N
print("Rainy: ", Rainy, " ", "Rainy_Y: ", Rainy_Y, " ", "Rainy_N: ", Rainy_N)
```

```
Rainy: 5 Rainy_Y: 3 Rainy_N: 2
```

```
In [19]: total_yes = len(data[data["Play Football"] == "Yes"])
total_no = len(data[data["Play Football"] == "No"])
total = total_yes + total_no
```

```
In [21]: def entropy(pos, neg):
total = pos + neg
if total == 0 or pos == 0 or neg == 0:
return 0
p_pos = pos/total
p_neg = neg/total
```

```
return -p_pos * math.log2(p_pos) - p_neg * math.log2(p_neg)
```

```
In [22]: entropy_total = entropy(total_yes, total_no)
print(entropy_total)
```

```
0.9402859586706311
```

```
In [24]: entropy_sunny = entropy(Sunny_Y, Sunny_N)
print(entropy_sunny)
entropy_overcast = entropy(Overcast_Y, Overcast_N)
print(entropy_overcast)
entropy_rainy = entropy(Rainy_Y, Rainy_N)
print(entropy_rainy)
```

```
0.9709505944546686
```

```
0
```

```
0.9709505944546686
```

```
In [25]: weight_sunny = Sunny/total
weight_rainy = Rainy/total
weight_overcast = Overcast/total
print(weight_sunny, weight_overcast, weight_rainy)
```

```
0.35714285714285715 0.2857142857142857 0.35714285714285715
```

```
In [26]: weighted_entropy_outlook = (weight_sunny * entropy_sunny) + (weight_overcast *
print(weighted_entropy_outlook)
```

```
0.6935361388961918
```

```
In [27]: gain_outlook = entropy_total - weighted_entropy_outlook
print(gain_outlook)
```

```
0.24674981977443933
```