Today's Agenda

Which ad campaign would a customer most likely to accept based on their profiles and shopping behavior?

1

Basic Understanding

- Business Content
- Data Detail
- Relationships in Data

2

Modeling

- Ad Recommendation
- Model Details

3

Evaluation

- Acceptance Probability Models
- Ad Recommendations

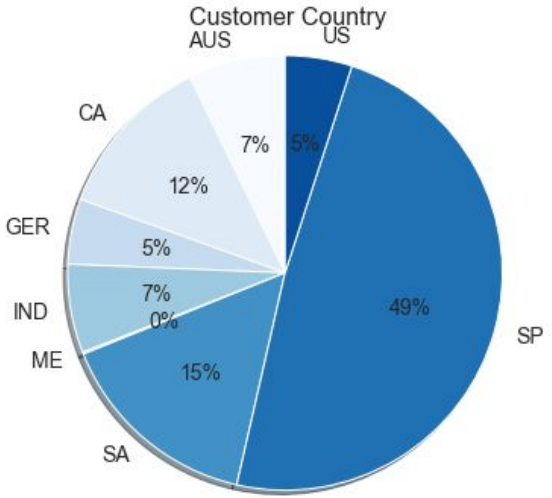
Context

- Our online grocery store is testing a marketing campaign that consists of 5 ads.
- We showed all 5 ads to a selection of customers and recorded their responses.
- Goal: develop a way to recommend one of the 5 ads to a customer, based on their profile and shopping data, to maximize the probability they act on (or "accept") the ad.
- Solution: we developed ad recommendation algorithms using machine learning models that roughly **doubled acceptance rates** compared to the best individual ad.
- Benefits:
  - Marketing: **improve effectiveness** in advertising
  - Financial: **reduce cost** on unnecessary advertisement
  - Client: improve client's **user experience**

Overall Acceptance Rates	
Ad 1	6.4%
Ad 2	1.3%
Ad 3	7.3%
Ad 4	7.5%
Ad 5	7.3%

Data Detail

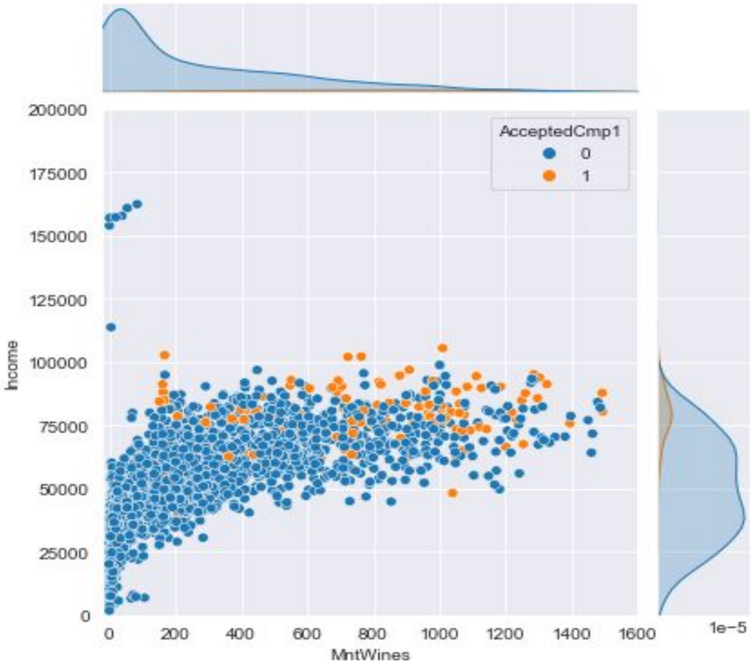
- Data collected for **2,240** unique customers
- 8 Demographic data fields, including:
  - Age, Income, Household Size, etc.
- 13 Shopping history fields, including:
  - Days since last purchase
  - Amount spent on various categories in last 2 years:
    - Fruit, Meat, Wine, Sweets, etc.
  - Number of purchases made with discount
- Data required minimal cleaning
- Data limitations:
  - Gathered over limited period of time



