The Technologies Used for Artwork Personalization and the Challenges

Zhuoyan Guo(✉)

Behavioral and Social Sciences College, University of Maryland, Maryland 20742, USA
zhuoyan6@terpmail.umd.edu

Abstract. For many years, many streaming companies' personalized recommendation systems mainly focused on how to employ the right algorithm to predict what the subscriber would be interested in based on their viewing history and preferences. By showing those content, companies believe this could provide efficiency to their subscribers and thus their content could reach better performance. However, since it is impossible to present all the details of that content on the homepage, the title shown is unable to contain enough information to trigger the user to click on that. Instead, the artwork which represents the content plays a significant role in the number of clicks the specific content could receive. Although few companies already realized this unprecedented aspect of the personalized recommendation system and started to work on the development of a certain algorithm using A/B testing and contextual bandits approach to improve the system, there are limited research methods that have been employed and they are still facing many challenges. By examining how the A/B testing and contextual bandits approach actually work and the logic behind these two basic research tools and at the same time dealing with those challenges, companies could come up with more comprehensive research designs to better study the users' reactions and inclination when provided with different artworks. Thus, fully understanding where those challenges lay could help companies to be sure about the future development direction of the personalized recommendation field.

Keywords: Artwork personalization · contextual bandits approach · A/B testing · predictive data analysis

1 Introduction

1.1 Background

Because of the overload of information, companies are working on an algorithm that could automatically pick the content or product that would best meet the specific user's interests and motivators to attract more users [1]. This is content personalization. Not only do the users have personal attributes, but so does the content. Content features may include descriptive information, categories, and sometimes keywords. And personalization do is gather and analyzes user attributes, and past and current behavior, then match them with the content features, and then provides the individual best content [2].

S. Kadry et al. (Eds.): BDEIM 2022, AEBMR 233, pp. 230–238, 2023.
https://doi.org/10.2991/978-94-6463-124-1_28

The most up-to-date algorithm used for personalization is "experimental transcoding" which applies techniques and mechanisms to screen content based on users' experience – including the viewing history, information within the account, and frequently browsed keywords – by both understanding and predicting it. Besides, companies could even change how the website looks for each user based on their preference and maybe physical disabilities to improve the experience. One of the essential components of the layout of a webpage is the artwork, and almost all online articles, news, and video are recommended to users with artwork to attract attention. In the past, companies are only able to find one single artwork and use it for all users to maximize content performance. However, with the development of technology and the competition in the markets, companies began to work on finding the best artwork for each subscriber to maximumly improve their experience and gain user loyalty, and the term for this process is called artwork personalization [3, 4]. Artwork personalization is built based on content personalization – an algorithm used to determine the appropriate content for specific users should be employed first, and according to that result another algorithm is used to choose among a set of artwork and determine the one that fits with the user's personal preference and most likely drive the users to click on that content. Since most of the content is only represented by a title and an artwork, it is essential to find the artwork that could best show the content with the trigger driving users to look into that content. In other words, the artwork works as the representation of a title and captures the features that are possible to attract users [5].

For example, on the streaming company's users' homepage, besides the title of the movie and TV show, what is available is the artwork. In the past, the artwork shown is the same for every subscriber, but it is inefficient to attract subscribers to click on the content and begin to watch since not every subscriber would be interested in the same elements shown in that artwork. In order to meet different interests of different users, the artwork may contain the figure of the actor that the users may be able to tell the name, capture an exciting moment or contain a dramatic scene that is predicted would fascinate specific users again based on their viewing history [6]. And the artwork personalization could also be employed on the news website in order to generate more click-rate for the articles.

Thus no matter how advanced the company's algorithm used to find the most suitable content for the users' interests, it won't be efficient enough to increase the overall performance of that website if they could not able to determine the most efficient artwork that conveys the main interest for the users. It says that if the content shown on the homepage could not attract users in 90 s, then they may never come back and provide a chance for that content again [7]. Besides the title, the only thing that could show the users what the content is the artwork, since different users may be interested in a different perspective of that specific content and the title is only able to provide a limited amount of information about the content, so how to employ specific artwork that could best attract the specific group of users determine how well the performance of that website could be.

