Ethical Considerations and Societal Impact of Personalized Advertising Algorithms

Anirudh Parasrampuria[1] and Katherine Williams[#]

[1]Mission San Jose High School, USA
[#]Advisor

ABSTRACT

Data mining and algorithmic decisions making have reshaped marketing by allowing businesses to present products best applicable to the user in the shape of personalized advertisements. Using algorithms with pre-coded biases, companies can exploit data in order to increase revenue. Companies will present inferior products or services to certain demographics with a low-household income or higher quality products to high income neighborhoods. This is problematic as it not only violates ethics but prolongs discrimination in marginalized communities. Indeed, businesses over the years have targeted specific audiences based on data gathered to recommend products or even sway a presidential election. However, in recent years, there has been a rise in government regulations and various regulatory frameworks. The European Union has forced companies to be more transparent about their data mining technologies. In addition, there are various ways companies can increase their own integrity by enforcing regulation within the company itself. This paper analyzes the ethics of personalized advertisements and its impact on society; it also investigates specific companies and highlights possible solutions to the problem.

Introduction

It is no secret that big data is king in the world that we live in today. It is a piece of technology that impacts virtually all aspects of the economy from traffic and navigation to advertisements (M. S., 2023). Currently, data mining is vital to Fortune 500 companies and big tech, including Google, Facebook, Amazon, and Microsoft. These companies rely on business models where data is a key component of their value propositions. Their methods allow them to attract and retain users using their platforms and increase revenue. For example, Amazon collects data on their users to learn and enhance the product offerings seen on their page. Microsoft Azure is one of the fastest growing products and revenue sources. They offer data analysis and data storage which can help in advertising, a key tactic for increasing market share and setting business apart from other competitors (Lorenzo et al., 2020). In recent years, personalized advertisements have risen dramatically because of data collection and algorithmic decision making, which has led to the use of data to specifically target interested customers (*Data Mining for Advertising,* 2022). Although this saves time for the users and prevents them from having to surf through the Internet to find products of their choice, personalized advertisements raise a variety of concerns. For example, it has shown to have ethical concerns, discriminatory effects, and has been proved to contain certain biases.

This paper will examine how customer data is exploited through algorithms and machine learning, and the impact this technology has on society. Ethical considerations are important in examining the impact as well as discovering possible frameworks companies and the government could use to address the issues. The following questions will be addressed:

1. What are some ethical concerns in using data mining and algorithmic decision making for personalized advertisements and how can the concerns be mitigated to maintain a high level of responsibility and integrity.
2. What are the effects of algorithmic biases in personalized advertising, and how might these biases reinforce societal injustices or discriminate against marginalized communities?

This paper will discuss the definition of personalized advertisements and algorithms used before moving on to the effects of this technology on specific demographics. Then it goes on to the idea of ethics and maintaining integrity to have this technology benefit society. Towards the end, case studies regarding companies are considered and possible regulatory frameworks that could be used to mitigate the harm this technology could present.

Personalized Advertising Algorithms

Advertising has been around for generations as a way for companies to promote their product and attract new customers. However, personalized advertising changes the entire idea of advertising itself as the content provided would be specific and relevant to the customer. For example, consumers would get appropriate products, services, and communicators (Instapage, 2023; Strycharz et al., 2019). Personalization can be described as a technology with positive and negative concerns. On one hand, users would decrease their time spent on the Internet by not having to browse through the Internet to find products or services applicable to their needs. In addition, research has shown that it would result in financial benefits through personalized discount coupons. On the other hand, personalized advertisements do have some concerns as well. These mainly involve discrimination and privacy of the user. In order to have personalized advertisements, companies must collect data from the user, which may be considered an invasion of privacy, in order to display targeted ads for them (Instapage, 2023; Setyani et al., 2019). The idea that personal data is being collected and used for personalization may make users feel targeted or maybe suppress their own autonomy.

Data mining is known as the activity and process of information collection from various data sources. These sources contain a mass amount of data on the user. Large databases, including CRMs, e-commerce sites, social media sites like LinkedIn, and a variety of other sources can all yield valuable data (*How is Data Mining Used in Marketing,* n.d.). Once all the data is gathered, it is then analyzed using algorithms and analytical methods. Once usable, it is compiled into easy-to-understand graphs. Meanwhile, algorithms related to advertisements would make sure that the data can be used widely by companies to promote products or ideas (*How is Data Mining Used in Marketing,* n.d.).

