Analysis of Algorithmic Bias in Customer Segmentation Models

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Harshil Anishkumar Patel

Spring 2023

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Joshua Earle, Department of Engineering and Society

HOW ARTIFICIAL INTELLIGENCE IS TRANSFORMING THE WORLD

In recent years, there has been a surge of interest in artificial intelligence and machine learning, which are rapidly transforming various aspects of our lives. Some of the world's biggest tech companies like Facebook, Meta, and Amazon have announced significant upgrades to their services, and in each case, the innovations are accompanied by new AI technology. AI technology extends and enhances the capabilities of the human body (Brey, 2000). By leveraging AI, businesses and other institutions can analyze large amounts of data about their customers and can create personalized experiences (Crafts, 2022). This is achieved through the use of various techniques and algorithms such as decision trees, k-means clustering, and neural networks. Different models such as RFM model, LRFM model, and CHAID model are also introduced to classify the shoppers. This is because it is easier for businesses to develop necessary marketing and technical strategies for their customers when they can understand the relation between their characteristics, needs, and behavior. In this paper, I explore some of the approaches to the customer segmentation models that define target markets and precisely identify customer segments. I argue that the models are influenced by biases throughout their development and implementation.

Artificial Intelligence and Machine Learning enables companies and organizations to find and hyper-target their ideal customers at the right time with the right messages. AI based customer segmentation models have revolutionized the way people shop and how businesses work. It is essential for the businesses to detect similarities among their customer base, predict their behavior, and present better options to them. Customer Segmentation is the process by which the businesses divide their customer base into different segments based on common characteristics, such as demographics, common interests, and needs, so that they can target and market to the customers effectively (Gupta, 2020). By doing so, companies increase their customer engagement and drive more positive outcomes for their businesses. Evaluating the limitations of the customer segmentation models within the lens of the Actor-Network Theory and Social Construction of Technology will help identify the strengths and weaknesses of AI models.

Moreover, Artificial Intelligence is transforming the way businesses interact with their customers. AI-powered customer segmentation models can help companies analyze data and gain insights into customer behavior and preferences, enabling them to tailor their marketing strategies and drive sales growth. As AI technology continues to be adopted by companies, it raises important ethical questions about the ethics of population segmentation and nudging people towards or away from certain actions, products, or beliefs. Here, it is essential to recognize the potential for biases to be introduced into these customer segmentation models, which can have negative consequences. To mitigate these risks, companies must prioritize diversity and inclusivity in their data collection and analysis processes. By doing so, they can ensure that their customer segmentation models provide accurate, actionable insights that drive meaningful business results (Daoud, Amine, Bouikhalene, & Lbibb, 2015). Ultimately, the responsible integration of AI in customer segmentation models can help companies build trust with their customers, create personalized experiences, and either positively or negatively impact growth in the increasingly competitive business landscape.

BIASES AND FLAWS IN THE CUSTOMER SEGMENTATION MODELS

It is not possible for medium to large size businesses to have intuition about each and every buyer because they have a large, diversified customer base. At such a stage, along with attracting new customers, it is important to pay attention to the current customers. Customer segmentation models are created as an answer to that. Customer segmentation is something that businesses have been doing for years (Kolodko, 2019). However, these modern customer segmentation models have been created to allow businesses to quickly find and precisely reach specific profit-making customers, and to do so customers are required to be grouped together based on factors, such as demographic, psychographic, ethnographic, needs-based, and value-based. By understanding each customer's preferences and needs, businesses tailor their products and services (Gupta, 2020). When implemented correctly, customer segmentation models can greatly benefit a company, as they are designed to identify specific groups of customers based on shared characteristics, behaviors, or preferences. Which indeed helps data scientists of companies to tailor their marketing strategies to specific target markets, ultimately increasing customer engagement and driving sales. However, although data scientists are trained in data handling, the bias can creep in at a number of stages of this process.

When businesses decide to implement a customer segmentation model, there are a few things that businesses need to consider. They need to figure out what they're trying to achieve - is it to understand their customers better, improve sales, or provide better customer service? Choosing the wrong goal for customer segmentation can lead to a biased model, as the focus may be placed on certain characteristics or behaviors of customers, rather than a holistic understanding of their needs and preferences. For example, if a company's goal is to improve sales, they may only segment customers based on their purchasing history, ignoring other important factors like demographics or psychographics. This can lead to a biased model that targets only a specific group of customers, potentially alienating others and hindering overall growth. It's crucial for companies to carefully consider their goals and ensure that their customer segmentation model is designed with a fair and comprehensive approach. While it may not be possible for a segmentation model to be completely unbiased, being aware of this potential pitfall can enable companies and data scientists to create an effective and relatively impartial customer segmentation model (Daoud, Amine, Bouikhalene, & Lbibb, 2015).

