The Future Ad

In Vehicle Coupon Acceptance Analysis

Group 16
Nirvi Shah Vy Nguyen Zichen Fan

The Future Ad

Table of contents

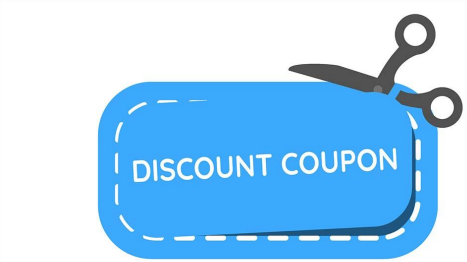
1. Executive Summary

Business Opportunity:

- Strategic analysis of coupon acceptance
- Optimize promotional strategies
- Personalize campaigns

Project Objective:

- Enhance customer targeting
- Understanding customer behavior and preferences related to coupon usage



2. Background

The Future Ad

3. Core Team & Stakeholders

R&D team

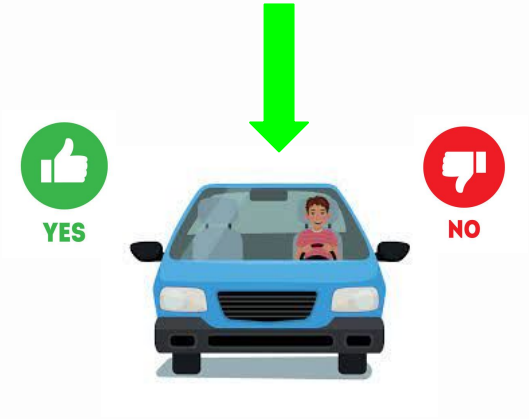
- Research on coupon acceptance
- Analyze & Understanding customer behavior and preferences related to coupon usage
- Provide the most efficient marketing strategy to our clients