It is not hard to see that content personalization is the most heated area in the past decade and the technology and system it employed are well-developed at this stage. However, the efficiency and profit that could be brought by artwork personalization could never be replaced by content personalization, that is to say, to improve the overall performance of the website, the importance of artwork cannot be ignored. The existing method and technology that are used to develop artwork personalization is limited and face many challenges, this article is going to provide the overall information about the most efficient methods that have been used to develop artwork personalization as well as the logic behind to point out the benefits, deficiencies, as well as the challenges that are waiting to be solved.

2 Methodology Development

Streaming companies like Netflix know that it is important to catch the attention of the users within 90 s and get them to click on the video content or there would be little chance that they will scroll back [7]. There are already well-developed methods to determine what content should be on the homepage based on subscribers' interests, but these technologies don't guarantee the click-through rate – the ratio of users who click on the link to the number of users who view a page [8] – as there is a possibility that the artwork shown for the content doesn't provide sufficient proof that subscribers should not miss that. Thus, the three goals that have been established while discovering the proper method are: 1) determine the artwork for users that enables them to determine whether the content they want to watch faster; 2) encourage users to watch more in aggregate; 3) don't use the artwork that misrepresents titles as the engineers evaluate multiple images [7]. The following will introduce how the A/B testing and contextual bandits approach help the companies to develop the artwork personalization feature while meeting these three goals.

2.1 A/B Testing

The connectivity between websites and online services enables the evaluation of ideas through controlled experiments. This involves an experiment with an experimental group and a comparative control group, changing only one variable factor in the experimental group, to see how changes in this variable affect the dependent variable [9]. The controlled experiment is also called A/B testing. Simply put, A/B testing helps companies determine if it is better than the current/default setting. The S/B testing pathway works by testing variables (in this setting, a particular artwork) separately in a random subset of members. By employing this, companies can form a hypothesis that whether a specific change introduced to the business, could improve the key metrics. By employing this test in the artwork personalization case, the company splits users randomly between variants (shown different artworks for the same content), then the data which shows subscribers' reaction to the artwork – whether to take a look at the detail -- is recorded to determine the overall engagement rate [10].

Using this method, different experiments could be developed to achieve the goal which is to figure out the most efficient artwork. The most simple setup is to first choose a default image, and the users that be randomly categorized are assigned different images relate to the same title. Then the technology group record the engagement with that specific title which includes the click-through rate, aggregate play/viewing duration, fraction of plays/views with short duration, and fraction of content viewed which measured how long the users have stayed at the webpage [7]. After fully evaluating those data, the technology group will decide which piece of artwork leads to the best engagement rate for a specific group of subscribers who share similar characteristics. And subscribers are usually grouped by their preferences and interests based on their viewing history and account information.

Though this single-cell experiment could help the company determine which piece of artwork works the best, it is time-consuming as developers need to wait until all of the results came out, then aggregate the data and make the final decision. After the release of the new content, without the best representative artwork to attract users, the company is losing a significant amount of potential subscribers every minute, so it is not efficient to employ the single-cell experiment. It works well with a small set of content, but apparently, it is not the case for many companies that generate tons of new content every day. This is when the multi-cell explore-exploit test came into the picture. The basic difference between this experiment and the single-cell experiment is it enables the company to run with a remarkably larger set of titles at a time. This test is constructed with two parts: the "explore" test which measured the engagement of each variant for a group of contents and the "exploit" test which determines the artwork with the highest engagement rate for future users and sees whether it could increase the aggregate viewing hours [7]. By employing this test, companies could obtain the measurement of performance with different artworks representing each title at the same time and immediately test how those most engaging artwork work.

While the experiments mentioned above are successful, still there are faster ways to learn the performance of a specific artwork. And the core idea of artwork personalization is how to find the most suitable artwork for different subscribers in the shortest amount of time to maximize the profit that the company can generate. Before a technological group can confidently decide on the greatest artwork for every user, it is desirable to draw the fewest amount of randomly picked members and receive the outcome as soon as possible [7].