A data mining algorithm is a collection of computations and techniques used to build data mining models with the use of data. However, a difficulty persists with finding the algorithm and model best suited for the data and solving a certain problem. Even though one may utilize many algorithms to complete the same objectives, each algorithm produces a unique set of outcomes (Joseph et al., 2016). There are three popular algorithms that are used in data mining. First is the K-mean Algorithm, a common clustering algorithm. K-means is an unsupervised learning algorithm according to typical implementations since it discovers the cluster on its own without the use of outside data. Although it may not be assured, group members will likely be more like one another than non-group members (upGrad, 2022; Pedamkar, 2023). Second, is the Naive Bayes Algorithm which contains multiple algorithms used for categorization and is based on estimations. It assumes that every parameter is indeed independent, and using this it can determine the data's classification (Pedamkar, 2023; upGrad, 2022). Third is the ANN Method. This is the most common and it uses estimation models on mass amounts of data to discover commonalities within the dataset. This system can be described as multiple operating parts that work together to take a value and produce another value (Pedamkar, 2023). It works by taking in a specific input and producing a relevant output with hidden layers in between. These hidden layers apply a weight to the input which then results in the specific output. However, it is these layers that lead to

discrimination as they are developed specifically to target certain demographics harmfully based on the input value (DeepAI, 2019; Dertat, 2017).
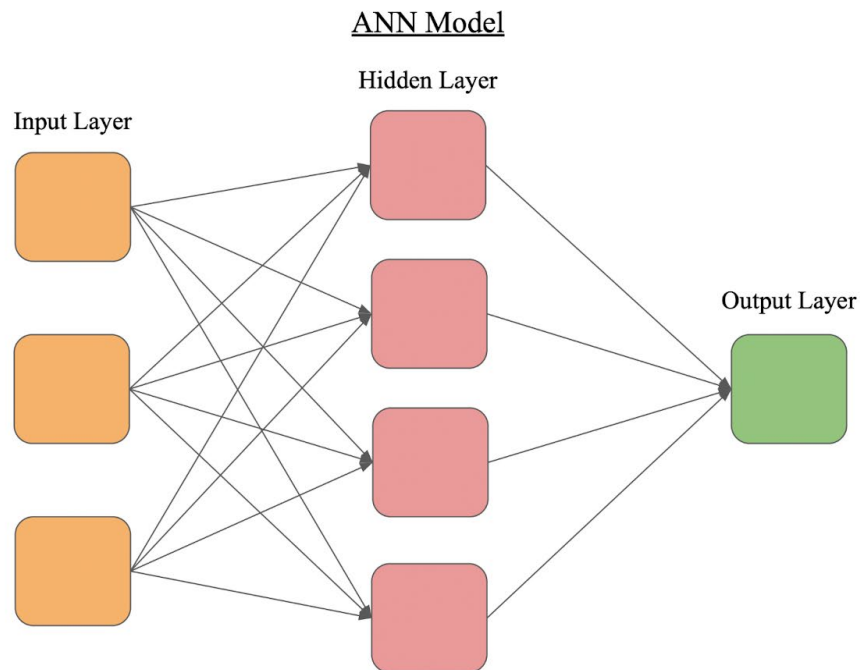
Figure 1. This is a basic visual representation for how the ANN method works.

In personalization specifically, two types of algorithms are used to display relevant information. First, are the segmentation-based algorithms that categorize users based on their location. Then, using the demographics or characteristics of that location, they can target users on specific attributes like age, race, and income (Invisibly, 2023). Second is the product recommendation algorithm which is based on customer data to offer the best products suited for other customers. Through these methods, companies are able to display products that are best suited for each user. Examples of this are Netflix offering something unique for each user to make the account more personalized or Amazon using machine learning to estimate and display the best products for its customers (Invisibly, 2023).

