Personalised Recommendation Systems and the Impact of COVID-19: Perspectives, Opportunities and Challenges

Rabaa Abdulrahman and Herna L. Viktor
School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Ontario, Canada

Keywords: Recommendation Systems, COVID-19, Machine Learning, Cold Starts, Grey Sheep.

Abstract: Personalised Recommendation Systems that utilize machine learning algorithms have had much success in recent years, leading to accurate predictions in many e-business domains. However, this environment experienced abrupt changes with the onset of the COVID-19 pandemic centred on an exponential increase in the volume of customers and swift alterations in customer behaviours and profiles. This position paper discusses the impact of the COVID-19 pandemic on the Recommendation Systems landscape and focuses on new and atypical users. We detail how online machine learning algorithms that are able to detect and subsequently adapt to changes in consumer behaviours and profiles can be used to provide accurate and timely predictions regarding this evolving consumer sector.

1 INTRODUCTION

Recommendation Systems have been widely utilized in e-commerce settings to guide users through their shopping experiences. A principal advantage of these systems is their ability to narrow down purchase options and market items to customers. Specifically, personalised recommendation systems based on collaborative filtering recommend items that have been rated by other users with preferences similar to those of the targeted users. Intuitively, the more information that is collected about users, the more accurate the recommendations put forth by such systems will be.

Creating accurate and timely recommendations for new or atypical users is an active area of research. In the literature, new users are referred to as cold starts, while atypical users are categorised as grey sheep. Recently, machine learning (ML) algorithms have had much success in improving the accuracy of recommendations for these user categories that are 'difficult to pinpoint'. For instance, (Abdulrahman et al., 2019) combines cluster analysis, deep learning and active learning, or the so-called user-in-the-loop system, to yield accurate recommendations for cold-start users. In another recent study, (Abdulrahman and Viktor, 2020) employs one-class learning in order to address the grey sheep challenge.

Although personalized recommendations have been discussed in the literature since the 1990s, they have only been widely adopted by e-businesses recently. According to (Chen et al., 2014), a personalized Recommendation System should include data collection, data warehousing, data mining, and data applications. Data mining techniques[1] can make predictions without accessing users' profile information and items; hence, they have been used to improve recommendation performances (Yoon-Joo, 2013) (Lucas et al., 2012). Many successful businesses have implemented personalized approaches. For instance, Amazon created a personalized recommendation list for each user and was followed by other businesses, such as Hotels.com, which helps the user come to a decision based on a pared down suggestion list (Oestreicher-Singer, 2013). Furthermore, studies have shown that using these approaches increases profits (Li and Karahanna, 2015). However, the e-business landscape changed abruptly with the onset of the COVID-19 pandemic. This position paper presents some thoughts on the current state of the field and suggests some perspectives with regards to the future.

This paper is organised as follows. Section 2 focuses on the above-mentioned challenges that personalised Recommendation Systems are currently facing. In Section 3, we discuss the current impact of

[1]We use the terms data mining technique and ML algorithm interchangeably. However, we wish to note that data mining focuses more heavily on the discovery of patterns, often by using ML algorithms.

Abdulrahman, R. and Viktor, H.
Personalised Recommendation Systems and the Impact of COVID-19: Perspectives, Opportunities and Challenges.
DOI: 10.5220/0010145702950301
In *Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2020) - Volume 1: KDIR*, pages 295-301
ISBN: 978-989-758-474-9

COVID-19 and highlight future research directions. Section 4 concludes the paper.

2 TOWARDS PERSONALISED RECOMMENDATION SYSTEMS

Consumers face information overload every time they access the Internet to make a purchase. In today's fast-paced world, they have neither the time nor the patience to explore all these suggestions. Therefore, the main idea behind Recommendation Systems is to address the above-mentioned problem and aid users in narrowing down their list of choices through an understanding of their preferences and personalising the experience, as depicted in Figure 1.