2.2 Contextual Bandits Approach

Although the advanced experiment using A/B testing could bring the company the useful information that they want, inevitably those tests require a lot of time. As mentioned above, after the release of the new content, the longer the time it takes for the test to figure out which artwork works the best, the more potential viewers the company is losing. Besides, the web service today is featured with dynamically changing pools of content and the content popularity also changes over time, thus it requires the system to consistently inject the new emerging data into the consideration to reach the best outcome. And the contextual bandit's approach is more about collecting the users' real-time feedback to evaluate a variant and simultaneously adapting content personalization based on the contextual information of the users and contents as an alternative to simply predicting what works the best based on the group of randomly selected users' data [2].

Rather than running groups of the test, contextual bandit asks for the injection of controlled randomization in the well-established model. The complexity of the randomization schemes can range from closed-loop systems that adaptively change the level of randomization in response to model uncertainty to uniform randomness. And this process is referred to as data exploration. The data exploration strategy is determined by the number of candidate artworks in total for specific content and the size of the testing subscribers. Recording the degree of randomness for each artwork pick is crucial during this data exploration phase. The firms may be able to compensate for skewed selection propensities and afterward conduct the impartial offline model evaluation by later evaluating those data [6].

With a multi-armed bandit formulation, the set of actions the users perform and the features that be taken into account may contain imprecision and the result generated could be suboptimal, thus people may argue that the artwork chosen leads to the cost of accuracy. When there is a really wide user base, the regret is typically not as great and is spread out among a lot of members, all of whom are providing reactions to the artwork in a continual manner. This exploration process can increase short-term regret since it cannot avoid suboptimal arms being chosen, but with more information on arms obtained and refine the estimate from the large base of users, the price of exploration is quite low per user. When deciding whether to use contextual bandits, controlling the regret is a crucial factor, and that is why contextual-based bandit is less suitable if exploration costs a lot.

To employ this online machine learning setting, the business may also train its contextual bandit model to choose each member's best piece of art based just on the content. Companies like Netflix may simply score pictures for a member independently across titles to assess the personalization process since they often have up to a few dozen possible pieces of art per video. Additionally, this sophisticated model may be used to forecast each (member, title, and image) tuple and establish the likelihood that the member would take a look at the content [6].

2.3 Related Research

2.3.1 Differences Between the Two Approaches

When the artwork personalization concept was first introduced, the method employed is A/B testing, users are persistently separated at random for the variants (in the event of artwork personalization, various artwork depicts the same title) in a persistent manner (each user receives the same webpage layout in multiple visits) [11]. Key metrics are recorded following the computation based on the instrumented subscribers' interactions with the website, and in this case, the metrics should be the engagement rate. Simply deciding which artwork is shown to the specific group of subscribers who share similar attributes provides the company with the highest amount of viewing, the company would decide to employ which artwork to generate the most profit out of content. Though this is already a great improvement from the past when companies are only able to determine which artwork should be employed for all of the users to maximize the profit, with the contextual bandits approach the company can do more.

While the data collection step of A/B testing is time-consuming, the contextual bandits approach directly inject the up-to-date data into the predicted model with varying degree of randomization. As a result, without running the entire algorithm beforehand, the contextual bandit's technique may quickly determine the best-tailored artwork selection for a title for each subscriber [12]. To put it simply, with the contextual bandit, a learning algorithm can test out different actions with the newly emerging data and automatically learn which action has the most rewarding outcome, then employ that directly without waiting for the testing result and analyzing the conclusion. By employing the contextual bandit's approach, the company could not only determine which artwork candidate could attract the subscribers the most, but they can also achieve the ultimate goal of artwork personification by indicating which candidate works the best for a specific user to maximize the click-through rate.

2.3.2 Pitfalls of A/B Testing

Besides what have mentioned earlier in the paper the A/B testing could take a long time to finish running the whole bunch of testing and provide a comprehensive testing result, it also faces the undetermined experiment length problem. The more data that has been used to run the test, the better the result to predict the reality, but how can a company determine how much data they need to collect before running the test? And the variation in the test length might negatively affect the results.

