A COLLABORATIVE FILTERING-BASED RECOMMENDATION SYSTEM FOR AN ONLINE HIGH-END RETAILER

by
BARDIA ALIZADEH MOGHTADER

Submitted to the Sabancı Graduate Business School
in partial fulfilment of
the requirements for the degree of Master of Science in Business Analytics

Sabancı University
June 2021

A COLLABORATIVE FILTERING-BASED RECOMMENDATION SYSTEM FOR AN ONLINE HIGH-END RETAILER

Approved by:

Date of Approval: June 29, 2021

ABSTRACT

A COLLABORATIVE FILTERING-BASED RECOMMENDATION SYSTEM FOR AN ONLINE HIGH-END RETAILER

BARDIA ALIZADEH MOGHTADER

Business Analytics M.Sc. THESIS, JUNE 2021

Thesis Advisor: Assoc. Prof. Dr. Ayşe Kocabıyıkoğlu
Thesis Co-Advisor: Asst. Prof. Dr. Burak Gökgür

Keywords: Collaborative Filtering, Recommendation Systems

In online retail platforms, consumers seek to find the products that are best suited for their needs while limiting their search efforts. With the growing trend in online shopping, retail companies utilize a range of tools to assist customers in their journey and improve their purchase experience. One of the tools that can minimize these exploration efforts is recommendation systems that suggest a tailored set of available product options to consumers based on their preferences. In this thesis, we focus on a high-end Turkish retailer that did not utilize such engines in its practice and study the value that these systems can provide to the company. To that end, we implement a collaborative filtering-based recommendation system that uses the similarity of the consumers to derive their preferences and suggest item sets for their next purchase. We evaluate the recommendation model with transactions data to acquire the hyper-parameters and test it on the transactions made in the last month and provide recommendations on three granularity levels. We also analyze the predicted preferences to suggest bundling options and derive empirical insights. We found that by generating 20 suggestions for each customer in their shopping session, the engine can reach an accuracy of 38% at brand-level and 7% at item-level while using cosine similarity as its similarity metric.

ÖZET

ÇEVRIMIÇI ÜST DÜZEY BIR PERAKENDECI IÇIN İŞBIRLIĞINE DAYALI FILTRELEME TABANLI ÖNERI SISTEMI

BARDIA ALIZADEH MOGHTADER

İş Analitiği YÜKSEK LİSANS TEZİ, Haziran 2021

Tez Danışmanı: Prof. Dr. Ayşe Kocabıyıkoğlu
İkinci Tez Danışmanı: Asst. Prof. Dr. Burak Gökgür

Anahtar Kelimeler: Collaborative Filtreleme, Öneri Sistemleri

Çevrimiçi perakende platformlarında tüketiciler, arama çabalarını sınırlandırırken ihtiyaçlarına en uygun ürünleri bulmaya çalışırlar. Online alışverişte artan trendle birlikte perakende şirketleri, müşterilere yolculuklarında yardımcı olmak ve satın alma deneyimlerini iyileştirmek için bir dizi araç kullanmaktadır. Bu keşif çabalarını en aza indirebilecek araçlardan biri, tüketicilerin tercihlerine göre uyarlanmış bir dizi mevcut ürün seçeneğini onlara öneren öneri sistemleridir. Bu tezde, uygulamalarında bu tür motorları kullanmayan üst düzey bir Türk perakendecisine odaklanmaktayız ve bu sistemlerin şirkete sağlayabileceği değeri incelemekteyiz. Bu amaçla, tercihlerini türetmek ve bir sonraki satın alımları için ürün setleri önermek için tüketicilerin benzerliğini kullanan, işbirlikçi bir filtrelemeye dayalı öneri sistemi uygulamaktayız. Hiper parametreleri elde etmek için öneri modelini işlem verileriyle değerlendirmekteyiz ve geçen ay yapılan işlemler üzerinde test ederek üç ayrıntı düzeyinde öneriler sunmaktayız. Paketleme seçenekleri önermek ve ampirik içgörüler elde etmek için tahmin edilen tercihleri de analiz etmekteyiz. Motorun, her müşterinin alışveriş oturumlarında 20 öneri oluşturarak, benzerlik metriği olarak kosinüs benzerliğini kullanırken marka düzeyinde %38 ve ürün düzeyinde %7 doğruluğa ulaşabildiğini gözlemledik.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to my family, who have been a great source of support, especially my brother. I am deeply grateful to Ayşe Kocabıyıkoğlu and Burak Gökgür, whose guidance and encouragement has been invaluable throughout this study.

Dedication page

I dedicate my dissertation work to my family and many friends.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

Over the past two decades, personalization has become a core component of online retailing settings. Considering the wide variety of goods found in online retail platforms, even seasoned customers become overwhelmed by the sheer number of products presented to them. This leads to a problem when a consumer's product choice is discontinued, and they need to find a substitute for it, causing them to be reluctant to purchase products. As a result, companies have invested heavily in diverse technologies and analytical tools, such as recommendation systems, to guide consumers in their purchases. These systems have proven beneficial not only for users but also for retailers. On the one hand, they assist consumers in finding interesting items by providing personalized recommendations and filtering a tailored set from a large pool of products based on their preferences. On the other hand, they boost the retailer's revenue by increasing its sales—resulted from recommending their customers with a personalized assortment that matches their preferences. Assisting customers in finding their desired products with more ease will also increase their satisfaction from the retailer and result in a raise in their loyalty (Ansari & Mela, 2003; Arora et al., 2008). Another way that these systems are considered necessary is the effects of long-tail phenomena. While the shelf capacity constraints on physical institutions make them provide only the most popular items, this phenomenon forces online institutions to recommend items to individual customers since it is impossible to present all available items to the user (Leskovec et al., 2014). Moreover, the benefit of these systems is not just limited to the retail industry, and they contribute to other online settings by either offering relevant topics in content-curation platforms or helping users navigate through web pages based on their interests.

There are numerous examples where online retailers use recommendation systems to provide their customers with personalized recommendations. For instance, Amazon.com utilizes collaborative filtering to provide their customers with customized product suggestions (Arora et al., 2008; Linden et al., 2003). Content curation websites like Digg and Reddit filter and find popular subjects based on user visits to the articles and suggest them to people who are likely to find them interesting. In social media networks such as Twitter and Facebook, a small subset of personalized posts are created for users based on their friend groups' feedback on the content post and their interests. In all these applications, the data from heterogeneous users' previous interactions are utilized to generate recommendations.

The potential value that these systems can provide to businesses makes them an ideal candidate for studying. The primary focus of research in this area has been on improving algorithmic design and system performance. However, in this paper, we study the value that these systems can provide to online businesses and the insights that can be derived from them. Specifically, we focus on Baskasindarama.com, an online retail company that did not previously benefit from such systems, making it a great candidate to study the effects of these systems. This Turkish-based retailer sells lifestyle products as well as perishable items and has provided us transactional data of their 2020 sales. In this study, we implement a recommendation engine for this retailer using a collaborative filtering approach to provide an optimal recommendation policy to their customers. We utilize two distinct similarity measures to derive customers' preferences and compare the engines' performance with these parameters. We also evaluate the performance of the models on a test dataset using a simple experiment in an offline setting and study the efficiency of these models based on the number of predictions they have been set to provide for different granularity levels. After building the models and obtaining the customer preferences, we analyze the correlation at various levels to find item pairs that can be bundled and provide insights. Our results illustrate models with the Pearson correlation as the similarity criteria and the inputs as the transactions' count would yield lower RMSE. However, in the experiment we conducted, models with cosine similarity as the distance measure provided a higher accuracy when predicting the items in customer baskets with fewer recommendations at the brand level on the test data. These findings show that various parameters affect the performance of models, such as normalization choice and distance measures. Moreover, the nature of data, sparsity, and duration of the captured data also influence the choice of parameters for building the models. This enforces that there is no one solution for all and the need to evaluate the models with different performance measures to access their performance.

Our study focuses on applying recommendation systems in online retail environments and the use of personalized item offerings to consumers to facilitate their search for products that, in return, have benefits both for customers and the retailers' yield. The use of data with this subject matter makes us contribute to the data-driven assortment personalization literature. Moreover, using these systems to provide bundling options contributes to the literature on product bundling. It goes without saying that this research builds upon the already existing body of literature concerned with recommendation systems. More specifically, knowledge discovery, information filtering, and information extraction from big-data

The organization of this thesis is as follows: Section 2 provides a background and literature review on the recommendation systems. Section 3 characterizes the empirical setting of the retailer and describes the data utilized for this thesis. In this section, we will also analyze the data and provide some descriptive information on the data. Section 4 discusses the methodology used for building the models. Section 5 presents the results derived from applying these models to the data and studies them. Finally, in section 6 the concluding remarks are presented, and the future directions for improvements are discussed. The auxiliary information is provided in the appendices.

2. Literature Review

In the past, people used to rely on word of mouth to make purchasing decisions. As the marketing industry grew and companies started to advertise their products, people became more familiar with the different products and made more conscious decisions about what they wanted. However, the sheer variety of products in the long-tail distribution made it impossible to become aware of all the products through advertisement. The online retailers tried to address this problem by providing the trending products as suggestions, which were the most sold items in the previous days. Yet, the heterogeneity in people's tastes creates the need for more personalized content. With the adoption of big data and predictive analytics in retail environments, companies utilized the system to dynamically provide suggestions to their customers to address customers' personalization needs and improve their experience (Bradlow et al., 2017). The emergence of these needs and the need for improvement also motivated researchers to design optimal performing recommender systems and improve upon the accuracy of suggestions.

The modern design of recommendation systems was first mentioned in a technical report as a "digital bookshelf" in 1990 by Jussi Karlgren at Columbia University. However, the first implementation of these systems dates back to 1992, where the researchers at Xerox Palo Alto Research Center (PARC) invented Tapestry (Goldberg et al., 1992) to address the employees' issue of becoming overwhelmed by the vast amount of incoming emails. Although some of the early recommendation system designs posed some challenges, such as sparsity, cold start, scalability, and overspecialization (Sharma & Singh, 2016), over the years, much research has been done to find solutions for these problems (Chen & Chen, 2019; Elahi et al., 2016; Jiang et al., 2014; Rubens et al., 2015; Schein et al., 2002). These solutions soughted to improve the engines' performance and decrease their time complexity, making them more suitable for fast-paced online retail environments to alleviate the customer purchase experience.

The study of personalization builds upon the body of research done on both management sciences and retail operations. The focus of research on personalization in management sciences is on three key stages: learning about consumer preferences, matching offerings to customers, and evaluation of the learning and matching processes (Murthi & Sarkar, 2003). In retail operations, personalization falls under the emerging research category of analytics and prediction and influence of the behavior of consumers, which is making an impact in both academia and industry. One aspect that the work on personalization in this area is focused on is providing customized assortments to the consumers, which is similar to the personalized product offerings displayed online through the use of recommendation systems (Caro et al., 2020). The focus of our study is to utilize customers' previous purchases to derive their preferences for products and provide them personalized recommendations based on their taste; Thus, relating this thesis to work done on both streams of literature. One of the most referred works in this literature stream is Amazon's use of collaborative filtering to personalize product offerings to their customers (Linden et al., 2003). In the recent research on this stream, Bernstein et al. (2015) propose a stylized model with the goal of revenue maximization to explore the revenue impact of dynamic assortment customization in the presence of heterogeneous customer segments and inventory constraints. They conduct their study on Beymen, a high-end fashion retailer in Turkey, and find the gain of implementing the model can be as much as 2.4% in optimal conditions. In another study, Bernstein et al. (2018), proposed a clustering-based recommendation system to maximize cumulative revenue, which adaptively segments customers as they make new transactions. Their proposed policy utilizes the MNL model to learn their preferences. They test their system on a Chilean retailer and find that it results in 27% more transactions than the second-best model they tested.

Moreover, our research utilizes the customer preference values learned for different granularity levels to suggest finding brands to be bundled. This makes this research related to another research topic that is concerned with product bundling. With a different approach, Ettl et al. (2020), construct two classes of approximation that recommends discounted product bundles to the consumers based on individual preference to maximize profit while managing the inventory constraints. They present their findings with two case studies in the retail sector and airline travel industries and an expected revenue improvement of 2%–7% on average over existing practices depending on the setting. The recommendation systems' potential for providing personalized advertisement and promotions are also studied in the literature. Zhang and Krishnamurthi (2004) considered micro-level promotion customization and developed an optimization procedure to derive the optimal discounted price for each visiting customer in an online environment. In a different research, Ghose et al. (2019) utilized the temporal duration, spatial dispersion, semantic information, and movement velocity of consumers to infer their preferences and found that applying trajectory-based mobile promotion strategies can provide more robust recommendations. They tested their proposed method in a field experiment conducted in a major shopping mall in Asia and found that using such a method can generate higher transaction amounts and increase revenues.

