Instacart Recommender System

Aaron Xiao, Kiran Sidhu, Paul Flemming, Rozi Hagos, Tara Singh

Executive Summary

The overall objective is to create a system that will make a recommendation based on the last order and what is currently in a user's cart. In order to achieve this goal, we have built a recommender system and tested five models to obtain the best result. Overall, the popular items model gave the best result apart from item-based collaborative filtering model which would be our suggested recommender model. The reason why we chose IBCF over popular is that the IBCF is the best performing model that is tailored to each individual user's cart.

Business Objectives and Problem Formulation

The online grocery shopping market is expected to grow 500% over the next decade. More than 70% of consumers will be shopping for groceries online by 2025. Instacart is planning on offering same day delivery of fresh groceries to 80% of American homes in 2018. Instacart earns revenue based on size of orders as well as through revenue based on advertisements. In addition, promotions of packaged goods are the now fastest growing revenue model for them. Delivery fees have been the traditional method of obtaining revenue for Instacart, however, they are currently focusing on customer experience and exploring additional revenue streams. Recently, they have been testing promoted search results on their site, using pay-per-click and add-to-cart ads which represents 20% of their advertising revenue.

Even with quick expansion, Instacart has not yet turned a profit. We are the consultants that Instacart has engaged to utilize their prior order transactional data to develop models that recommend new products to new and existing users. We will use this anonymized data on customer orders over time to recommend products as a customer builds their cart and their user profile. The goal is to increase user engagement, satisfaction, and loyalty over time. The business objective is to build a recommender system for Instacart that will recommend products that the consumer will be more inclined to try and purchase based on their previous order history.

Aaron Xiao, Kiran Sidhu, Paul Flemming, Rozi Hagos, Tara Singh

Methodology

The datasets provided are a relational set of files describing customers' orders over time. The dataset contains samples of over 3 million grocery orders from more than 200,000 Instacart customers. The dataset contains 4 to 100 orders from each customer, with the add-to-cart sequence of products purchased in each order. The week, hour of day the order was placed, and a relative measure of time between orders are also included in the datasets. In the sections that follow, we outline our preliminary analysis and descriptive analysis of the data, the models are developed and tested, along with our insights and final recommendation.

Preliminary Analysis, Data Manipulation & Descriptive Analysis

In our preliminary analysis, we found that close to 60% of the orders were from customers who have previously shopped on Instacart. We discovered that customers tended to reorder at around the seventh day after their initial order (See Appendix A: Figure 2) and they normally had at least 3 prior orders.

We determined that the majority of orders had a size of 5-10 orders, so we focused on that order segment. The number of items per order (See Appendix A: Figure 4) starts to decline at 10 items and trails off after 20 items have been placed in the cart.

Fresh produce was the most popular department with fresh fruits and vegetables being the most popular aisle (See Appendix A: Figures 1, 2, 3 & 5). Also, the top 10 products per hour ordered (See Appendix A: Figure 3), clearly confirms that the top products consist of fresh produce. As opposed to fresh produce, pantry items have a longer shelf life and therefore do not need to be repurchased as frequently which is reflected in the data.

Model Building and Evaluation

We developed five models for this recommender: Association Rules, Random, Popular, Item Based Collaborative Filtering (IBCF) and User Based Collaborative Filtering (UBCF). Due to processing limitations and technical complications, we had to filter and reduce the size of the data in order to conduct the analysis. Since the median cart size contains 8 products and the mode cart size has 5 products, we targeted a basket of 5-10 items for our recommender as a way to reduce to the size of the data and better focus on that segment (See Appendix A: Figure 4). It represents the majority of the data with respect to the order size and removes any outliers that exist.

Insights and Results Summary

We found that the popular items model had good results, but would likely not be useful in practice, as these items are already being purchased. It would however be useful to suggest items for new customers. Overall, ICBF model would be our suggested recommender model. IBCF was chosen over popular items because it had the best performance along with being customized to each individual user's cart.

The ROC curve was created to compare the models. A ROC curve is constructed by plotting the true positive rate (TPR) against the false positive rate (FPR). The area under the ROC curve is a measure of the usefulness of the test. Classifiers that give curves closer to the top-left corner indicate a better performance. Two ROC curve analyses were conducted for the models (See Appendix B: Figures 8 & 9). Popular items ranked as the best performing model, followed by IBCF in second, UBCF is third, fourth is the AR model and the least effective model was the random items model.