After identifying the purpose of the model, the next step is to define the basis of segments. These segments can be based on factors like age, gender, location, shopping habits, and more. Based on these segments, companies collect necessary data. Businesses can collect information from sales transactions, customer surveys, and website interactions. Here comes the tricky part - sometimes the data itself can be biased. It can perpetuate stereotypes and perpetuate inequality. For example, if a company's sales data only includes purchases made by wealthy individuals, then the segmentation model created from that data will only be representative of the wealthy population, leaving out the lower-income customers. This flawed model can result in a company missing out on potential customers, alienating existing customers, and ultimately hurting their bottom line. Besides, as the algorithms learn from the data they receive, if the data is biased, the algorithm will produce biased results. As revealed in the ProPublica investigation, it found that a widely used algorithm to predict the recidivism of criminal defendants was biased against black defendants (Angwin et al., 2016). The algorithm was trained on data that was biased toward a black defendant, so it retained that bias in its predictions.

After collecting the data, it's time to analyze it and identify patterns and characteristics within each segment (Kirkpatrick, 2017). This is where businesses can really understand the unique needs and preferences of their customers, which allows them to create targeted marketing efforts that resonate with each segment. This can let biases in via the individuals who design and interpret the models, as these people may have their own biases, which have an impact on the results. Once segments are identified and analyzed, businesses can implement the segmentation model by creating personalized marketing campaigns, developing targeted products or services, and improving customer service (Hao, 2019). For instance, a clothing store may segment customers based on age, gender, and shopping behavior. The store can collect relevant data, analyze it to understand each segment's unique characteristics, and create personalized marketing campaigns and customer service experiences that cater to each segment's specific needs and preferences. The store can offer exclusive discounts to loyal customers and create email campaigns targeted at younger customers interested in fast fashion. Later on in the paper, we will discuss more on how these biases are encoded in AI and how they are perpetuated in targeted ads.

The problem often lies in the fact that these models use data to categorize and target customers, and if the data is biased, the resulting model will be biased too (Kirkpatrick, 2017). To tackle this issue, businesses need to prioritize carefully evaluating the data quality, identifying potential biases, and testing algorithms to make sure they are unbiased and effective. As machine learning technologies are advancing at a great rate in today's world, the design errors in these algorithms and mathematical models leading to discrimination and biases are becoming more and more significant. It is important to recognize that bias can be introduced at different stages of the data lifecycle, including data collection, algorithm design, and implementation (Kolodko, 2019). A comprehensive approach that considers each of these stages is necessary to effectively address the issue of algorithmic bias in customer segmentation models. Therefore, by further evaluating this issue of algorithmic bias that comes with these modern customer segmentation models, through the lens of two different STS frameworks: the Actor-Network Theory and the Social Construction of Technology, I argue that by analyzing and improving the customer segmentation process and models, we can prevent targeting the wrong or incomplete group of customers.

EVALUATION OF BIAS IN CUSTOMER SEGMENTATION MODELS

Customer segmentation models powered by artificial intelligence have become increasingly prevalent in today's business landscape (Crafts, 2022). While these models offer numerous benefits to companies, such as targeted marketing and increased customer engagement, as seen in the above sections, they also raise concerns about potential biases and ethical issues. In this section, we will evaluate the limitations and biases present in AI-based customer segmentation models through the lens of the Actor-Network Theory and Social Construction of Technology. By examining the various actors and social forces involved in the development and implementation of these models, we can gain a deeper understanding of the complex socio-technical systems at play and identify ways to address the limitations and biases of these models. Specifically, we will explore the role of data collection and analysis, algorithmic decision-making, and the human factors involved in designing and deploying these systems. Through this evaluation, we aim to provide insights and recommendations for more responsible and equitable use of AI in customer segmentation.

