Ex No. 2 **1905121**

07/03/2022

Naïve Bayes Classification

Aim:

To construct a Naïve Bayes Classifier for the following Problem Statement.

Consider the following training data set given below. The data tuples all together contribute to the class C1 as mammals and C2 as non-mammals. Consider the following tuple combination as interest of classification.

X = (Give birth= yes, Can fly= no, lives in water = yes, have legs= no)

Y = (Give birth= no, Can fly= yes, lives in water = no, have legs= yes)

Z = ((Give birth= yes, Can fly= no, lives in water = no, have legs= yes)

Code:

In [1]:
```
import pandas as pd
```

In [2]:
```
df = pd.read_csv("animals_classification.csv")
```

In [3]:
```
print(df)
```

```
             Name Give Birth Can Fly Live in Water Have Legs        Class
0           human        yes      no            no       yes     mammals
1          python         no      no            no        no  non-mammals
2          salmon         no      no           yes        no  non-mammals
3           whale        yes      no           yes        no     mammals
4            frog         no      no     sometimes       yes  non-mammals
5          komodo         no      no            no       yes  non-mammals
6             bat        yes     yes            no       yes     mammals
7          pigeon         no     yes            no       yes  non-mammals
8             cat        yes      no            no       yes     mammals
9   leopard shark        yes      no           yes        no  non-mammals
10        turtle         no      no     sometimes       yes  non-mammals
11       penguin         no      no     sometimes       yes  non-mammals
12     porcupine        yes      no            no       yes     mammals
13           eel         no      no           yes        no  non-mammals
14    salamander         no      no     sometimes       yes  non-mammals
15  gila monster         no      no            no       yes  non-mammals
16      platypus         no      no            no       yes     mammals
17           owl         no     yes            no       yes  non-mammals
18       dolphin        yes      no           yes        no     mammals
19         eagle         no     yes            no       yes  non-mammals
```

In [15]:
```
l = len(df)
l
```

Out[15]: 20

In [18]:
```
tm = df[df['Class'] == "mammals"]['Name'].count()
gbym = df[(df['Give Birth']=="yes") &(df['Class']=="mammals")]['Name'].count()
gbyn = df[(df['Give Birth']=="yes") &(df['Class']=="non-mammals")]['Name'].count()
print(tm,gbym,gbyn)

cfym = df[(df['Can Fly']=="yes") &(df['Class']=="mammals")]['Name'].count()
cfyn = df[(df['Can Fly']=="yes") &(df['Class']=="non-mammals")]['Name'].count()
print(cfym,cfyn)

lym = df[(df['Live in Water']=="yes") &(df['Class']=="mammals")]['Name'].count()
lyn = df[(df['Live in Water']=="yes") &(df['Class']=="non-mammals")]['Name'].count()
print(lym,lyn)

hlym = df[(df['Have Legs']=="yes") &(df['Class']=="mammals")]['Name'].count()
hlyn = df[(df['Have Legs']=="yes") &(df['Class']=="non-mammals")]['Name'].count()
print(hlym,hlyn)
```

```
7 6 1
1 3
2 3
5 9
```

In [23]:
```
xm,xn = ((gbym*(tm-cfym)*lym*(tm-hlym))*tm)/((tm**4)*l),((gbyn*(tn-cfyn)*lyn*(tn-hlyn))*tn)/((tn**4)*l)
xm,xn = xm/(xm+xn),xn/(xm+xn)
print("Probabilty of x being a mammal:",xm)
print("Probabilty of x not being a mammal:",xn)
if(xm>xn):
    print("x is Mammal")
else:
    print("x is Non-mammal")
```

```
Probabilty of x being a mammal: 0.8848761495603142
Probabilty of x not being a mammal: 0.11512385043968584
x is Mammal
```

In [24]:
```
ym,yn = ((tm-gbym)*cfym*(tm-lym)*hlym)/(tm**4),((tn-gbyn)*cfyn*(tn-lyn)*(hlyn))/(tn**4)
ym,yn = ym/(ym+yn),yn/(ym+yn)
print("Probabilty of y being a mammal:",ym)
print("Probabilty of y not being a mammal:",yn)
if(ym>yn):
    print("y is Mammal")
else:
    print("y is Non-mammal")
```

```
Probabilty of y being a mammal: 0.0840695539347942
Probabilty of y not being a mammal: 0.9159304460652058
y is Non-mammal
```

In [25]:
```
zm,zn = (gbym*(tm-cfym)*(tm-lym)*hlym)/(tm**4),(gbyn*(tn-cfyn)*(tn-lyn)*hlyn)/(tn**4)
zm,zn = zm/(zm+zn),zn/(zm+zn)
print("Probabilty of z being a mammal:",zm)
print("Probabilty of z not being a mammal:",zn)
if(zm>zn):
    print("z is Mammal")
else:
    print("z is Non-mammal")
```

```
Probabilty of z being a mammal: 0.92245332988825
Probabilty of z not being a mammal: 0.07754667011174989
z is Mammal
```

Output:

X is a Mammal

Y is a Non-mammal

Z is a Mammal

Result:

The model has been successfully designed and tested for the given 3 testing data.