The Precision-Recall (PR) curve method was used to evaluate the models and measure the relevance. Precision is the measure of result relevancy and recall is a measure of how many truly relevant results are returned. The PR curve shows the trade-off between precision and recall for different thresholds. A great PR curve result would show the classifier in the farthest upper right corner of the graph where precision and recall equal 1. Overall, the popular items ranked as the best performing model, followed by IBCF in second, and UBCF in third. The association rules model and random items ranked poorly (See Appendix B: Figures 8, 9,10 & 11).

Recommendations

Instacart needs to focus on personalization and improving the user experience. Since the customers will lack the hands-on experience of buying in person, Instacart needs to focus on the strengths of buying online. By understanding their customers likes and preferences in order to customize the online experience, it allows Instacart to tailor the experience to suit each individual. An additional benefit of using our recommender system, is that the Instacart shoppers picking the items in person can use the app to find replacements using the algorithm instead of having to make an ad-hoc choice on behalf of the customer.

It might also be useful to add sale items, niche products or less frequently purchased items to the recommender for novelty, price competitiveness, to secure ongoing customer satisfaction and loyalty. Additionally, it could be a way to discover new popular items and an opportunity to remove items where sales are lagging.

Performance Management

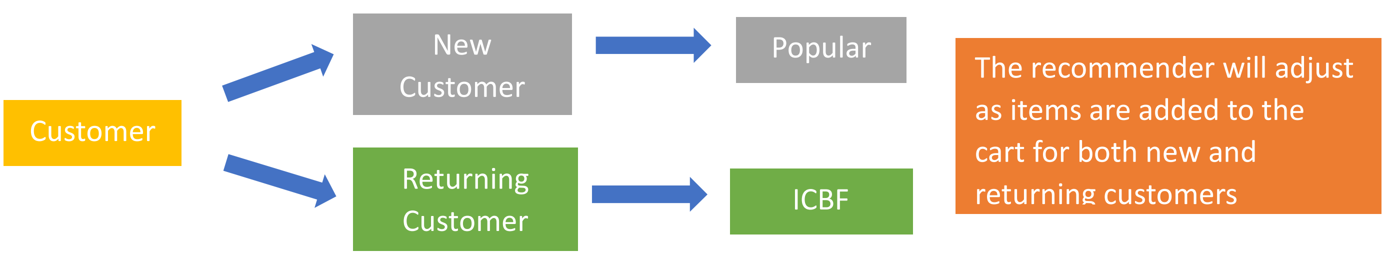
This is not a static model, as more orders are added to your dataset with different items ordered, the output of the recommender model will change so you will need to run the model regularly to obtain the best results. Since customers order on a weekly basis, it would be useful to update the recommender on a weekly basis as well to align with the frequency which customers normally place their orders. By updating the model frequently, the recommendations that the system will provide will be continually refreshed and more aligned with what customers are currently purchasing.

Cost Benefit Analysis

Data management costs would include data storage, maintenance, processing, and data governance. Model management costs include model development, testing, evaluation, and maintenance. Since Instacart order data is internal, continually increasing and automatically updated, there is minimal cost for them to acquire the data. Any cost that they incur to analyze the data has the potential to contribute to increase sales, create greater revenue and profits as well as increased customer loyalty. As their customer base increases, there would also be an incremental increase in the above noted benefits as each customer tends to spend an average of $100 – 120 per order.

Conclusion

The business objective is to build a recommender system for Instacart that will recommend products that the consumer will be more inclined to try and purchase based on their previous order history. The goal is to increase user engagement, satisfaction, and loyalty over time. Although, the popular items model was the best performing model, IBCF tailored suggestions to what each customer was adding to their cart (See Appendix B: Figures 8 & 9). The Popular items model fills in the gap for new customers by suggesting the most popularly purchased items to them (See the diagram below). IBCF was ultimately the model that we suggested as it creates a more personalized customer experience. This strategy aligns with the business objectives and goals for future business growth and expansion.