The Future Ad

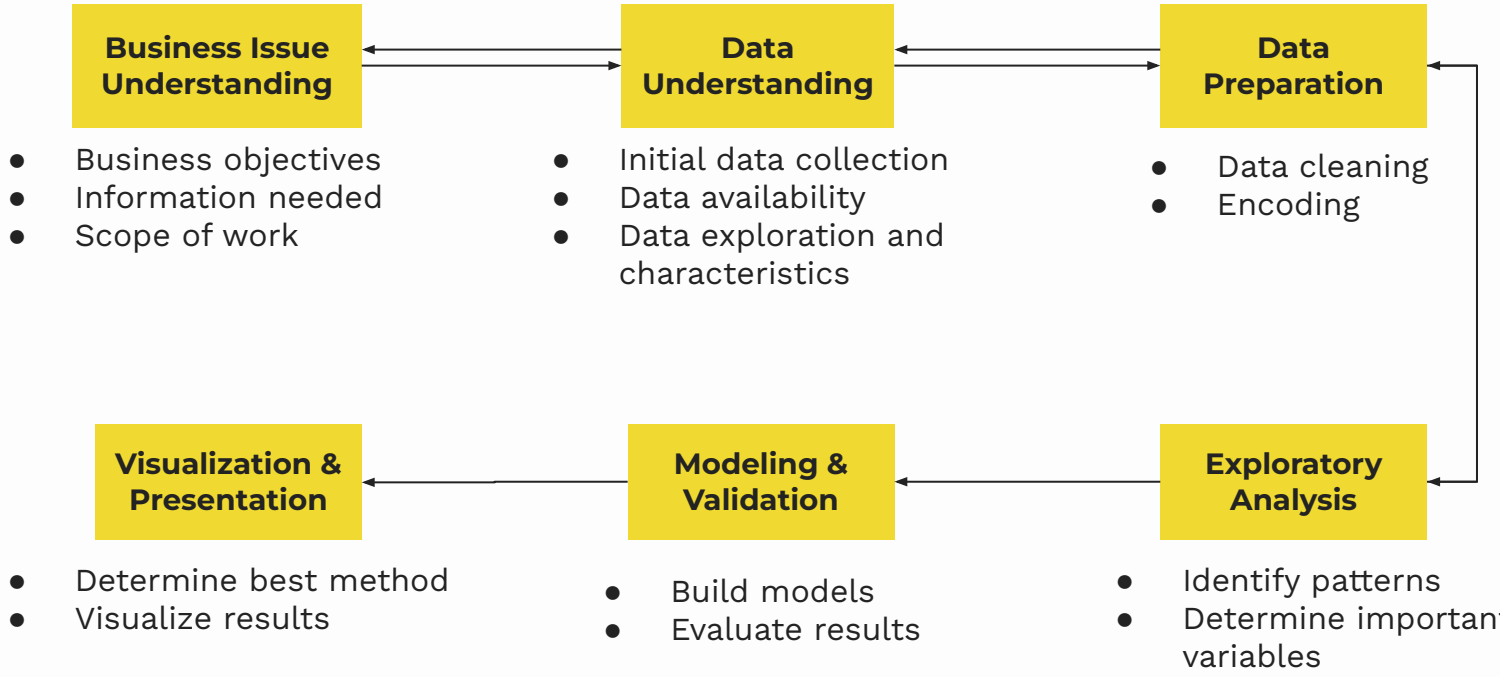
4. Scope & Objectives

The objective of the prediction task is to **anticipate whether a customer will approve or decline a coupon** for a particular venue, considering demographic and contextual characteristics.

Will this person accept the 20% off Restaurant coupon?
- age : 22
- weather: sunny
- etc



5. Project Development Timeline



6. Development Strategy

Machine learning techniques can be employed to develop an enhanced **coupon recommendation system.**

Personalization and Targeting

Handling Complexity

Adaptability to Changing Preferences

Scalability and Efficiency

Optimized Resource Allocation

Iterative Improvement

7. Development Details

Dataset information:

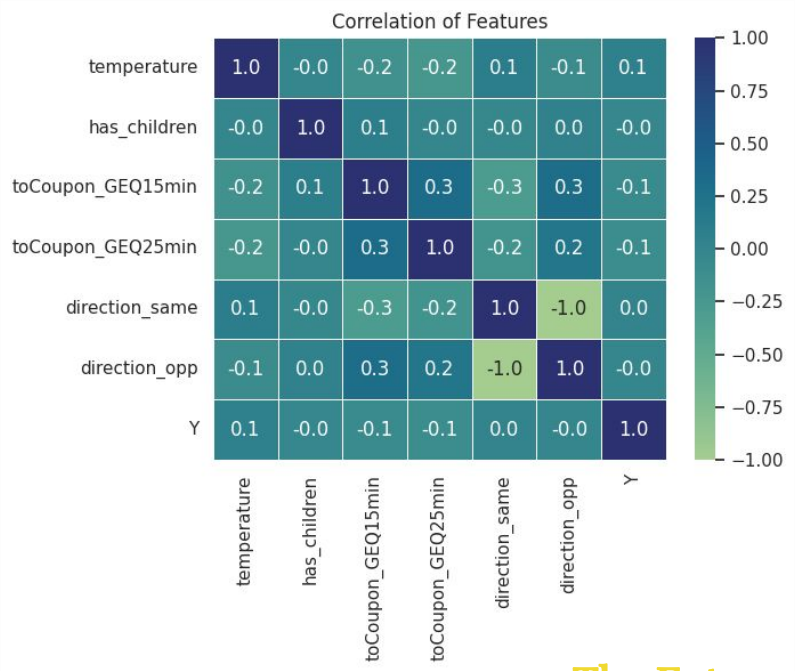
- in-vehicle coupon recommendation. (2020). UCI Machine Learning Repository. https://doi.org/10.24432/C5GS4P
- **12684** instances, **24** features
- **Data composition**: personal information, consumption habits, environment, coupon restrictions
- **Data cleaning**: binary feature with 1 value, feature no value, highly correlated
- **MIssing Values**: filled with -1 as a new category