Societal Impact

Data analytics have reshaped many industries as part of our economy. This new system can disrupt civil rights protections and change the way personal information is used in various aspects including loans, medicine, education, and the workforce. Data mining is a rational and statistical form of discrimination. Its purpose is to correctly assign to the person the characteristics held by those who are statistically similar, while also providing a reasonable foundation for differentiating across people (Gerlick & Liozu, 2020). There have been many theories for why data mining discrimination exists however there have been few papers to narrow down the exact reason for discrimination. However, two theories have been getting support. First, there could already be hints of bias and discriminatory consequences on disadvantaged groups in the data that has been inputted into the system. The second explanation is that bias was pre-programmed into the algorithm by its creators. In any case, intolerance and discrimination would show itself in some way. This is problematic as it gives way to various

decision-making models used in data mining that could have an adverse impact on marginalized communities and protected classes (Barocas et al., 2016; Gironda & Korgaonkar, 2018).

The use of algorithmic decision making in advertisement has raised concerns of possible discrimination based on the user's geographic area. Decision making models factor in the region's demographics, average wages, and region's characteristics to offer the best recommended products. However, some groups may be presented with lower quality products; this would arise from business using the algorithms to judge a regions' characteristics and offer the best product they see fit. Although there are a few platforms that are building policies to prevent this from happening, advertisers will still try to find loopholes and evade the polices. In order to put this to an end, we would need stronger enforcement of such policies (Ali, 2021).

Another problem is the way ads are delivered, which requires the advertiser to pick the best candidates from a wider audience. This issue could be the results of bias implemented into an algorithm. This could also spread misinformation targeted to change the mindset of people in a region to influence a political agenda or campaign for an official running for office (Ali, 2021; Häußler, 2021).

However, there have been recent developments in recommender systems research and, more broadly, machine learning techniques for personalization. These are frequently the driving forces behind these ad delivery algorithms and advertising systems in general. In order to mitigate the effects, it is important for data systems to account for user harm and discrimination as they are being built. On websites like Facebook and Twitter, where advertisements are included into browsing feeds and the line between organic material and advertising is essentially blurred, the potential for such harm is particularly highlighted. These platforms combine several customization mechanisms, including feed classification, topic selection, and commercial auctions. Therefore, it is important to understand the potential harms of such algorithms to society and mitigate them during system construction (Ali, 2021).

Ethical Considerations

When it comes to advertisements in general, there are a multitude of ethical considerations that may be violated when trying to promote a product or service. For example, many businesses use images which lack empirical verification. Images do not necessarily contain text or words which means they cannot be checked for statements such as false promises, misinformation, or political promises (Borgerson et al., 2007).

Furthermore, various ethical frameworks play a key role when examining the ethics of personalized advertisements and data mining specifically. When four different frameworks were examined, Data Ethics Decision Aid, Digital Decisions Tool, Data Ethics Canvas, and Data Ethics Workbook, the results showed the same parameters to be present in all the frameworks. However, the stress of each parameter varied through the different frameworks. The welfare, justice, accountability, respect for privacy, transparency, and autonomy were the most essential ideals emphasized (Häußler, 2021; Borgerson et al., 2007).

Table 1. Important Ethical Consideration Presented in Frameworks

Ethical Considerations	Definition
Welfare	Welfare refers to the advantages the user would receive through personalized advertisements like viewing recommended products. This is the key factor to be considered during system development and should be the primary reason for creating the technology.

Justice	Justice refers to equality in the treatment of customers. The data model should avoid discrimination, bias, and recommending inferior products to certain demographics. For example, companies would target specific populations to promote specific products. This also raises the idea of whether or not the intentions behind personalized advertisements are good. Some may say companies use it to exploit users and increase profits while others say it is built to aid the users by presenting specific products best fit for them (Häußler, 2021; *The Ethics of Data Mining,* 2022).
Accountability	Accountability refers to responsibility such that companies take accountability for negative outcomes and wrongdoing (Häußler, 2021).
Privacy	Privacy refers to the protection of personal information. It is about adhering to laws, such as the Europeans General Data Protection Regulation, not purposefully collecting information private to the user. Having a key distinction between what and what not to collect is critical is maintaining ethics (Tucker, 2014).
Transparency	Transparency refers to companies releasing methods of data models, methodology, and code used for personalized advertisements. It is vital to maintain communication between companies and users. Furthermore, it is important for companies to ask for permission in data collection before mining for data (Häußler, 2021; *The Ethics of Data Mining,* 2022).
Autonomy	Autonomy refers to pursuing your own thoughts while not being purposely manipulated through personalized advertisements. Self-consciousness should still be present in order to guide the user from right and wrong. Companies may limit customers' knowledge on how data mining and personalized advertisements work, which makes them vulnerable to being exploited (Chen et al., 2019; Häußler, 2021; Nwachukwu et al., 1997; Stephen, 2016).

