Practical 5 ： Logistic Regression

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import confusion_matrix, accuracy_score, precision_score, recall_score

df = pd.read_csv("Social_Network_Ads.csv")
df = df.drop(columns=["User ID"])
label_encoder = LabelEncoder()
df["Gender"] = label_encoder.fit_transform(df["Gender"])
X = df[["Gender", "Age", "EstimatedSalary"]]
y = df["Purchased"]
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)
log_reg = LogisticRegression()
log_reg.fit(X_train_scaled, y_train)
y_pred = log_reg.predict(X_test_scaled)
cm = confusion_matrix(y_test, y_pred)
TN, FP, FN, TP = cm.ravel()
accuracy = accuracy_score(y_test, y_pred)
error_rate = 1 - accuracy
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)
print("Confusion Matrix:")
print(cm)
print(f"True Negatives (TN): {TN}")
```

```
print(f"False Positives (FP): {FP}")
print(f"False Negatives (FN): {FN}")
print(f"True Positives (TP): {TP}")
print(f"Accuracy: {accuracy:.4f}")
print(f"Error Rate: {error_rate:.4f}")
print(f"Precision: {precision:.4f}")
print(f"Recall: {recall:.4f}")
y_probs = lor.predict_proba(X_test_scaled)[:, 1]
auc = roc_auc_score(y_test, y_probs)
print("AUC-ROC Score:", round(auc, 4))
fpr, tpr, _ = roc_curve(y_test, y_probs)
plt.plot(fpr, tpr, label=f'AUC = {auc:.2f}')
plt.plot([0,1], [0,1], linestyle='--', color='gray')
plt.title('ROC Curve')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.show()
```

Output :

```
Confusion Matrix:
[[50  2]
 [ 7 21]]
True Negatives (TN): 50
False Positives (FP): 2
False Negatives (FN): 7
True Positives (TP): 21
Accuracy: 0.8875
Error Rate: 0.1125
Precision: 0.9130
```

Recall: 0.7500

AUC-ROC Score: 0.9705