	destination	passanger	weather	temperature	time	coupon	expiration	gender	age	maritalStatus	...	CoffeeHouse	CarryAway	RestaurantLessThan20	Restaurant20To50	toCoupon_GEQ5min	toCoupon_GEQ15min	toCoupon_GEQ25min	direction_same	direction_opp	Y
0	No Urgent Place	Alone	Sunny	55	2PM	Restaurant(<20)	1d	Female	21	Unmarried partner	...	never	NaN	4~8	1~3	1	0	0	0	1	1
1	No Urgent Place	Friend(s)	Sunny	80	10AM	Coffee House	2h	Female	21	Unmarried partner	...	never	NaN	4~8	1~3	1	0	0	0	1	0
2	No Urgent Place	Friend(s)	Sunny	80	10AM	Carry out & Take away	2h	Female	21	Unmarried partner	...	never	NaN	4~8	1~3	1	1	0	0	1	1
3	No Urgent Place	Friend(s)	Sunny	80	2PM	Coffee House	2h	Female	21	Unmarried partner	...	never	NaN	4~8	1~3	1	1	0	0	1	0
4	No Urgent Place	Friend(s)	Sunny	80	2PM	Coffee House	1d	Female	21	Unmarried partner	...	never	NaN	4~8	1~3	1	1	0	0	1	0

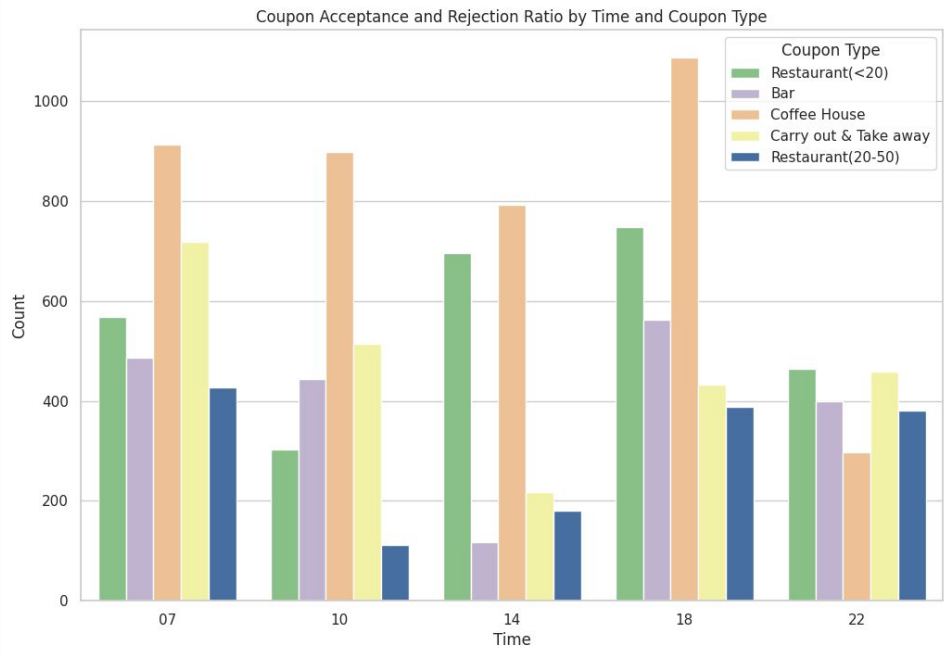
7. Development Details

Insights exploration

- Feature Selection for models
- Eliminating one of the highly correlated columns, 'direction_opp,' simplified the dataset without sacrificing information
- Improved Decision-Making

