


```
In [3]: import numpy as np
import matplotlib as plt
import pandas as pd
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix, accuracy_score
from sklearn.feature_selection import RFE
from sklearn.inspection import permutation_importance
```

```
In [4]: df=pd.read_csv("usa_rain_prediction.csv")
df.head(5)
```

Out[4]:

	Date	Location	Temperature	Humidity	Wind Speed	Precipitation	Cloud Cover	Pre
0	2024-01-01	New York	87.524795	75.655455	28.379506	0.000000	69.617966	1026.0
1	2024-01-02	New York	83.259325	28.712617	12.436433	0.526995	41.606048	995.9
2	2024-01-03	New York	80.943050	64.740043	14.184831	0.916884	77.364763	980.7
3	2024-01-04	New York	78.097552	59.738984	19.444029	0.094134	52.541196	979.0
4	2024-01-05	New York	37.059963	34.766784	3.689661	1.361272	85.584000	1031.7



```
In [7]: df.dropna()
df
```

Out[7]:

	Date	Location	Temperature	Humidity	Wind Speed	Precipitation	Cloud Cover
0	2024-01-01	New York	87.524795	75.655455	28.379506	0.000000	69.617966
1	2024-01-02	New York	83.259325	28.712617	12.436433	0.526995	41.606048
2	2024-01-03	New York	80.943050	64.740043	14.184831	0.916884	77.364763
3	2024-01-04	New York	78.097552	59.738984	19.444029	0.094134	52.541196
4	2024-01-05	New York	37.059963	34.766784	3.689661	1.361272	85.584000
...
73095	2025-12-27	Washington D.C.	40.614393	65.099438	28.778327	0.000000	54.168514
73096	2025-12-28	Washington D.C.	52.641643	30.610525	12.282890	0.871000	22.068055
73097	2025-12-29	Washington D.C.	56.492591	96.740232	2.894762	1.191956	52.336048
73098	2025-12-30	Washington D.C.	65.748956	63.900004	24.632400	0.483421	76.785280
73099	2025-12-31	Washington D.C.	54.648609	80.812021	22.722505	0.151903	19.674960

73100 rows × 9 columns



```
In [8]: x = df[["Temperature","Humidity","Precipitation","Cloud Cover"]]
y = df["Rain Tomorrow"]
xtrain,xtest,ytrain,ytest = train_test_split(x, y, test_size=0.4, shuffle = True)
```

```
In [9]: model = LogisticRegression()
model.fit(xtrain,ytrain)
ypred = model.predict(xtest)
```

```
In [10]: print("Coefficients = ",model.coef_)
print("Intercepts = ",model.intercept_)
cf = confusion_matrix(ytest,ypred)
print("Confusion Matrix: \n",cf)
ac = accuracy_score(ytest,ypred)
print("Accuracy Score: ",ac)
```

```
Coefficients = [[-2.34244199e-03  1.20780521e-01  4.82435217e+00 -3.00868157e-04]]
```

```
Intercepts = [-12.04384071]
```

```
Confusion Matrix:
```

```
[[21928  860]
```

```
[ 1937 4515]]
```

```
Accuracy Score: 0.904343365253078
```