Relationships in Data

- Relationships exist between customer data and ad acceptance.
- ML models represent these relationships and predict the probability that a customer will act on each ad based on their profile and shopping history.
  - Example: customer who accept Ad 1 tend to buy at least some wine and be in upper half of income spectrum.
- Different ML models represent the relationships differently.
- We consider 3 alternatives:
  - Logistic Regression
  - Naive Bayes
  - Random Forest

Income and Wine Purchases Distribution, colored by Ad 1 response (accept in orange)



Ad Recommendation

- No matter which ML alternative we choose, the ad recommendation approach is the same.
- Acceptance probability: for each of the 5 ads we create a binary classification model to predict the probability that a customer accepts the ad.
- Ad recommendation: show the customer the ad where the predicted acceptance probability is highest among all 5 ads.

Predicted Acceptance Probability

Customer ID	Ad 1	Ad 2	Ad 3	Ad 4	Ad 5
0	23%	0.1%	6%	0.4%	31%
1	5%	2%	8%	9%	0.9%

Ad Shown to Customer

Max Predicted Prob.
Ad 5
Ad 4

Model Details

Description / Advantages:

Logistic Regression

LR: Linear relationship between customer data and odds of acceptance is ***easy to understand***.

Naive Bayes

NB: Independence relationship between customer data is ***easy to understand and performs*** well in other contexts.

Random Forest

RF: Does ***not make assumptions*** about relationships between variables (can be nonlinear or dependent); ***reduces overfitting*** by averaging multiple decision trees.

Customer Data used in Modeling:

1. Year of birth
2. # kids
3. # teens
4. Recency
5. **Wine purchases**
6. Fruit purchases
7. **Meat purchases**
8. Fish purchases
9. Sweets purchases
10. Gold purchases
11. Purchases with discount
12. Web purchases
13. Catalog purchases
14. In-store purchases
15. Web visits per month
16. Complaints
17. **Income range**

Evaluation of Acceptance Probability Models

Average ROC AUC	Ad 1	Ad 2	Ad 3	Ad 4	Ad 5
Random Forest	**91.85%**	**86.11%**	**80.98%**	**90.66%**	**96.95%**
Logistic Regression	91.26%	81.56%	74.15%	87.38%	96.36%
Naive Bayes	86.03%	70.25%	64.29%	76.16%	91.95%

- We use the training data for cross-validation with **k=5 folds**; take the average ROC curve AUC over the 5 folds to assess model performance.
- In overall, **Random Forest** has the highest accuracy, and Naive Bayes has the lowest.

Evaluation of Ad Recommendations

Methodology Summary: Test ad-picking success on the customers held out of the training data.
For comparison, show the acceptance rate of each ad individually.

	No Model					Ad Recommendation Models		
	Ad 1	Ad 2	Ad 3	Ad 4	Ad 5	**Random Forest**	Logistic Regression	Naive Bayes
Acceptance	8.04%	1.56%	8.71%	8.04%	7.14%	**17.41%**	16.07%	15.85%

- Model-based ad-picking algorithms have much better acceptance rates than any single ad (***about 2 times higher***).
- The ad-picking algorithm based on ***Random Forest models has the highest acceptance rate***.

Key Takeaways

- Model-based ad recommendations can greatly increase the rate of ad acceptance, ***potentially doubling the acceptance rate*** of any individual ad.

- We recommend the ***Random Forest*** ML alternative, based on:
    - Its recommendations had the ***highest acceptance rate*** on the test dat
    - Its 5 underlying acceptance probability models were the ***most accurate*** compared to other alternatives

- Model should be ***re-trained over time*** using new customer data

- Directions for future analysis:
    - Recommend ads based on expected profit
    - Develop models for which ad to show second, third, etc.

Thank You!