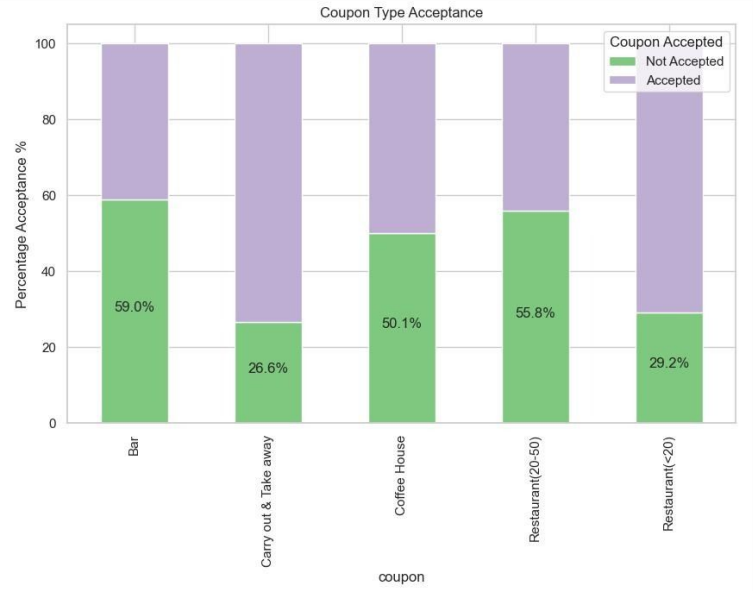


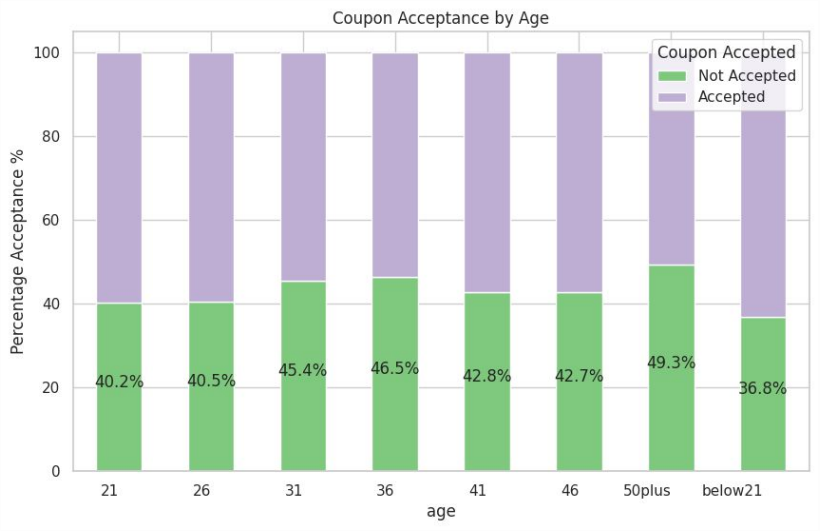
Insights exploration



- Across all time throughout the day, coupon for coffeehouse is the most popular coupon type, especially at time 6pm, but the acceptance dropped at time 10pm
- Coupon for bar has lowest acceptance at 2pm
- Coupon for restaurant has highest acceptance at time 2pm and 6pm
- Coupon for bar is more popular around 6pm or after, indicating interviewees could be working class.
- People accept all types of coupon at 7am, 6pm in average.
- Afternoon time around 2pm, it's not a good time for bar, carry out, and restaurant($20-50)
- At night time around 10 pm, all types of coupon has similar acceptance except for coffee house

Insights exploration





Unveiling coupon acceptance dynamics through visualizations, these graphs offer a nuanced perspective. The first, a bar chart, showcases the percentage acceptance of coupons categorized by types, illuminating preferred choices. Meanwhile, the second graph delves into age-based patterns, providing crucial insights for targeted marketing strategies.

Model Implementation

- **Data Preparation:**
  - Data partition: 70% train, 30% test
  - Data preparation: Ordinal, frequency, one-hot
- **Models:**
  - Logistic Regression, Categorical Naive Bayes, KNN, SVM, Decision Tree, Random Forest, ANN
- **Process:**
  - SKlearn models on three data structure
  - Insights analysis on predictions matrix:
    - odds of features: subcategory with the highest chance to accept coupon
    - true positive rate(recall): correctly predicting positive class
  - Cross validation for best model: SVM, Random Forest, KNN
  - Model summary and selection

