Stimulation Exercise 3

You are a data scientist working for an e-commerce company. Your task is to develop a machine learning model to predict customer churn (i.e., the likelihood of customers leaving the platform). You will need to use various machine learning techniques and concepts to achieve this goal.

Questions

1. **Data Collection: What type of data would you need to collect to train your machine learning model for predicting customer churn?**

A) Customer purchase history

B) Customer interaction logs (e.g., website visits, clicks)

C) Customer demographic information

D) All of the above

2. **Data Preprocessing**: **Before training your machine learning model, what preprocessing steps would you take to ensure the data is ready for analysis?**

A) Data cleaning (removing missing or incorrect data)

B) Data normalization (scaling data to a standard range)

C) Feature engineering (creating new features from existing data)

D) All of the above

3. Model Selection: **Which type of machine learning model would be most suitable for predicting customer churn?**

A) Linear Regression

B) Decision Trees

C) K-Means Clustering

D) Support Vector Machines (SVMs)

4. Training the Model: What is the purpose of splitting your data into training and testing sets?

A) To evaluate the model's performance on unseen data

B) To reduce the amount of data needed

C) To increase the model's complexity

D) To improve data storage

5. Model Evaluation: After training your model, which metrics would you use to evaluate its performance in predicting customer churn?

A) Accuracy

B) Precision

C) Recall

D) All of the above

Answers

1. D) All of the above
2. D) All of the above
3. B) Decision Trees
4. A) To evaluate the model's performance on unseen data
5. D) All of the above

Reflection

- **Data Collection**: Collecting comprehensive customer data is crucial for training an accurate machine learning model.
- **Data Preprocessing**: Proper preprocessing ensures the data is clean, normalized, and ready for analysis, which improves model performance.
- **Model Selection**: Choosing the right model, such as Decision Trees for classification tasks like predicting customer churn, is essential for accurate predictions.
- **Training the Model**: Splitting data into training and testing sets helps evaluate the model's generalization ability.
- **Model Evaluation**: Using multiple evaluation metrics provides a holistic view of the model's performance.