Aaron Xiao, Kiran Sidhu, Paul Flemming, Rozi Hagos, Tara Singh

Appendix A:

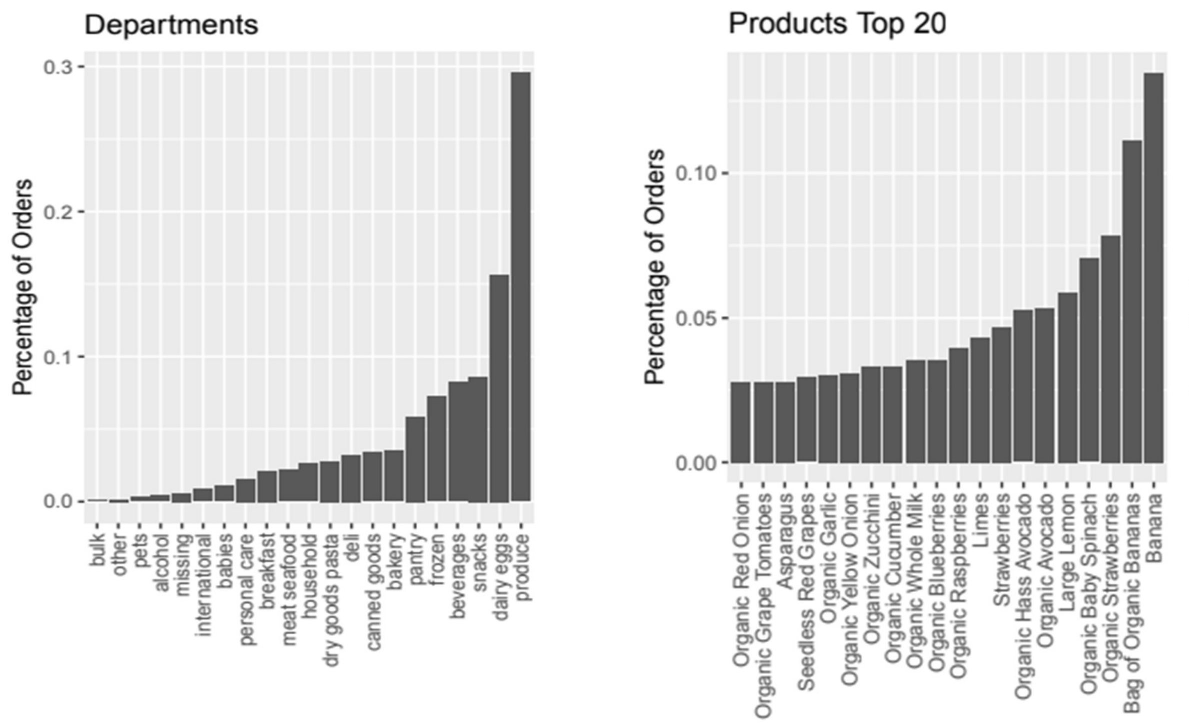


Figure 1: On the left, Percentage of orders by departments and on the right percentage of the Top 20 Products