Furthermore, we utilize collaborative filtering-based methods to implement a recommendation system to predict customer preferences. Using these methods in e-commerce settings to provide product recommendations based on only previous sales transactions has been beneficial (Huang et al., 2007). However, the performance of recommendation systems is dependant on data, and the choice of the algorithm affects the accuracy of these systems. One of the papers that studied the effect of data on recommendation accuracy was Chen and Chen (2019). This study evaluated the accuracy of user-level recommendations with different measures and demonstrated the impact of rating-scheme characteristics on such systems' performance, namely rating value, structure, and neighborhood network embeddedness. In another research, Huang and Zeng (2011) discussed an approach for model/algorithm validation and selection in recommendation systems based on the characteristics of the setting. Moreover, the use of traditional collaborative filters is associated with a decrease in aggregate sales diversity. This finding was demonstrated in a randomized field experiment by Lee and Hosanagar (2019). They also found that using these systems is not always accompanied by a decrease in individual-level consumption diversity. This indicates that although retailers can benefit from increased sales under recommendations, using these systems for retailers whose target is to offer greater product variety will not be consistent with their strategy. In another study, Song et al. (2019) proposed a multicategory utility model that can dynamically diversify the suggested items while maintaining the recommendation accuracy.

Additionally, one of the goals in developing new algorithms for recommendation systems is to improve their performance and address their limitations. Different approaches for designing recommendation systems have been proposed over the years. For example, in early works of designing such systems, Ansari and Mela (2003) developed a scalable content-based stochastic variational Bayesian framework to generate recommendations. Using a joint space map based on past purchase behavior, Moon and Russell (2008) suggested a predictive recommendation model that derived the purchase probability of a product by calculating the customer's relative distance to other customers on the map. Bruyn et al. (2008) proposed a framework in which they used preference models from conjoint analysis to create a questionnaire-based recommendation system that predicted online user's preferences with minimum prior information. Another well-known approach in designing these engines is the use of association rule mining to provide suggestions. In a research conducted by Ghoshal and Sarkar (2014), they adopt associative rule mining technique to develop an information mining algorithm that obtains disjunctive consequent rules. Following this study, Ghoshal et al. (2015) proposed a maximum likelihood recommendation framework to improve the quality of offerings made by combining association rules to provide alternative choices as product combinations. Using a graph-based approach, Banerjee et al. (2016) design a collaborative filtering recommendation engine for content-rich settings where the number of items and the number of item-views by users are of a similar order. In a more recent study, Farias and Li (2019) proposed the use of a slice learning algorithm to learn customer preferences with side information. Following their methodology, they used the slicing algorithm to recover specific slices of a three-dimensional tensor from a noisy customer observations tensor.

When utilizing recommendation systems, researchers should also consider the behavioral effects of the recommendation systems on consumers. Although recommendation engines are designed to provide suggestions by learning preferences for products, the predictions generated by these systems also impact consumers. In one study, it was found that the recommendations presented by a recommender system significantly influence costumers' preference and introduce anchoring biases (Adomavicius et al., 2013). These suggestions can also affect consumers' willingness-to-pay (Adomavicius et al., 2018). Furthermore, As consumers search through the products or purchase them, they update and change their valuation for product attributes and features. Dzyabura and Hauser (2019) empirically demonstrate that consumers learn preference weights during search and propose a recommendation system that encourages preference-weight-learning. They suggest that these systems can be beneficial in a setting where product or service categories are multi-attributed, infrequently purchased, costly to experience without buying, and sufficiently valuable to justify an extensive and expensive search.

As discussed throughout this chapter, this thesis contributes to the research focused on personalization, boundling, and the implementation of collaborative filtering recommendation systems. Next, we will describe the empirical setting and the data used for this thesis. In the following chapter, first, we will provide some background on the retailer. Then, we will focus on the data they have provided us and discuss the changes made. Finally, we will analyze the data and provide some descriptive insight.

3. Empirical Setting & Data

In this chapter, we will discuss the retailer understudy. In the first section, we will provide a description of their background, and explore their operations, objectives, selling strategy, and their focus for future improvement. Next, we will talk about the data they have provided for the purposes of this thesis and delineate its structure and the modifications made to draw out the information. Finally, we dedicate the last section to the analysis of the raw data and present descriptive information gathered from it.

3.1 Baskasindarama

Baskasindarama, which will be denoted as the retailer from now on, is an online retail company founded in 2016 (See Appendix A.1 for an illustration of retailer's environment). This retailer is based in Izmir and provides its services from there to people anywhere in Turkey. It works with over 100 brands, and the number of the categories it offers can change between $6-9$ based on the season and the situation. The co-founders of this retailer, Esra and Melis Sarıhan, started their business by selling a few local and international brands on their website without significant investment and initial capital. Initially, they entered the E-commerce world, intending to provide services that were a bit different from those that already existed at the time period, and ever since, they have been striving to satisfy their target consumers by offering and delivering high-quality products. On top of that, they prioritize creating a friendly relationship and warm environment with their customers to answer their questions and help them with the emerged problems. They also promote female entrepreneurship skills by providing women farmers and producers with opportunities to sell their products on their retail platform and supporting them by including their brands in their roster, especially people who use domestic and native goods as raw input in their practices or use organic material to produce cosmetics products. The brands that fall under the retailer's market category are the best representative of its support for women entrepreneurs who have developed their brands and have the highest importance.

Beacause of their efforts, aims, and objectives, they were recognized as one of the top 20 successful women entrepreneurs by Startup Turkey event among participants and business owners from 63 different countries for their services. Gaining this achievement boosted the retailer's brand as well as their objective to promote women's power to create and manage businesses. In 2019, they were nominated for the Stevie Awards, one of the prestigious international awards given to businesses, for the best small-scale retailer company of the year and became the winner of the Bronze Stevie Award. Moreover, in 2020, they again won another Stevie award, making them the only retailer receiving this award in 2 consecutive years in Izmir.

What makes this retailer different from other retailers and its competitors is the choice of brands with premium quality products to be marketed and traded on its platform. Selecting these goods carefully that are either of high quality, or use pioneer technology, or have trending features, or have a unique design, provides this retailer an edge on delivering its customers merchandise special in its field or category. The care for quality satisfaction is also reflected in packaging as well, where they present customers with their unique packaging on top of the brand's original box. The other attribute differentiating them is that they are not a marketplace; this makes it so to strive to sell products with more honest prices that best represent the quality of the brands it is collaborating with compared to its competitors in the retail industry.

3.1.1 Purchase Process

In order for the retailer to procure the customer orders, it follows the subsequent processes; after the company makes a contract with the brands, the products are uploaded to the website, and the search engine is optimized to show it to prospective customers. The customer comes to the retailer's website and places an order on the product by adding it to their basket and purchases it either by paying it through secure payment gateways or bank transfers. Depending on the product's shelf life and its inventory status, the goods will be procured either from the stock or the producing company itself and it will be sent to the retailer. Since this retailer does not keep an inventory of its own and practices dropshipping, it acts as a hub and procures customer orders from specified warehouses to prepare customer orders. After collection and quality control, the product will be put into the retailer's packaging with its original package and will be sent to the customer with the invoice. The retailer uses UPS services for the delivery of their products. In case of any damages to the merchandise during the delivery process, they will resend the product without any additional charges. Building trust and safely while providing high quality services and products to customers is one of the essential goals for this retailer. For instance, due to the pandemic, they avoid doorstep payment methods, which can be unsafe and diminish the brand quality. Instead, they offer online and long-term payment methods to accommodate the situation.

3.1.2 Brands

When the retailer started its business, it promoted and featured lesser-known imported products, an example of which could be "Roli Doner", a designer that was partnered with Beymen and Vakko up until 2015. In 2016, they signed a contract with this firm and started working with them; this made the company to be perceived as a retailer of luxurious products. Another example of these products that fit the objective and aim of Baskasindarama was Bugatti, which produced kitchen and household appliances with unique designs. However, as described before, their aim was not just to sell the most expensive and luxurious items, but rather products with high quality for all people that best fit their vision and goal. Because of that, they expanded their product range to include more affordable, high-quality goods, such as Neutrogena and Simple, which were better known and had a much broader consumer base. This conversion aided with capturing more of the market. Also, it helped circumvent some problems caused by Turkey's 2018 financial crisis. These problems, such as consumer distrust in brands and diminishing return value, were due to the discontinuity of some companies in sustaining their imports and replenishing their stocks. Adding the well-known and trusted brands in its practice not only secured their customers' trust it also provided the opportunity to include other international brands in their roster safely and introduce them to the consumer base. It added up to a point where now the retailer works with over 50 brands and hosts over 500 products in its platform in 6 – 9 main categories. To work with a brand, they check its target market to be aligned with them; In addition to that, they also investigate the brand's objectives to consider including them with other retailed brands in their platform. In case of stopping sales of a specific product or discontinuing a particular brand, they would freeze its sales by showing the products as being out of stock. This way, if the number of click-throughs on that product page increases or reaches a threshold, the retailer would bring them back and start selling again.

3.1.3 Price & Promotions

The retailer is in constant communication with the brands they work with, and the companies provide and determine the retailer's prices. This retailer knows the finished product cost and the markups. The companies tell them their desired selling price and the limits, and when the company wants to create promotions for their products; Thus, the retailer does not have a pricing policy of its own and follows the policies determined by the brands. Due to the set margins by these companies, the number of discounted products offered by the retailer is minute and limited. Most of the promotions offered on the website are from well-known brands that can afford a reduction in their prices. Also, since the number of visits to pages with discounted prices is low, this company tries to focus away from bulk promotions and differentiate itself from competitors that offer discounts with high frequency.

Moreover, since the retailer follows the brands' pricing scheme and only includes the prices' tax margins, it provides the advantage to work with a plethora of brands and companies. When the retailer introduces discounts for the products, it includes them from its own profit margins. Moreover, the retailer obtains the updated prices, campaigns, and promotions biweekly, issued regularly by the brands. In case of any interruptions, they ask directly to these companies for the newest changes in prices. For highly competitive products with rapidly fluctuating prices, they tend to check the prices on other platforms and compare and adjust them by considering their set margins. Perishable products with a very limited shelf life, such as food, dairy, or baked goods, usually have a small profit margin, and their prices rarely have a significant change; thus, to keep the competitive advantage, the prices of these products are kept close to the price in the market and their competitors

3.1.4 Selling Strategy

The retailer's marketing strategy revolves around offering affordable and high-demand products to the customers at the first step. After introducing these products and catching their attention, it offers lesser-known, more expensive products. The exposure to these new products provides customers the opportunity to compare them and learn about their advantages and choose from a more extensive product variety. Also, as a way of promoting new and seasonal products, the retailer shows on top of websites search results, and in return, it tries to increase the customers' focus and attention.

Moreover, the retailer has provided some features in their platform, where customers can ask questions regarding the availability dates of out-of-stock directly. Asking questions from the retailer not only helps them with the objective of creating a friendly relationship and warm environment with their customers but also aids them in learning the interests of their customers and their desires. This way, they can better target the customer and promote the related products by providing them with samples while realizing their orders. The added sample size is determined by considering the order quantity and purchase quota. This helps increase awareness of the new merchandise and acts as a token of appreciation for people's purchases, resulting in sales growth.

The retailer also utilizes a scoring system as a way to track and reward customers for their purchases. These scores are given to the customers based on specific activities they perform, such as referring the retailer's platform to their friends, shopping up to a certain quota or amount, or purchasing on particular occasions. For instance, they gain 5000 points when they become a member. When these scores reach a predetermined threshold, the customers will be provided discount codes for their subsequent purchases.

For the products whose expiry date is near, the retailer tries to sell them by promoting them to customers who were interested in similar products before while considering their basket size and their overall spending's. For instance, if the customer is purchasing an oven or its total purchase quota is high, they include a product whose expiry date is near as a complimentary gift. This will motivate the customer to order more and also become aware of other product types in its product range. This tactic is specially used for products such as flour as caramelized onion, which can be gifted in small sample sizes, and helps a great deal in liquidating products that are nearing their expiry date. This retailer's strategy for short-lived consumables is not storing them; however, after the Covid pandemic, they started allocating much more space on their stocks of perishable goods.

One of the markets that the retailer tries to enter is the eastern regions of Turkey. Penetrating this market is challenging since the number of women per capita in these areas who use or own any financial cards is low, and the utilization of e-commerce businesses is scarce. Also, people in these regions are generally using social media to get their needs; Thus, penetrating this market becomes very hard. However, capturing this market can be beneficial for them since the number of competitors in these regions is small, and also, people in these regions mainly consist of women farmers, which fit into the category they try to promote.

Currently, the retailer's suggestion strategy follows two schemes: first, fixed suggestions, and second, the season-related suggestions. In the first scheme, products are bundled and mapped to one another based on category and relevancy beforehand. Then they are uploaded to the platform and provided on the page of the products. This way, when customers visit the product page, they are presented with some suggestions. However, the introduced products are not dynamic nor based on customers' behavior, interests, or activity. Sometimes, these suggestions can be coupled with the best-selling products on the website for special occasions. The other scheme involves introducing customers to season-specific products to raise awareness. For instance, during the new year, they would include merchandises that are likely to be sold during this period and relevant to the winter in their recommendation section, such as socks, calendars, and cosmetics. Like the other one, this scheme is fixed and requires mapping the categories and products manually before uploading them on the platform. Having a more sophisticated recommendation system can help to deduce customer interests and preferences. It can also help with offering present bundles to the correct target audience, such as sending sport sets for sports enthusiasts or stationery gift sets for people who love writing. Thus, we focus on implementing such system in their practice.

