Module 6: Assignment 7 – Confusion Matrix & Accuracy

Problem Statement:

You work in XYZ Corporation as a Machine Learning Engineer. After training the logistic model, and predicting the values on top of the test set, now find out the confusion matrix and accuracy score of the models.

Tasks to be performed for Simple Logistic Regression:

1. After building a simple logistic model on top of the train set for the city_temperature data frame, now find the prediction values of the model1 on top of the test dataset.

    a. Now combine the actual values and predicted values of the model and store that combined data in a new data frame named 'Result1'.

    b. Find out the errors in the predicted values.

    c. Now calculate the confusion matrix from the actual and the predicted values.

    d. From the confusion matrix, calculate the accuracy of the model.

2. After building a simple logistic model on top of the train set for the Customer_Churn data frame, now find the prediction values of the model2 on top of the test dataset.

    a. Now combine the actual values and predicted values of the model and store that combined data in a new data frame named 'Result2'.

    b. Find out the errors in the predicted values.

    c. Now calculate the confusion matrix from the actual and the predicted values.

    d. From the confusion matrix, calculate the accuracy of the model.

3. After building a simple logistic model on top of the train set for the Pharmacovigilance_audit_Data data frame, now find the prediction values of the model3 on top of the test dataset.

    a. Now combine the actual values and predicted values of the model and store that combined data in a new data frame named 'Result3'.

    b. Find out the errors in the predicted values.

    c. Now calculate the confusion matrix from the actual and the predicted values.

    d. From the confusion matrix, calculate the accuracy of the model.

Tasks to be performed for Multiple Logistic Regression:

1. After building a multiple logistic model on top of the train set for the city_temperature data frame, now find the prediction values of the model1 on top of the test dataset.

   a. Now combine the actual values and predicted values of the model and store that combined data in a new data frame named 'MultiResult1'.

   b. Find out the errors in the predicted values.

   c. Now calculate the confusion matrix from the actual and the predicted values.

   d. From the confusion matrix, calculate the accuracy of the model.

2. After building a multiple logistic model on top of the train set for the Customer_Churn data frame, now find the prediction values of the model2 on top of the test dataset.

   a. Now combine the actual values and predicted values of the model and store that combined data in a new data frame named 'MultiResult2'.

   b. Find out the errors in the predicted values.

   c. Now calculate the confusion matrix from the actual and the predicted values.

   d. From the confusion matrix, calculate the accuracy of the model.

3. After building a multiple logistic model on top of the train set for the Pharmacovigilance_audit_Data data frame, now find the prediction values of the model3 on top of the test dataset.

   a. Now combine the actual values and predicted values of the model and store that combined data in a new data frame named 'MultiResult3'.

   b. Find out the errors in the predicted values.

   c. Now calculate the confusion matrix from the actual and the predicted values.

   d. From the confusion matrix, calculate the accuracy of the model.