(Insurance prediction)

In this the classification perfume in following steps-

(1) Import pandas module and import the file.

```
import pandas as pd
```

```
df=pd.read_csv("insurance_data.csv")
```

(2) Show the head

```
df.head()
```

	age	bought_insurance
0	22	0
1	25	0
2	47	1
3	52	0
4	46	1

(3) In next step we split the data in dependent and independent variable.

```
x=df[["age"]]

y=df.bought_insurance
```

```
x
```

	age
0	22
1	25
2	47
3	52
4	46
5	56
6	55
7	60
8	62
9	61
10	18
11	28

```
y
```

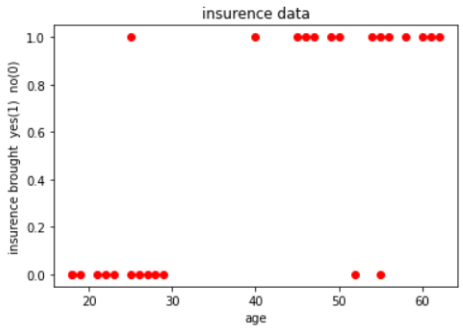
```
0    0
1    0
2    1
3    0
4    1
5    1
6    0
7    1
8    1
9    1
```

(4) In next step, we visualized the data by import (from matplotlib import pypolt)

```
from matplotlib import pyplot as plt
```

```
plt.scatter(x,y,color="red")
plt.title("insurence data")
plt.xlabel("age")
plt.ylabel("insurence brought  yes(1)  no(0)")
```

```
Text(0, 0.5, 'insurence brought  yes(1)  no(0)')
```



(5) In next step we have to train and test the data.by import the modules Import(from sklearn.model_selection import train_test_split).

```
: from sklearn.model_selection import train_test_split
```

```
: train_x,test_x,train_y,test_y=train_test_split(x,y,test_size=0.2,random_state=10)
```

```
: len(train_x),len(test_x)
```

```
: (21, 6)
```

(6) After that we have to apply the correct algorithm. We can take algorithm by import(from sklearn.linear_model import LogisticRegression).

```
from sklearn.linear_model import LogisticRegression
```

```
model=LogisticRegression()
```

```
model.fit(train_x,train_y)
```

```
▾ LogisticRegression
LogisticRegression()
```

Logistic regression-basically logistic regression is used in those data which are in categorical data distribution. Like- fraud detation, disease ,spam mail etc.
The formula of logostic regression is Y=1/1+e^{-x}. *e=2.713*

- After that we apply the module by using fit functions.

(7) After in next step we have to know the score and predict the values..

```
model.predict([[21]])
```

```
c:\users\hp\appdata\local\programs\python\python39\lib\site-packages
ature names, but LogisticRegression was fitted with feature names
  warnings.warn(
```

```
array([0], dtype=int64)
```

```
model.score(test_x,test_y)
```

```
1.0
```

```
model.predict(test_x)
```

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

This is the binary classification..

- **LLD**