In the next section, we will discuss the data that the retailer provided for the purposes of this thesis. We will examine the general structure of the supplied reports, the information they contain, and their relevancy.

3.2 Data

The retailer mainly accumulates the data generated by Google Analytics as a means to control and analyze their resources. This data is also used to prepare the reports provided for the purposes of this thesis. These reports provide information on the company's operations and performance, and the reports encapsulate information at different levels. The first report is daily sales data which contains the purchase information for each customer. It depicts what was purchased, when it has was returned, and the amount of turnover from transactions each period. The second report is the monthly traffic acquisition summary, which shows the sales channels that the retailer used to promote its products and gain its website's traffic each month. The third report illustrates the main categories of goods and brands sold by the retailer. Next, the page visits report outlines the number of daily views of each webpage in their websites. Lastly, the demographic data depicts the demographic of their customers for the first 6 months of 2020. A snapshot of these reports has been presented in Appendix A.2-A.6.

3.2.1 Daily Sales Data

The sales report provided by the retailer illustrates all transactions and customer purchases throughout 2020. The sales data is scattered in 11 documents, each covering monthly purchase information from January to December. Due to the disruptions caused by the Covid Pandemic, the retailer had no transaction during March; thus, the file for this month is not included in the report. The sheets in each report summarize the daily customer purchases. The days that the retailer had no transactions were excluded from the document.

Each sheet contains 8 columns. The first column shows the ID of the customer who purchased a product. The IDs are coded by appending 'C' at the beginning of the customer number. The second column is the page URLs of the purchased products. Next, we have the situation column that represents the status of the placed order. Each transaction takes three different statuses: completed, pending for payment completion, and canceled. These values are denoted by numbers 1 to 3, respectively. The fourth column indicates how many units have been ordered for each product. It is worth noting that transactions with multiple items are split into consecutive rows, and each row depicts the order amount of one product and not the overall order amount in the transaction. The product price column indicates the unit price of each ordered merchandise, and the turnover column is the revenue generated from that product in the transaction. The values in turnover can also be derived by multiplying the unit price by the number of orders. The profit margin column displays the profit percentage for each item in the transaction, and it can take values starting from 10 up to 50+%. The values in this column are separated by increments of 10 and are presented with a lower and upper bound. They are also coded with values 1 to 4. For instance, a transaction item with a profit margin of 1 shows that the retailer earned a profit of 10 up to 20% of its original price by selling this product. On special occasions, such as charity campaigns, their yield can decrease as much as 0 percent. The last column illustrates which items were promoted with a discount campaign at the purchase time. Additionally, the sheets also include information about the visitors to their platform; specifically, it counts the daily number of new visitors along with returning customers.

Working with the data in these documents in their base form is challenging and can not be used as an input for predictive models. As a result, we combined all the documents and the sheets into one. To separate the transactions, we also included a column representing the date of the purchase and another column representing the transaction ID. The transaction IDs are created by concatenating the date with customer IDs. Furthermore, we used the links provided in the second column to access product pages and gathered the product information directly from their platform. A Python script was developed to automatically access the retailer's platform and scarp data from each URL to gather this information. We collected the product name, its main category, subcategory, and brand from each product page and included each in a separate column.

3.2.2 Monthly Traffic Acquisition

The monthly traffic acquisition report is a single document that illustrates the different channels' contribution in guiding the visiting individuals to the platform. Six sheets are included in this document, and each shows the monthly incoming traffic to the website from January to June for each channel. In each sheet, two columns are differentiating the new visitors from returning customers. Also, the rows show the different channels that customers can gain access to the platform. In the first four months of 2020, these channels are only composed of direct access, search engines such as Google, referrals such as blogs, social media such as Instagram, Facebook, or Twitter. For May and June, the retailer included two other channels; paid advertisement, and display such as YouTube. The values in the document are expressed in percentages. Also, the total number of visitors is also included for each month.

3.2.3 Brand Category Lists

The retailer provides information about the brands they are working with and their assigned categories in the brand list report. The report depicts assignments in a hierarchical manner, where main categories are at the top, and brands are listed under each, similar to a flow chart. The report indicates that the retailer sells its products in 7 different categories: apparel, jewelry, baby care, cosmetics, lifestyle, stationery, and groceries; However, recently, there has been a change. Specifically, the lifestyle category has been split into life and kitchen product classes. The stationary category has been merged with the life category, and jewelry has been combined with apparel. Also, they launched a hygiene category to respond to their customer needs emerging from the Covid pandemic. Moreover, the report shows 63 different brands that are working with the retailer. Additionally, the retailer has also built partnerships with new brands throughout the end of 2020. The brand information in this report combined with the data retrieved from the retailer's platform is used to populate the transaction data, which is the primary input of the developed model.

3.2.4 Page Visits

The page visit report shows the daily product pages visit on the retailer's platform, and it is provided for the first five months of 2020 in 5 documents. The daily information is shown in separate sheets in each Microsoft Excel document and contains the URL link of the product webpage in the first column and the number of individual visits in the second one. The sheets also indicate the total number of visits as well as the frequency with which visitors leave immediately and the average duration for which they spend on the pages. The information in this report can be used as the means to see the popular items along with the products that people find interesting.

3.2.5 Demographic Data

The demographic report is presented in a single document in which it provides an overview of the retailer's customer demographics for the first 6 months of 2020. Each sheet is dedicated for every month and includes monthly user information in five categories: their approximate age, gender, country of origin, devices they use to access the website, and their interests. The user ages are split into 6 ranges starting from 18 with increments of 10, and the last range is for people 65 years of age or more. The values in this category are expressed in percentages. In the third category, country of origin, only the top 9 countries with the highest visit percentage are shown. The device category depicts the percentage of different equipment used for accessing their platform, specifically whether it was a mobile, desktop computer, or tablet. Finally, the last category illustrates the visitors' interests by dividing them into 30 different groups based on their searches in other websites.

3.3 Descriptive Analysis

This section will analyze and provide descriptive insights on the information gathered from the reports and the engineered data (See Appendix A.7). As described in subsection §3.2.1, the primary input engineered for the recommendation system comes from changing the data in the sales report. The encoded values in this data are replaced with their original values, such as order status and profit margin. Also, new information is scrapped from the retailer's platform and appended to each transaction record, specifically, product name, category, subcategory, and brand name.

The sales data consists of 2125 records and depicts the retailer's sales over the 12 months of 2020. It contains 1281 unique transactions made by 841 distinct customers during this period. There are 599 products sold in these transactions in 98 different subcategories from 59 brands, and their prices differ from 2.25 up to 9300 TL. The sales data only includes the completed or pending transactions; otherwise, they were excluded. This dataset contains transactions with up to six unique products. Each transaction consists of between 1 and 100 products, with an average of 3.2. The highest purchase frequency among their products belongs to stone ovens from Tasoven brand and barbeque subcategory; However, the highest purchase rate among their main categories belongs to items in grocery. This is due to having much more variant product roaster and brands selling merchandise in this category. Figure 3.1 illustrates the aggregate count of categories that were purchased by customers in each transaction. If customers bought multiple of the same product in one transaction, they were counted as one instance.

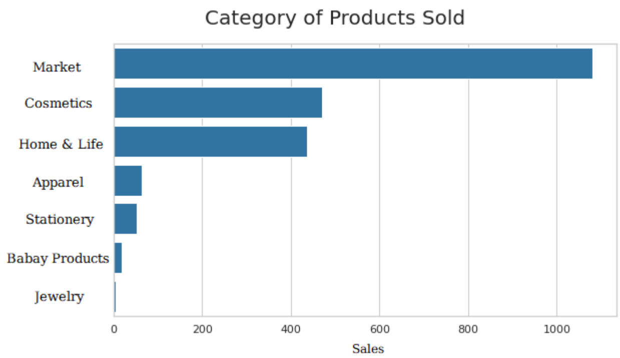


Figure 3.1 Aggregate Count of Purchased Items In Each Categories

Furthermore, the retailer hosted 12 different promotional campaigns in 2020. They promote products on several occasions, for instance, when they include new merchandise in their roster or for charity events. The promotions were spread out across the first two months and the second half of the year and, on average, lasted for two to three days, except for their Christmas campaign that lasted a month.

Looking at the data, we can see that the retailer's revenue can vary from 10.5 to 12589 TL for each transaction. Based on the data, the retailer had a comparably low revenue during the first 6 months of 2020 due to the circumstances created by the Covid pandemic. During the second half of the year, its revenues had a noticeable increase, starting from summer. A large portion of this increased yield is selling products from barbeque subcategories, specifically stone ovens with a large unit price. This item had the highest purchase frequency throughout this period. Their revenues had a minor decrease towards the end of the year; however, they increased at the last month following their Christmas promotional campaign.

Additionally, based on the information provided in the sales data, we can see that their average monthly new users are 37674, while their average monthly returning customers are 12439. Likewise, the daily average number of new customers and returning customers are 1504 and 498, respectively. Figure 3.2 shows the number of monthly visitors on the platform. Values indicated in orange shows the returning customers visiting the website, while the blue bars illustrates the new customers. As seen in the plot, in the last two months, the retailer attracted a large number of visitors over the last two months of 2020. During this period, the retailer profited from a high number of purchased products with lower profit margins. Also, their campaigns, especially ones in December, boosted these sales and helped increase the visitor to the website. Compared to the last two months, the visitors to the website were much lower during June. However, the products sold during this month were much more expensive with higher profit margins that made the retailer generate the highest revenue.

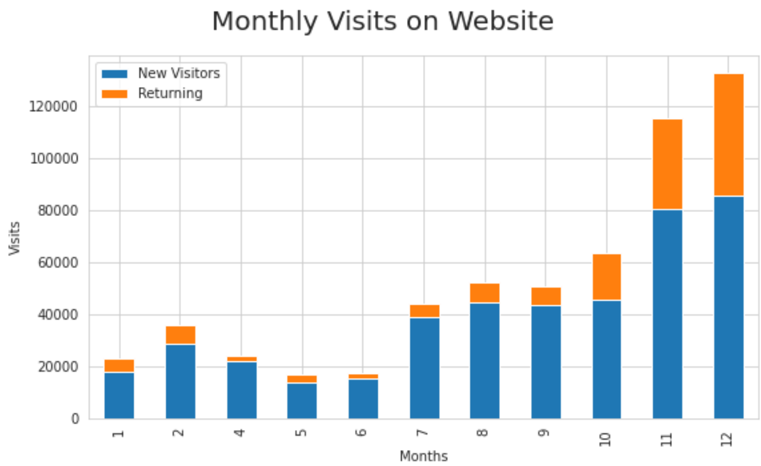


Figure 3.2 Number of monthly visitors to the platform

Based on the page visits report, the rate of immediate exit from the platform after visiting one page is 71.66%, and the average time spent on the page is 2 minutes and 7 seconds. Additionally, according to the traffic acquisition report, in the first three months of 2020, the retailer mostly gained its visitor through search engines. However, In the second quarter, they mainly attracted new visitors through social media platforms such as Instagram and Facebook. Also, the second channel that contributed to this was YouTube and the posted videos. Furthermore, the age range of people visiting the website is mostly between 25 and 34, with an average percentage of 38.06%. Next, people aged 35 to 44 comprise the following group that mostly visits their platform, comprising 24.46% of the total visitors. The website users are mainly female, and users primarily use their mobile phones to access the website, precisely 75.61% of users on average.

In the next chapter, we will discuss the developed model and provide specifications of how it provide recommendations.

4. Methodology

As discussed in the previous chapter, the retailer utilizes two types of suggestions for their platform's visitors: fixed and season suggestions. Employment of these methods provides value by introducing new and interrelated products to prospective consumers by guiding them through the platform. However, since they are static and they do not change from customer to customer, they do not capture the heterogeneity among visitors and their preferences. One way to increase the added value of the suggestions is to incorporate recommendation systems, denoted as RS from now on, which provide personalized recommendations by considering past consumer purchases and their individual behavior. There are two approaches for making RS; collaborative filtering and content-based filtering. Due to the nature of the provided data-set by the retailer, employing the former approach for implementing the model is thought to be more feasible since the latter requires much more information on the products themselves. Utilizing collaborative filtering RS can be helpful in these situations since it can provide recommendations to customers requiring little to no features about the items or consumers to be known. Moreover, we focus on application domains of recommendation systems for experience goods where consumer preferences are taste-driven as opposed to being vertical where customers are more aware of the product quality (Adomavicius et al., 2018). By focusing on this application domain providing personalized recommendations to individual customers becomes more relevant. Furthermore, as mentioned before, the retailer offers fixed suggestions on each webpage which can sometimes result in overspecialization of offers, for instance, suggesting a pair of sneakers to someone who just bought another similar pair of sneakers. Using collaborative filtering can reduce this by providing completely different items from what consumers have already seen based on other consumers' purchases. These characteristics were some of the reasons that derived us to use the collaborative filtering approach in our model.