8. Results - Odds

	odds
occupation_Construction & Extraction	2.387616
coupon_Carry out & Take away	2.216430
occupation_Healthcare Practitioners & Technical	2.149012
education_Some High School	2.056463
coupon_Restaurant(<20)	1.981822
RestaurantLessThan20_-1	1.946451
direction_same	1.763865
destination_No Urgent Place	1.618358
expiration_24.0h	1.569936
CoffeeHouse_4~8	1.530253
CoffeeHouse_1~3	1.509735
occupation_Healthcare Support	1.449617
weather_Sunny	1.401641
occupation_Architecture & Engineering	1.380484
CarryAway_-1	1.349410
CarryAway_never	1.301967
passanger_Partner	1.301792
Bar_1~3	1.289228
time_18	1.283508
occupation_Protective Service	1.258615
income_50000−62499	1.250686
occupation_Farming Fishing & Forestry	1.247989
passanger_Friend(s)	1.226461
toCoupon_GEQ25min	1.207421
maritalStatus_Single	1.197659
Restaurant20To50_4~8	1.189964
Restaurant20To50_1~3	1.187270
occupation_Life Physical Social Science	1.177818
Bar_4~8	1.172880

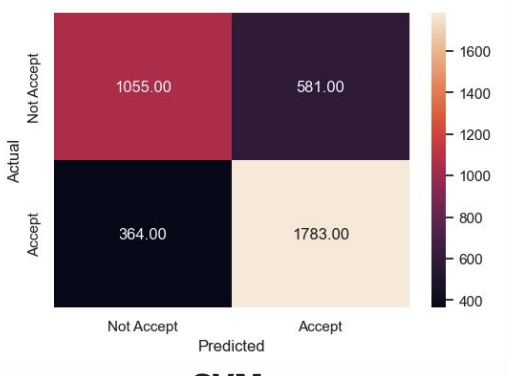
- **Occupation:**
  - _construction & Extraction: 2.4
  - Healthcare Practitioners & Technical: 2.1
- **Coupon**
  - _Carry out & Take away: 2.2
- **Education**
  - Some high school: 2
- **Direction**
  - Same direction: 1.7
- **Destination**
  - No urgent place: 1.6
- **CoffeeHouse**
  - 4~8: 1.5
- **Expiration**
  - 24 hours: 1.6

8. Results - Prediction

	precision	recall	f1-score	support
0	0.75	0.66	0.71	1636
1	0.77	0.83	0.80	2147
accuracy			0.76	3783
macro avg	0.76	0.75	0.75	3783
weighted avg	0.76	0.76	0.76	3783

Random Forest

	precision	recall	f1-score	support
0	0.74	0.64	0.69	1636
1	0.75	0.83	0.79	2147
accuracy			0.75	3783
macro avg	0.75	0.74	0.74	3783
weighted avg	0.75	0.75	0.75	3783



SVM

	precision	recall	f1-score	support
0	0.70	0.65	0.68	1636
1	0.75	0.79	0.77	2147
accuracy			0.73	3783
macro avg	0.72	0.72	0.72	3783
weighted avg	0.73	0.73	0.73	3783

ANN

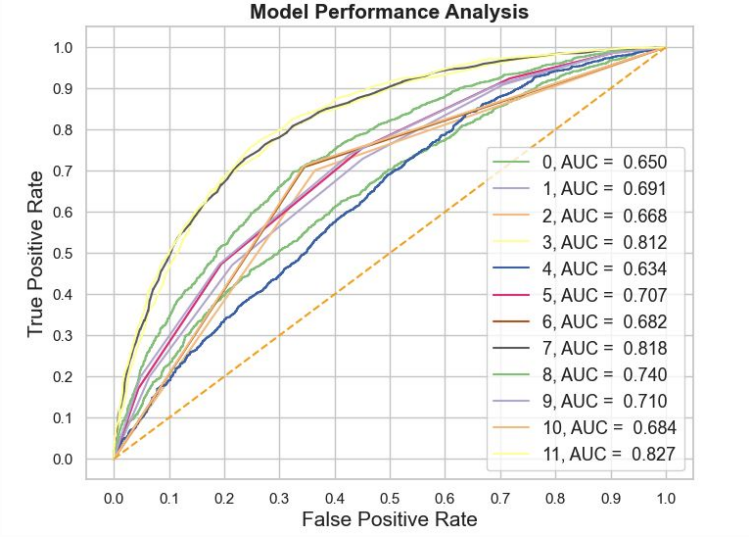
- Prediction Error: incorrectly predicting rejection when one wants coupons(FN)
- Average accuracy: 0.75
- Correctly identify both negative and positive classes
- Limitation: size of instances is not large enough for ANN