Overall, it is important to consider ethical considerations, such as, transparency, communication, and respecting human privacy and autonomy, when collecting and using personal data for targeted marketing. Companies must endeavor to gain the trust of consumers by being open and honest about the ways in which they gather and utilize personal data, as well as by giving customers the information they need to make well-informed choices (*The Ethics of Data Mining,* 2022).

Case Studies

Cambridge Analytica Scandal

Founded in 2013, Cambridge Analytics is a consultancy group revolving around politics. They use personalized advertisements and data mining to trick users into voting for a certain candidate, such as Ted Cruz or Donald Trump (Boerboom, 2020). The company obtained access to personal data through Facebook. Developers originally designed the app as a personality test, where users answer a certain number of questions to gauge their personality. However, in order to access the quiz, the user must login, and most did throughout their Facebook account, thus giving Cambridge Analytica access to their account ID. This yields various qualities about the user such as location, interests, gender, etc. In addition, Facebook gave Cambridge Analytica access to certain information for research purposes. Using this information, the company used machine learning and algorithms

to manipulate a user into voting for a particular candidate (Boerboom, 2020; Schneble et al. 2018; Confessore, 2018).

This scandal clearly violated the ethical consideration. There was a lack of communication between Cambridge Analytica and its users when the company failed to disclose how it truly obtained access to information and their data mining models. This also violated the idea of privacy where users should have the right to keep personal information private, as well as autonomy, since users did not have access to proper information and were not able to make thoughtful decisions.

Target Scandal

Target used a new way of customer tracking and analytics that was discovered once a pregnant teen girl started receiving maternity related teen advertisements before she told her close ones that she was pregnant. With the use of algorithms, predictive analytics, and machine learning, Target identified 25 products that are usually purchased together when someone is pregnant. Using such a model, Target can display the best recommended products for the user and enhance the user experience while also using psychological strategies to influence consumer spending (Lubin, 2012). This scandal as well is in violation of the ethical considerations. It invades privacy when Target gets personal details about a user's life without their consent. They determine that a user is pregnant and prompt products to boost revenue at the expense of user privacy and relationships. Furthermore, this violates transparency and communication as Target was not entirely clear on its analytical methods and failed to disclose it to the user (Kuhn, 2023; Hill, 2022).

Overall, these scandals are an example of wrongdoing. They both obtain access to personal data through wrongdoing and violate simple ethics in order to increase profits. They fail to communicate to users and extort personal information for their own benefit. However, these misdoings have been identified and various governments have been doing more to enforce the ethical considerations in personal advertisements.

Regulatory Policies

In recent years, new laws have been enforced around the world to protect privacy for the users. For example, in Europe, the new General Data Protection Regulation and in California, the California Consumer Privacy Act, force companies to communicate with users about their data mining methods and how they use it. Consumer protection against unethical targeted marketing tactics can be greatly aided by government legislation. Legislation provides companies with a clear pathway to follow, preventing any gray lines for which companies may exploit. This ensures users privacy and makes sure data is collected and used ethically. More importantly, it is important to create a culture. Instead of constantly keeping certain information private, we need to create a new normal of increasing transparency between companies and users.

Having a distinct board with the sole purpose of maintaining ethics with enhancing data policies will aid in solving problems pertaining to ethics. Various department heads can together determine the effects of various data mining models and methods on society and the company. The board would also be responsible for increasing transparency and enforcing rules on the company.

Accountability should be enforced by having a specific person in the lead of methods, algorithms, policies, etc. and having a specific chain of command. When an issue arises related to ethics or discrimination, the lead would be held accountable. Managers won't disobey ethics out of fear of repercussions, which stops problems from developing in the first place (Kirk, 2022).

Furthermore, companies should prioritize users over profit. The purpose of data mining technology is to aid the user by showing recommended products and services, making it easier to find necessities. However, this technology can also be used to access data unethically and display ads to increase revenue. Companies

should change the ideology that data mining's purpose is to improve customer relations, not to increase revenue (Raitaluoto, 2023).

