

Logistic Regression

1. Objective

Build a Logistic Regression model to classify whether a student passes or fails based on study hours.

2. Install Required Libraries (One Time Only)

```
pip install numpy pandas matplotlib scikit-learn
```

3. Import Libraries

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
```

4. Create Sample Dataset

```
# Study Hours
X = np.array([1,2,3,4,5,6,7,8,9,10])

# Result (0 = Fail, 1 = Pass)
y = np.array([0,0,0,0,0,1,1,1,1,1])

mydata = pd.DataFrame({
    "StudyHours": X,
    "Result": y
})

print(mydata)
```

5. Separate Feature and Target

```
X = mydata[["StudyHours"]]
y = mydata["Result"]
```

6. Split Train and Test Data

```
X_train, X_test, y_train, y_test = train_test_split(  
    X, y, test_size=0.2, random_state=42  
)
```

7. Create Model

```
model = LogisticRegression()
```

8. Train Model

```
model.fit(X_train, y_train)  
print("Model Trained Successfully")
```

9. Make Predictions

```
y_pred = model.predict(X_test)  
print("Actual:", list(y_test))  
print("Predicted:", list(y_pred))
```

10. Evaluate Model

```
accuracy = accuracy_score(y_test, y_pred)  
print("Accuracy:", accuracy)
```

```
print("Confusion Matrix:")  
print(confusion_matrix(y_test, y_pred))
```

```
print("Classification Report:")  
print(classification_report(y_test, y_pred))
```

11. Predict New Student Result

```
new_hours = [[7]]  
prediction = model.predict(new_hours)
```

```
if prediction[0] == 1:  
    print("Student will PASS")  
else:  
    print("Student will FAIL")
```

12. Visualization

```
plt.scatter(X, y)  
plt.plot(X, model.predict(X), linewidth=2)  
plt.xlabel("Study Hours")  
plt.ylabel("Pass(1) / Fail(0)")  
plt.title("Logistic Regression Demo")  
plt.show()
```

13. Key Learning

- Logistic Regression is used for classification
- Output is probability between 0 and 1
- Uses sigmoid function
- Good for binary classification

Logistic Regression is a supervised learning algorithm used for binary classification that models probability using a sigmoid function.