Data algorithm: the underlying logic behind shopping behavior

Dexin Sun, Yuanjie Zhu, Zhuoying Li, and Ziliang Huang
Department of Informatics, University of Washington
INFO 300: Research Methods
Professor Clarita Lefthand
May 26, 2022

Abstract

Knowing that online shopping is becoming a popular trend, we decided to research how recommendation algorithms in shopping sites affect customers' behavior and attitudes towards online shopping. Many customers are getting addicted to online shopping, causing financial problems. We utilize two research methods: expert interviews and an online survey to study this phenomenon. And we expect to conduct statistical analysis through graphs and data calculation. Hope we can reduce people's concerns and provide more insight into the development and restriction of recommendation algorithms.

Data algorithm: the underlying logic behind shopping behavior

With the development of information technology and the wide application of big data, recommendation algorithms make information spread in a more personalized, customized, and intelligent way, especially in the field of e-commerce. And due to this creation, online shopping has become part of people's daily lives. While people enjoy the convenience of online shopping, which significantly facilitates the access to information between different merchants, people's behaviors are also influenced by this algorithm implicitly. If customers are exposed to certain repeating information very frequently, their attitudes and behaviors may become different than they previously were. And even some customers are addicted to online shopping, making unreasonable purchases.

So, our research question is how this algorithm is applied to eCommerce and what the logic behind it is, and to determine how customers' behaviors are impacted by the recommendation algorithm in either a positive or negative way. The customers care about their privacy, mental health, and other relevant consequences, but it is hard to observe themselves and make a systematic analysis. Through our research, we want to help customers build up a healthier and more rational attitude toward online shopping. Our research will provide a more professional and objective analysis through online surveys and expert interviews. And through this process, we look forward to bringing more insights into the development of big data that help find the balance to deal with the relationship between humans and big data.

Literature Review

Data algorithm is a broad topic; the subset of the data algorithm that relates to our study is the recommendation algorithm. When customers frequently browse information for similar products while shopping online, the recommendation system uses algorithms and data analysis techniques to provide customers with products that are related to their search history. Gao(2021) points out that, nowadays, many online shopping websites will predict the merchandise that customers may be interested in by analyzing their order, searching, and browsing records. In Gao's study, they propose a "dynamic network intelligent hybrid recommendation algorithm distinguishing time periods (DIHR) model" (Gao et al., 2021), which joins the shopping behavior and buying preference of customers and takes the customers' preference as a factor and sorts by order. This model solves the problem of the accuracy of the recommendation caused by the customer's numerous and miscellaneous searching records. This study shows that the algorithm's accuracy largely influences the products that appear on customers' recommendation pages. However, we still want to figure out how the customer's preferences are influenced and how the algorithm can affect shopping behavior.

Lo et al.(2016) and Yang et al.(2020) both mentioned in their study the quantity and duration of clicks. Only a few of the variables can influence a customer's preferences. The most essential aspects that influence a customer's purchase are the customer's personality and the item's characteristics. Yang mentioned a very important point: "For that, some studies took the neutral sentiments into account and calculated the sentiment orientation of multidimensional attributes of products based on the product attribute classification framework of the website to support online shopping."(Yang et al., 2020) However, data on customers' historical click behavior on shopping sites are abundant when compared to customers' purchase history or ratings. Similarly, the same reaction to Lo et al.(2016) paper also proposed a framework for predicting customers' preferences based on past click history, enhancing purchase prediction accuracy. Customer click data has been demonstrated to be useful in predicting customers' preferences and avoiding the problem of data sparsity. However, the answer to the question of how the algorithm affects customers' shopping behavior is still ambiguous.

While shopping online, customers leave ratings and reviews of the products, which are also a form of shopping behavior. There could be some relations between these shopping behaviors and the algorithm. The sentiment orientation of product opinion pairs from different platforms interacts with customers' purchase decisions, according to Yang et al(2020). However, due to a large number of product reviews and people's limited ability to comprehend information, accurately describing the sentiment orientation of reviews and their interaction processes are difficult. Lim et al.(2016) made a good point that the conclusion can be depicted that subjective norms and perceived usefulness significantly positively influence online purchase intention but subjective norms insignificant influence shopping behavior negatively. To support customers' buying decisions, a method for opinion pair mining and interaction operators based on web reviews is required. (Yang et al., 2020) These papers provide us with a very good example of a side-by-side illustration of the fact that algorithms for internet shopping impact a large part of customers and how behavior and algorithms can influence each other. However, this study still lacks how exactly the customers are affected and what their behavior might have changed because of the algorithm.