Throughout this chapter, we will discuss the specifics of the first approach and the model of the RS used in this thesis. We will also explore the derived results from implementing such a model and applying it to the retailer's sales data. In the first section, we will start by providing some background on collaborative filtering. Next, we will move on to describing the model and the characteristics of the RS used in this thesis.

4.1 Collaborative Filtering

Collaborative filtering is the most common way to provide product recommendations. It is a method of making automatic data-driven predictions on a given customer's preferences on the basis of other customers' purchase history. The suggestions are specific to the user but use information gleaned from many users. This differs from the more straightforward approach of recommending the frequently purchased products based on the overall purchase volume of items. One of the most trivial applications of such systems is in an online retailing environment where the retailer can suggest a tailored set of products to individuals based on their behavior.

The underlying assumption of the collaborative filtering approach is that if visitor U_A share the same purchase history as visitor U_B on a product, U_A is more likely to share U_B's opinion on other products than that of a randomly chosen person. The preference of each customer for a product is then teased out to construct a utility matrix from the data where rows represent the customers, and the columns represent the items itself (Leskovec et al., 2014). Values for each customer-item pair in the matrix represent the known information about the degree of preference of that user for items. These values can either be explicitly derived from customer ratings of the product or implicitly calculated based on their past purchases. The matrix is assumed to be sparse, meaning that most values are "unknown." An unknown value implies that there is no explicit information about the customer's preference or purchase history for the item. The goal of the RS now would be to determine and predict the unknown values (blanks) in the utility matrix. There are different methods for finding these values; The first method is a classification-based method and uses the principles of k-nearest neighbor. This method aims to calculate the similarity of each customer U's preference vector with others in the utility matrix based on a distance measure and select the top nearest-neighbors K with a size of n which have the highest similarity values. After finding the closest neighbors, the predicted preference value of each customer U for an item z, which has a blank value, will be the average of preference values of all n closest neighboring customers for that item. Given that each item z is among all items Z found in the utility matrix and customer i is not a part of its nearest-neighbor ($i \notin K$), Formula 4.1 depicts how the unknown preference value m of customer i for specific item z is calculated by averaging the known preference values of neighbors j for item z in utility matrix M:

$$m_{iz} = \frac{\sum\limits_{j \in K} M_{jz}}{n}, \quad z \in Z, \quad j \neq i. \tag{4.1}$$

Recommendation for a customer U is then made by looking at the most similar customers to U in this sense. Given U_{iz} as the customer i which has an undisclosed valuation for item z, customer U_j is deemed to be similar if the similarity measure of their vectors yields a high value. The criteria used in practice are Pearson correlation;

$$sim(U_i, U_j) = \frac{\sum\limits_{z=0}^{|Z|} \left[(U_{iz} - \bar{U_i})(U_{jz} - \bar{U_j}) \right]}{\sqrt{\left[\sum\limits_{z=0}^{|Z|} (U_{iz} - \bar{U_i})^2 \right] \left[\sum\limits_{z=0}^{|Z|} (U_{jz} - \bar{U_j})^2 \right]}}. \tag{4.2}$$

Jaccard distance;

$$sim(U_i, U_j) = \frac{|U_i \cup U_j|}{|U_i \cap U_j|}. \tag{4.3}$$

or Cosine similarity;

$$sim(U_i, U_j) = \frac{A.B}{\|A\|\|B\|} = \frac{\sum\limits_{z=0}^{|Z|} (U_{iz} U_{jz})}{\left[\sqrt{\sum\limits_{z=0}^{|Z|} U_{iz}^2} \right] \left[\sqrt{\sum\limits_{z=0}^{|Z|} U_{jz}^2} \right]}. \tag{4.4}$$

The other approach is using UV-decomposition, which is based on matrix factorization of the utility matrix and decomposing it into low-dimensional matrices. This method focuses on reducing the dimensionality of the matrix and finding the two matrices that their product closely resembles the known values in the utility matrix.

The assumption in this method is that there exist two U and V matrices with latent features that their product closely approximates the original sparse matrix. The product of these two matrices can then derive the targeted unknown values in the original matrix.

$$M = U^T V. \tag{4.5}$$

The goal here is to reconstruct $r^{'}$ from the product of customer vector u and item vector v such that $r^{'}$ captures most of the variance within the original utility matrix M; thus, Formula 4.6 will be an approximation for Formula 4.5.

$$r_{ij}^{'} = u_i^T v_j. \tag{4.6}$$

A typical choice of measuring the closeness of $r_{ij}^{'}$ to M, is the root-mean-square error, where this measure is applied to all non-blank entries in M and the corresponding entry in the product of $u_i^T v_j$. Finding the UV-decomposition with the least RMSE involves starting with some arbitrarily chosen u and v, and repeatedly adjusting them to make the RMSE smaller. Here, different starting values should be tested to decrease the chances of finding global minima. Adjusting values in each matrix only affects certain entries in the product matrix M; this makes it easy to apply changes to the matrices concurrently and minimize the RMSE. Given that u_i and v_j are matrices of dimensions n-by-d and d-by-m, Formula 4.7 depicts the optimal value for the entry u_{iz}, denoted as x, that minimizes the overall RMSE if m_{ij} is a is a non-blank entry of the matrix M (full derivation of this formula is provided in Leskovec et al. (2014));

$$x = \frac{\sum_j \left[v_{zj}(m_{ij} - \sum_{\substack{k=0 \\ k \neq z}}^{d} u_{ik} v_{kj}) \right]}{\sum_j v_{zj}^2}, \quad \sum_j = \text{sum over all } j | m_{ij} \text{ is nonblank.} \tag{4.7}$$

The repetitive adjustment of values to optimize the RMSE starts after prepossessing and normalizing the utility matrix and continues row-by-row in a round-robin fashion until the RMSE reaches a certain threshold. Ideally, the intent is to reach RMSE of 0; however, since there are normally many more non-blank elements in M than combined elements found in matrices U and V, reducing the RMSE to 0 would be an indication of overfitting. The technique utilized for finding a UV-decomposition is gradient descent, where the changes that most decrease the error function is chosen, and if further improvement falls below a threshold, the optimization is stopped.

The mentioned methods in this section each have their advantages and disadvantages. The K-NN-based method, which used the similarity measures to recommend the products to customers, provides a much interpretable approach in finding preference values and produces results strictly based on the similarity of the individuals. However, it performs poorly in highly sparse matrices, and its performance drops significantly in big datasets, making it not scalable (Al-Bashiri et al., 2017). Also, choice of similarity measures along with the input source of model can produce significantly different predictions for the unknown values in the utility matrix. On the other hand, the Matrix Factorization-based methods perform quite well in big datasets, and it is scalable. They can generate results even for highly sparse datasets. The drawback of these methods is that they are complex and prone to overfitting if they are not optimized. In the case of sparse datasets, since there are many matrix combinations that their product fits the known values of utility matrix, they are susceptible to generating matrices with the RMSE of local minima.

Several other methods are used for creating RS, such as clustering and Principal Component Analysis based algorithms. These methods apply different model-based approaches to find the utility matrix and provide the preference values of target customers. However, all of them aim to calculate the preference scores and, based on these scores, sort the corresponding items for each customer and provide the k high-ranking components as the recommended list to them.

4.2 Model

The two described methods mentioned in the previous section are employed in the RS used for this thesis. The principle used here is to generate recommendations by employing the models and evaluating it with the ground truth. First, the similarity-based method will be utilized to create the unknown values in the utility matrix. Then, the matrix factorization approach will be used to predict the values that all of their nearest neighbors are unknown. This approach benefits from the interpretability of utilizing the preference of similar users while taking advantage of the performance of the matrix factorization method on predicting values for items with no preference entry from any customer.

Before utilizing this model and providing the recommendations, the input of the model needs to be set. As discussed in the section 3.2, the data used as the input of the model is the preprocessed data from the daily sales report, which was augmented with additional information from the retailer's website with the provided links in the report. This data represents customers' transactions and their purchase history. The additional information contained the brand, category, subcategory, and the name of the items. After cleaning up the data and manually imputing the missing information, the user-item table, which is the utility matrix, is created, which is the primary input of the model. The rows and columns in this matrix represent the customers and products, respectively. We tested two different ways to fill out the matrix values: utilizing the purchase frequency of the products by the customers and using their purchase amount. For example, in the first approach, if the customer A bought item 1 in two transactions throughout the time span of the data, the corresponding value in the utility matrix would be 2. Further, we tested the normalized version of these matrices to analyze their performance on three different granularity levels: Category, brand, and item level. After constructing the input, we use it for building the model.

Additionally, we have to choose similarity criteria for comparing the customers for the k-nearest neighbor. We chose Pearson correlation and Cosine distance for comparing customers from the three measures introduced in the previous section. The Jaccard similarity is disregarded since it requires the input to be filled only with Boolean values indicating whether a purchase happened or not. Choosing this metric would result in loss of information and is not suitable for substantially sparse data. The two metrics were considered as hyper-parameters of the model to be tuned. The model is then set to find the k most similar customers and average out their values in the utility matrix to find their scores and use them to sort the top n products and recommend them. The choice of nearest neighbors is set to the square root of number of customers in the input data. The model's output is a table with four columns; each indicates users, item codes, and expected items. This output will then be transformed into a user-item table where the values depict the preference percentage for each item.

We construct different models with the introduced inputs and evaluate their performance with the baseline to choose between the input parameters. The parameters of the best-performing model with the lowest RMSE are then selected as the final model. We designate the transactions in the final month as the test set to test the model's prediction accuracy and performance. The final model is then built on the training dataset and executed on the test data to obtain the predicted values and recommendations. The predicted products are then compared with the customer choices from the final month's purchases and analyzed. The results of this analysis and hyper-parameter tuning are presented in Chapter 5. Although the model can recommend all products with respect to their ranking, there is a threshold on the number of items to be shown on online platforms and a limit for each customer to retain information. We also analyze the optimal number of items to be shown in online settings.

The models constructed for predicting customer scores and recommending products that are expected to be bought by the user are built with the help of Python libraries. Specifically, to create and evaluate the recommendation systems used in this research, we utilized the methods and functions offered by Turicreate's Python library. The modules implemented in this library provide a wide variety of functionality for making, testing, and evaluating systems such as RS. The following chapter is dedicated to analyzing the results gathered from applying this model to the test data and deriving managerial insights.

5. Analysis & Results

In this chapter, we first characterize the optimal hyper-parameters used along with the designed RS and determine the performance of this system. Per the description provided in Section 4.2, we first construct the models with different introduced parameters and compare their prediction power on the known values against the baseline. The lower the value of the RMSE, the closer they are to the baseline. Since this comparison would not be enough, we also analyze their prediction power on the last month's data as well. To achieve that, we also conduct a simple experiment on this data. This experiment will also help us show the effect of the number of recommendations for an online setting. Further, after finding these parameters, we use them as inputs to create the RS engine and obtain predictions on all users. Finally, we analyze the output to derive managerial insights.

5.1 Performance & Hyper-parameters

Before finding the optimal parameters, we need to split the dataset into two parts. This segmentation defines a test set and obtains a final, unbiased performance measure of the entire model building process. The first segment will be used in training the model and its evaluation. The other portion is used to provide the results of this research. We used the data from the last month as our test set. From the remaining segment of the data, we used 10% of it to evaluate the model, and the rest was used to build the model. All models were constructed according to the methodology described in the previous chapter. Further, the hyper-parameter selection was made by evaluating the models with the function provided by Turicreate's Python library. The models were analyzed in three different levels:

- Main Category level
- Brand level
- Product level

These models, as described in the literature, tend to be in a high order of complexity. The use of Turicreate's python library for implementing these models helped us reduce the computation time of the models. The functions implemented in this library are optimized to decrease the time complexities of RS. Moreover, to choose the similarity criteria discussed in the previous section, we created the model both with Pearson correlation and cosine similarity criteria. The results presented in Table 5.1 compares the RMSE performance of RS engines that were made for the category, product, and brand level while considering different similarity criteria. The RS models used here were set to predict the known values in the normalized utility matrix and compare them with the baseline. Note that these values are the RMSE of predicted values against the known values and are derived by comparing with the baseline. Also, we only considered normalizing the user-item matrix while building the engines to analyze and compare the performance difference of them on the evaluation set. The choice of normalization is to improve the stability and prediction accuracy of the RS and decrease the variance in the output(Leskovec et al., 2014).

	Transaction Counts			Purchased Amount		
Similarity Criteria	Brand	Category	Item	Brand	Category	Item
Pearson	0.27390	0.15397	0.29066	0.29625	0.17983	0.43101
Cosine	0.34430	0.18071	0.29059	0.29627	0.18558	0.47417

Table 5.1 RMSEs of models build with different parameter choices.

