IE 7374 Project:
In-vehicle Coupon Recommendation

Krishna Hemant, Anushka Hedge, Yanming Liu
durgasharan.k@northeastern.edu
nush.1997@gmail.com
liu.yanmi@northeastern.edu

Data Analytics Engineering, College of Engineering, Northeastern University

June 5, 2021

Yan. *et al.*

Northeastern University
College of Engineering

Project Background

Coupons are an important tool for increasing corporate revenue and retaining customers. However, the issue is recommending the coupon to those who have a high acceptance rate so that the company can reduce costs and gain many advantages. In the context of an in-vehicle setting, personalising eatery recommendations for users or drivers can result in higher footfall. Machine learning, which assigns computers the ability to learn in the context of mining flexible, complex, and numerous business data records, provides a fresh perspective on how to solve this problem.

Problem Statement

To predict if a customer will accept the coupon or not, under learning algorithms, and find patterns or relationships between the predictors and the dependent variable.

Dataset Description

The dataset originates from the UCI Machine Learning Repository. By addressing surveys on Amazon Mechanical Turk, 12684 records with 26 attributes are acquired along with their missing values. All attributes are categorical variables, even age, which is usually encoded as a numeric. To be specific, the content for features can be split into four types:

1. **Travel info (7)**, such as destination, passenger, weather, temperature, time, direction_same, direction_opp.
2. **Demographic info (8)**, like gender, age, marital status, child status, education, occupation, income, car possession type.
3. **Coupon related info (2)**, coupon type, expiration.
4. **Purchase behavior info(7)**, bar, coffeeHouse, carryAway, restaurant Less Than20, restaurant Less Than20, restaurant 20To 50, to Coupon GEQ15min, toCoupon GEQ25min.
5. **Target (1)**, Y: (1,0).

Here is a sample record for the in-vehicle coupon recommendation data set.

Table 1. A Sample of Coupon Recommendation Dataset

Feature Name	Value	Feature Name	Value
income	37500–49999	destination	No Urgent Place
car	NaN	passanger	Alone
Bar	never	weather	Sunny
CoffeeHouse	never	temperature	55
CarryAway	NaN	time	2PM
RestaurantLessThan20	4~8	coupon	Restaurant(<20)
Restaurant20To50	1~3	expiration	1d
toCoupon_GEQ5min	1	gender	Female
toCoupon_GEQ15min	0	age	21
toCoupon_GEQ25min	0	maritalStatus	Unmarried partner
direction_same	0	has_children	1
direction_opp	1	education	Some college - no degree
Y	1	occupation	Unemployed

For more details, check the link below.
[UCI Machine Learning Repository: in-vehicle coupon recommendation Data Set](#)

<u>Potential Methods</u>

As the output is a discrete variable, the task is a classification. As a starting point, we intend to use traditional classification methods such as logistic regression as a benchmark. The tuning and testing of naive bayes, decision trees, SVM, random forests, and neural networks will follow. We may also explore adopting some of the Bayesian frameworks mentioned in the paper, like [Bayesian Rule Set](#) (BRS1, BRS2), and apply the mathematical principles to coding implementation.