AI in Personalized Product Recommendations

Kartik Hosanagar
The Wharton School, University of Pennsylvania

Dokyun Lee
Questrom School of Business, Boston University

Consumers today are presented with a vast wealth of product offerings in digital marketplaces and storefronts. It is essential that managers consider how to ensure that shoppers find their company's products. Kartik Hosanagar and Dokyun Lee explain the power of recommender systems, fueled by AI, to do just that.

Given the explosion in product offerings in digital marketplaces and storefronts, managers must consider how consumers will find products of interest among seemingly endless alternatives. Recommender systems are an important solution to the problem. These systems combine data drawn from sources including clickstream, purchases, product ratings, user profiles, and social networks to predict which products are best suited to a particular customer. They help consumers to become aware of new products and to select desirable ones from a myriad of choices. For firms, recommender systems have the potential to convert browsers into buyers, cross-sell products, and increase customer loyalty. Introduced nearly two decades ago, recommender systems are becoming ever more relevant and their impact ever stronger as available data and product assortments increase.

> Recommender systems exert a significant influence over consumer choice.

Recommenders are known to exert a significant influence over consumer choice. According to a McKinsey & Co report, 75 percent of viewing hours streamed on Netflix and nearly 35 percent of sales at Amazon originate in automated recommendations.[1] In 2016, Netflix estimated that their recommendation engine is worth $1 billion or more per year.[2] In e-commerce, recommender systems have been conservatively estimated to drive a 15 percent increase in product views and a 7.5 percent increase in total conversion, from click to purchase.[3] Firms such as TikTok have also found personalization technologies to be a source of competitive advantage.[4]

Recommender Systems Design Practices

The design of recommender systems is an active area of research, with many articles available on its specific facets.[5] More generally, recommender designs can be classified as either *content-based* or *collaborative filter-based* systems. Content-based systems use either product metadata (e.g., author, genre, musical attributes) or deep-learning-processed raw content (e.g., soundwave data for songs) to recommend items similar to those that a user rated highly. Collaborative filters, in contrast, recommend what similar customers bought or liked. For example, the classical customers "who bought item X also bought item Y."

Let us consider the design of three music recommendation services: Pandora, Last.fm, and Spotify. These services use three different approaches to recommender system design.[6]

Pandora's online radio service emerged from the Music Genome Project, a research effort that uses musical attributes to describe songs. Musicologists listened to tracks and assigned over 450 attributes to each. These range from the extent of instrumentation in the music to more esoteric attributes such as "rhythmic syncopation." Once a user indicates that they like a song on Pandora, the algorithm finds other songs that have similar musical attributes. For example, if a user chooses "Thunder" by Imagine Dragons, Pandora might recommend "Ride," a pop song by Twenty One Pilots, which features reggae beats. Pandora indicates that the song was recommended because it features "a dub production, a reggae feel, acoustic rhythm piano, use of a string ensemble, and major key tonality."

Pandora's approach depends on having the detailed attributes of every product (in this case, musical tracks), and is therefore called a content-based recommender system. Collecting this data manually is incredibly time-consuming and expensive. Moreover, knowing these attributes isn't all that useful when a retailer sells a wide range of products. If you've only ever bought books on Amazon.com, your taste for post-war thrillers will do little to help the company know what music to recommend, let alone which shoes, couches, or cars.

Collaborative filtering is an alternative approach that you might recognize through the "People who bought this also bought…" and "People like you also liked…" recommendations that we often see on Amazon and other websites.[7] The collaborative filtering approach used by Last.fm identifies the users who listen to the song "Thunder" and notes the other songs they have been listening to. This approach does not depend on rich metadata and as a result is unlikely to come with detailed domain-specific explanations like those provided by Pandora.

What collaborative filters lack in depth of knowledge, they make up for in simplicity. They are easy to implement and roll out in a short time. They have therefore become the most popular class of automated product recommenders on the Internet. They work very well in practice and the social pull of knowing what others are listening to or buying likely helps to drive consumers' choices.