To analyze the results in Table 5.1, we should compare the models made for the same level. The first column of this table depicts the models created by counting the number of transactions that the customer had purchased an item. The second column is for data that count the amount that the customer had bought from that item. The data were also normalized before creating the models. All of the models from the first column obtained a lower RMSE than the models in the second column, except for the model made for brand-level data with cosine similarity criteria. Some of the models in the second column yielded an RMSE of 1.5 higher than their counterparts. This shows that having the input as the count of the transactions and then normalizing it would deliver better results. Moreover, by comparing the results for similarity measure, we can see that in all of the instances, models with Pearson Correlation as their similarity criteria yielded a lower RMSE in all of the cases. This result is aligned with the literature (Albadvi & Shahbazi, 2009) and findings in one study (Cho & Kim, 2004) which showed that Pearson correlation as a similarity measure performs better when suggesting items.

However, the experiment we conducted showed a higher accuracy provided by choosing the cosine similarity as our metric at the brand level. This was also the case when creating the models for transactions made in shorter periods where the results were reversed and showed that the RMSE of models with Cosine similarity as their distance measure was lower and performed better at this level. This observation depicts that the components that affect the choice of similarity metric should be the type of the data along with the amount of sparsity and the length of data understudy; also, it is worth mentioning that this finding does not mean that the choice of similarity metric for all sparse brand-level data should always be cosine distance, but rather the choice is affected by a plethora of elements such as duration and nature of data. We should consider these elements while making an RS. The other important factor here is the normalization of the user-item matrix, which drastically affects the performance of the recommendation system. Because of close competition between engines with different similarity measures, we considered both measures for the final model and tried them both on the test dataset to see which performs better in a dataset that has not been seen before.

With the choice of normalized utility matrix with the number of transactions for each item, we conducted a simple experiment to determine the model's performance and capture the system's accuracy. In this experiment, to observe the performance of the recommendation system in a real-world scenario, we separated the last month's transactions based on their date. Then we set the engines to recommend items based on the customers on brand and item level (the category level was not selected since the number of categories is minute, and suggesting one item from each category would yield an accuracy of 100 in this experiment). If the visiting customer were not in the historical transaction database, the system would recommend the popular items to that customer. However, if a returning customer made the order, then based on the algorithm, the engine would compare the customer to the others, provide top N recommendations, and append the transaction information to the data used to create the user-item matrix. If even one of the recommendations were in the respective customer's transaction, the system would have created value. We track the number of times this system created value and divide it by the total number of transactions made by returning customers to compute its effectiveness. In this setting, each customer visits the store sequentially, and their transactions are captured and added to the data set to create the recommendation system based on the transaction observed until then.

The final month's data contained 132 unique transactions, 86 of which were returning customers. We iteratively applied this procedure for all of the returning customers, and the rest of the 46 customers were recommended popular items. To further analyze how the number of recommendations provided to customers can create value, we also set the system to create 0-30 recommendations for each individual. It is imperative that as the number of suggestions increases, the possibility of them being in the transaction also increases. However, we have to keep in mind that there are about 500 items provided by this retailer. Second, the sparsity of the customer transaction data confines us from accurately predicting their next purchase. Also, the limitations on the number of items processed by the customers restrict us from showing too many recommendations. Providing too many recommendations can overwhelm customers. Therefore, the number of shown recommendations should be optimized based on the feedback of the customers and the capacity of the website for the number of items it can display. Figures 5.1 and 5.2 summarizes the results we gathered from the experiment. The y-axis in these figures is the percentage of correctly suggesting an item that ended up in the customer's basket with respect to the number of provided suggestions. This percentage can also be seen as the probability that an item from the suggested options will end up in the customer's basket, given that the customer will buy products from the platform.

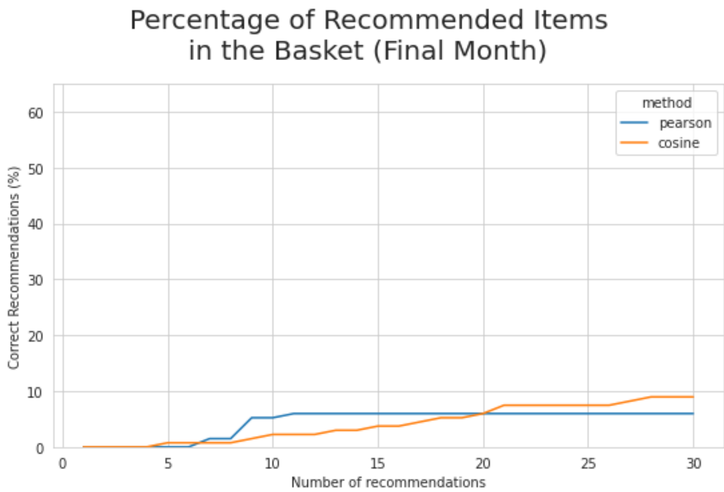


Figure 5.1 Item-level Purchase Accuracy with Respect to the Number of Suggestions

As seen in Figure 5.2, the curve for the correctness percentage of the recommended items became flat after 8 to 10 recommendations for item-level suggestions and provided up to 6% accurate recommendations to customers for items in their basket. This is to be expected since the number of products in the dataset is around 500 items. Also, cosine similarity converges to the 10% accuracy with a slower pace than its similarity measure counterpart. Moreover, looking at the brand-level results (Appendix A.8 shows the predictions of recommendation system for brand-level

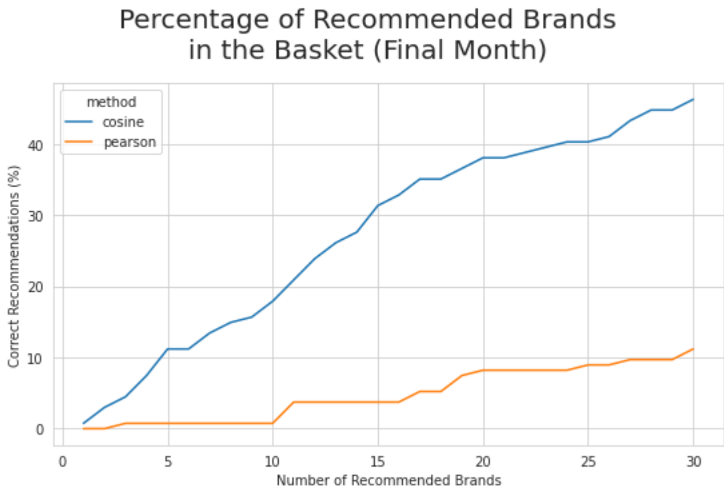


Figure 5.2 Brand-level Purchase Accuracy with Respect to the Number of Suggestions

utility matrix), it is evident that as the number of recommendations increases, there is a higher possibility of accuracy, and we can see this increasing trend in both considered similarity measures. However, using the cosine similarity measure, we can see that the system reached higher levels of accuracy, much faster, and much higher compared to its Pearson Correlation. By recommending 10-11 brands to the customer who wants to make a purchase with this similarity metric for the RS, we can expect that a product from these suggested brands will end up in the customer's basket with an accuracy of 20%.

It can be seen that the implementation of an RS on a retailer proved to be quite valuable. Although the results derived here are highly dependant on the size of the dataset provided by the retailers, the accuracy of these recommendations can be improved by more transactions captured from customer activity.

5.2 Managerial Insights

Next, we will explore the results and output of this engine to derive insights. These insights are drawn by analyzing the final engine outputs created using Cosine Similarity with different granularity levels. As depicted in the previous section, the recommendation system's accuracy in providing suggestions for item-level granularity was very low. Thus, we will not analyze the model's output for this level and only explore the output generated for main-category and brand levels.

We derive our insights from a model that was trained on all the data. This model was set to provide suggestions for all users and all the granularity level classes, not just the first few suggestions with the highest rank. The outputs depict the prediction of what each user would buy next with what probability. These probabilities can also be conveyed as the customer's desirability percentage of classes in each granularity level. As mentioned before, rows represent individual customers in each of these reports, and the columns are unique brands or main categories offered to them. The values are expressed in percentages which add up to 100% per row. They depict the customers' preference for buying from each class of granularity in their next purchase based on their previous and other customers' behavior.

5.2.1 Main-Category Level

The first granularity that we are going to analyze is the main category level. This company has 7 classes at this granularity level, and they are representative of all items with similar subcategories. For example, all the items in the Coffee or Cooking Oils subcategories, which are consumable products, fall under the main Market category.

To obtain an overall preview of preferences in each class of the main category, we sum the values of each class and normalize them by the total value. Since the number of brands in each main category was variable, the values were also normalized by the number of brands in each class to provide a more accurate measure of desirability. This will provide us the percentage score of category recommendations with respect to overall suggestions. Figure 5.3 summarizes the overall desirability of all customers for each category based on the recommendations made by the model. The higher percentages indicate that customers have more overall preference customers have for that category. The plot in Appendix A.9 provides the normalized version of this plot where values for each category are divided over the number of brands.

Recently there have been some changes in these categories. Mainly, the Stationery category has been merged with Home, and Home & Life has split into two different categories. Also, Jewelry is combined with Clothing. By considering all these changes and looking at Figure 3.1, we can derive that:

- Beauty products, Apparel, Market and Life & Home categories have the high-

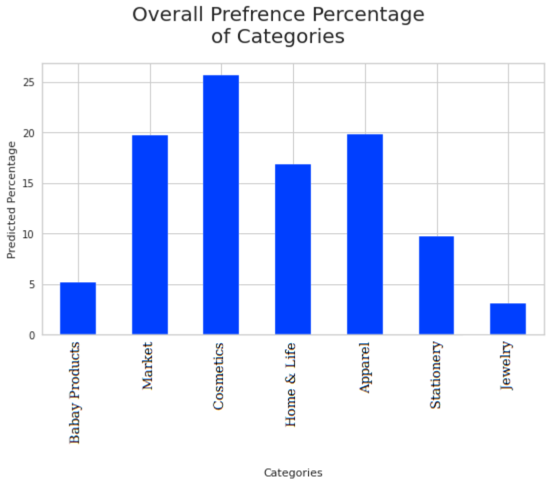


Figure 5.3 Overall Preference Percentage for each Category

est overall percentage of being recommended and having desirability among other categories. However, based on Appendix A.9, per brand desirability of items in Apparel and Life & Home categories are highest [By considering recent changes in the in their categories]. Although items in these two categories have the lowest median profit margin, the brands in these categories can be a basis for attracting customers.

- The Market category is growing and captures 15% of overall customer preference. Because of this category's profit margin, more advertisement in this category can also increase their profit more.
- Baby & Mom has the lowest overall predicted preference percentage share among consumers; however, since per brand desirability of this category is average and has a high median profit margin, investing in this category and advertising can increase their profits. Based on the median profit margin of the categories, presented in table 5.2, investing in this category along with the Market will increase their profits.

Categories	Profir Margin
Market	%31-%40
Jewelry	%21-%30
Apparel	%10-%20
Stationery	%31-%40
Cosmetics	%21-%30
Home & Life	%10-%20
Baby Products	%31-%40

Table 5.2 Median profit margin of each category

Additionally, analyzing the trend between main category classes can provide useful information. Since the model's output is derived from the customers' purchase behavior and depicts each category class's appeal, finding trends and relationships in each class can show which classes are behaving the same and which are not. Investing in one of the class pairs with the same trend can boost the sales of the other one. Moreover, category classes that either do not show any or have minimal trends may require more attention and investment to provide profitability. To understand these trends, we employ Pearson correlation to purchase probabilities in each category class pair and observe their relation. The results from using this method are outlined in Table 5.3. Based on this table Baby & Mom category is highly correlated with other categories, especially the market category, except Life & Home. Life & Home category has a negative correlation with the market and Life & Home categories. Apparel is also correlated with the market category. Products from the correlated categories can be bundled together for promotions.

	Baby Products	Market	Cosmetics	Home & Life	Apparel
Baby Products	1.000	0.769	0.589	-0.349	0.588
Home & Life	-0.349	-0.210	0.181	1.000	0.225
Market	0.769	1.000	0.439	-0.210	0.532
Cosmetics	0.589	0.439	1.000	0.181	0.360
Apparel	0.588	0.532	0.360	0.225	1.000

Table 5.3 Correlation between different categories offered by the retailer

5.2.2 Brand Level

Next, we will analyze the brand-level recommendations. Just like the previous section, an RS was created with cosine similarity that provides suggestions for all of the customers in which it ranks the brands they will place a purchase next from this company. This ranking is based on the probability of the next purchase generated by this model, which considers the similarity of their previous purchases with other customers' transactions for the duration of the data provided. In the output, the rows distinguish different users, and the 57 columns indicate different brands. The values in the dataset are the future period purchase probabilities, which add up to 100% for each row.

The recommendations provided by this engine show with what probability the customers are interested in each brand and will buy from them. To capture the overall desirability of each brand, we summed up the purchase probability of each brand among all users and normalized it by their total. The figure in Appendix A.11 paints the picture of the overall brand desirability by order. The highly desirable brand among all is shown to be Gourmezz.