Second, since the recommendation system should highly fit the specific users' personal preferences to create the most value, in reality, in many cases the system may not have a sufficient amount of data about the group of users due to the cold-start problem [13]. Thus they may have an insufficient amount of input for the company to run a test and make a credit decision based on the result.

Third, A/B testing is mainly about building a causal relationship between the variant and the measurable change in users' behaviors. It is essential to maintain the premise that alternations are well controlled by randomization. However, artwork personalization specifically, focuses on identifying the differences between users and providing different artwork for them. In addition, some artwork may have better performance for a part of the population than for another part, thus generating differences in the population cannot always be weakened by the randomization of the test [14].

2.3.3 Pitfalls of Contextual Bandits Approach

Although the contextual bandit approach saves a lot of time for the company to determine which artwork to provide to the specific users, this efficiency is costed by the need to collect training data for a model without rejection. Since the attributes and actions that have been captured may unable to best represent a specific user, the result may be suboptimal. Although the regret could be amortized as time passes, the cost of obtaining and gathering those data is significant [6]. People should be clear that although this approach saves time and cost in many situations, it is costly under specific setups [15].

2.3.4 Other Existing Approach

Besides A/B testing and the contextual bandit's approach, there are not yet no other well-developed methods that could be used for artwork personalization because the logic behind the personalization is well-fitted into these two approaches.

3 Challenges in this Field

No matter whether use A/B testing or a contextual bandit approach, one essential step is to collect data. Businesses must gather feedback data consistently across devices, at scale, and over time, which presents a barrier for them to develop an effective system to better assist with subsequent data processing, and the system not only requires a well-established structure but also a remarkable huge space, which is itself already be challenging. Besides these technical challenges, there are still other challenges that impede the artwork personification become more accurate.

First, as there is only one selected piece of artwork that will be shown on the homepage, and whether the subscribers going to click on it or not entirely depends on this single piece of artwork, this means that if a subscriber plays a specific video it only depends on the image that the algorithm decided to present to that member. Thus the core problem the company needs to figure out is under what situation the user clicks because of the attraction of the artwork and under what situation the user clicks on the content regardless of which image they presented. And to properly determine this process, companies must gather a lot of data to identify the signals that reveal the circumstances in which a work of art has a major influence on consumers' actions.

Second, it is important to understand the effect of changing artwork that is shown to a member for a title that represents a series of stories or seasons of TV shows. It is hard to measure whether changing artwork to reduce the recognizability of the title would make people unable to be attracted by the title again or will it oppositely impact which is to lead people to reconsider because of the different elements contained in the artwork. Intuitively people may think continuous changes could confuse users, but it cannot be denied that changes provide people with refreshments. It is therefore unclear which artwork prompts a user to be attracted to a certain title due to this conflict.

Another difficulty to determine the effectiveness of a piece of artwork is how understanding the relationship between a specific artwork and other artwork the algorithm selected for a member on the same page or session. Since the page is shown as a whole to the users, it may not be enough to only look at each piece of artwork individually, and to achieve a better outcome, companies should think about how to arrange a variety of artworks across titles on a single page. Since the feelings of the user are so abstract, it is hard to find the determinants and thus hard to get a measure. Besides, there are other factors that could also influence the performance of a single artwork which are hard to measure. For example, the performance may also relate to other types of evidence and assets, such as synopses, trailers, font and so on which could represent the content of a title. Since the process of how people make decisions is complicated, it is really a challenge for the company to analyze the effect comprehensively.

Besides, how to create a high-quality group of artwork for every content is also a challenge for companies. First of all, each title's pool of artwork must be diversified enough to appeal to a wide range of possible audiences since they would be attracted by only specific aspects of content. Although this could be helped with the algorithm that analyzes the different subscribers' viewing history in order to see what element could attract them the most how to create artwork that efficiently triggers their interest is not as simple as how interesting and instructive a piece of art is really dependent how the person seeing it, which is difficult to count and difficult to evaluate.

