BRAIN STATION 23

Machine Learning Engineer Test

Dataset: TATA: Online Retail Dataset

Part 1: Machine Learning Model Implementation

Dataset Description: You are provided with a dataset containing information about [add dataset information]

Tasks:

1. Perform Data Preprocessing:

   Handle missing values (e.g., imputation, deletion), Encode categorical variables, Feature scaling, dimensionality reduction (e.g., PCA, t-SNE)

2. Perform Exploratory Data Analysis (EDA) on the dataset. Include statistical analysis and meaningful visualization techniques to understand the distribution of features, correlations, and any patterns in the data.

3. Choose an appropriate machine learning algorithm for predicting customer purchase behavior based on the given features with justification. Try to use ensemble algorithms. Implement the algorithm from scratch without using any machine learning framework.

4. Split the dataset into training and testing sets. Train your model on the training set and evaluate its performance on the testing set using appropriate metrics.

5. Visualize the results of your model. Include plots or charts to illustrate the performance metrics and any insights gained from the analysis.

Part 2: Chatbot Development

Dataset Utilization: You will continue to use the same dataset provided in Part 1.

Tasks:

1. Design a relational database schema to store the dataset. Choose an appropriate DBMS (Database Management System) and implement the dataset into the database.

2. Develop a chatbot using a framework of your choice (Flask/FAST API). The chatbot should interact with users based on the customer information stored in the database.

3. Implement features such as:

   - Customer inquiry handling (e.g., product recommendations, order status)
   - Natural language processing for understanding user queries
   - Personalization based on customer profiles

4. Deploy the chatbot as a service, ensuring scalability and reliability in the backend design.
5. Provide documentation and explanations for your backend design choices and the features implemented in the chatbot.

Submission Requirements:

- Create a private GitHub repository for your code.
- Share the repository link with the examiners for evaluation.
- Include clear instructions on how to run and test your code.
- Provide documentation explaining your approach, design decisions, and any challenges faced during the implementation process.

Note: Examiners will evaluate the completeness of your solution, the accuracy and efficiency of your machine learning model, the effectiveness of your chatbot's features, and the quality of your code and documentation. Good luck!