Scalable Ad-Targeting Technology using AWS

The Fairness of Corporate Artificial Intelligence

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
Nitin Maddi

November 1, 2021

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.

ADVISORS

MC Forelle, Department of Engineering and Society

Rosanne Vrugtman, Department of Computer Science

Introduction

Multi-billion dollar corporations including but not limited to Google, Microsoft, and Apple have an incredible amount of power, as they possess massive amounts of personal information and are an unseen but critical part of our everyday lives. Although these companies rely on us and our information, they are not always looking out for the best interest of the people who interact with them and their products. These corporations have been plagued by controversies due to their constant pursuit of more money, and have been willing to do nefarious things with the data that they have collected about their users and the general population. With the advent of Artificial Intelligence and its widespread use in our everyday lives, we have come to see these same corporations utilize AI in almost every aspect of their business. (Jovanovic, 2022) Hidden behind walls of legal jargon, businesses have found a way to keep their AI completely proprietary even though it directly impacts millions of people (Knight, 2017). While it has been developed into one of the most powerful and effective tools of this era, AI is still far from perfect. One of the major flaws that have been recognized about AI is its dependence on data without having the power to recognize biases. AI relies on data and algorithms to learn, which in turn means that AI also gains biases from many different sources (Mehrabi et al., 2021). If these huge businesses are using bad or dirty data to power their algorithms, AI could be corrupt and lead to discrimination against certain social groups. Facial recognition software is an example of this corruption that has perpetuated racist stereotypes, as the amount of facial data collected on people of color is lacking in comparison with data that has been collected on Caucasian people (Kaur et al., 2020). This discrepancy in the diversity of data that is available has led to reduced accuracy for facial recognition for these groups, and dire consequences such as incorrectly matched mugshots (Najibi, 2020).

Currently, companies collect data on their users in any way, shape, or form that they can. For example, E-Commerce companies track and keep records of items that a customer adds to their cart or items that the person just hovered over (Subramanian, 2022). They can even track data that represent a customer's current behavior by analyzing downgrades in subscriptions, order cancellations, and much more. With this data, they detect which customers are most likely to buy certain products, and the company can send targeted advertisements to them through platforms such as Google and Facebook. Although this analysis is extremely powerful, the amount of data the company generally collects can be on the scale of petabytes which is an impossible amount of information to work with. To evaluate that amount of data, the company would need to invest millions of dollars into a data analysis platform so that they would have the resources needed to do in-depth data analysis on such a large scale. Additionally, separate pipelines would need to be developed for each advertisement platform. Although this is a serious problem that large companies face, the development of ad-targeting software holds the power to negatively impact certain communities in ways that we should be cognizant of. For example, it is a fact that people who come from low-income communities are more likely to smoke cigarettes (Casetta, 2017). If a company were to use this information to send targeted advertisements to people with lower incomes, this would exacerbate this issue and could harm more people. In today's world, data is the most powerful tool, but if it is not properly managed and applied, the people with access to this data can use it to exploit communities for their profit using ad-targeting and AI.

Technical Topic

Businesses try to be as efficient with their money as possible, which drives them to use data-driven insights to make every decision the best they possibly can. Collecting this data is not a problem, as there are hundreds of sources from which to extract information. The more difficult part is being able to utilize the information in their decisions. The amount of data that can be collected is on the order of petabytes, which is an unfathomable amount. If we were to try to do anything with that amount of data on a regular computer it would take years to finish computing (Spurlock, 2019). To handle it, companies would need to invest many millions in infrastructure, as well as develop efficient code to conduct their in-depth data analysis. This would mean companies would need to hire engineers to constantly monitor the system and always be ready to fix unexpected breakages that may bring the system down for days at a time. Also, if an unexpected amount of data comes into the system, the servers may not be capable of accommodating this burst as they only have a fixed amount of speed and storage. (Catteddu, 2010).

At Salesforce, I worked on a scalable solution to this problem that allowed any company to use its data to push targeted advertisements based on the results of the data analysis. My team utilized Amazon Web Services(AWS) to act as the frame of our solution. AWS is a cloud service provider, meaning it can provide as much power and storage as necessary for our system to run, while we pay on an hourly basis based on how many resources are used (Barney & Gillis, 2022). No matter how much data is sent through this system, my team was able to automatically scale up resources on a need-based basis. From this, my team utilized AWS Simple Queue Service (SQS) to simplify and orchestrate the data analysis workflow. This process started by splitting the original process into hundreds of smaller-sized processes. With the help of Java, these smaller processes are examined and each user is matched to the best possible advertisement based on their information. After the data analysis is conducted, the system matches these data points to real accounts on an advertisement platform. This is done by creating an Audience, which is essentially a list of email addresses to that a specific advertisement will be sent. From this Audience, a request is made to Google, Facebook, or any other platform, and the corresponding advertisement is delivered to all of the accounts.