8. Results - Summary

	Datasets	Model Name	AUC	Recall	Precision	F1
0	dataset_ordinal	LogisticRegression	0.649792	0.756404	0.634871	0.690329
1	dataset_ordinal	KNeighborsClassifier	0.690852	0.72939	0.679688	0.703662
2	dataset_ordinal	DecisionTreeClassifier	0.668397	0.698649	0.716332	0.70738
3	dataset_ordinal	RandomForestClassifier	0.811613	0.819283	0.756885	0.786849
4	dataset_freq	LogisticRegression	0.634034	0.797857	0.630939	0.704648
5	dataset_freq	KNeighborsClassifier	0.70728	0.755007	0.687739	0.719805
6	dataset_freq	DecisionTreeClassifier	0.682027	0.706102	0.729548	0.717633
7	dataset_freq	RandomForestClassifier	0.818018	0.80857	0.766108	0.786766
8	dataset_OneHot	LogisticRegression	0.7404	0.770377	0.706838	0.737241
9	dataset_OneHot	KNeighborsClassifier	0.710204	0.739637	0.694359	0.716283
10	dataset_OneHot	DecisionTreeClassifier	0.684142	0.712156	0.73088	0.721397
11	dataset_OneHot	RandomForestClassifier	0.826616	0.825803	0.769197	0.796496

	Accuracy	AUC	Recall	F1
svm_Ordinal_enco	0.676183	0.726871	0.767117	0.728922
svm_Freq_enco	0.702881	0.702881	0.763391	0.744662
svm_OneHot_enco	0.750198	0.750198	0.830461	0.790512
knn_Ordinal_enco	0.652921	0.686315	0.775966	0.717330
knn_Freq_enco	0.659265	0.691363	0.792734	0.725336
knn_OneHot_enco	0.706318	0.766186	0.838845	0.764269
rf_Ordinal_enco	0.746497	0.814423	0.824872	0.786936
rf_Freq_enco	0.748612	0.821618	0.815557	0.786436
rf_OneHot_enco	0.758128	0.829371	0.851421	0.799825

RandomSerachCV



- **Best performance:** Random Forest
- **Accurate prediction:** Random Forest, SVM, ANN
- **Best Data Structure:** One-hot encoded

9. Values & next steps

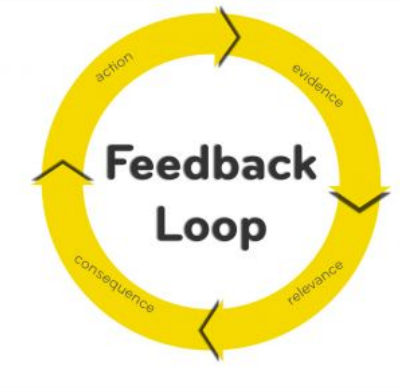
Values

- **Personalized Marketing**
  Targeting right customers
- **Increase customer engagement**
  i.e. Coffeehouse coupon subscription
  Restaurant coupon with conditions...
- **Driving Sales and Revenue**
  Encouraging consumptions
- **Customer Retention and Loyalty**
  Personalized experience
- **Marketing budget management**
  Right customers, less marketing costs
- **Data-driven insights**
  Valuable insights like customer preferences

Next Steps

- **Data Collection**
  Diversity, randomness, feature selection
- **Model improvement**
  Refine algorithm, incorporate more data
- **Feature Engineering**
  Apply feature engineering for better fitting
- **Domain knowledge**
  Research knowledge on marketing strategy and consumer behavior
- **Scalability**
  Assess the scalability of marketing plans
- **Continuous Updating**
  Regularly check model performance

10. Socialization & Distribution



Thank You

Do you have any questions?