These two design approaches can be used to recommend not only music but also videos (Netflix, YouTube), news (Google News) or, really, any product on retail websites (Amazon).

The advantage of a collaborative filtering design is that there is no need for detailed metadata and it is better able to capture the social benefits of shared consumption. However, the approach has its drawbacks. One major drawback is that collaborative filters create a sales concentration bias whereby popular products are more likely to be recommended. Because they recommend products based on what others have consumed, collaborative filters cannot recommend products that are not yet popular even if they would have been rated favorably by the consumer.[8] These systems also cannot explain the reasons for their recommendations, since the algorithm has no knowledge of product characteristics. Finally, collaborative filters also have a cold-start problem with new products; they are unlikely to recommend new products because they have no previous customer purchase or ratings data. These three drawbacks are well addressed by content-based designs. But content-based recommenders are difficult to build because they require detailed metadata which may not be available.

Because the two design approaches have different strengths and weaknesses, hybrid designs that combine the simplicity of collaborative filters with the impartiality of content-based designs are highly appealing. But how do we extract detailed metadata about products in a way that does not involve significant manual effort? To avoid the time and effort involved in having musicologists listen to and collect musical attributes for each and every track, we can use AI to automate the work. Spotify's hybrid design does precisely that.[9]

Spotify crawls the web to look at blog posts and other online discussions about music, figuring out the kind of descriptive language that listeners use to discuss different songs and artists. These terms then become attributes of the songs. But new or niche songs aren't discussed as much online, so the data this process finds about such songs is insufficient. Spotify compensates for such deficits by using a machine learning algorithms to analyze each song and extract audio characteristics such as tempo, loudness, key, and tonality.[10] Algorithmic approaches that take raw audio data and find interesting patterns to be used as attributes in recommender systems are called feature engineering algorithms. At the end of this feature engineering, Spotify's algorithms have both Last.fm's understanding of music listening patterns and Pandora's deep understanding of the music itself.

The end result is Spotify's Discover Weekly, an algorithmically generated personalized weekly playlist. According to the company, as many as 8,000 artists get over half their streams from users listening to their Discover Weekly playlists.

Recent advances in deep neural networks, a type of machine learning algorithm, now enable automated feature engineering, or feature learning. Beyond just the audio data as mastered by Spotify, deep neural networks can natively handle and process multi-modal data, that is any combination of both structured and unstructured data like text, photo, video, audio, to automatically engineer features and utilize content-based recommendations for any products including images, videos, people, and even firms.[11]

But as with any algorithmic solution, the broader impact of recommender systems on consumer choice needs close attention. It is worth noting that while some of the aforementioned design tradeoffs, such as specific explanations of content or ease of cold starting, are well understood, we know very little about how these different designs affect consumer engagement. This could be a fertile area of research. Specifically, while recent research has shown the positive impact of recommenders on short-term consumer engagement, including product views, purchases, and consumption diversity, the long-term impact of recommenders on consumer engagement and retention calls for more investigation.[12] Furthermore, different types of recommenders have been shown to influence consumption diversity differently. While collaborative filters have been shown to decrease aggregate consumption diversity, hybrid systems do not seem to have the same affect.[13] This difference is important because a recent study from Spotify provides correlational support linking the diversity of consumption to long-term user engagement, including conversion and retention.[14]

Lee and Hosanagar found that recommenders work with other e-commerce features, such as review ratings and descriptions, to influence consumers' purchase decisions. In addition, Karim et al found that focusing only on recommendation accuracy at the expense of other objectives results in the recommendation of harmful content and even negative affects on mental health.[15] Poorly designed recommendation systems, especially in social news feeds, can also create filter bubbles in which consumers are exposed to a narrow set of content instead of a range of perspectives, fueling the fragmentation of society.[16] However, Hosanagar et al analyzed the impact of a content-based music recommender design and did not find evidence that it fragments users. Instead, they found that the design increases users' overlap in consumption. Still, over-representation of subgroups of customers with specific taste could introduce bias into the recommender system, perpetuating feedback loops in the system and giving rise to unfair distribution of attention. For example, the system might only recommend items preferred by the over-represented group at the expense of other subpopulations within the system.[17] This bias can then marginalize or shift the preferences of the affected subpopulations. This erosion occurs because the algorithm's parameters are estimated from the initial starting data, which then influences recommendations with both selection bias and a preference-shifting effect, in which the user's preference is changed simply because the recommendations are made.[18] Deldjoo et al. provides a survey of research on fair recommender systems.[19]