This product is a great answer for any company that wishes to send targeted advertisements, but can not spend millions of dollars in capital and even more to maintain the infrastructure necessary for an operation of this size. However, this software is the perfect example of how companies can further propagate bias against certain communities. In the data analysis process, there were safety measures taken to how the data was used; some collected information would be blocked to be omitted from the data analysis step. However, information such as gender and age that didn't break any legal boundaries was realistically fair game to be used in the advertisement process. While this may lead to increased profits, it could be detrimental to the communities targeted by these advertisements.

STS Topic

As companies strive for greater efficiency and automation, they will inevitably accept AI and machine learning into their arsenal of tools. It is estimated that 43% of companies are currently using machine learning, and this number is estimated to double shortly as cloud-based applications continue to rise in popularity (Stewart, 2019). This is an astounding number and it highlights the amount of control that AI holds over our everyday lives. However, for machine learning models to learn accurately and efficiently, they require data as input (Brownlee, 2016). Sadly though, there are countless ways for bias to enter this data. For instance, bias can arise through the sampling process, manipulation and cleaning stage, and more (Gevaert, 2021). While machine learning is made to be a humanless learning process, when humans interact with the process, a form of bias is implicitly added. Furthermore, explicit bias can be directly introduced into algorithms when engineers choose to deliberately influence how the algorithm learns concerning a particular social group (Fosso Wamba, 2020).

One instance of a machine learning model which has been greatly impacted by bias is Amazon's ML-based recruitment algorithm. After observing resumes submitted to Amazon over a 10-year period, the model deduced that male applicants were more qualified than female applicants due to the distribution of overall applicants being heavily skewed toward male applicants (Parikh, 2021). Other AI in the field of criminal justice, facial recognition, COVID-19 detection, and much more are all plagued by the biases that exist in the real world that are reflected within the data that we collect and feed to the models (Rayhan, 2022). Proper collection of data may alleviate some of these issues, but may still not be enough to reduce the bias enough in certain sectors where problems are more deeply rooted than just the data collection process (Floridi et al., 2020). Without measures to combat these biases and ensure the fairness of AI, companies can veer off into creating technology that is detrimental to humanity (Butterworth, 2018).

To analyze this problem, I will utilize the Actor-Network Theory (ANT) to better understand these corporations and their interactions with consumers and AI. An important characteristic of ANT is the fact that both human and non-human actors are treated the same within the network. This abstraction would allow for the many non-human components of AI, such as the data and the algorithms to have the same influence as the human actors. Additionally, this model reduces the biases that arise from the assumptions that humans and non-humans are fundamentally different within a social network (Cressman, 2009). While this helps ANT reduce bias, it is also the source of its flaw. Actors do not hold completely equal weights within a network, and by considering each actor as equivalent, this model makes a generalization that may remove important information. However, this model stands to be a great fit for understanding AI, as these algorithms are entirely non-human but try to imitate humans in their learning and decision-making process.

An additional analysis of the infrastructure behind corporate AI may prove to be useful, as AI is traditionally used as the infrastructure behind a feature. A typical customer-facing feature does not indicate that artificial intelligence is being used under the hood and is invisible to a typical user. By inspecting the infrastructure, I will be able to get a better understanding of the overall architecture and discover integral design choices which may lead to sources of bias (Star, 2016).

Methodologies

Research Question: To what extent have consumer technology companies taken measures to develop AI that is socially conscious and doesn't inherently discriminate against certain social groups?

With the amount of control that companies have over the lives of billions, it is necessary to keep them accountable for the impacts of the technology they employ in their systems. This research question will answer if companies are taking the proper measures to ensure fairness in their AI. I will analyze this topic by conducting discourse analysis. I plan to closely read eight research papers by prominent AI companies in the field of consumer technology including

Apple, Google, and Microsoft, all of which have widespread use of AI in their business and interact with billions of people daily. Conducting in-depth research into these major companies and understanding their stance on fairness in Artificial Intelligence will provide key insight into how they are currently utilizing resources to reduce bias. Data will be collected by creating a template that includes sections corresponding to certain fairness indicators, such as AIF360 and GDPR framework (Mehrabi, 2021). These indicators and tools have been developed by the community and AI experts and provide a great baseline for measuring fairness. After collecting this data, each AI will be scored against the amount of impact it has. In other words, AI will then be given a rating based on if its fairness level is high enough relative to the number of consumers it interacts with. An AI that interacts with more people, such as Apple Facial Recognition, will require a higher fairness rating in order to be considered fair enough to be utilized by a company (Stanley, 2017).

Conclusion

The technical portion of this paper describes an innovation in data analysis that allows other companies to utilize Salesforce to send advertisements based on the massive amounts of data that they collect on their customers. This system is one that is scalable and robust, allowing for any amount of data to be processed for low costs. While very beneficial for companies, this innovation may cause biases deeply rooted within communities to be further exacerbated.

On the other hand, this paper will also analyze the fairness of artificial intelligence and machine learning models in large consumer technology companies. These companies hold tremendous influence over our everyday lives and it is imperative to ensure that they are upholding the highest level of fairness within their AI. This paper will understand and score the impartiality of these algorithms and compare them to the amount of influence they hold. By the end of this paper, the reader will be left with an understanding of the current research and resources being implemented for fair artificial intelligence, if companies should be doing more, as well as the consequences for not following proper guidelines.

