Navigation-Based Personalized Recommender System

Millie Huang, Vicky Liu

Advised by: Professor Taufid Zaman, Lei Zhang
With support from: Anthony Accardi, Joelle Angel, Saurabh Chadha, Abhishek Katiyar, Ronny Mathew, Megan Wellons

I. Data Description and Data Processing

- Training: 9M observations (4/2019–7/2020); testing: 500K observations
- Transform historical user viewing data into product transitions

![Diagram showing data processing](image)

II. Clustering Algorithm Design

Based on historical site browsing data, cluster users:
1. Randomly assign users to groups, calculate transition probability for all product transitions for a given user group
   - Transition probability within group: \( \frac{N_{ij}}{N_j} \)
2. Adjust group assignment to maximize sum log of transition probabilities
3. Recalculate in-group transition probability
4. Iterate until number of changes for entire epoch ≤ 5% of users

![Diagram showing clustering process](image)

III. Recommendation Algorithm Design

When a user logs on to Rue Gilt website, provide a list of product recommendations based on group assignment
- Rank recommendations based on a calculated score
  - Score = Transition probability from product i to product j (transition probability from current product to the recommended product) x Probability of ending the browsing session after product j (recommended product)
  - \( E[U] \) = Expected length for product j (recommended product)

![Diagram showing recommendation process](image)

IV. Evaluation Mechanism Design

- Recommendation accuracy: Percentage of hitting the next browsed product within the first 15 tries
- Reduce the probability of user immediately exiting the site
- Increase expected length of browsing

![Distribution of Accuracy](image)

![Distribution of Expected Length](image)

V. Business Impact

- Use user engagement as the primary business metric
- Generated by the extension of user length of stay on the website
- Assumptions:
  - Each user will view 13 more products
  - Historical data on order rates vs. number of products viewed, average revenue per order, and average daily visitors (with length > 1 filter)
  - 10-20% customer adoption rate of new feature
- Business impact calculated as:
  - Expected increase in revenue = revenue with new feature - revenue without new feature + daily visitors \( \times \) order rate at new number of products viewed per user \( \times \) new feature adoption rate of new feature \( \times \) order rate at current number of products viewed per user \( \times \) new feature adoption rate of new feature \( \times \) average order value
  - Daily visitors \( \times \) order rate at current number of products viewed per user \( \times \) new feature adoption rate of new feature \( \times \) average order value

![Business impact calculation](image)

1-3% expected annual revenue lift across Rue and Gilt

Special thanks to the entire Rue Data Science team for being incredible sponsors and mentors and giving us an amazing experience, to our MIT advisor for his continued help and support in this project, and to all others who made this fun, challenging, and impactful project possible!