According to Actor-Network Theory, society is made up of many actors, which can either be human or non-human, that interact with each other to form a network of interconnected parts (Crawford, 2020). In the context of customer segmentation models, these actors include the company, its employees, the customers, the algorithms used to analyze customer data, and the technology involved in the process. The relationships between these actors can vary, with some actors holding more power or influence than others. For instance, the algorithms and technology used in customer segmentation may have a significant impact on the results and the way the company interacts with its customers. The segmentation model itself serves as a central point where all the actors come together and interact and plays a vital role against the competition that may disrupt the network by identifying the best targets. By applying Actor-Network Theory to evaluate the customer segmentation models, we can gain a better understanding of the actors involved, how they relate to each other, and how the network functions as a whole. This can help us identify potential weaknesses or biases in the network and work towards improving it for more effective customer segmentation.

There are numerous ways, such as company's segmenting factors, customer's data, and mathematical models, through which the bias can enter into the customer segmentation models (Hao, 2019). Segmentation can lead to bias if the segments that are created are not representing the overall population, that is, the mathematical algorithm will include only those customers who have already purchased a certain product in the past and be biased against those who haven't by excluding them from the target customer base (Kirkpatrick, 2017). Customer data is another source that produces biases in the customer segmentation models. As the customer segmentation models need to work for various types of customers, rather than the ideal type of customers defined by the data used in training, the data used to develop the mathematical algorithm needs to be unbiased.

When looking for the solution to prevent the algorithmic bias associated with the customer segmentation models, the addition of new actors into the network who will control and direct the process by a set of rules is required. The purpose of these new actors would be to search for the possible source of biases and prevent the potential biases. An advisory board should be introduced that can provide ethical knowledge to both engineers and businesses when framing the problem and creating the customer segmentation models. Besides, it is better to tackle the biases in the developing stage rather than the production stage (Amini et al., 2019). A new set of actors who test these models in the developing stage should be introduced, as they could prevent the biases from coming onto the market. This is because the companies that market these algorithms and models have less motivation to solve issues once the software is in production and widely accepted.

As described in SCOT, a technical artifact is defined differently by different user bases and the involved social groups (Bijker et al., 2012). Instead of the common notion that technology determines human actions, SCOT suggests that human decisions and their use of a technology shapes that technology. It highlights that AI technology, in our case, is shaped by social processes and is highly influenced by the values and beliefs of our current society. By including people from different backgrounds, experiences, and perspectives, biases can be identified and addressed before the model is implemented. Additionally, involving end-users in the process can help ensure that the model is designed with their needs and preferences in mind, leading to a more accurate and fair segmentation. Here, it will be essential to have an open and collaborative approach to model development. This includes actively seeking out feedback and input from a diverse group of stakeholders, being transparent about the data and algorithms used, and regularly reviewing the model to ensure that it remains unbiased and effective. By doing this, businesses can create customer segmentation models that are not only accurate and effective but also equitable and fair for all customers.

In short, a responsible approach here would be to have a diversified team and continual monitoring along with established governance and control.

PREVENTING BIAS IN CUSTOMER SEGMENTATION MODELS

Now that we have discussed the biases and shortcomings of customer segmentation models, we explore ways to avoid such biases in the models. In this section, we discuss various strategies and measures that can be implemented to prevent bias from entering these models. According to Actor Network Theory, in the case of customer segmentation models, the network includes software developers, data scientists, marketers, advertisers, consumers, and algorithms.

Each of these actors has their own interests, values, and goals, which can influence the design, implementation, and outcomes of the technology. For example, data scientists may be interested in optimizing the accuracy of the model, while advertisers may be more concerned with maximizing profits. Consumers may have different preferences and needs that are not adequately represented in the model. These conflicting interests can lead to biases in the model, as different actors seek to promote their own interests. Thus, ANT can help identify potential sources of bias by examining the interactions between the developers of the algorithms, the data sets used to train the algorithms, and the algorithms themselves. By understanding these interactions, it may be possible to prevent or mitigate potential biases that may have been inadvertently encoded into the algorithms.

