Write code logic for the given Scenarios:

1. A healthcare startup wants to build a model to predict the onset of diabetes. The dataset has 50 features, but the team wants to use only the most relevant ones. How can they apply feature selection effectively?

Answer:

1. Import the dataset
2. Do preprocessing the dataset and remove unwanted data, null data
3. Train the model
4. Apply SelecKBest with take a reduce columns of the inputs
5. Find the best models using ROC-AUC and F1-Score values.
6. Save the best model and deploy it

---

2. You're creating a new project for managing student data. The project name is student_portal. What is the step-by-step command to start it?write logic for this django project

Answer:

1. Open the Anaconda Prompt
2. Open the Project contains folder path
3. And give the following cmd “django-admin startproject student_portal”
4. This will be creates folder.

---

3. A company is analyzing customer purchase patterns with 200+ behavioral features. How can they reduce dimensionality without losing predictive power?

Answer:

1. Import dataset
2. Preprocessing the dataset, delete unwanted columns and null data
3. Use PCA (Principal Component Analysis) to reduce dimensions
4. Train dataset
5. Evaluate using RMSE (if regression) or Accuracy (if classification).

4.A digital library wants to recommend books to readers based on what similar readers liked. How should they design this system?

Answer:

1. Import dataset and preprocessing
2. Find users or items based similarity using Cousin function
3. Recommend the book to the user who is not reader yet

---

5. A bank wants to assess the risk level of credit applicants using only the most important financial indicators. How can they reduce the number of features?

Answer:

1. Import applicants details
2. Preprocessing the dataset
3. Use tree based model like RF to compute feature importance
4. Evaluate using ROC-AUC, Precision, and Recall.

---

6. A news app wants to recommend articles based on both article similarity and user reading history. How can they implement a hybrid system?

Answer:

1. Collect user using Article data
2. Build content-based recommendation
3. Build collaborative filters using form metrix
4. Train the model
5. Evaluate using ROC-AUC, Precision, and Recall

7.**Question:**
You're building a spam detection model and have thousands of text features from emails. How do you identify the most useful ones?

Concept: Feature selection

1. Split the mail with specified words like Wow, amazing etc.,
2. Apply Feature selection Chi-squre test to select the correlated with the label
3. Train the model
4. Evaluated with F1-Score

---

8. An ed-tech platform wants to recommend courses based on what similar learners have enrolled in. What steps would you take?

Answer:

Concept: User or Item Based Recommendation System (Collaborative Filtering)

1. Build matrix with user data
2. Use user or item based filter to find the similar user who are all enrolled same courses
3. Filter highly rated courses and recommending to the user
4. Train and Evaluate the model

9.You're developing a car price prediction tool. With 100+ features (e.g., brand, mileage, engine type), how do you reduce complexity?

Answer:

Concept: This is Feature selection

1. Collect all the required details like cars name, model, milage, type, etc.,
2. Encoding categorical features and normalize the data
3. Apply correlation feature to reduce the dimension
4. Evaluate the model with R2 score.

10.
How do you recommend products to new users who haven't interacted with anything yet?

Answer:

This is Popularity Based Recommendation System

1. Analyze the dataset and import as required details as we need based on this recommendation system
2. Use content-based filter as per the user selection when their interests

3. Filtered the products which was highly demand/bought/viewed by other customers and group by with top 5 or 10 products and display to the user