Finally, we applied Pearson correlation to this output to observe the trend between each brand pairs. We set a threshold of minimum 40% and maximum -40% to capture any strong relations between them. This method can reveal any inclination between the brand pairs and provide valuable results for bundling products based on customer preferences. Table 5.4 illustrates the captured trend between the set bounds. As depicted in this table, there are five brands that are highly correlated; Züber, Cleanwynd, Happy Folks, Trunki, and Ordina Bag. The retailer can bundle products from these brands to increase their sales. The complete work of this method is presented in Appendix A.10.

Brand 1	Brand 2	Correlation Value
Cilt1 Beauty	Kombuçça	0.608282
Cleanwynd	Neutrogena	0.500194
Cleanwynd	Ordina Bag	0.629374
Cleanwynd	Züber	0.761100
Deep Fresh	Filtr Café	0.459283
Deep Fresh	Muscle Cheff	0.402172
Fabooks	Güzel Gıda	0.455964
Freshbak Crispy	Güzel Gıda	0.452624
Gourmezz	Tashoven	0.437588
Happy Folks	Real Techniques	0.488457
Happy Folks	Trunki	0.743980
Italtrike	Soul 2 Seven	0.698004
Le Nouveau	Shaman's Secret	0.425933
Neutrogena	Ordina Bag	0.514075
Neutrogena	Züber	0.623827
Ordina Bag	Züber	0.785764
Real Techniques	Trunki	0.489345

Table 5.4 Brands that have an absolute correlation higher that 0.4.

6. Conclusion & Discussion

This research studies the value of recommendation systems in online retail settings. We present our work by implementing a collaborative filtering recommendation engine for an online retailer that does not take advantage of such technology. The recommendation system used here utilizes cosine similarity to provide suggestions to the customer putting orders from this retailer's website. Also, we used these preference values derived from implementing the recommendations systems as a proxy to predict what customers will buy at the next step. The results show that adopting this system can be valuable for the retailer, and with 20 recommendations, the system can provide up to 7% correct product predictions in customers purchasing baskets. Also, at the brand level, the system can be as accurate as 38% with only 11 suggestions on the website. Furthermore, we used the predicted customer preferences to suggest bundling options for the retailer and derive insights on the setting.

There is an argument to be made about the feasibility of using collaborative filtering to suggesting recommendations. Intuitively, items tend to be classifiable in simple terms. It is easier to discover items that are similar because they belong to the same category than it is to detect that two users are similar because they prefer one category in common, while each also likes some categories that the other does not care for (Leskovec et al., 2014). On the other hand, although it is easier to find and recommend similar items because they belong to the same category, finding similar users with the same purchase behavior overcomes the relevancy limitations of items from different categories. For instance, people who want to enjoy going to the beach during summer buy sunscreen to protect themselves from sunburn. They also purchase use sunglasses and hats to protect their eyes and keep themselves cool. However, the categories of these items are different and not related. This irrelevancy shows the potential that user purchase similarity holds and justifies the use of collaborative filtering in these settings. However, utilizing these systems can have some limitations. For example, when two users both like a category, they may not have bought any items in common. In these cases, using more sophisticated and complex approaches such as hybrid recommendation systems that combine the predictive power of collaborative and content-based filtering methods can result in higher accuracy and better recommendations. However, to use this class of systems, we need to have access to both transactional data as well as product attributes. Moreover, similarity-based collaborative filtering approaches are relatively simple in the way they derive customer preferences compared to alternative methods, making them more interpretable. Furthermore, the size of the dataset and the number of records in the data negatively affect the rate at which they can provide recommendations, and there is a trade-off between the number of records included for building the model and the time performance of the system. In other words, these methods are not scalable. In this thesis, since the dataset used here was small and including less data was at the cost of losing valuable information, we included all of the records. For datasets containing more records and more volume, this problem can be addressed using the rolling window technique and considering a limited time frame for each customer to provide recommendations. Also, applying this approach creates an opportunity to suggest items based on more recent transactions and making the suggestions more up-to-date and relevant to the trends.

The outbreak of Covid-19 affected many businesses worldwide in terms of their sales and revenues. This outbreak also impacted the retailer understudy and caused a decrease in their sales. The reduction in their sales is also reflected in the monthly transaction data they have provided us. Since recommendation systems are highly affected by the amount of utilized data and their sparsity, we attribute the scarcity of sales in some months and its sparsity in general as one limitation of this study. Having more data is always beneficial in these settings and can improve the accuracy of recommendation engines. One of the contributions of this work is to utilize a small dataset where sparsity is a factor in analyzing the information derived from recommendations systems to study the value they can provide to retailers and customers. Moreover, the number of repeat transactions made by returning customers was low, which was one of the data attributes that contributed to its sparsity. Having this characteristic in the data limited our ability to track each customer's purchases and analyze the difference in their per transaction expenditure to observe the monetary effect of the recommendations systems on the retailer and study these engines' value. One way to address this issue is to increase the time span of the data analyzed to include more transactions by these customers. Although having more transactions in the data can influence the performance of these systems, studying these effects is an interesting topic for research.

The brand-level suggestions generated by the recommendations engine can be used as a lever for product substitution recommendations. In future studies, we can consider the exploration-exploitation trade-off of the suggestions to offer customers items that are more relevant to their taste and introduce them to new sets of products to encourage variety and diversity. To that end, we can follow the work of Jiang et al. (2014) and incorporate a metric that measures the quality of recommendations suggested to customers. This way, we can provide suggestions based on customers' preferences and introduce new products to enable substitution. By providing more diverse options, we can also address the problem of decreased aggregate sales diversity that entails traditional collaborative filters (Lee & Hosanagar, 2019). Improving upon these areas can be a direction of future studies that can help with increasing sales, and as a result, customer loyalty. Another focus could be studying the effects of using revenue maximization techniques in sorting the suggested products and the impact of order in item sets provided to customers. Studying this approach can help us understand the revenue impact of offering customized products to individual customers while maximizing the revenue gain of the retailer. Another avenue for further research is the use of other metrics to properly evaluate the performance of recommendation systems and assist in choosing correct hyper-parameters that can provide the highest accuracy.

BIBLIOGRAPHY

Adomavicius, G., Bockstedt, J. C., Curley, S. P., & Zhang, J. (2013, December). Do recommender systems manipulate consumer preferences? a study of anchoring effects. *Information Systems Research*, *24*(4), 956–975. Retrieved from `https://doi.org/10.1287/isre.2013.0497` doi: 10.1287/isre.2013.0497

Adomavicius, G., Bockstedt, J. C., Curley, S. P., & Zhang, J. (2018, March). Effects of online recommendations on consumers' willingness to pay. *Information Systems Research*, *29*(1), 84–102. Retrieved from `https://doi.org/10.1287/isre.2017.0703` doi: 10.1287/isre.2017.0703

Albadvi, A., & Shahbazi, M. (2009, November). A hybrid recommendation technique based on product category attributes. *Expert Systems with Applications*, *36*(9), 11480–11488. Retrieved from `https://doi.org/10.1016/j.eswa.2009.03.046` doi: 10.1016/j.eswa.2009.03.046

Al-Bashiri, H., Abdulgabber, M. A., Romli, A., & Hujainah, F. (2017, September). Collaborative filtering recommender system: Overview and challenges. *Advanced Science Letters*, *23*(9), 9045–9049. Retrieved from `https://doi.org/10.1166/asl.2017.10020` doi: 10.1166/asl.2017.10020

Ansari, A., & Mela, C. F. (2003, May). E-customization. *Journal of Marketing Research*, *40*(2), 131–145. Retrieved from `https://doi.org/10.1509/jmkr.40.2.131.19224` doi: 10.1509/jmkr.40.2.131.19224

Arora, N., Dreze, X., Ghose, A., Hess, J. D., Iyengar, R., Jing, B., ... Zhang, Z. J. (2008, September). Putting one-to-one marketing to work: Personalization, customization, and choice. *Marketing Letters*, *19*(3-4), 305–321. Retrieved from `https://doi.org/10.1007/s11002-008-9056-z` doi: 10.1007/s11002-008-9056-z

Banerjee, S., Sanghavi, S., & Shakkottai, S. (2016, June). Online collaborative filtering on graphs. *Operations Research*, *64*(3), 756–769. Retrieved from `https://doi.org/10.1287/opre.2016.1508` doi: 10.1287/opre.2016.1508

Bernstein, F., Kök, A. G., & Xie, L. (2015, October). Dynamic assortment customization with limited inventories. *Manufacturing & Service Operations Management*, *17*(4), 538–553. Retrieved from `https://doi.org/10.1287/msom.2015.0544` doi: 10.1287/msom.2015.0544

Bernstein, F., Modaresi, S., & Sauré, D. (2018, September). A dynamic clustering approach to data-driven assortment personalization. *Management Science*. Retrieved from `https://doi.org/10.1287/mnsc.2018.3031` doi: 10.1287/mnsc.2018.3031

Bradlow, E. T., Gangwar, M., Kopalle, P., & Voleti, S. (2017, March). The role of big data and predictive analytics in retailing. *Journal of Retailing*, *93*(1), 79–95. Retrieved from `https://doi.org/10.1016/j.jretai.2016.12.004` doi: 10.1016/j.jretai.2016.12.004

Bruyn, A. D., Liechty, J. C., Huizingh, E. K. R. E., & Lilien, G. L. (2008, May). Offering online recommendations with minimum customer input through conjoint-based decision aids. *Marketing Science*, *27*(3), 443–460. Retrieved from `https://doi.org/10.1287/mksc.1070.0306` doi: 10.1287/mksc.1070.0306

Caro, F., Kök, A. G., & Martínez-de Albéniz, V. (2020, January). The Future of Retail Operations. *Manufacturing & Service Operations Management*, *22*(1), 47–58. Retrieved 2021-06-15, from http://pubsonline.informs.org/doi/10.1287/msom.2019.0824 doi: 10.1287/msom.2019.0824

Chen, H.-H., & Chen, P. (2019, January). Differentiating regularization weights – a simple mechanism to alleviate cold start in recommender systems. *ACM Transactions on Knowledge Discovery from Data*, *13*(1), 1–22. Retrieved from https://doi.org/10.1145/3285954 doi: 10.1145/3285954

Cho, Y. H., & Kim, J. K. (2004, February). Application of web usage mining and product taxonomy to collaborative recommendations in e-commerce. *Expert Systems with Applications*, *26*(2), 233–246. Retrieved from https://doi.org/10.1016/s0957-4174(03)00138-6 doi: 10.1016/s0957-4174(03)00138-6

Dzyabura, D., & Hauser, J. R. (2019, May). Recommending products when consumers learn their preference weights. *Marketing Science*, *38*(3), 417–441. Retrieved from https://doi.org/10.1287/mksc.2018.1144 doi: 10.1287/mksc.2018.1144

Elahi, M., Ricci, F., & Rubens, N. (2016, May). A survey of active learning in collaborative filtering recommender systems. *Computer Science Review*, *20*, 29–50. Retrieved from https://doi.org/10.1016/j.cosrev.2016.05.002 doi: 10.1016/j.cosrev.2016.05.002

Ettl, M., Harsha, P., Papush, A., & Perakis, G. (2020, May). A data-driven approach to personalized bundle pricing and recommendation. *Manufacturing & Service Operations Management*, *22*(3), 461–480. Retrieved from https://doi.org/10.1287/msom.2018.0756 doi: 10.1287/msom.2018.0756

Farias, V. F., & Li, A. A. (2019, July). Learning preferences with side information. *Management Science*, *65*(7), 3131–3149. Retrieved from https://doi.org/10.1287/mnsc.2018.3092 doi: 10.1287/mnsc.2018.3092

Ghose, A., Li, B., & Liu, S. (2019, November). Mobile targeting using customer trajectory patterns. *Management Science*, *65*(11), 5027–5049. Retrieved from https://doi.org/10.1287/mnsc.2018.3188 doi: 10.1287/mnsc.2018.3188

Ghoshal, A., Menon, S., & Sarkar, S. (2015, September). Recommendations using information from multiple association rules: A probabilistic approach. *Information Systems Research*, *26*(3), 532–551. Retrieved from https://doi.org/10.1287/isre.2015.0583 doi: 10.1287/isre.2015.0583

Ghoshal, A., & Sarkar, S. (2014, August). Association rules for recommendations with multiple items. *INFORMS Journal on Computing*, *26*(3), 433–448. Retrieved from https://doi.org/10.1287/ijoc.2013.0575 doi: 10.1287/ijoc.2013.0575

Goldberg, D., Nichols, D., Oki, B. M., & Terry, D. (1992, December). Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, *35*(12), 61–70. Retrieved from https://doi.org/10.1145/138859.138867 doi: 10.1145/138859.138867

Huang, Z., & Zeng, D. D. (2011, February). Why does collaborative filtering work? transaction-based recommendation model validation and selection by analyzing bipartite random graphs. *INFORMS Journal on Computing*, *23*(1), 138–152. Retrieved from https://doi.org/10.1287/ijoc.1100.0385 doi: 10.1287/ijoc.1100.0385

Huang, Z., Zeng, D. D., & Chen, H. (2007, July). Analyzing consumer-product graphs: Empirical findings and applications in recommender systems. *Management Science*, *53*(7), 1146–1164. Retrieved from `https://doi.org/10.1287/mnsc.1060.0619` doi: 10.1287/mnsc.1060.0619

Jiang, H., Qi, X., & Sun, H. (2014, October). Choice-based recommender systems: A unified approach to achieving relevancy and diversity. *Operations Research*, *62*(5), 973–993. Retrieved from `https://doi.org/10.1287/opre.2014.1292` doi: 10.1287/opre.2014.1292

Lee, D., & Hosanagar, K. (2019, March). How do recommender systems affect sales diversity? a cross-category investigation via randomized field experiment. *Information Systems Research*, *30*(1), 239–259. Retrieved from `https://doi.org/10.1287/isre.2018.0800` doi: 10.1287/isre.2018.0800

Leskovec, J., Rajaraman, A., & Ullman, J. D. (2014). *Mining of massive datasets* (2nd ed.). USA: Cambridge University Press.