Generally speaking, Recommendation Systems use content-based filtering (CBF), collaborative filtering (CF) or hybrid approaches (Abdulrahman et al., 2019). These systems rely on two basic inputs: the set of users in the system, U (also known as customers), and the set of items to be rated by the users, I (also known as the products) (Kumar and Thakur, 2018). The systems employ matrices based on past purchase patterns. With CBF, the system focuses on item matrices, whereby it is assumed that if a user liked an item in the past, he or she is more inclined to prefer a similar item in the future. Therefore, these systems study the attributes of the items. On the other hand, CF systems focus on user-rating matrices, recommending items that have been rated by other users with preferences similar to those of the targeted user. Thus, these systems rely on historic data consisting of user ratings and similarities across the user network. As hybrid systems employ both the CBF and CF approaches, they concurrently consider items based on users' preferences and on the similarity between the items' contents. In recent years, research has trended toward hybrid systems. Another growing trend is the use of ML algorithms to identify patterns in users' interests and behaviours, including supervised, unsupervised and one-class learning algorithms.

In essence, the recommendation process consists of three main phases, namely, **information collection**, **learning** and **recommending** (Kumar and Thakur, 2018). During information collection, as the name suggests, the aim is to learn more about the users. As many authors have noted, the accuracy of the recommendation is highly related to the quality of information about the users in the system. This information enters the system in the form of users' feedback. There are three types of feedback that could exist in the system: explicit feedback, where the user provides a rating through the system interface; implicit feedback, where the system monitors user behaviour, history, and purchases; and hybrid feedback, which is a combination of explicit and implicit feedback. During the learning phase, an algorithm is applied to learn the users' preferences. Finally, the system turns out predictions in the form of prediction scores, where a particular prediction score measures how likely it is that user U_i will be interested in item I_o, or recommendations, each of which list the top N items that might be of interest to a specific user.

As noted above, recommendation systems have been highly successful in tracking existing customers with typical profiles. However, when clients are first-time users of e-business systems, as is the case with the surge in online shopping during the COVID-19 pandemic, their preferences are unknown. Furthermore, an increasing number of users have unique and exotic tastes, which makes it harder for the system to match their interests with the current customer base. These two categories of users may also overlap, leading to inaccurate or nonsensical recommendations.

2.1 Cold-start Users

Recall that a cold-start user refers to a new user with unknown preferences. In recommendation systems, users' preferences, historic data regarding what they like and dislike and their item ratings and reviews are used to match them with other users. In cold-start situations, such information does not exist, which makes it difficult for the system to calculate similarity scores. Indeed, the tremendous increase in the use of e-commerce websites during the current COVID-19 pandemic has highlighted the importance of, and difficulty in, providing accurate recommendations to many first-time users (Argaman, 2020).

To address this problem, some researchers use CBF systems, in which information about the items is used to find the best match (Lu, 2015). Other systems simply present these users with a predefined recommendation list. Although these solutions may be successful with some users, they often result in redundant lists being presented, which causes these users to lose interest. Another solution is to use conversational learning models, where the new user is presented with a list of questions to build a preference profile (Lamche et al., 2014). Doing so might also drive them away due to the time it takes to build the profile or privacy concerns. Recently, (Abdulrahman et al., 2019) proposed the Popular User Personalized Prediction (PUPP-DA) framework, which combines active learning, ML, and deep learning algorithms to