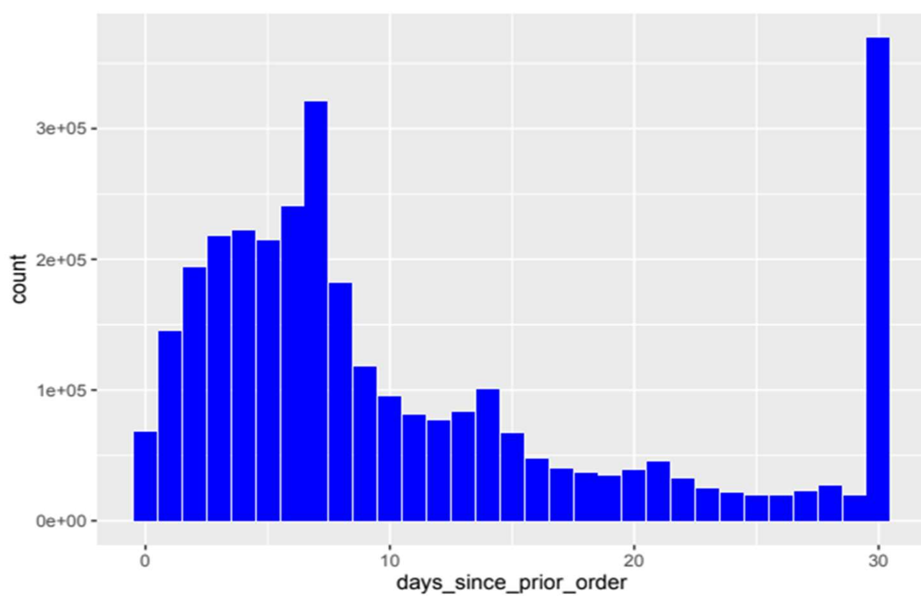


Figure 2: Number of days since the last order placed

Figure 3: Daily Top 10 Products Ordered by DOW

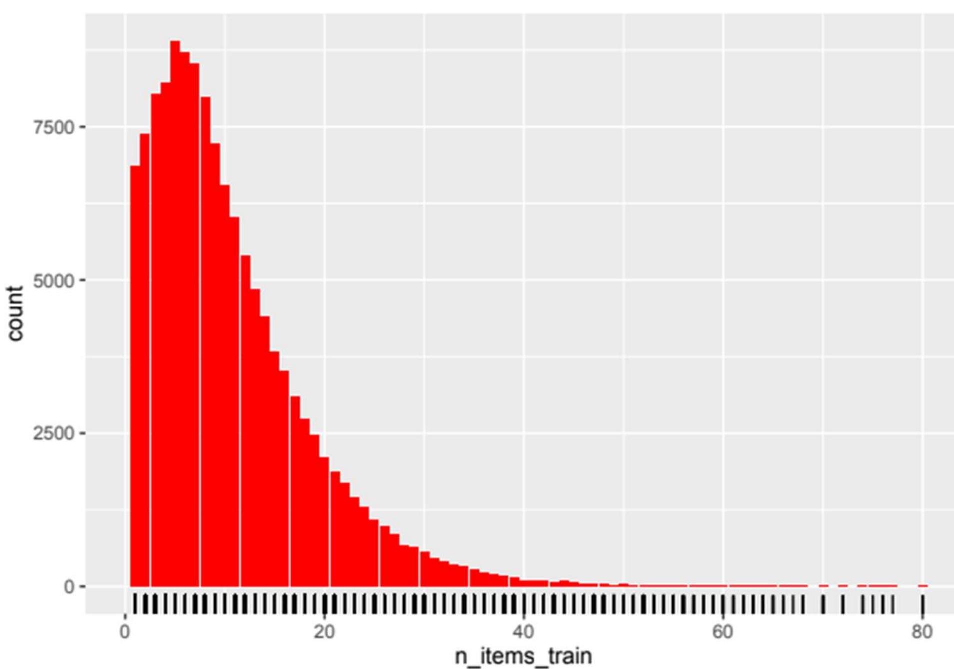


Figure 4: Number of items in each order

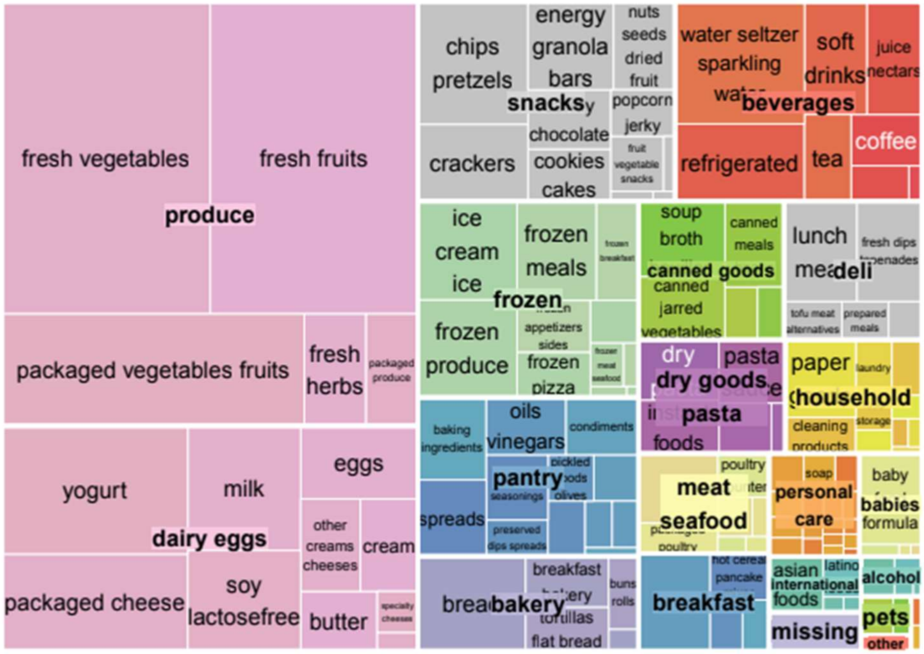


Figure 5: Treemap indicating the departments and aisles

```
sample_order
```

```
##     product_id                      product_name
##          <int>                            <fctr>
## 1:       39821 S.O.S Reusable Steel Wool Soap Pads
## 2:       31562                      Sweet Onions
## 3:        6844        Organic Unsulphured Molasses
## 4:       30960          Regular Pork Sausage Tube
## 5:       13914        Cheez-It Baked Snack Crackers
```

Figure 6: Sample order that was input to the recommender models

```
##     product_id                        product_name aisle_id department_id
##          <int>                              <fctr>    <int>         <int>
## 1:        1215 Kidz All Natural Baked Chicken Nuggets      129             1
## 2:        1940              Organic 2% Reduced Fat Milk       84            16
## 3:        4037                   Blackberry Preserves       88            13
## 4:        4658                  Imported Mineral Water      115             7
## 5:        4913                    Table Water Crackers       78            19
```

Figure 7: Sample output from the user-based recommender model

Aaron Xiao, Kiran Sidhu, Paul Flemming, Rozi Hagos, Tara Singh

Appendix B:

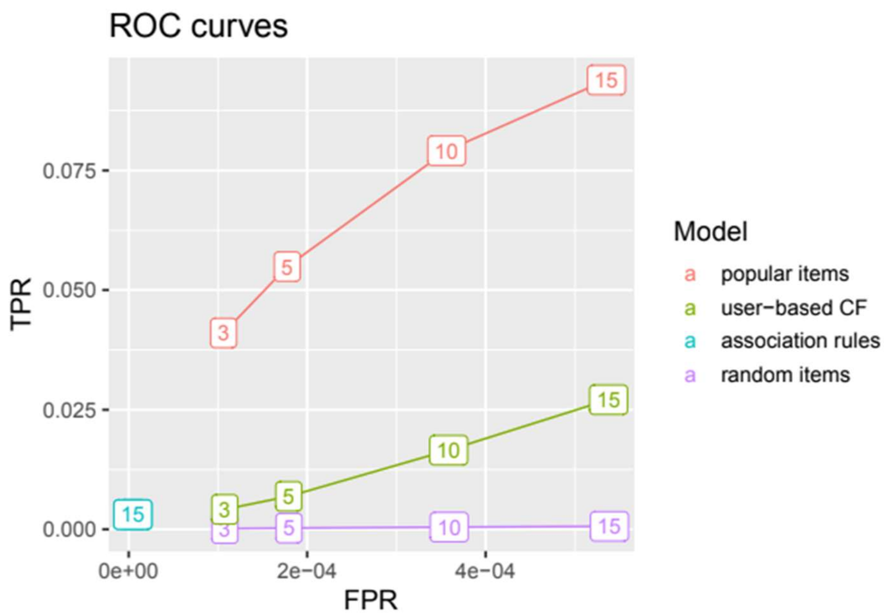


Figure 8: Classification Models Performance Comparison

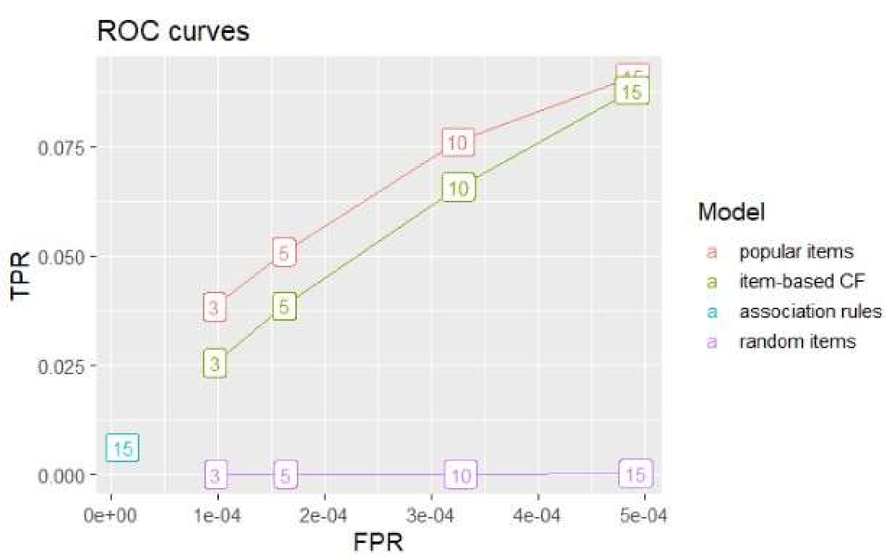


Figure 9: Model Evaluation Comparison

In the first analysis (Figure 8:

- the popular items ranked as the top model
- the second ranked model is the user-based CF model,
- third ranked is association rules model
- fourth ranked is the random items model.

The second ROC curve analysis (Figure 9):

- the popular items ranked as the top model,
- second ranked model is the item-based CF model,
- third ranked is association rules model
- fourth ranked is the random items model.
- Item-based CF model (Figure 9) is plotted higher left on the plot than user-based CF model (Figure 8).

Overall, the popular items ranked as the top model, second ranked model is the item-based CF model, third ranked is the user-based CF model, fourth ranked is the association rules model and the fifth ranked is the random items model.