With regard to the articles we looked for above, we can learn that from their point of view, they all tell us about the impact that different algorithms have on the customers when shopping online. Even though they talk about the importance of algorithms, they are not specific

to the customer's shopping behavior. In our study, we want to see through the phenomenon and know how algorithms actually change the shopping behavior of customers.

Research Methods

Survey

Lester et al. (2005) indicated that more than 95% of the college-age market accesses the Internet, and over 91 percent of that cohort makes online purchases, as projected. Nearly a quarter of purchasers spend more than $500 each year on Internet items. Since the main demographic of online shoppers is young people, we will conduct an online questionnaire survey of about 200 young people. Half of these people are men, and half are women, and we will draw from the undergraduate students at the University of Washington, about 50 people each grade from freshman to senior. The questionnaire will ask the participants about their online shopping experience, their opinion of the website's recommendation page, and whether their spending habits, such as budget planning, have changed due to shopping online. In order to avoid bias that probably exists, the sample size should be large enough, and we will ask the participants if they understand what a recommendation algorithm is and if they are aware of the existence of recommendation algorithms on online shopping sites. In this way, we can eliminate some of the untrustworthy or worthless responses while maintaining a sufficient amount of data to complete the data analysis.

The reason for choosing online surveys is that online surveys are easy to conduct compared with traditional research methods such as paper surveys. It is faster: the response time is instant, all the data is being gathered automatically, and we don't need extra time and energy to organize the data. Besides, it is financial-friendly because we can save costs by moving to online platforms from a paper format. But most importantly, online surveys are easy to be analyzed at any time due to their easy access to data, which will provide convenience for researchers and improve their productivity.

We plan to post the survey on school forums such as Reddit to gain support from the students in this way. Suppose the questionnaire posted on the forum does not collect enough data. In that case, we will consider contacting the professors of our Info major and ask them to help us send our questionnaire to the students in their classes via email and ask them to fill it out voluntarily.

Expert interview

Inspired by Tobias et al. (2018), an expert interview aims to obtain data from a purposely selected group of individuals rather than from a statistically representative sample of a broader population; we chose an expert interview as our research method. After conducting the survey and analyzing its data, we will conduct an expert interview of approximately 10 people who have experience with recommendation algorithms (e.g., software development engineers, Algorithm Engineer, and professors at a university who focuses on algorithm research). We will lead to discussions about their experiences with recommendation algorithms. This strategy will allow us to gain a deeper understanding of how recommendation algorithms work with customers' shopping behavior, and we want to narrow the gap in professional background knowledge that the general survey left open. To locate participants, we'll send emails via LinkedIn to software engineers who work in tech companies, asking whether they'd be interested in participating in a short discussion regarding recommendation algorithms. We can also look for participants with relevant experience among UW alumni or inquire if any professors from related majors are interested in our topic. After we have sufficient participants, we will host a 15 to 20 minutes

conversation. The group will be half male and half female from the Seattle area with at least a Bachelor's degree.

The interview will begin with a broad question about how their employment or past experience ties to algorithms and then ask them how recommendation algorithms aid in enhancing business sales. Then we'll talk about the issues or worries they have with this algorithm, as well as how they feel about the algorithm's contribution/effect on their customers. After we've answered all of these questions, we'll ask more detailed questions to help me better comprehend the data collected from the online survey.

Data Analysis

First, we clean up the online survey we receive to eliminate some possible bias and fabricated answers. Secondly, we will analyze those who think they often do online shopping individually and select participants who think their shopping behavior has been influenced by the recommendation algorithm. Through these analyses, we are able to obtain the percentage of people who have been influenced and their perceptions of the recommendation algorithm.

For the expert interview, we will try to find more insights from their side to get insiders' opinions on how the algorithm can affect customers and their concerns about the impact of the algorithm. And we could categorize the samples into different groups based on a factor such as the frequency of online shipping. Then we will use tools such as RStudio and Python to analyze how other factors may affect it. Graphs such as pie charts, histograms, scatter plots, etc., will be applied to compare the differences and find out which factor has the most impact. And through this comparison, we will form some new questions prepared for the following expert interview.

<div align="center">

Discussion

</div>

Threats To Validity

One of the drawbacks of online surveys is that the participants may not be aware of this algorithm and lack background knowledge. And this potential issue may result in meaningless answers or even misleading results if our questionnaire contains questions that are hard to understand. And a small portion of biased data will largely affect the analysis. Besides, our questionnaire generates questions that require descriptive and subjective answers. These answers are hard to be converted to digital data and are difficult to analyze. And to improve our survey question, we may use a more visualized/digitized answer such as a rank from 1 to 5 that participants could choose to represent their preference. And the subjective answers will have bias, especially when we only have one group of samples.