SCOT, on the other hand, views technology as a social construct that is shaped by the individuals and organizations that develop and use it. To prevent bias in customer segmentation models using SCOT, it is necessary to examine the social context in which the models are developed and used. By understanding the social norms and values of the individuals and organizations involved in the development and use of these models, it may be possible to identify potential sources of bias and develop strategies to prevent or mitigate that bias. In the case of customer segmentation models, biases can arise from the data used to train the algorithms, the decision rules embedded in the model, and the assumptions made about the target audience. For example, if the data used to train the model reflects historical inequalities, such as discrimination against certain groups, the model may perpetuate these biases. Similarly, if the decision rules are based on outdated or inaccurate assumptions about consumer behavior, the model may misrepresent the preferences and needs of the target audience (Janet, 2021). To prevent bias in customer segmentation models, it is essential to consider the interests, values, and goals of all actors involved in the development, deployment, and use of the technology. This requires a collaborative and interdisciplinary approach that involves data scientists, marketers, advertisers, consumers, and other stakeholders.

Diversifying the development team can ensure that the model reflects the interests and needs of diverse populations. It's essential to have a diverse team of developers, including individuals from different racial, ethnic, gender, and socioeconomic backgrounds. A diverse team can help identify and mitigate biases in the data, decision rules, and assumptions made about the target audience. Besides, using diverse and representative data can prevent biases in the data used to train the algorithm. It's necessary to use data from multiple sources, such as social media, surveys, and public records. Additionally, the data should be representative of the target audience and should not reflect historical inequalities (Uzzi, 2020).

In addition to this, incorporating transparency and accountability into the development and use of AI algorithms is crucial to ensuring that the model is fair and unbiased. This can be achieved by documenting the data used, the decision rules embedded in the algorithm, and the assumptions made about the target audience. Such documentation can be made available to stakeholders, such as consumers, regulators, and watchdog groups. In addition to this, Ethical considerations must also be incorporated into the design, development, and use of the technology to ensure that the model is ethical (Silberg & Manyika, 2019). This includes considering the potential impacts of the model on different groups of people, potential harms, and benefits, as well as the potential for discrimination, stigmatization, and exclusion.

In addition to internal governance frameworks, external audits can also be conducted to identify and mitigate potential biases. These audits can be conducted by third-party organizations or by internal teams with diverse perspectives and expertise. External audits can provide an objective view of the algorithm and identify potential biases that internal teams may have overlooked. Moreover, preventing bias in customer segmentation models is crucial for promoting fairness and inclusivity. By using strategies such as external audits, increased transparency and accountability, and theoretical frameworks like ANT and SCOT, it is possible to develop customer segmentation models that are free from bias and contribute to a more equitable society.

ROLE OF TRANSPARENCY & ACCOUNTABILITY IN CUSTOMER SEGMENTATION MODELS

Customer segmentation models are not prejudice-free, and if they are not developed and put into operation in a transparent and accountable manner, they may result in unfair and discriminatory behaviors. Transparency in customer segmentation refers to the clear and open communication of the criteria and data used in the segmentation model. It means that the model's inputs, algorithms, and decision-making processes are visible and understandable to stakeholders. Accountability, on the other hand, relates to the responsibility of those involved in designing, implementing, and using the model. It means that they are answerable for their actions and decisions and are held liable for any harm caused by the model.

Transparency and accountability play a significant role in preventing biases in customer segmentation models. When the model's inputs and algorithms are transparent, stakeholders can identify any biases and correct them. For instance, if a company uses demographic data to segment its customers, it may inadvertently exclude or include certain groups based on their gender, race, or ethnicity, leading to discrimination. If the company is transparent about the criteria used, stakeholders can identify such biases and work towards correcting them. In addition to that, accountability ensures that those responsible for designing and implementing the customer segmentation model are aware of the potential biases and take necessary steps to prevent them. It also ensures that those affected by the model can hold the responsible parties accountable for any harm caused by the model.

Lastly, transparency and accountability are also essential in preventing biased interactions resulting from customer segmentation models. As discussed earlier, customer segmentation models can lead to discrimination and unfair treatment of certain groups. By designing and implementing transparent and accountable models, companies can ensure that their marketing efforts are fair and ethical. Ultimately, transparency and accountability are essential in customer segmentation models. They help prevent biases and ensure that marketing efforts are fair and ethical. By designing and implementing transparent and accountable models, companies can prevent discriminatory practices and build trust with their customers.