Conclusion

Personalized advertisements are on track to change the landscape of advertisements. It will impact how companies and users work together and it can affect society in a great way. With the rise of new technology, such as artificial intelligence and machine learning, personalized advertisements are becoming more specific to various users. Personalized advertisements may impact customer loyalty by using data mining models to predict behaviors of users. However, this technology may be harmful if it is used to target vulnerable communities and prolong stereotypes. This technology possesses the ability to disrupt the way humans think by suppressing thought to manipulate users into purchasing goods or services. This technology can benefit society as well through offering discounts or even save time for the users by offering recommended products, resulting in less time spent in Internet browsing. However, this technology has been used in the wrong ways before but with new government involvement and regulatory systems, the harms of this technology will be mitigated, raising customer confidence in personalized advertisements.

Acknowledgments

References

Ali, M. (2021, September 13). *Measuring and mitigating bias and harm in personalized advertising*. Northeastern University. https://ccs.neu.edu/~mali/papers/recsys21-26.pdf

Barocas, Solon and Barocas, Solon and Selbst, Andrew D., Big Data's Disparate Impact (2016). 104 California Law Review 671 (2016), Available at SSRN: https://ssrn.com/abstract=2477899 or http://dx.doi.org/10.2139/ssrn.2477899

Boerboom, C. (2020). Cambridge Analytica: The Scandal on Data Privacy. In *Augustana Center for the Study of Ethics Essay Contest* (pp. xx-xx). Retrieved from https://digitalcommons.augustana.edu/ethicscontest/18

Borgerson, Janet L. and Schroeder, Jonathan E., Identity in Marketing Communications: An Ethics of Visual Representation. MARKETING COMMUNICATION: NEW APPROACHES, TECHNOLOGIES, AND STYLES, Allan J. Kimmel, ed., pp. 256-277, Oxford University Press, 2005 , Available at SSRN: https://ssrn.com/abstract=969079

Chen, Q., Feng, Y., Liu, L., & Tian, X. (2019). Understanding consumers' reactance of online personalized advertising: A new scheme of rational choice from a perspective of negative effects. International Journal of Information Management, 44, 53-64. ISSN 0268-4012. https://doi.org/10.1016/j.ijinfomgt.2018.09.001.

Confessore, N. (2018, April 4). *Cambridge Analytica and Facebook: The scandal and the fallout so far*. The New York Times. https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

Data Mining for Advertising. (2022, October 28). Edubirdie. Retrieved August 17, 2023, from https://edubirdie.com/examples/data-mining-for-advertising/

DeepAI. (2019, May 17). *Hidden layer*. DeepAI. https://deepai.org/machine-learning-glossary-and-terms/hidden-layer-machine-

learning#:~:text=In%20neural%20networks%2C%20a%20hidden,inputs%20entered%20into%20the%20network.

Dertat, A. (2017, October 9). *Applied deep learning - part 1: Artificial Neural Networks*. Medium. https://towardsdatascience.com/applied-deep-learning-part-1-artificial-neural-networks-d7834f67a4f6

The Ethics of Data Mining. Texas A&M International University Online. (2022, November 23). https://online.tamiu.edu/articles/information-science/ethics-of-data-mining.aspx#:~:text=Ethical%20Concerns%20in%20Data%20Mining&text=Transparency%3A%20Customers%20should%20have%20a,for%20forgiveness%20after%20the%20fact

Gironda, J. T., & Korgaonkar, P. K. (2018). iSpy? Tailored versus Invasive Ads and Consumers' Perceptions of Personalized Advertising. Electronic Commerce Research and Applications, 29, 64-77. ISSN 1567-4223. https://doi.org/10.1016/j.elerap.2018.03.007.

Häußler, H. (2021). The Underlying Values of Data Ethics Frameworks: A Critical Analysis of Discourses and Power Structures. *Libri*, *71*(4), 307-319. https://doi.org/10.1515/libri-2021-0095

Hill, K. (2022, October 12). *How target figured out a teen girl was pregnant before her father did*. Forbes. https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/?sh=7f7413c76668

How is data mining used in marketing: Comptia. Default. (n.d.). https://www.comptia.org/content/articles/how-is-data-mining-used-in-marketing#:~:text=Data%20mining%20makes%20it%20possible,fuels%20data%20mining%20in%20marketing

Instapage. (2023, January 1). *Advertising personalization and landing pages: The next wave in digital marketing*. Instapage. https://instapage.com/blog/what-is-personalized-advertising/#:~:text=Personalized%20advertising%20is%20the%20act,intent%2C%20and%20even%20behavioral%20patterns