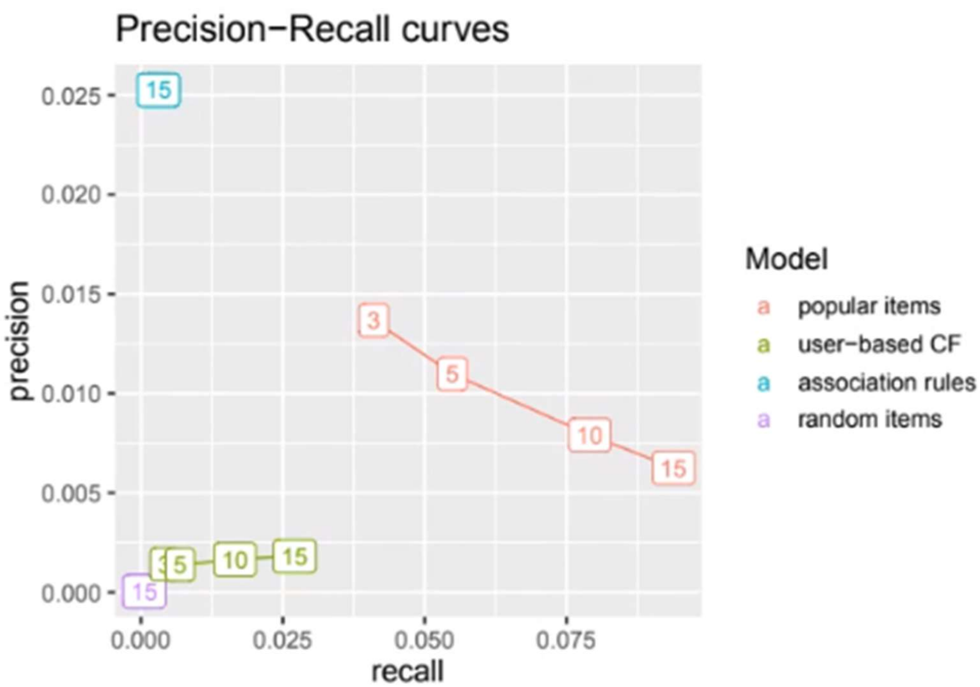


Figure 10: Precision Recall curves showing the relationship for each model

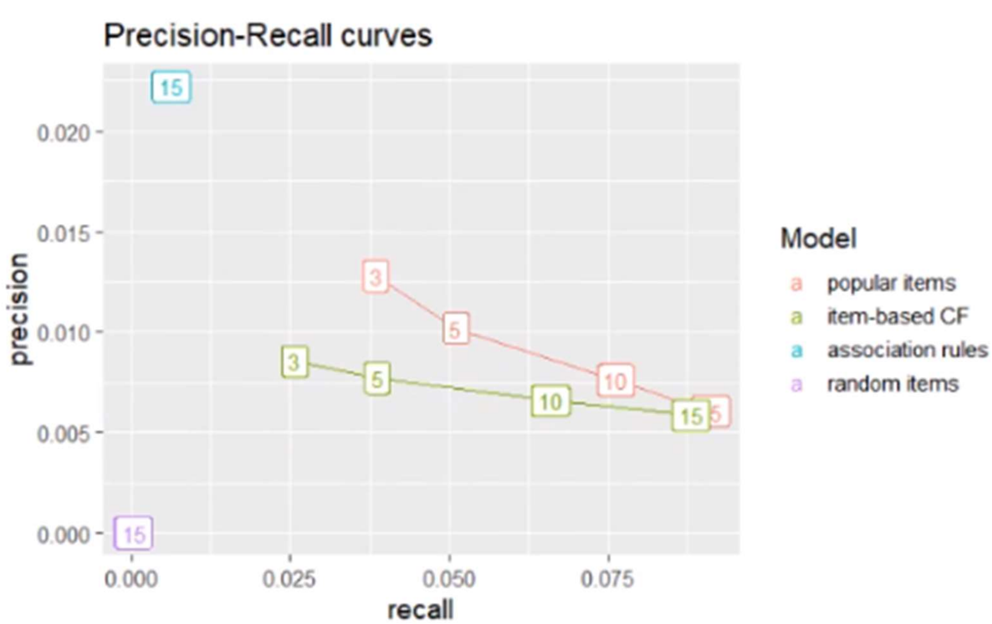


Figure 11: Precision Recall curve

In the Precision-Recall (PR) curve (Figure 10):

- the random items model is not relevant with a precision-recall of zero.
- The Association Rules (AR) model shows an imbalance with very high precision and low recall.
- The popular items model has the highest recall and precision.
- The user-based CF model (UBCF) would come in third.

In the second PR curve (Figure 11):

- the random items is not relevant with a precision-recall of zero.
- The AR model shows an imbalance again with very high precision and low recall.
- Popular items model has the highest recall and precision.
- The item-based CF model (IBCF) would come in second.
- When comparing where the IBCF and the UBCF model fall on the plot it is the IBCF model that has the higher precision and recall between the two models.

Overall, the popular items ranked as the top model, second ranked model is the IBCF model, third ranked is the UBCF model. The AR model and random items ranked poorly.

Aaron Xiao, Kiran Sidhu, Paul Flemming, Rozi Hagos, Tara Singh

References

"The Instacart Online Grocery Shopping Dataset 2017", Accessed from *https://www.instacart.com/datasets/grocery-shopping-2017 on September 12, 2020.*

Instacart is a spleen and it needs to be a kidney
https://techcrunch.com/2017/10/01/instacart-needs-to-deeply-invest-in-grocers-and-user-personalization-to-survive/

2017 IPO Prospects: Instacart Getting Ready to Deliver on the Stock Exchange
https://www.sramanamitra.com/2017/04/05/2017-ipo-prospects-instacart-getting-ready-to-deliver-on-the-stock-exchange/

Instacart AI helps personal shoppers buy groceries faster
https://venturebeat.com/2017/07/12/instacart-ai-helps-personal-shoppers-buy-groceries-faster/

Instacart itemizes its costs, even as it adds $400 million in funding
https://www.digitalcommerce360.com/2017/03/08/instacart-itemizes-its-costs-even-as-it-adds-400-million-in-funding/

How Instacart's big funding round aims to boost online grocery sales
https://www.digitalcommerce360.com/2017/03/09/instacarts-400-million-funding-round-will-fuel-expansion/

Exploring "The Instacart Online Grocery Shopping Dataset 2017"
https://medium.com/@dimitrilinde/exploring-the-instacart-online-grocery-shopping-dataset-2017-86db5156da2c