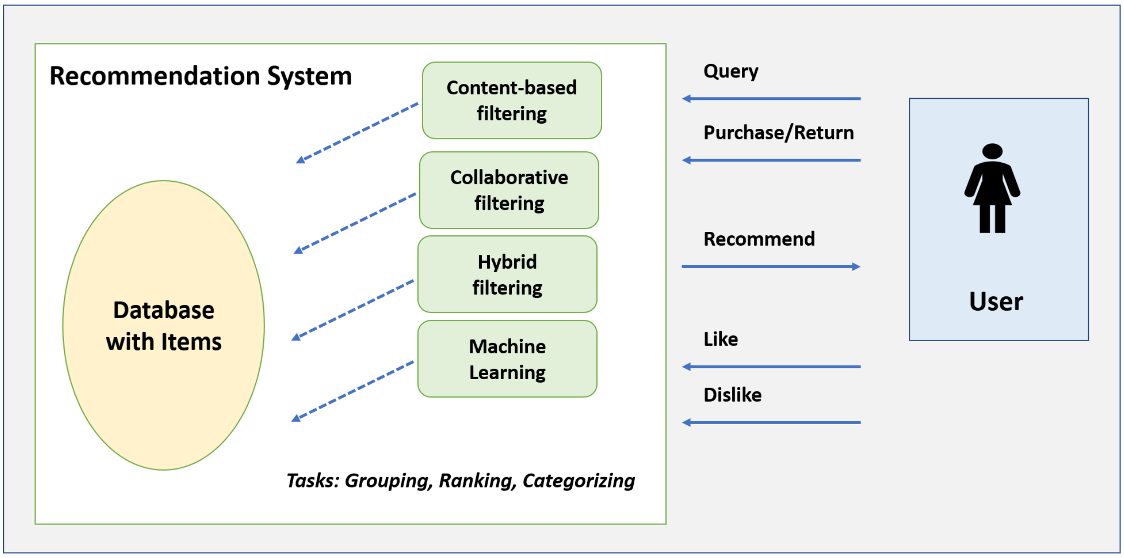

Figure 1: Personalised Recommendation System.

accurately recommend items to new users.

2.2 Grey-sheep Users

As mentioned above, grey-sheep users are difficult to identify or characterize. These users are often willing, to some degree, to share their feedback. However, their preferences typically do not match the majority of preferences in the system. In contrast to cold starts, the system might have the information it needs to calculate similarity scores and produce recommendation lists. However, such lists may not be accurate due to their unique tastes and characteristics. Typically, grey-sheep users are treated as outliers and removed from the system (Srivastava et al., 2020). Other researchers move them to a separate system where their preferences can be better matched with those of others (Zheng et al., 2017). However, doing so is not a realistic option in large and online systems, as identifying and moving users/items to a secondary system is time consuming. (Abdulrahman and Viktor, 2020) presents the Grey-Sheep One-Class Recommendation (GSOR) framework, which is designed to create accurate prediction models while considering both regular and grey-sheep users. The GSOR framework utilizes one-class classification, whereby the learning process is accomplished with information from the majority class, while predictions are made for the minority class, i.e. the grey-sheep users.

3 COVID-19 AND PERSONALISED RECOMMENDATION SYSTEMS

Recently, with the onset of the COVID-19 pandemic, many businesses have turned to e-commerce solutions in an attempt to not only survive but also thrive in the post-pandemic world (Goldstein, 2020). When COVID-19 appeared in late 2019, governments were forced to develop plans for facing the virus when it arrived in their countries. In many countries, lockdown procedures were implemented immediately, leaving citizens with little choice but online shopping.

As the World Bank Group (Ungerer et al., 2020) notes, e-commerce is emerging as a major pillar in the COVID-19 crisis. Before the pandemic, for many users, e-commerce was used to import unique items unavailable in local markets or to provide the luxury of shopping from the warmth and comfort of home during cold winters. However, for many, the pandemic transformed e-commerce into a tool for survival. In many countries, even if a complete lockdown was not enforced, physical distance measures were encouraged. Thus, as infection rates climbed, people started to turn to online ordering to avoid contact with other people. In addition, the movements of the vulnerable and elderly were restricted, leading a large portion of these individuals to turn to e-commerce for the first time. Furthermore, in many countries, most non-essential businesses closed until further notice. In order for these businesses to survive, they needed to reach out to customers through web-based or social media stores, for instance, the Instagram and Facebook markets.

In terms of general e-commerce, online shopping has shifted from being a convenience in terms of time and location to being a necessity. In fact, as the United States started lifting its partial lockdown and opening up the economy again, a survey of consumers' intentions regarding the return to old shopping practices was conducted (Post, 2020). The results showed that 24% of those surveyed did not intend to shop in a mall during the next six months, while another 16% stated that they did not intend to do so for the next three months. We believe that the same observations hold true in Canada.

The current shift in consumers' habits stresses the importance of meeting customers' demands. Furthermore, it confirms the significance of catering the right products to the right customers, including cold starts and grey sheep, to avoid losing them to other businesses and to also streamline supply chains. Online competition is at its peak, and a significant percentage of businesses must address this challenge. Several studies have shown the importance of e-commerce, along with personalised Recommendation Systems, during the pandemic across all sectors. For instance, this shift is also relevant to the health care sector, where healthcare providers have moved to e-commerce to provide tailor-made care and treatments (Ungerer et al., 2020).

