In [1]:
```python
#load dependencies
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
```

In [2]:
```python
dataset = pd.read_csv('logisticregressioncsv.csv')
```

In [3]:
```python
#preprocess your data
y=dataset['PumpActivity']#target variable
features = ['PumpDevice','PumpDeviceAge']
x=dataset[features]
```

In [4]:
```python
#split and train the dataset
x_train,x_test,y_train,y_test = train_test_split(x,y)
```

In [5]:
```python
#Let the model predict results
log = LogisticRegression()
log.fit(x_train,y_train)
y_pred = log.predict(x)
y_prob =log.predict_proba(x)[:,1]
```

In [6]:
```python
#Lets add the columns back to the dataframe
dataset['predictions'] = y_pred
dataset['probabilities'] = y_prob
```

In [7]:
```python
#results of model predict and probability results
dataset
```

Out[7]:

	PumpActivity	PumpDevice	PumpDeviceAge	predictions	probabilities
0	0	0	0	0	0.007381
1	1	4	1	1	0.640374
2	0	4	0	1	0.590299
3	0	4	0	1	0.590299
4	1	4	2	1	0.687568
...	...	...	...	...	...
995	0	0	0	0	0.007381
996	1	4	0	1	0.590299
997	0	0	0	0	0.007381
998	1	4	0	1	0.590299
999	0	0	0	0	0.007381

1000 rows × 5 columns