4 Conclusion

Since people can't get the whole idea of the content just through the title, it is particularly necessary to build an intelligent algorithm for the personalization of artwork. Art personalization algorithms can better enhance the subscriber experience and deliver content more efficiently. To improve the performance of the content, it is essential to provide users with artwork that could better attract their attention with the visual elements. Since the basic logic behind this personalization is well-developed, for example, the A/B testing and the contextual bandit approach, there are many available alternative ways to carry out specific testing for different goals. By paying attention to the pitfalls of each approach, the further development prospect in the artwork personalization field is very broad and countless possibilities could be achieved in the near future. As there are still many challenges faced by the companies because this is a newly emerged field, solving the data collection limitation and studying how people react to the features of artwork could help companies to achieve better performance with their business.

References

1. Ferretti S, Mirri S, Prandi C, & Salomoni P.: On personalizing web content through reinforcement learning. Universal Access in the Information Society, International Journal, Vol. 16(2), 2017, pp. 395–410. https://doi.org/10.1007/s10209-016-0463-2
2. Li L, Chu W, Langford J, & Schapire R E.: A contextual-bandit approach to personalized news article recommendation. In Proceedings of the 19th international conference on World wide web, 2010, April, pp. 661–670.
3. Mirri S, Prandi C, & Salomoni P.: (2013, October). Experiential adaptation to provide user-centered web content personalization. In Proc. IARIA Conference on Advances in Human oriented and Personalized Mechanisms, Technologies, and Services (CENTRIC2013) (pp. 31–36).
4. Fouladirad M, Neal J, Ituarte J V, Alexander J, & Ghareeb A.: Entertaining data: business analytics and Netflix. Int J Data Anal Inf Syst, Vol. 10(1), 2018, pp. 13-22.
5. Chandrashekar, A. C. (2018, September 12). Netflixartwork personalization through multi-armed bandit testing. Reforge. Retrieved September 16, 2022, from https://www.reforge.com/brief/netflix-artwork-personalization-through-multi-armed-bandit-testing#-I84JSKg8bhzPl5wTnn3Ug
6. Chandrashekar, A., Amat, F., Basilico, J., & Jebara, T. (2017, December 7). Artwork Personalization at Netflix. Retrieved October 31, 2018, from The Netflix Tech Blog: https://medium.com/netflixtechblog/artwork-personalization-c589f074ad76
7. Gopal Krishnan (2017, April 19). Selecting the Best Artwork for Videos Through A/B Testing. Retrieved from The Netflix Tech Blog: https://netflixtechblog.com/selecting-the-best-artwork-for-videos-through-a-b-testing-f6155c4595f6
8. Wikimedia Foundation. (2022, August 23). Click-through rate. Wikipedia. Retrieved September 16, 2022, from https://en.wikipedia.org/wiki/Click-through_rate
9. Merriam-Webster. (n.d.). Controlled experiment definition & meaning. Merriam-Webster. Retrieved September 16, 2022, from https://www.merriam-webster.com/dictionary/controlled%20experiment#:~:text=Definition%20of%20controlled%20experiment,that%20is%20changed%20or%20altered%20%E2%80%A6
10. Kohavi R, Longbotham R. Online Controlled Experiments and A/B Testing. Encyclopedia of machine learning and data mining, Vol. 7(8), 2017, pp. 922-929.

11. Kohavi, R., & Thomke, S. (2017). The surprising power of online experiments: getting the most out of a/b and other controlled tests. Harvard Business Review, 2017(September-october).
12. Pilani A, Mathur K, Agrawald H, Chandola D, Tikkiwal V A, & Kumar A.: Contextual bandit approach-based recommendation system for personalized web-based services. Applied Artificial Intelligence, Vol. 35(7), 2021, pp. 489-504.
13. Mary, J., Gaudel, R., & Philippe, P. (2014). Bandits warm-up cold recommender systems. arXiv preprint arXiv:1407.2806.
14. Esteller-Cucala, M., Fernandez, V., & Villuendas, D. (2019, June). Experimentation pitfalls to avoid in A/B testing for online personalization. In Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization (pp. 153–159).
15. Agarwal, A., Bird, S., Cozowicz, M., Hoang, L., Langford, J., Lee, S., ... & Slivkins, A. (2016). Making contextual decisions with low technical debt. arXiv preprint arXiv:1606.03966.