In January 2020, the U.S. Census Bureau of the Department of Commerce reflected on the growth of e-commerce and noted that, in the United States alone, sales were expected to top $4.2 trillion USD in 2020 and that 2.1 billion customers would have shopped online by the end of the year (Winkler, 2020). These numbers and expectations were based on data previously collected for 2020. However, on April 30, 2020, Amazon released their first-quarter financial results, which described their total earnings as "exceptionally" strong, as Amazon had made an estimated $33 USD million an hour in sales for the first three months of the year (Kaplan, 2020). In North America alone, sales increased by 29%, i.e. by about $46.1 billion, compared to the same period in 2019.

3.1 Addressing Cold Starts and Grey Sheep

The implications of this trend for the Recommendation System research community are manifold. Indeed, as the number of users increased exponentially, many new users were added to systems. Another aspect of note is that, even for existing users, there has been a shift in their preferences. Since the pandemic started, many users have switched preferences from "what to buy" to "what is needed," which has resulted in previously popular and frequently rated items being ignored. Furthermore, considering the current situation we live in, many businesses have decided to maintain work-from-home practices until the end of 2020. Consequently, many consumers have changed their clothing preferences from formal dress, for instance, to comfortable lounge wear.

Another challenge centres on cold-start users, as many of these individuals have turned to e-commerce for the very first time. This influx poses a challenge for Recommendation Systems, since there substantial gaps exist in what is known about these users. It may well be that a substantial portion of these new users are indeed grey sheep who typically would not use e-businesses during normal times.

Considering these challenges, let us now illustrate the current situation with some examples from the Canadian perspective. As discussed earlier, the shift in preferences causes data sparsity, which is a principal challenge for Recommendation Systems. According to Statista (2019), the lowest two categories by household type who shopped online in Canada prior to the pandemic were singles who cohabitated with other adults (e.g. parents or roommates) and single parents (Statista, 2019). These groups represented 12% and 3%, respectively, of all users. Within these groups, there are users that have never shopped online before or are currently using e-commerce now for different types of demands. In Canada, Millennials and Baby-Boomers produced the highest percentage of online consumer sales during 2019 (Post, 2020). Today, preferences have turned towards ordering what is necessary for homeschooling or entertaining children. A 2019 report by Canada Post indicated that 62% of Canadians shopped for clothing apparel, whereas 41% shopped for computers and electronics using e-commerce. After the pandemic hit, a report by Cision (2020) showed that all e-commerce sales increased, except for clothing (which had the lowest increase of 21%). Meanwhile, the sales of electronics increased by 160% (Absolunet, 2020).

In 2019, it was reported that Pre-Boomers, i.e. those aged 73 and older, as well as Gen Z, i.e. customers in the 18–23 age group, constitute the lowest percentages, 5% in each category, of online shoppers in Canada, as depicted in Figure 2. These customers represent two very different generations and are thus often difficult to target. For a business to thrive online, it must understand its customers' behaviours and characteristics in order to expand its customer base. Gen Z, for instance, is considered to exert the main

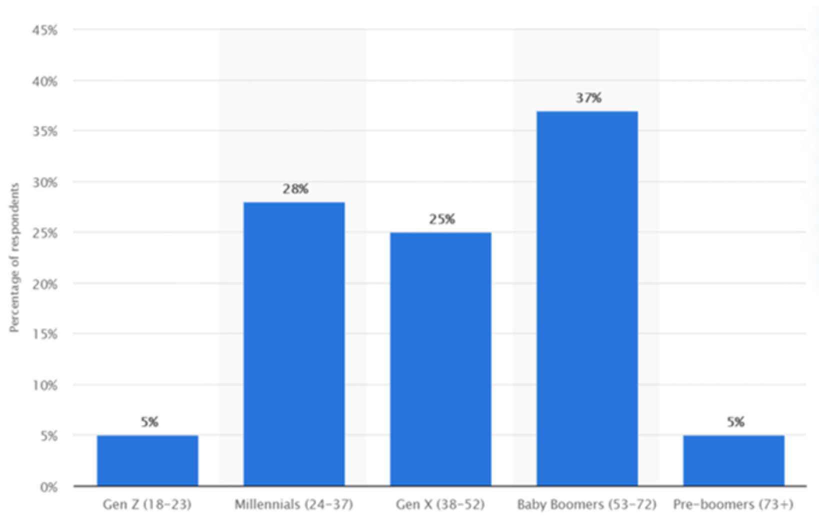

Figure 2: Online Shopping by Age in Canada in 2019 (Statista, 2019).

influence over buying decisions for families (Goldstein, 2020). According to Forbes (2020), technology is crucial for enhancing the Gen Z shopping experience and providing them with instant and quality services (Goldstein, 2020). In Canada, it is expected that by 2026, 21% of the population will fall into the 73 years old and older category. Older customers often fall in the grey-sheep category and, as discussed in (Insider, 2020), they have a preference for products that provide them with improved quality of life. As noted by Retail Insider (2020), this group of customers prefers to view products physically before buying them (Insider, 2020). For instance, as the pandemic lockdown started in Canada in mid-March 2020, many grocery stores dedicated special hours to senior shoppers. However, a recent study by Statistics Canada indicated that a large portion of such customers turned to online shopping, with 45% of people aged 75 and older indicating that they did so (Post, 2020). The question here is how to target these customers and, as many have turned to e-commerce for the first time, how to keep them in the customer base when life returns to normal. Next, we explain how online ML algorithms can be used to address this challenge.

3.2 Adaptive Machine Learning

ML algorithms have been used successfully in increasing the accuracy of personalised Recommendation Systems. The earliest works focused on using the k-nearest neighbours technique (Kumar and Thakur, 2018), where a recommendation is provided by calculating the distance between a user U_i and all others in the database in terms of user characteristics that are described by a number of features $F = \{f_1, .. f_l\}$. Next, the k nearest neighbours $\{U_1, .. U_k\}$ of user U_i are determined using some distance measure, such as Euclidian distance, and their preferred items are suggested to user U_i. This approach is based on the assumption that consumers are easily grouped into neighbourhoods, and the accuracy of the approach is highly dependent on the available features and the distance measure. Other recent methods employ advanced algorithms, such as ensembles, cluster analysis and deep learning algorithms, to improve the quality of predictions (Abdulrahman et al., 2019).

A major drawback of most the above-mentioned algorithms is that they are unable to detect and adapt to changes. Such changes can be gradual, incremental, re-occurring, seasonal or abrupt, as illustrated in Figure 3. Abrupt change occurs when customer behaviours and/or customer profiles change over a very short time period. Gradual change occurs more slowly and less radically. Gradual drift can be incremental, with many intermediate steps between the extremes, or dispersed, whereby new trends appear in increasingly more instances. It is also possible for previous patterns to reoccur through time. For example, seasonal patterns might reoccur each year but not necessarily at exactly the same time. Formally, let us assume that a set of features $F = \{f_1, .. f_l\}$ is utilised to recommend an item I_o from the item set $I = \{I_1, .. I_p\}$ to user U_i. A concept drift has occurred if there is a change in the probability $P(I_o|F_i)$, i.e. the probability that item I_o will be preferred by user U_i, who is described by a feature set F_i.

Indeed, the COVID-19 pandemic constituted a major and abrupt change in consumer behaviours. In addition, the presence of numerous new, and atypi-

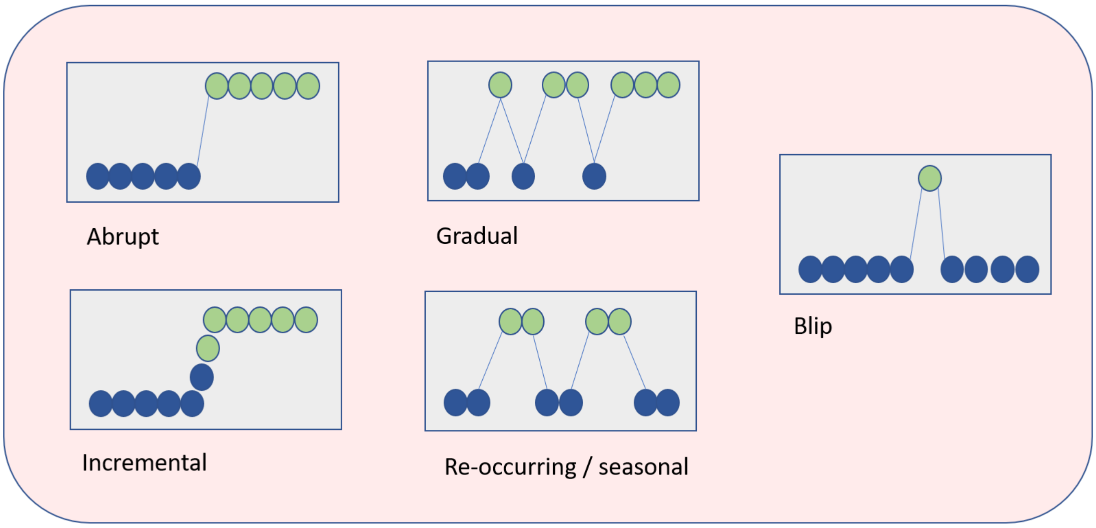

Figure 3: Drifts in consumer behaviour (adapted from (Gama et al., 2014)).

cal, users led to a further recommendation challenge. Traditional ML algorithms are not able to automatically detect and handle such changes in user preferences and profiles. Rather, a decrease in the accuracies of their predictions will indicate that the models are incorrect, leading to the realisation that new models need to be built; this typically happens after some delay. On the other hand, online or incremental learners, such as adaptive trees and ensembles, are highly suitable for learning in such changing environments (Bifet et al., 2018). These incremental learning algorithms update their models upon the arrival of new data and may "forget" old concepts using local replacement, in which irrelevant subsections of the model are discarded and replaced with subsections trained on recent data. This process made possible by the incorporation of drift detection into their designs; thus, these algorithms are able to dynamically and seamlessly adapt their models to changes in user preferences. Although such explicit concept drift detection is not necessary for incremental algorithms to adapt to drifting concepts, as they often do so naturally by continually updating and forgetting, it does afford several advantages. For example, if concept drift occurs abruptly, the model can detect and adapt to it more quickly. Concept drift detection also provides insights into the mechanics of the generation process in order to facilitate the modelling of future re-occurring or seasonal changes in customer profiles or purchase patterns.

In terms of the COVID-19 pandemic, it is too early to say whether a second or third wave will occur. It is also not possible to predict consumer behaviours in the unfortunate event that these waves occur. However, the authors are of the opinion that any such event may potentially lead to another abrupt drift or the recurrence of the patterns observed in mid-March 2020. The use of online learning algorithms that incorporate drift detection algorithms appears to be a promising research direction, helping to ensure that e-businesses are able to adapt rapidly and efficiently to changes in their customer bases and purchase patterns (Ungerer et al., 2020), while facilitating interactions with cold starts and grey sheep.

4 CONCLUSIONS

Recommendation Systems are crucially important for the economic growth of businesses engaged in e-commerce. With the recent abrupt shift in their lives, many consumers currently depend on e-commerce for essential items. The challenge is accommodating this entire customer base, including loyal customers, new users, and those with unique tastes, as the pandemic continues to ebb and flow. This position paper illustrated how online ML, by incorporating change detection in the design, can be potentially utilised to address these challenges.

The COVID-19 pandemic has been a shocking, yet eye-opening experience, with a wide impact on e-commerce, technology, and Recommendation Systems. This impact has, however, not been all negative. For instance, Shopify, a well-known Ottawa-based e-commerce business, recently became the most valuable publicly traded company in Canada in May 2020, even topping the stock value of the Royal Bank of Canada (Simpson, 2020). Shopify's financial results for the first quarter of 2020 increased by 47%, an increase in total revenue of $470 million USD compared to the same period last year. In the merchant solutions component of Spotify, which houses its Recommendation System, there was growth of 57%, as reported by (Simpson, 2020). Indeed, the Shopify case study reconfirms the value, importance and growth of intelligent personalised Recommendation Systems.

Incremental ML approaches will continue to offer crucial insights into evolving consumer bases, and our current research focuses on this aspect of ML. We plan to utilize drift detection algorithms from the online learning research community (Gama et al., 2014) to build adaptive predictive models. Our future work will also include a study of the world-wide impact of shifting consumer habits on the Recommendation Systems landscape, with a focus on cold starts and grey sheep.