THE FUTURE OF CUSTOMER SEGMENTATION MODELS

Customer segmentation models will continue to evolve and improve, incorporating more advanced technologies related to machine learning and artificial intelligence. These advancements will allow for even more precise and accurate customer targeting and personalization. However, it will also be important for businesses to prioritize ethical considerations and prevent biases from creeping into the models. As these models are increasingly becoming a part of our everyday lives, the dark side of it is left unchecked. It is important to question that side, as the bias is difficult to detect and fix. Fixing the discrimination and biases in the systems based on mathematical models and algorithms is a long, ongoing process, which is not going to end that easily. However, with careful evaluation and testing, companies can ensure that their customer segmentation models are both effective and fair. Since artificial intelligence has the capacity to improve technology greatly yet also discriminate immensely, preventing biases from growing, or even beginning, will have a positive impact on new technology introduced to society in future. Overall, it is essential to take a responsible approach when implementing a customer segmentation model that is both fair and accurate. To do so, a diverse team of professionals will be required, who can offer different perspectives and spot potential biases. In addition to this, it is also necessary to continuously monitor the data and algorithms used in the model to make sure it stays fair and accurate as time goes on. Having an established governance and control framework can also help make sure that the model's decisions align with the company's values and ethical standards.

BIBLIOGRAPHY

Amini, A., Soleimany, A., Schwarting, W., Bhatia, S., & Rus, D. (2019). *Uncovering and mitigating algorithmic bias through learned latent structure.* Proceedings of the 2019 AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society (AIES) (pp. 1-7). Honolulu, HI: Association for the Advancement of Artificial Intelligence/Association for Computing Machinery. Retrieved from http://hdl.handle.net/1721.1/121101

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine Bias: Risk Assessments in Criminal Sentencing. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Bijker, Wiebe., et al. *The Social Construction of Technological Systems New Directions in the Sociology and History of Technology*. MIT Press, 2012.

Daoud, R.A., Amine, A., Bouikhalene, B., & Lbibb, R. (2015). *Combining RFM model and clustering techniques for customer value analysis of a company selling online*. *2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA)*, 1-6.

Gupta, S. (2020, July 9). *Is RFM still king? A data science evaluation*. ReSci. Retrieved December 16, 2022, from https://www.retentionscience.com/blog/rfm-king/

Hao, K. (2019, February 4). *This is how AI bias really happens—and why it's so hard to fix.* MIT Technology Review. Retrieved from https://www.technologyreview.com/s/612876/this-is-how-ai-bias-really-happensandwhy-its-so-hard-to-fix/

Janet. (2021, May 14). Overcome and prevent bias in AI. Retrieved May 8, 2023, from https://f8federal.com/overcome-and-prevent-ai-bias/

Kirkpatrick, K. (2017). *It's not the algorithm, it's the data*. Communications of the ACM, 60(2), 21–23. doi:10.1145/3022181

Kolodko, J. (2019, October 24). *Customer segmentation: "when and where" instead of "who" - blog*. Julia Kolodko. Retrieved December 16, 2022, from https://www.juliakolodko.com/segmentation

Miller, A. P., & Hosanagar, K. (2019, November 8). *How targeted ads and dynamic pricing can perpetuate bias*. Harvard Business Review. Retrieved February 15, 2023, from https://hbr.org/2019/11/how-targeted-ads-and-dynamic-pricing-can-perpetuate-bias

Papa, V. (2018, October 11). Amazon's sexist hiring algorithm could still be better than a human. IMD. Retrieved from https://www.imd.org/research-knowledge/articles/amazons-sexist-hiring-algorithm-could-still-be-better-than-a-human/

Petrone, J. (2019, July 18). AI Embeds Bias in Social Media. Pulitzer Center. https://pulitzercenter.org/blog/ai-embeds-bias-social-media

Polovin, G. (2020, December 15). *Market research bias and the customer experience*. Sogolytics Blog. Retrieved December 16, 2022, from https://www.sogolytics.com/blog/market-research-bias-and-the-customer-experience/

Silberg, J., & Manyika, J. (2019, June 06). Tackling bias in artificial intelligence (and in humans). Retrieved May 8, 2023, from https://www.mckinsey.com/featured-insights/artificial-intelligence/tackling-bias-in-artificial-intelligence-and-in-humans

Sweeney, L. (2013). *Discrimination in online ad delivery.* ACM Queue, 11(3), 1-19. https://arxiv.org/abs/1301.6822

Uzzi, B. (2020, November 04). A simple tactic that could help reduce bias in AI. Retrieved May 8, 2023, from https://hbr.org/2020/11/a-simple-tactic-that-could-help-reduce-bias-in-ai