In addition, since we only send out the survey on the campus, the limited size of samples and the result of the sample can only represent the student group and may not apply to other groups outside of the school zone. In other words, respondents with biases may choose themselves into the sample of online surveys used on the population to which they are dispersed. Only when the findings from a sample can be applied to a larger population is research worthwhile. Findings from online surveys cannot be generalized and may therefore mislead when the population addressed by the survey cannot be identified, and the sample is tainted by respondents with biases.

Our expert interview will seek professional profiles that provide us with a specific level of expertise and accuracy. However, it will also bring us a lot of bias at the same time. As practitioners of different companies, they have the responsibility and obligation to keep certain business secrets and to hide them from competing companies due to conflicts of interest. Some information is very sensitive, such as the company's composition, and they have the right not to disclose the truth to us. This can lead to a certain bias in the conclusions we draw. Likewise, as

practitioners at the same level, we are concerned about the homogeneity of their answers. There is no doubt that the world is diverse, but many things tend to be easily homogenized in the human mind because people have a purpose. Because of the profit-making purpose of enterprises and the rendering of society's money-earning atmosphere, it is easy to form an alienation of people, leading to their homogenization and bias in our research results.

Expectation

Based on our research and speculation, we already have certain expectations about our results. Customers tend to purchase products relevant and recommended by the website, and they are also more likely to purchase the products recommended many times. This behavior may put more financial pressure on customers' budgets by making unnecessary purchases, and merchants' sales will increase. It is also possible that customers feel annoyed by certain recommendations if they appear excessively, which means the recommendation algorithm isn't always effective.

Nowadays, recommendation algorithms are widely used in major online shopping platforms, and the impact of related technologies on customers urgently needs to be popularized. Through the online surveys and expert interview research methods, we expect to understand the impact of recommendation algorithms on people, get insights from relevant practitioners, and promote better improvements for practitioners. Besides, we hope the customers can establish a healthier and rational attitude toward online shopping, making a better balance on their budget. Generally, our research provides better insight for both customers and companies on the future development of recommendation algorithms.

Conclusion

Our study as a whole explores the two-way influence of information recommendation algorithms on producers and consumers. Information recommendation algorithms not only influence individual access to information but also the recommendation distribution across the platform, which in turn influences the orientation of producers, algorithms, and data to establish a feedback loop between content consumers and producers. Forming the logic of online shopping platforms and bringing about a new kind of producer-consumer relationship allows producers to have a more intuitive and direct understanding of customers' needs, enabling producers to understand the needs and evaluations more intuitively and directly of their customers.

We have studied various existing studies, summarized them, and found that there are still questions to be solved: how customers' behaviors are impacted by the recommendation algorithm positively or negatively. The online survey was used to find out more about customers' perceptions of recommendation algorithms and the impact they felt they had on their shopping experience. The expert interview approach involved asking industry professionals to get an inside look at the industry. Data algorithms will be more widely used in people's lives in the near future, and our research is therefore of great importance.

References

Gao, Y., Liang, H., & Sun, B. (2021, January 1). *Dynamic Network Intelligent hybrid recommendation algorithm and its application in online shopping platforms*. Journal of Intelligent & Fuzzy Systems. Retrieved May 25, 2022, from https://doi.org/10.3233/JIFS-201579

C. Lo, D. Frankowski et al., "Understanding Behaviors that Lead to Purchasing: A Case Study of Pinterest," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 531-540, 2016.

Zoli Yang, Tianxiong Ouyang, Xiangling Fu, Xindong Peng. *A Decision‑Making Algorithm for Online Shopping Using Deep‑Learning ...* https://onlinelibrary.wiley.com/doi/10.1002/int.22225.

Lester, D. H., Forman, A. M. & Loyd, D. (2005). Internet Shopping and Buying Behavior of College Students. Services Marketing Quarterly, 27(2), 123-40.

O.Nyumba, T., Wilson, K., Derrick, C. J., & Mukherjee, N. (2018). The use of focus group discussion methodology: Insights from two decades of application in conservation. *Methods in Ecology and Evolution*, *9*(1), 20–32. https://doi.org/10.1111/2041-210x.12860

Lim, Yi Jin, et al. "Factors Influencing Online Shopping Behavior: The Mediating Role of Purchase Intention." *Procedia Economics and Finance*, Elsevier, 4 Mar. 2016, https://www.sciencedirect.com/science/article/pii/S2212567116000502.