In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sea
```

In [2]:
```
from sklearn.linear_model import LogisticRegression
```

In [3]:
```
df = pd.read_csv(r"C:\Users\user\Downloads\C5_health care diabetes.csv")
df
```

Out[3]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunctio
0	6	148	72	35	0	33.6	0.62
1	1	85	66	29	0	26.6	0.35
2	8	183	64	0	0	23.3	0.67
3	1	89	66	23	94	28.1	0.16
4	0	137	40	35	168	43.1	2.28
...	...	...	...	...	...	...	.
763	10	101	76	48	180	32.9	0.17
764	2	122	70	27	0	36.8	0.34
765	5	121	72	23	112	26.2	0.24
766	1	126	60	0	0	30.1	0.34
767	1	93	70	31	0	30.4	0.31

768 rows × 9 columns

In [4]:
```
df.columns
```

Out[4]:
```
Index(['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin',
       'BMI', 'DiabetesPedigreeFunction', 'Age', 'Outcome'],
      dtype='object')
```

In [8]:
```
feature_matrix = df.iloc[:,0:8]
target_vector = df.iloc[:,-1]
```

In [9]:
```
feature_matrix.shape
```

Out[9]:
```
(768, 8)
```

In [10]:
```
from sklearn.preprocessing import StandardScaler
```

In [11]:
```
fs = StandardScaler().fit_transform(feature_matrix)
```

In [12]:
```
logs = LogisticRegression()
logs.fit(fs,target_vector)
```

Out[12]:
```
LogisticRegression()
```

In [13]:
```
observation = [[1.4,1.5,1.6,2,2.1,1,1,3]]
prediction = logs.predict(observation)
```

In [14]:
```
print(prediction)
```

```
[1]
```

In [15]:
```
logs.classes_
```

Out[15]:
```
array([0, 1], dtype=int64)
```

In [16]:
```
logs.predict_proba(observation)[0][0]
```

Out[16]:
```
0.09599090836764246
```

In [ ]: