import Libraries

In [46]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

import Linear Regression

In [47]:

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

In [139]:

```
lgr=LogisticRegression()
```

Select Required data from certain columns

In [140]:

```
a=pd.read_csv("bot.csv")
a
```

Out[140]:

	User ID	Username	Tweet	Retweet Count	Mention Count	Follower Count	Verified	Bot Label	
0	132131	flong	Station activity person against natural majori...	85	1	2353	False	1	
1	289683	hinesstephanie	Authority research natural life material staff...	55	5	9617	True	0	S
2	779715	roberttran	Manage whose quickly especially foot none to g...	6	2	4363	True	0	H
3	696168	pmason	Just cover eight opportunity strong policy which.	54	5	2242	True	1	Ma
4	704441	noah87	Animal sign six data good or.	26	3	8438	False	1	Cai
...	...	...	...	...	...	...	...	...	
49995	491196	uberg	Want but put card direction know miss former h...	64	0	9911	True	1	Kiml
49996	739297	jessicamunoz	Provide whole maybe agree church respond most ...	18	5	9900	False	1	(
49997	674475	lynncunningham	Bring different everyone international capital...	43	3	6313	True	1	D
49998	167081	richardthompson	Than about single generation itself seek sell ...	45	1	6343	False	0	St
49999	311204	daniel29	Here morning class various room human true bec...	91	4	4006	False	0	I

50000 rows × 11 columns

In [141]:

```
c=a.dropna()
c
```

Out[141]:

	User ID	Username	Tweet	Retweet Count	Mention Count	Follower Count	Verified	Bot Label	
1	289683	hinesstephanie	Authority research natural life material staff...	55	5	9617	True	0	S
2	779715	roberttran	Manage whose quickly especially foot none to g...	6	2	4363	True	0	H
3	696168	pmason	Just cover eight opportunity strong policy which.	54	5	2242	True	1	Ma
4	704441	noah87	Animal sign six data good or.	26	3	8438	False	1	Cai
5	570928	james00	See wonder travel this suffer less yard office...	41	4	3792	True	1	
...	...	...	...	...	...	...	...	...	
49995	491196	uberg	Want but put card direction know miss former h...	64	0	9911	True	1	Kiml
49996	739297	jessicamunoz	Provide whole maybe agree church respond most ...	18	5	9900	False	1	(
49997	674475	lynncunningham	Bring different everyone international capital...	43	3	6313	True	1	D
49998	167081	richardthompson	Than about single generation itself seek sell ...	45	1	6343	False	0	St
49999	311204	daniel29	Here morning class various room human true bec...	91	4	4006	False	0	N

41659 rows × 11 columns

In [142]:

```
c.columns
```

Out[142]:

```
Index(['User ID', 'Username', 'Tweet', 'Retweet Count', 'Mention Count',
       'Follower Count', 'Verified', 'Bot Label', 'Location', 'Created A
t',
       'Hashtags'],
      dtype='object')
```

In [143]:

```
fm=c[['User ID', 'Retweet Count', 'Mention Count',
       'Follower Count']]
tv=c[['Bot Label']]
```

Shape

In [144]:

```
fm.shape
```

Out[144]:

```
(41659, 4)
```

In [145]:

```
tv.shape
```

Out[145]:

```
(41659, 1)
```

To make the data in order (feature matrix)

In [146]:

```
from sklearn.preprocessing import StandardScaler
```

In [147]:

```
fs=StandardScaler().fit_transform(fm)
```

Imply Logistic Regression

In [149]:

```
lgr.fit(fs,tv)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\utils\validation.py:63:
DataConversionWarning: A column-vector y was passed when a 1d array was ex
pected. Please change the shape of y to (n_samples, ), for example using r
avel().
  return f(*args, **kwargs)
```

Out[149]:

```
LogisticRegression()
```

Prediction

In [150]:

```
ab=[[3,90,543,34]]
```

In [151]:

```
pre=lgr.predict(ab)
```

In [152]:

```
print(pre)
```

```
[0]
```

To check the output var we have got

In [153]:

```
lgr.classes_
```

Out[153]:

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

Prediction in Probablity value

In [154]:

```
lgr.predict_proba(ab)[0][1]
```

Out[154]:

```
0.0013247604361351912
```