References

Barney, N., & Gillis, A. S. (2022, October 6). *What is AWS (Amazon Web Services) and how does it work?* SearchAWS. Retrieved October 30, 2022, from https://www.techtarget.com/searchaws/definition/Amazon-Web-Services

Brownlee, J. (2016, March 11). *How machine learning algorithms work (they learn a mapping of input to output)*. Machine Learning Mastery. Retrieved October 31, 2022, from https://machinelearningmastery.com/how-machine-learning-algorithms-work/

Butterworth, M. (2018). The ICO and artificial intelligence: The role of fairness in the GDPR framework. *Computer Law & Security Review*, *34*(2), 257–268. https://doi.org/10.1016/j.clsr.2018.01.004

Casetta B, Videla AJ, Bardach A, Morello P, Soto N, Lee K, Camacho PA, Hermoza Moquillaza RV, Ciapponi A. (2017). Association Between Cigarette Smoking Prevalence and Income Level: A Systematic Review and Meta-Analysis. Nicotine Tob Res. 2017 Nov 7;19(12):1401-1407. doi: 10.1093/ntr/ntw266. PMID: 27679607.

Catteddu, D. (2010). Cloud Computing: Benefits, Risks and Recommendations for Information Security. In: Serrão, C., Aguilera Díaz, V., Cerullo, F. (eds) Web Application Security. IBWAS 2009. Communications in Computer and Information Science, vol 72. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-16120-9_9

Cressman, D. (2009). *A Brief Overview of Actor-Network Theory: Punctualization, Heterogeneous Engineering & Translation*. Simon Fraser University. https://summit.sfu.ca/item/13593

Floridi, L., Cowls, J., King, T. C., & Taddeo, M. (2020). How to design AI for social good: Seven essential factors. *Science and Engineering Ethics*, *26*(3), 1771–1796. https://doi.org/10.1007/s11948-020-00213-5

Fosso Wamba, S., Bawack, R., Guthrie, C., Queiroz, M., & Carillo, K. (2020). Are we preparing for a good ai society? A Bibliometric Review and research agenda. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3735322

Gevaert, C. M., Carman, M., Rosman, B., Georgiadou, Y., & Soden, R. (2021). Fairness and accountability of AI in disaster risk management: Opportunities and challenges. *Patterns*, *2*(11), 100363. https://doi.org/10.1016/j.patter.2021.100363

Jovanovic, B. (2022, March 8). *55 fascinating AI statistics and Trends for 2022*. Dataprot. Retrieved October 29, 2022, from https://dataprot.net/statistics/ai-statistics/#:~:text=37%25%20of%20businesses%20and%20organizations,million%20new%20ones%20by%202025.

Kaur, P., Krishan, K., Sharma, S. K., & Kanchan, T. (2020). Facial-recognition algorithms: A literature review. *Medicine, Science and the Law*, *60*(2), 131–139. https://doi.org/10.1177/0025802419893168

Knight, W. (2020, April 2). *The dark secret at the heart of ai*. MIT Technology Review. Retrieved October 29, 2022, from https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, *54*(6), 1–35. https://doi.org/10.1145/3457607

Najibi, A. (2020, October 26). *Racial discrimination in face recognition technology*. Science in the News. Retrieved October 31, 2022, from https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/

Parikh, N. (2021, October 14). *Council post: Understanding bias in AI-enabled hiring*. Forbes. Retrieved October 30, 2022, from https://www.forbes.com/sites/forbeshumanresourcescouncil/2021/10/14/understanding-bias-in-ai-enabled-hiring/?sh=135922427b96

Rayhan, M. D., Alam, M. D. G., Dewan, M. A., & Ahmed, M. H. (2022). Appraisal of high-stake examinations during SARS-COV-2 emergency with responsible and transparent AI: Evidence of fair and detrimental assessment. *Computers and Education: Artificial Intelligence*, *3*, 100077. https://doi.org/10.1016/j.caeai.2022.100077

Spurlock, R. (2019, October 19). *Petabyte - how much information could it actually hold? - cobalt iron*. Cobalt Iron. Retrieved October 30, 2022, from https://info.cobaltiron.com/blog/petabyte-how-much-information-could-it-actually-hold

Stanley, J. (2017, September 14). *Apple's use of face recognition in the new iphone: Implications: News & commentary*. American Civil Liberties Union. Retrieved October 31, 2022, from https://www.aclu.org/news/privacy-technology/apples-use-face-recognition-new-iphone

Star, S. L. (2016). The ethnography of Infrastructure. *Boundary Objects and Beyond*. https://doi.org/10.7551/mitpress/10113.003.0030

Stewart, M. (2019, July 29). *The limitations of machine learning - towards data science*. Retrieved October 31, 2022, from https://towardsdatascience.com/the-limitations-of-machine-learning-a00e0c3040c6

Subramanian, R. (2022, October 20). *How is customer data used in ecommerce industry*. TheCommerceShop. Retrieved October 29, 2022, from https://www.thecommerceshop.com/blog/customer-data-and-ecommerce-how-important-is-it/