

A3: Demo of a Supervised Learning System Application

1. Value: Who would pay for or benefit from an application based on this supervised learning demo?

This data shows how people's age, gender, education, and review feedback can affect their buying decision. Any online shopping platform could use this kind of model to understand which type of customers are more likely to buy a product. It helps in making marketing and product recommendation decisions and automating customer analysis.

2. Data or knowledge source: What is the data, knowledge or both that you used for this demo?

Source of Data: Customer Review Patterns and Buying Decisions

<https://www.kaggle.com/datasets/ayeshaimran123/customer-review-patterns-and-buying-decisions>

Description: This dataset contains customer information (age, gender, education), review quality (Good, Average, Poor), and purchase status (Yes or No). I used it to study how these features influence whether a customer makes a purchase. The data is already labeled, so it fits a supervised learning problem.

3. AI complex task: Indicate the AI task in your application demo and provide a set of three examples of inputs and outputs.

Supervised learning - binary classification to predict whether a customer will purchase a product (Yes/No). The model is not a full recommendation system, but it could be used to target likely buyers.

Examples of Inputs and Outputs:

Age	Gender	Review	Education	Predicted Purchased
56	Male	Poor	PG	Yes
46	Female	Good	School	Yes
32	Female	Good	PG	No

4. AI method: Which AI method you utilized, provide source library and a link to your code with required instructions to run it.

AI Method: Decision Tree Classifier (supervised learning, scikit-learn library)

Library Used:

- pandas
- scikit-learn
- seaborn
- matplotlib

Code File Name: purchase_prediction_decision_tree.py

Instructions to Run:

1. Keep Customer_Review.csv in the same folder as the Python file.
2. Create a virtual environment: `python3 -m venv myenv`
3. Install dependencies: `pip install pandas scikit-learn seaborn matplotlib`
4. Run the script: `python3 purchase_prediction_decision_tree.py`
5. The output will print the training and testing sizes, accuracy, classification report, confusion matrix, and predictions for new cases.

5. Testing and evaluation: How did you test or evaluate your model? Describe the evaluation process and results for selected metrics.

I split the data into 80% for training and 20% for testing, using stratified sampling to keep the same class balance. The model was trained with a decision tree (max depth 3, minimum 15 samples to split, and 5 per leaf). I evaluated it using accuracy, precision, recall, F1-score, and a confusion matrix.

Results:

- Model Accuracy: around 55%
- Precision (No): 0.60
- Precision (Yes): 0.50
- Recall: around 0.55 for both
- F1-score: around 0.55

Confusion Matrix: Visualized using a heatmap showing true positives, true negatives, false positives, and false negatives.

Testing: I added two new customer examples for prediction:

Age	Gender	Review	Education	Predicted Purchased
45	Female	Good	PG	No
25	Male	Poor	UG	Yes

The results show that the model can make logical predictions, but accuracy is not high because the dataset is small and only has a few features.