Predicting Consumer Preferences and Personalized Recommendations using Machine Learning

Team Members
Yarden Ochana – 315321851
Amit Sasson - 322759648

Dataset
This dataset from [info.data.gov.il] LINK includes 5,012 rows and 34 columns. It captures demographic data (age, sex, marital status) and preferences (restaurants, pharmacies, entertainment). After cleaning, 24 columns remain.

Motivation
Businesses aim to understand customer needs for personalized services, enhancing user experience and identifying market trends. The project goal is to predict preferences and build a recommendation system.

Methodology

1. Preprocessing
- Clean irrelevant/missing values in the data.
- Engineer features like socio-economic indices.

2. Classification
- Train Decision Trees and Random Forest models to predict categories (e.g., restaurants).
- Evaluate using Accuracy, Precision, Recall, F1 Score.

3. Clustering
- Use K-Means and Hierarchical Clustering to group users.
- Evaluate with Silhouette Score.

4. Recommendation System
- Develop systems with Content-Based and Collaborative Filtering.
- Evaluate using Precision@K, Recall@K.

Planned Experiments
1. Train/Test Split (80/20): Evaluate with Accuracy and F1 Score.
2. Clustering: Segment users by demographics and preferences; validate with Silhouette Score.
3. Recommender System: Align recommendations with true preferences, assessed by Precision@K and Recall@K.

In some of the above texts, we used AI in order to better formulate and correct spelling errors.