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing*, *7*(1), 76-80. doi: 10.1109/MIC.2003.1167344

Moon, S., & Russell, G. J. (2008, January). Predicting product purchase from inferred customer similarity: An autologistic model approach. *Management Science*, *54*(1), 71–82. Retrieved from `https://doi.org/10.1287/mnsc.1070.0760` doi: 10.1287/mnsc.1070.0760

Murthi, B. P. S., & Sarkar, S. (2003, October). The role of the management sciences in research on personalization. *Management Science*, *49*(10), 1344–1362. Retrieved from `https://doi.org/10.1287/mnsc.49.10.1344.17313` doi: 10.1287/mnsc.49.10.1344.17313

Rubens, N., Elahi, M., Sugiyama, M., & Kaplan, D. (2015). Active learning in recommender systems. In *Recommender systems handbook* (pp. 809–846). Springer US. Retrieved from `https://doi.org/10.1007/978-1-4899-7637-6_24` doi: 10.1007/978-1-4899-7637-6_24

Schein, A. I., Popescul, A., Ungar, L. H., & Pennock, D. M. (2002). Methods and metrics for cold-start recommendations. In *Proceedings of the 25th annual international ACM SIGIR conference on research and development in information retrieval - SIGIR '02.* ACM Press. Retrieved from `https://doi.org/10.1145/564376.564421` doi: 10.1145/564376.564421

Sharma, R., & Singh, R. (2016, May). Evolution of recommender systems from ancient times to modern era: A survey. *Indian Journal of Science and Technology*, *9*(20). Retrieved from `https://doi.org/10.17485/ijst/2016/v9i20/88005` doi: 10.17485/ijst/2016/v9i20/88005

Song, Y., Sahoo, N., & Ofek, E. (2019, August). When and how to diversify—a multicategory utility model for personalized content recommendation. *Management Science*, *65*(8), 3737–3757. Retrieved from `https://doi.org/10.1287/mnsc.2018.3127` doi: 10.1287/mnsc.2018.3127

Zhang, J., & Krishnamurthi, L. (2004, November). Customizing promotions in online stores. *Marketing Science*, *23*(4), 561–578. Retrieved from `https://doi.org/10.1287/mksc.1040.0055` doi: 10.1287/mksc.1040.0055

APPENDIX A

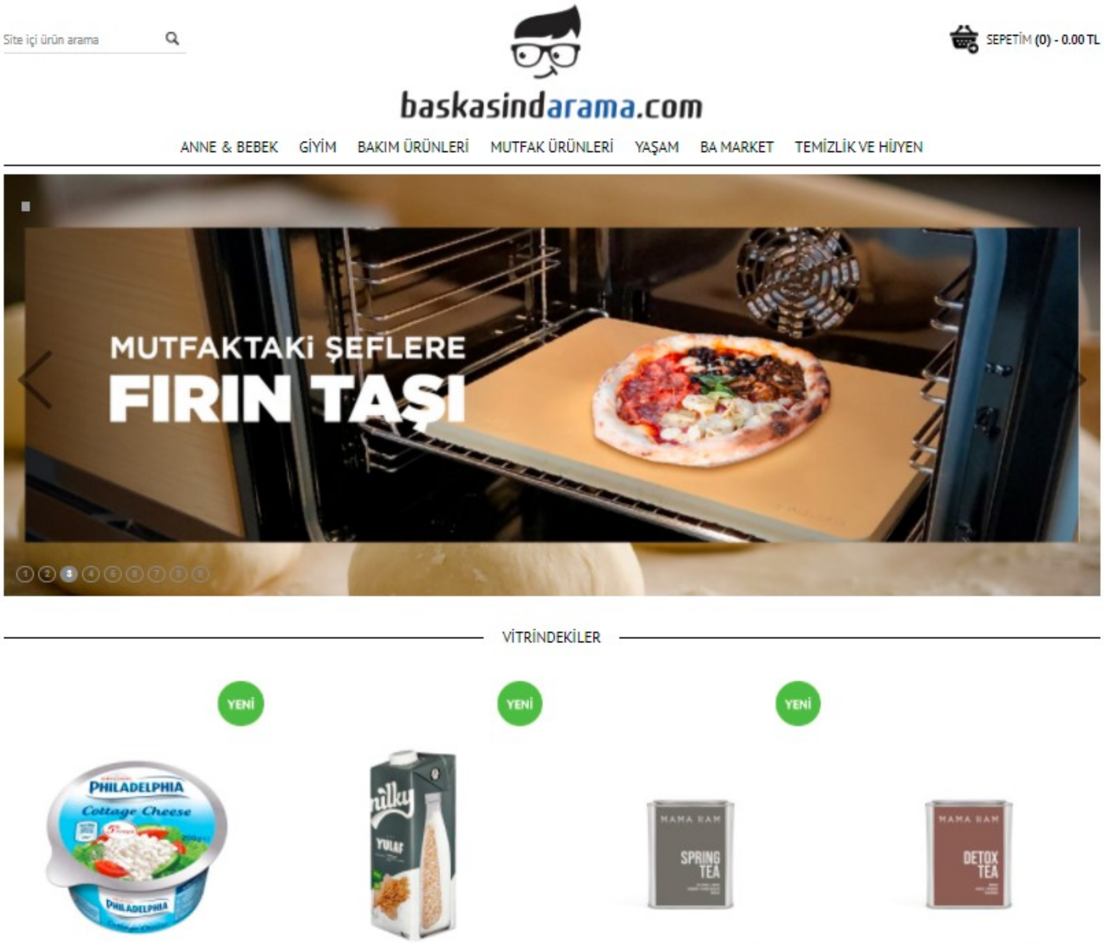


Figure A.1 Snapshot of retailer's environment

			NEW VISITOR	RETURNING VISITORS		
			540	63		
CUSTOMER ID	PAGE VIEWED	CHART SITUATION	NUMBER OF ORDER	PRODUCT PRICE	TURNOVER	PROFIT MARGIN
C62	https://www.baskasindarama.com/urun/biyik-sekillendirici-wax	1	1	59	59	2
C62	https://www.baskasindarama.com/urun/vintage-biyik-fircasi-ve-taragi	1	1	42	42	2
C63	https://www.baskasindarama.com/urun/rosece-vucut-losyonu-yasemin-ve-sandal-ag	1	1	68.75	68.75	2
C64	https://www.baskasindarama.com/urun/fabooks-perfect-is-boring-spiral-bloknot	1	1	19	19	3
C64	https://www.baskasindarama.com/urun/fabooks-you-are-super-duper-amazing-tale	1	1	15	15	3
C64	https://www.baskasindarama.com/urun/fabooks-i-love-london-mini-defter	1	1	6.9	6.9	3
C65	https://www.baskasindarama.com/urun/flip-alarm-saat	1	1	265	265	2

Figure A.2 Snapshot of Daily Sales repot provided by the retailer

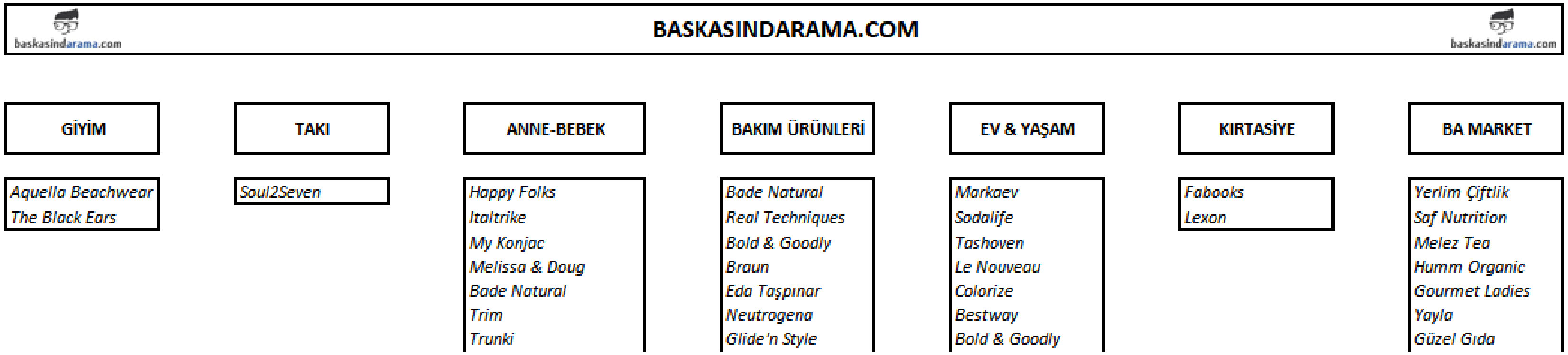


Figure A.3 Snapshot of Brands and their association to each respective category

TOTAL NUMBER OF PAGES VIEWED	RATE OF IMMEDIATE EXIT	AV. DURATION ON ONE PAGE
5138	75.26%	0:02:20

PAGE LINK	VIEW NUMBER
https://www.baskasindarama.com/kategori/defter-ajanda	316
https://www.baskasindarama.com/urun/market-alisveris-listem	745
https://www.baskasindarama.com/kategori/ba-market?marka=gourmezz	526
https://www.baskasindarama.com/kategori/kadin-bakim-urunleri	726
http://www.baskasindarama.com/yeni-urunler	896
http://www.baskasindarama.com/populer-urunler	755
https://www.baskasindarama.com/kategori/kadin-bakim-urunleri?marka=bade-natural	354
https://www.baskasindarama.com/kategori/makyaj-temizleme-urunleri	455
https://www.baskasindarama.com/kategori/saglik-ve-hijyen-urunleri	365

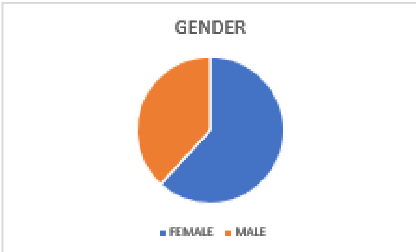
Figure A.4 Snapshot of daily Page Visit report

	NEW VISITORS	RETURNING VISITORS
	17456	5259
Organic Search	61.40%	61.90%
Direct	27.20%	23.80%
Social	9.30%	7.20%
Referral	2.10%	7.10%

Figure A.5 Snapshot of Traffic Acquisition report

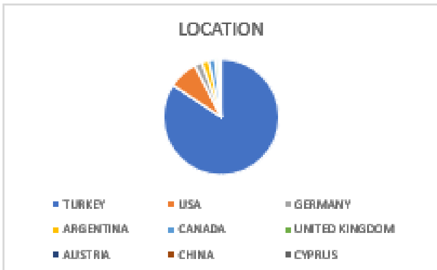
2. GENDER

ALL USERS	
FEMALE	61.80%
MALE	38.20%



3. LOCATION

ALL USERS	
TURKEY	82.09%
USA	7.93%
GERMANY	2.06%
ARGENTINA	1.91%
CANADA	1.76%
UNITED KINGDOM	0.73%
AUSTRIA	0.29%
CHINA	0.29%
CYPRUS	0.29%



