Module End Activities

Activity 1: Customer Churn Prediction:

Predict whether a customer will churn (stop using a service) based on various features such as usage patterns, customer feedback, and tenure.

a. **Dataset:** Utilize a dataset with information on customer behaviours, feedback, and churn status. (kaggle.com/datasets/blastchar/telco-customer-churn)

b. **Steps:**

- Explore and understand the dataset.
- Handle missing values and encode categorical variables.
- Split the data into training and testing sets.
- Build a binary classification model (e.g., logistic regression) to predict customer churn.
- Interpret important features contributing to customer churn.

Activity 2: Fraud Detection:

Detect fraudulent transactions in a dataset containing information about financial transactions.

a. **Dataset:** Utilize a dataset with details about transactions, including features like transaction amount, time, and type.

b. **Steps:**

- Explore and preprocess the dataset, considering the imbalance in classes (fraudulent vs. non-fraudulent).
- Split the data into training and testing sets.
- Build a fraud detection model, such as a binary classification algorithm (e.g., Random Forest, XGBoost).
- Adjust model parameters to optimize for performance.

Activity 3: Predicting Customer Lifetime Value

Predict the Customer Lifetime Value (CLV) based on various features such as purchase history, average transaction value, frequency of transactions, and customer demographics.

- **Dataset:** Utilize a dataset with information on customer transactions, purchase history, and corresponding Customer Lifetime Values.
- **Steps:**
  1. **Explore and understand the dataset:** Analyze features related to customer transactions, purchase behavior, and the target variable (Customer Lifetime Value).
  2. **Handle missing values and encode categorical variables:** Address any missing data and convert categorical variables into a format suitable for regression modeling.
  3. **Split the data into training and testing sets:** Divide the dataset into training and testing subsets.
  4. **Build a regression model:** Choose a regression model, such as Linear Regression, to predict Customer Lifetime Value based on the identified features.
  5. **Evaluate the model:** Assess the model's performance using regression metrics like Mean Squared Error (MSE) or Root Mean Squared Error (RMSE).
  6. **Interpretation:** Examine the coefficients of the regression model to understand the impact of each feature on the predicted Customer Lifetime Value. Identify which features contribute the most to the overall lifetime value of a customer.