These issues suggest that while recent applications of machine learning address some of the technical roadblocks to recommender design, a broader social science perspective is urgently needed. Incorporating such a perspective will recognize that recommender algorithms cannot be evaluated merely on the accuracy of their recommendations, the proportion of recommended products that users find relevant. Instead, managers should conceptualize customer satisfaction or engagement more holistically, incorporating the recommendations' relevance with their impact on consumer well-being (avoid recommending content which is potentially addictive or otherwise harmful to mental health), social fragmentation (tending to narrow users' perspectives), breadth of consumption (tending to expose users to new topics and product categories), and long-term customer retention.

While recent applications of machine learning address some of the technical roadblocks to recommender design, a broader social science perspective is urgently needed.

Conclusions

Personalized product recommendations help consumers to discover products of interest and sort through the myriad of choices available online. They also help firms to improve customer retention and to cross-sell and upsell products. Despite their value, early recommender designs have many limitations, ranging from the need of content-based recommenders for rich product metadata to collaborative filters' popularity bias, cold-start problems, and inability to explain the details of their recommendations. In recent years, machine learning has been applied to create hybrid systems that combine the best of both approaches, as Spotify has done. The method shows great promise in terms of incorporating new kinds of unstructured data to generate personalized recommendations. With the number of people online continuously increasing[20] and the rapid approach of an AI-augmented creator economy and metaverse, we expect recommender systems to have an even greater impact on consumer choices and engagement in the coming years.[21] However, we caution managers to not just focus on short-term customer engagement metrics, but also to monitor and evaluate long-term customer engagement and the societal impacts of large-scale personalization. Recent advances in recommender systems have demonstrated that it is not necessary to sacrifice the accuracy of recommendations in order to increase their diversity, which could circumvent the filter bubble effect without causing a drop in customer engagement.[22]

Author Bios

Kartik Hosanagar is the John C. Hower Professor of Technology and Digital Business and a Professor of Marketing at The Wharton School. Kartik's research focuses on the digital economy, particularly the impact of analytics and algorithms on consumers and society, Internet media, Internet marketing, and e-commerce.

Kartik was named one of the world's top forty business professors under forty, and has received many teaching awards. He has consulted for Google, American Express, Citi, and more.

Dokyun "DK" Lee is a Kelli Questrom Chair Associate Professor of Information Systems Management and Computing and Data Sciences at Boston University. He studies the responsible application, development, and impact of AI in e-commerce and the digital economy, focusing on the economic impact of textual data and content extraction, understanding, and engineering. His research, which has won several awards, is supported by Adobe, Bosch Institute, Google Cloud, Marketing Science Institute, McKinsey & Co, Nvidia, and Net Institute.

