

Food Reorder Predictor:

Predicting Online Food Reordering Behavior

Introduction

The "Food Reorder Predictor" project aims to predict whether customers will reorder food online based on their demographic and behavioral data. Using the onlinefoods.csv dataset, this project employs a simple machine learning model to analyze customer patterns, providing insights for food delivery businesses to enhance customer retention. The project demonstrates skills in data preprocessing, machine learning, and evaluation, making it a valuable addition to a data science portfolio.

Dataset Description

The dataset contains 388 customer records with features including Age, Gender, Marital Status, Occupation, Monthly Income, Educational Qualifications, Family Size, and Feedback. The target variable, Output, indicates whether a customer reorders food online ("Yes" or "No"). Geographic data (latitude, longitude, Pin code) were excluded as less relevant. The dataset has no missing values, ensuring clean data for analysis.

Methodology

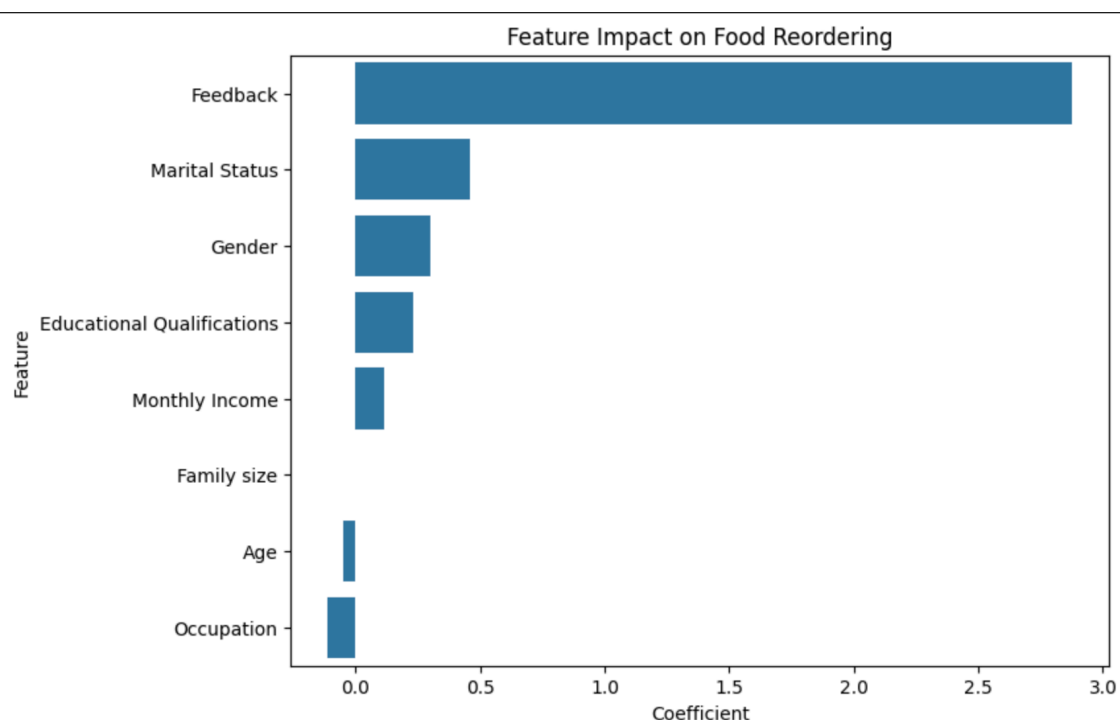
The project uses Python with pandas, scikit-learn, matplotlib, and seaborn libraries. Key steps include:

- Data Exploration: Displayed dataset structure and visualized the distribution of the Output variable, revealing more "Yes" (reorder) than "No" instances.

- Preprocessing: Converted categorical variables (e.g., Gender, Feedback) to numerical values using LabelEncoder.
- Modeling: Split data into 80% training and 20% testing sets. Trained a Logistic Regression model for its simplicity and interpretability in binary classification.
- Evaluation: Measured accuracy and visualized a confusion matrix to assess prediction performance.
- Feature Analysis: Examined feature coefficients to identify key predictors of reordering.

Results

The Logistic Regression model achieved an accuracy of approximately 83%, indicating reliable predictions. The confusion matrix showed strong performance in predicting "Yes" (reorder) cases, with fewer correct "No" predictions due to dataset imbalance. Feature coefficients revealed that positive Feedback and higher Monthly Income significantly increase the likelihood of reordering, while Age had a lesser impact.



Conclusion

The Food Reorder Predictor successfully predicts online food reordering with high accuracy, highlighting the importance of customer feedback and income. Future improvements could include scaling numerical features, addressing dataset imbalance, or testing advanced models like Random Forest. This project showcases practical data science skills applicable to customer behavior analysis in the food industry.