REFERENCES

Abdulrahman, R. and Viktor, H. (2020). *Catering for Unique Tastes: Targeting Grey-Sheep Users Recommender Systems through One-Class Machine Learning*. Technical Report IDEAL1-2020, University of Ottawa.

Abdulrahman, R., Viktor, H., and Paquet, E. (2019). Active learning and user segmentation for the cold-start problem in recommendation systems. In *Proceedings of the Eleventh International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2019*. ScitePress.

Absolunet (2020). COVID-19 driving consumers online: Canadian ecommerce sales double in 2 weeks. In *CISION*.

Argaman, O. (2020). Why customer re-engagement has become more valuable during COVID-19. In *https://www.destinationcrm.com/Articles/Web-Exclusives/Viewpoints/Why-Customer-Re-engagement-Has-Become-More-Valuable-During-COVID-19-141234.aspx*.

Bifet, A., Gavaldà, R., Holmes, G., and Pfahringer, B. (2018). *Machine Learning for Data Streams with Practical Examples in MOA*. MIT Press.

Chen, M., Chen, K., and Chen, C. (2014). The CRM-based digital exhibition system for clothing industry. In *International Journal of Electronic Business Managementg*.

Gama, J., Zliobaite, I., Bifet, A., Pechenizkiy, M., and Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4).

Goldstein, J. (2020). What does Gen Z want from retail? In *https://www.forbes.com/sites/forbesbusinessdevelopmentcouncil/2020/02/25/what-does-gen-z-want-from-retail/#578ce084661b*. Forbes.

Insider, R. (2020). How does aging change Canada's consumer behavior? In *Available from: https://www.retail-insider.com/articles/2020/2/how-does-aging-change-canadas-consumer-behavior*. Retail Insider Press.

Kaplan, M. (2020). Amid covid-19, Amazon's Q1 2020 earnings confirm e-commerce dominance. In *https://www.practicalecommerce.com/amid-covid-19-amazons-q1-2020-earnings-confirm-ecommerce-dominance*.

Kumar, P. and Thakur, R. (2018). Recommendation system techniques and related issues: a survey. In *International Journal of Information Technology*.

Lamche, B., Trottman, U., and Worndl, W. (2014). Active learning strategies for exploratory mobile recommender systems. In *Proceedings of the fourth Workshop on Context-Awareness in Retrieval and Recommendation*.

Li, S. and Karahanna, E. (2015). Online recommendation systems in a B2C e-commerce context: A review and future directions. In *Journal of the Association for Information Systems*.

Lu, J. (2015). Recommender system application developments: A survey. In *Decision Support System*.

Lucas, J., Segrera, S., and Moreno, M. (2012). Making use of associative classifiers in order to alleviate typical drawbacks in recommender systems. In *Expert Systems with Applications*.

Oestreicher-Singer, G. (2013). The network value of products. In *Journal of Marketing*.

Post, C. (2020). The 2020 Canadian e-commerce report. In *Canada Post Press*.

Simpson, M. (2020). Shopify sees 47 percent revenue growth in q1 2020 amid impacts of covid-19. In *https://betakit.com/shopify-sees-47-percent-revenue-growth-in-q1-2020-amid-impacts-of-covid-19*.

Srivastava, A., Bala, P., and Kumar, B. (2020). New perspectives on grey sheep behavior in e-commerce recommendations. In *Journal of Retailing and Consumer Services*.

Statista (2019). Distribution of online shoppers in Canada as of April 2019, by household type. Cited 2020, Available at https://www.statista.com/statistics/1044443/canada-online-shoppers-by-household-type/.

Ungerer, C., Portugal, A., Molinuevo, M., and Rovo, N. (2020). Recommendations to leverage e-commerce during the covid-19 crisis. In *World Bank Group*.

Winkler, N. (2020). Quarterly e-commerce sales. In *https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf*.

Yoon-Joo, P. (2013). The adaptive clustering method for the long tail problem of recommender systems. In *IEEE Transactions on Knowledge and Data Engineering*.

Zheng, Y., Agnani, M., and Singh, M. (2017). Identifying grey sheep users by the distribution of user similarities in collaborative filtering. In *Proceedings of the sixth Annual Conference on Research in Information Technology*.