MACHINE INTELLIGENCE LABORATORY

WEEK – 3

NAME : SANMAT SANJAYAKUMAR PAYAGOUDAR

SRN : PES1UG20CS385

CODE :

PES1UG20CS385.py

```
import numpy as np
import pandas as pd
import random


def get_entropy_of_dataset(df):
 entropy = 0
 target = df[[df.columns[-1]]].values
 _, counts = np.unique(target, return_counts=True)
 total_count = np.sum(counts)
 for freq in counts:
    temp = freq / total_count
    if temp != 0:
        entropy -= temp * (np.log2(temp))
 return entropy


def get_avg_info_of_attribute(df, attribute):
 attribute_values = df[attribute].values
 unique_attribute_values = np.unique(attribute_values)
 rows = df.shape[0]
 entropy_of_attribute = 0
 for current_value in unique_attribute_values:
     df_slice = df[df[attribute] == current_value]
     target = df_slice[[df_slice.columns[-1]]].values
     _, counts = np.unique(target, return_counts=True)
     total_count = np.sum(counts)
     entropy = 0
     for freq in counts:
        temp = freq / total_count
        if temp != 0:
             entropy -= temp * np.log2(temp)
     entropy_of_attribute += entropy * (np.sum(counts) / rows)
 return (abs(entropy_of_attribute))
```

```python
def get_information_gain(df, attribute):
 information_gain = 0
 entropy_of_attribute = get_avg_info_of_attribute(df, attribute)
 entropy_of_dataset = get_entropy_of_dataset(df)
 information_gain = entropy_of_dataset - entropy_of_attribute
 return information_gain


def get_selected_attribute(df):
 information_gains = {}
 selected_column = ''
 max_information_gain = float("-inf")
 for attribute in df.columns[:-1]:
    information_gain_of_attribute = get_information_gain(df, attribute)
    if information_gain_of_attribute > max_information_gain:
        selected_column = attribute
        max_information_gain = information_gain_of_attribute
    information_gains[attribute] = information_gain_of_attribute
 return (information_gains, selected_column)
```

OUTPUT :

```
Test Case 4 for the function get_selected_attribute PASSED
PS C:\Users\Sanam\Desktop\Study\MI\Lab\Week 3>  c:; cd 'c:\Users\Sanam\Desktop\Study\MI\Lab\Week 3'; & 'C:\P
ython310\python.exe' 'c:\Users\Sanam\.vscode\extensions\ms-python.python-2022.14.0\pythonFiles\lib\python\de
bugpy\adapter/../..\debugpy\launcher' '61913' '--' 'c:\Users\Sanam\Desktop\Study\MI\Lab\Week 3\PES1UG20CS385
.py'
PS C:\Users\Sanam\Desktop\Study\MI\Lab\Week 3> python SampleTest.py --SRN PES1UG20CS385
Test Case 1 for the function get_entropy_of_dataset PASSED
Test Case 2 for the function get_avg_info_of_attribute PASSED
Test Case 3 for the function get_avg_info_of_attribute PASSED
Test Case 4 for the function get_selected_attribute PASSED
PS C:\Users\Sanam\Desktop\Study\MI\Lab\Week 3>
```