4. DEVICE

ALL USERS	
mobile	54.33%
desktop	44.64%
tablet	1.03%

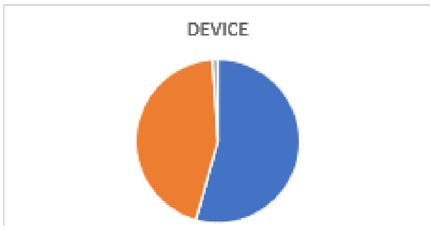


Figure A.6 Snapshot of Customer Demographics report provided by the retailer

	transaction	customer	customer code	item	category	Category Class	brand	day	month	date	situation	number_ordered	price	turnover	profit_margin	special campaign	url
0	20200101C9	C9	9	ıbian Filtre Kahv	Kahveler	BA MARKET	Filtr Café	1	1	2020-01-01	RDER COMPLET	1	35	35	%21-%30		https://www.baskas
1	20200101C8	C8	8	ıral Yoğun Besle	ilt Bakım Seruml	AKIM ÜRÜNLE	Bade Natural	1	1	2020-01-01	RDER COMPLET	2	74.95	149.9	%31-%40	ISCOUNT CAM	https://www.baskas
2	20200101C7	C7	7	Kaş ve Kirpik Be	ilt Bakım Seruml	AKIM ÜRÜNLE	Bade Natural	1	1	2020-01-01	RDER COMPLET	6	49.95	299.7	%31-%40	ISCOUNT CAM	https://www.baskas
3	20200101C6	C6	6	Kaş ve Kirpik Be	ilt Bakım Seruml	AKIM ÜRÜNLE	Bade Natural	1	1	2020-01-01	RDER COMPLET	2	49.95	99.9	%31-%40	ISCOUNT CAM	https://www.baskas
4	20200101C5	C5	5	Kaş ve Kirpik Be	ilt Bakım Seruml	AKIM ÜRÜNLE	Bade Natural	1	1	2020-01-01	RDER COMPLET	5	49.95	249.75	%31-%40	ISCOUNT CAM	https://www.baskas
5	20200101C4	C4	4	ıe Eğilimli Ciltler i	nizleme Jelleri ve	AKIM ÜRÜNLE	Bade Natural	1	1	2020-01-01	RDER COMPLET	1	89.9	89.9	%31-%40		https://www.baskas
6	20200101C8	C8	8	Karşıtı ve Canlan	ilt Bakım Seruml	AKIM ÜRÜNLE	Bade Natural	1	1	2020-01-01	RDER COMPLET	2	74.95	149.9	%31-%40	ISCOUNT CAM	https://www.baskas
7	20200101C4	C4	4	Bentonit ve Aktif	skeler ve Peeling	AKIM ÜRÜNLE	Bade Natural	1	1	2020-01-01	RDER COMPLET	1	109.9	109.9	%31-%40		https://www.baskas
8	20200101C4	C4	4	iltler için Hızlı Etk	ilt Bakım Seruml	AKIM ÜRÜNLE	Bade Natural	1	1	2020-01-01	RDER COMPLET	1	79.9	79.9	%31-%40		https://www.baskas
9	20200101C3	C3	3	şekersiz Fındık E	ık ve Fındık Ezme	BA MARKET	Rosta Fındık	1	1	2020-01-01	RDER COMPLET	4	21	84	%10-%20		https://www.baskas

Figure A.7 Processed data used as an input for the recommendation system

customer code	Aquella Beachwear	Bade Natural	Bestway	Bold & Goodly	Braun	Cilt1 Beauty	Cleanwynd	Colorize	Deep Fresh	Doğalsan	EcoTools	Eda Taşpınar	Era Gurme	Fabooks	Filtr Café	Freshbak Crispy
1	0	0	0	2.03106	0.61229	1.36763	0	1.26289	0	5.20252	0.95604	3.42169	0	2.20618	2.72934	0
2	2.76622	2.19302	1.03123	1.08148	1.60394	0.72452	0.80658	0.79957	0	0	1.52152	1.96162	2.28636	1.33493	1.56505	0
3	3.37314	0.54434	2.30943	1.97951	0	0	0	1.21297	0	5.09163	0	0	6.65649	1.39264	1.63301	0
4	8.48212	0	0	1.99108	0	1.43374	0	0	0	3.84104	0	1.75596	2.97572	0	4.92767	0
5	4.11325	0	0	1.32278	1.11391	0.95251	0	0	0	0	0	1.95424	2.31063	0.62833	2.53694	0
6	8.48213	0	0	1.99108	0	1.43373	0	0	0	3.84104	0	1.75596	2.97571	0	4.92768	0
7	4.14297	0	1.09796	1.39601	1.76066	0.66641	0	0	0	6.02154	0.78739	1.23117	2.08638	0.33105	2.67858	0
8	8.48212	0	0	1.99108	0	1.43374	0	0	0	3.84104	0	1.75597	2.97572	0	4.92767	0
9	1.66894	3.54447	0.14472	1.40061	0	0.17865	0.40934	1.01026	2.80829	3.401	0.78122	0.2188	3.32526	3.76692	0	0

Figure A.8 A snapshot of the recommendation system's output at brand-level

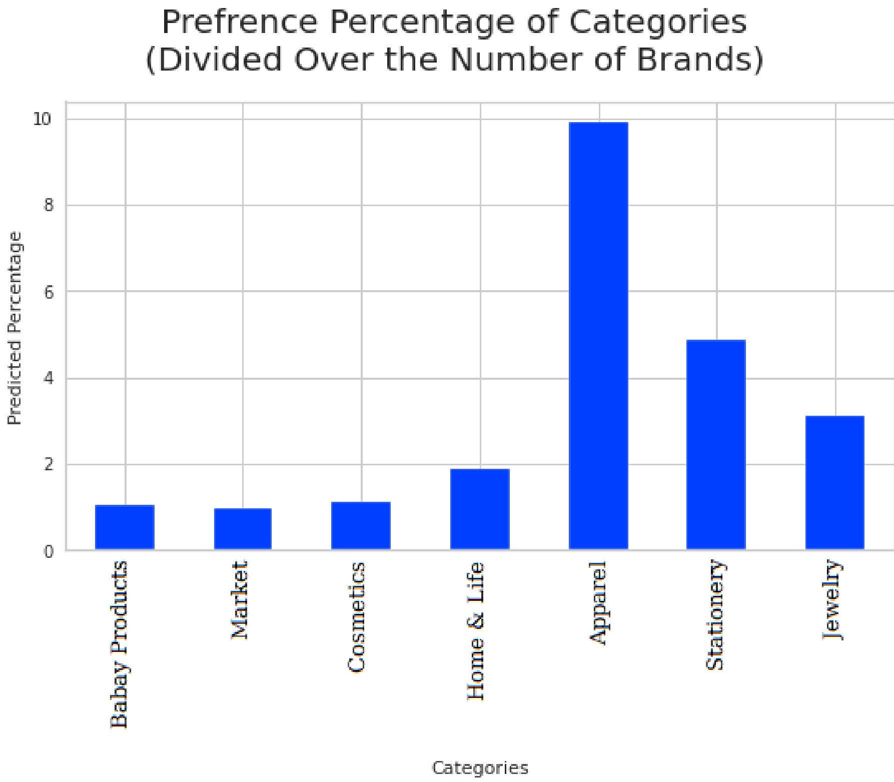


Figure A.9 This plot shows the overall predicted preference of all customers for each category per number of brands. The values are divided over the total number of brands in each category.

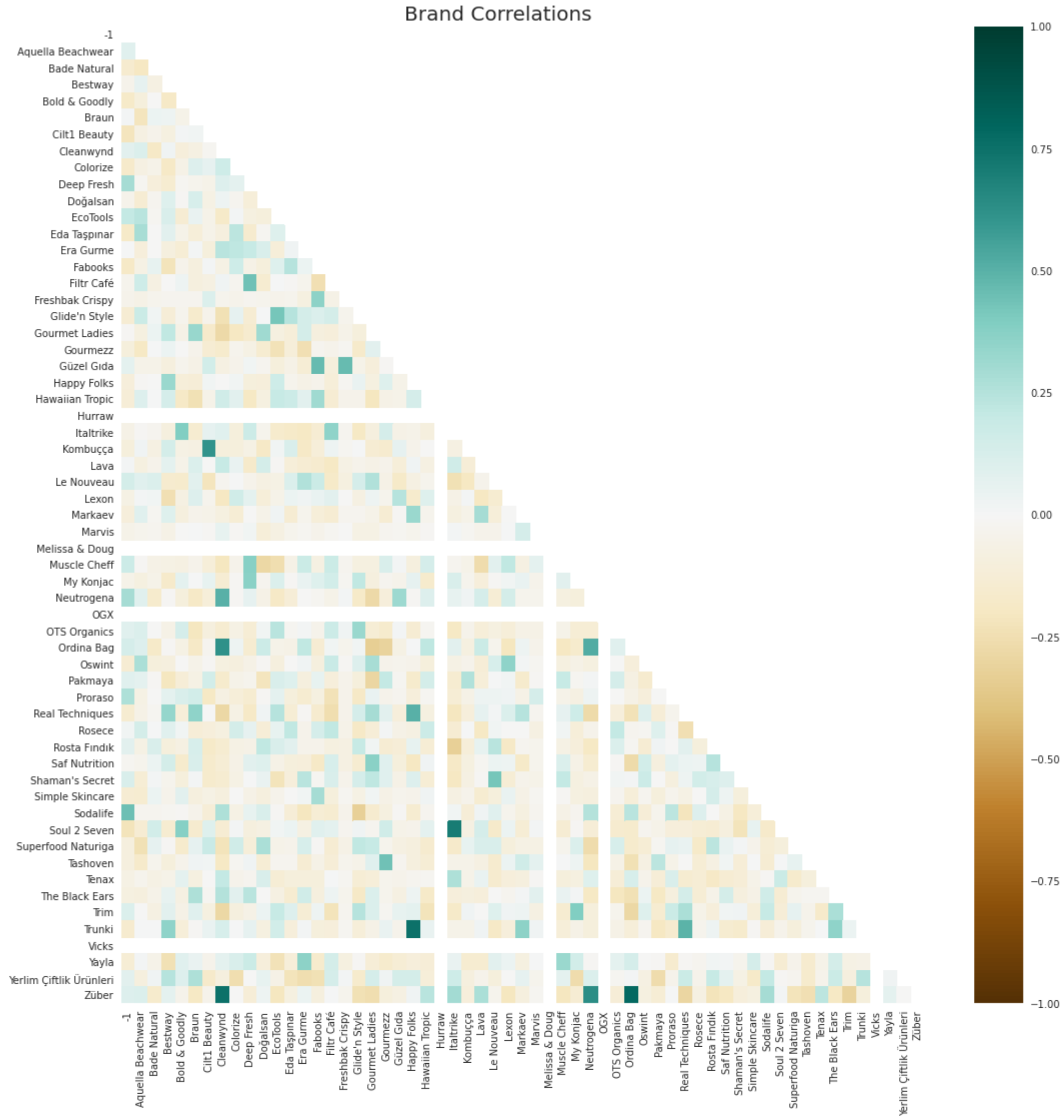


Figure A.10 The plot provide a heat map of correlation between all the predicted preference percentage for all brands for the next period

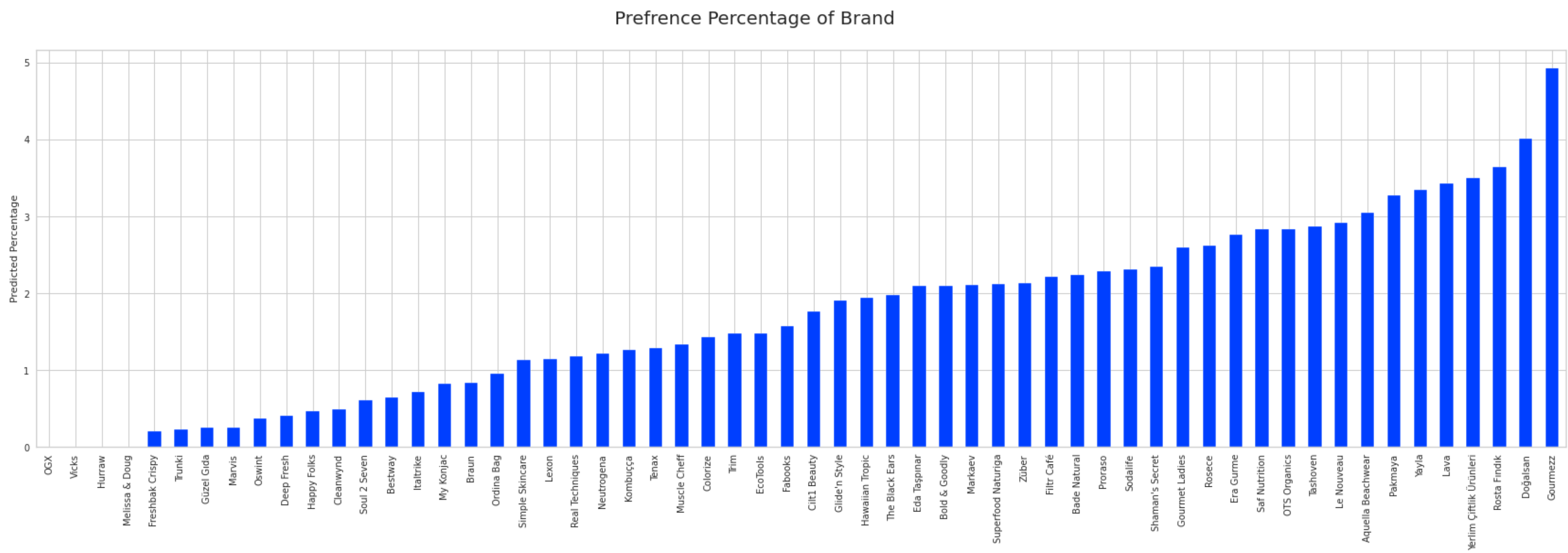


Figure A.11 The plot depicts the overall predicted preference percentage of all customers for each brand. The higher the percentage, higher preference overall customers have for that brand.