Joseph, S. R., Hlomani, H., & Letsholo, K. (2016, April). *Data Mining Algorithms: An overview - researchgate*. ResearchGate. https://www.researchgate.net/profile/Sethunya-Joseph/publication/309211028_Data_Mining_Algorithms_An_Overview/links/5805e8e008aeb85ac85e3708/Data-Mining-Algorithms-An-Overview.pdf

Kirk, E. (2022, November 10). *Council post: Why data privacy should matter to advertisers*. Forbes. https://www.forbes.com/sites/forbesbusinessdevelopmentcouncil/2022/11/09/why-data-privacy-should-matter-to-advertisers/?sh=36cdb5654e16

Kuhn, G. (2023, January 8). *How target used data analytics to predict pregnancies*. Market Research Company New York. https://www.driveresearch.com/market-research-company-blog/how-target-used-data-analytics-to-predict-pregnancies/

Invisibly. (2023, February 2). *Personalization algorithms: Why it matters and how it impacts people*. Invisibly. https://www.invisibly.com/learn-blog/personalization-algorithms/

Lorenzo, Patricia and Padilla, Jorge and Requejo, Alejandro, Consumer Preferences for Personal Data Protection in Social Networks: A Choice Modelling Exercise (October 21, 2020). Available at SSRN: https://ssrn.com/abstract=3716206 or http://dx.doi.org/10.2139/ssrn.3716206

Lubin, G. (2012, February 16). *The incredible story of how target exposed a teen girl's pregnancy*. Business Insider. https://www.businessinsider.com/the-incredible-story-of-how-target-exposed-a-teen-girls-pregnancy-2012-2#:~:text=Pole%20identified%2025%20products%20that,Pole%27s%20formula%20was%20a%20beast

Nwachukwu, S. L. S., Vitell, S. J., Gilbert, F. W., & Barnes, J. H. (1997). Ethics and Social Responsibility in Marketing: An Examination of the Ethical Evaluation of Advertising Strategies. Journal of Business Research, 39(2), 107-118. ISSN 0148-2963. https://doi.org/10.1016/S0148-2963(96)00146-4.

Setyani, V., Zhu, Y.-Q., Hidayanto, A. N., Sandhyaduhita, P. I., & Hsiao, B. (2019). Exploring the psychological mechanisms from personalized advertisements to urge to buy impulsively on social media. International Journal of Information Management, 48, 96-107. ISSN 0268-4012. https://doi.org/10.1016/j.ijinfomgt.2019.01.007.

Stephen, A. T. (2016). The role of digital and social media marketing in consumer behavior. Current Opinion in Psychology, 10, 17-21. ISSN 2352-250X. https://doi.org/10.1016/j.copsyc.2015.10.016.

Tucker, C. E. (2014). Social Networks, Personalized Advertising, and Privacy Controls. Journal of Marketing Research, 51(5), 546–562. https://doi.org/10.1509/jmr.10.0355

M, S. (2023, July 27). *14 big data examples & applications across industries*. Simplilearn.com. https://www.simplilearn.com/tutorials/big-data-tutorial/big-data-applications

Pedamkar, P. (2023, March 23). *Models in data mining: Techniques: Algorithms: Types*. EDUCBA. https://www.educba.com/models-in-data-mining/

Raitaluoto, T. (2023, May 11). *The ethical considerations of Personalized Marketing*. Markettailor. https://www.markettailor.io/blog/ethical-considerations-of-personalized-marketing

Schneble, C. O., Elger, B. S., & Shaw, D. (2018). The Cambridge Analytica affair and Internet-mediated research. EMBO reports, 19(8), e46579. https://doi.org/10.15252/embr.201846579

Strycharz, J., van Noort, G., Smit, E., Helberger, N. (2019). Consumer View on Personalized Advertising: Overview of Self-Reported Benefits and Concerns. In: Bigne, E., Rosengren, S. (eds) Advances in Advertising Research X. European Advertising Academy. Springer Gabler, Wiesbaden. https://doi.org/10.1007/978-3-658-24878-9_5

upGrad. (2022, September 21). *Top 10 most common data mining algorithms you should know*. upGrad blog. https://www.upgrad.com/blog/common-data-mining-algorithms/