Endnotes

1. MacKenzie, Ian, Chris Meyer, and Steve Noble. "How retailers can keep up with consumers." McKinsey & Company 18 (2013): 1.
2. https://www.businessinsider.com/netflix-recommendation-engine-worth-1-billion-per-year-2016-6
3. Lee, Dokyun, and Kartik Hosanagar. "How do product attributes and reviews moderate the impact of recommender systems through purchase stages?." Management Science 67.1 (2021): 524-546.
4. Wall Street Journal, "Inside TikTok's Highly Secretive Algorithm." July 21, 2021 https://www.wsj.com/video/series/inside-tiktoks-highly-secretive-algorithm/investigation-how-tiktok-algorithm-figures-out-your-deepest-desires/6C0C2040-FF25-4827-8528-2BD6612E3796
5. For a review of early design approaches, see: Adomavicius, Gediminas, and Alexander Tuzhilin. "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions." IEEE transactions on knowledge and data engineering 17.6 (2005): 734-749. A review of more modern deep-learning approaches is available in: and Zhang, Shuai, et al. "Deep learning based recommender system: A survey and new perspectives." ACM Computing Surveys (CSUR) 52.1 (2019): 1-38. To understand the potential problems of the latest recommender models, see: Ferrari Dacrema, Maurizio, Paolo Cremonesi, and Dietmar Jannach. "Are we really making much progress? A worrying analysis of recent neural recommendation approaches." Proceedings of the 13th ACM conference on recommender systems. 2019. Meanwhile, an examination of the latest trends and frontiers in recommender systems may be found in: Jannach, Dietmar, et al. "Recommender systems: Trends and frontiers." AI Magazine (2022).
6. Hosanagar, Kartik. "A Human's Guide to Machine Intelligence: How Algorithms Are Shaping Our Lives and How We Can Stay in Control." New York: Penguin Viking (2019).
7. Breese, John S., David Heckerman, and Carl Kadie. "Empirical analysis of predictive algorithms for collaborative filtering." arXiv preprint arXiv:1301.7363 (2013).
8. Fleder, Daniel, and Kartik Hosanagar. "Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity." Management Science, Vol. 55, No. 5, pp. 697-712 (2009). and Lee, Dokyun, and Kartik Hosanagar. "How do recommender systems affect sales diversity? A cross-category investigation via randomized field experiment." Information Systems Research 30.1 (2019): 239-259.
9. http://benanne.github.io/2014/08/05/spotify-cnns.html
10. https://hackernoon.com/spotifys-discover-weekly-how-machine-learning-finds-your-new-music-19a41ab76efe (october 10, 2017)
11. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521, no. 7553 (2015): 436-444.
12. Kumar, Anuj and Kartik Hosanagar. "Measuring the Value of Recommendation Links on Product Demand." *Information Systems Research*, vol. 30, No. 3, pages 819-838 (2019).
13. Hosanagar, Kartik, Daniel Fleder, Dokyun Lee, and Andreas Buja. "Will the global village fracture into tribes? Recommender systems and their effects on consumer fragmentation." Management Science 60, no. 4 (2014): 805-823.
14. Anderson, Ashton, et al. "Algorithmic effects on the diversity of consumption on Spotify." Proceedings of The Web Conference 2020. 2020.
15. Karim, Fazida, Azeezat A. Oyewande, Lamis F. Abdalla, Reem Chaudhry Ehsanullah, and Safeera Khan. "Social media use and its connection to mental health: a systematic review." Cureus 12, no. 6 (2020).
16. Pariser, Eli. The filter bubble: How the new personalized web is changing what we read and how we think. Penguin, 2011.
17. Chen, Jiawei, et al. "Bias and debias in recommender system: A survey and future directions." arXiv preprint arXiv:2010.03240 (2020).
18. Adomavicius, Gediminas, Jesse C. Bockstedt, Shawn P. Curley, and Jingjing Zhang. "Do recommender systems manipulate consumer preferences? A study of anchoring effects." Information Systems Research 24, no. 4 (2013): 956-975.
19. Deldjoo, Yashar, et al. "A Survey of Research on Fair Recommender Systems." arXiv preprint arXiv:2205.11127 (2022).
20. Statista, "Global digital population", April, 2022, https://www.statista.com/statistics/617136/digital-population-worldwide/
21. Facebook IQ, "Digital Research and Insights How AI Will Elevate the Future of Creativity" October 26, 2021, https://www.facebook.com/business/news/insights/how-ai-will-elevate-the-future-of-creativity?ref=search_new_1
22. Kunaver, Matevž, and Tomaž Požrl. "Diversity in recommender systems–A survey." Knowledge-based systems 123 (2017): 154-162.