

1. A real estate company wants to develop a system that predicts house prices based on square footage, number of bedrooms, and location.

Q: Identify the problem type and outline the step-by-step logic to solve it.

Identify the problem type: Regression

Step by step logic:

Collect data: Collect the data as follows square footage, number of bedrooms, location of the house

Preprocess data: Encode the categorical variables and split the training data and test data

Choose algorithm: Use the regression and decision tree

Train model: Train the model using the training data

Evaluate Model: Calculate the r-score value

Make predictions: Now use the data to predict the new house

2. A bank wants to build a model to detect fraudulent transactions by analyzing customer spending behavior and transaction history.

Q: Identify the problem type and outline the step-by-step logic to solve it.

Identify the problem type: clustering

Step by step logic:

Collect data: Get the data for the transaction is fraudulent or non-fraudulent.

Preprocess data: Encode the categorical features, and watch the transaction behavior

Choose algorithm: Use logistic Regression, Random Forest

Train model: Training the labeled transaction data.

Evaluate Model: Calculate the accuracy, precision, recall, and F1-score

Make predictions: Implement real-time fraud detection

3. A supermarket wants to segment its customers based on their shopping patterns to provide personalized promotions.

Q: Identify the problem type and outline the step-by-step logic to solve it.

Identify the problem type: clustering

Step by step logic:

1. Collect the customer frequency of purchases, how much amount spend and purchases' history
 2. Use K-Means Clustering and find the elbow method to find the best number of clusters.
 3. Apply the cluster algorithm to group the customers, analyze the clusters to separate the customer for his promotions
-

4. A company wants to estimate an employee's salary based on their years of experience, job title, and education level.

Q: Identify the problem type and outline the step-by-step logic to solve it.

Identify the problem type: Regression

Step by step logic:

Collect data: Collect the employee data like current salary, role, and experience

Preprocess data: Convert needed fields from categorical data into numerical data, and split the training data and test data

Choose algorithm: Use linear regression or random forest

Train model: Train the model with the training data

Evaluate Model: Calculate the r-score value

Make predictions: Now estimate the salary based on the new employee data

5. An email provider wants to automatically classify incoming emails as spam or not spam based on their content and sender details.

Q: Identify the problem type and outline the step-by-step logic to solve it.

Identify the problem type: classification

Step by step logic:

Collect data: Collect the data for the email is span or non-span

Preprocess data: Convert the email text to numerical format and split the training data and test data.

Choose algorithm: Use the support vector algorithm and naive bayes

Train model: Train the model using labeled data

Evaluate Model: Find the precision, recall and f1-score

Make predictions: Now classify the email is spam or not.

6. A business wants to analyze customer reviews of its products and determine whether the sentiment is positive or negative.

Q: Identify the problem type and outline the step-by-step logic to solve it

Identify the problem type: classification

Step by step logic:

Collect data: Collect the customer data to check if it's a positive or negative comments

Preprocess data: Remove the unwanted words, convert the text into to numerical format and split the training set and test set

Choose algorithm: Use logistic regression and naive bayes

Evaluate Model: Find the f1-score for the model performance

Make predictions: classify the review is positive or negative

7. An insurance company wants to predict whether a customer is likely to file a claim in the next year based on their driving history and demographics.

Q: Identify the problem type and outline the step-by-step logic to solve it.

Identify the problem type: Classification

Step by step logic:

Collect data: Get the past claim history, driving behavior, and customer demographics

Preprocess data: Handle missing values and encode categorical features

Choose algorithm: Separate the training data and test data and Use Logistic Regression, Decision Tree

Train model: Train the model using data

Evaluate Model: Use Precision-Recall, AUC-ROC score.

Make predictions: Predict claim for new customer.

8. A streaming platform wants to recommend movies to users by grouping them based on their viewing preferences and watch history.

Q: Identify the problem type and outline the step-by-step logic to solve it.

Identify the problem type: clustering

Step by step logic:

Collect data: Get the user watched movies genres and get their rating

Preprocess data: Change the movie name genres categorical data into numerical data

Choose regression algorithm: Use K-mean or hierarchical clustering algorithm and use elbow method

Train model: Apply the clustering algorithm to group the users

Evaluate Model: identify users based on genres

Make predictions: Suggest movie based on the watched list.

9. A hospital wants to predict the recovery time of patients after surgery based on their age, medical history, and lifestyle habits.

Q: Identify the problem type and outline the step-by-step logic to solve it.

Identify the problem type: Regression

Step by step logic:

Collect data: Collect the patient's data like age, medical history, and lifestyle habits.

Preprocess data: Change the needed values as numerical value and fill the missing value

Choose algorithm: Use random forest and linear regression

Evaluate Model: find the RMSE for check the accuracy

Make predictions: Now find the new patient recovery time.

10. A university wants to predict a student's final exam score based on study hours, attendance, and past academic performance.

Q: Identify the problem type and outline the step-by-step logic to solve it.

Identify the problem type: Regression

Step by step logic:

Collect data: Collect the student data like study hours, attendance, and past academic performance

Preprocess data: Handle the missing value and change the needed numerical value, split the training data and test data

Choose algorithm: Use linear regression and SVM

Train model: Fit the training data to train the model

Evaluate Model: Calculate the r-score value

Make predictions: Now predict the new student of the performance