Teaching an Intersectional Data Analysis on Affirmative Action

Olivia Dias
Massachusetts Institute of Technology
Media Lab
Cambridge, MA, USA
omdias77@mit.edu

Raechel Walker
Massachusetts Institute of Technology
Media Lab
Cambridge, MA, USA
raechelw@media.mit.edu

Cynthia Breazeal
Massachusetts Institute of Technology
Media Lab
Cambridge, MA, USA
cynthiab@media.mit.edu

ABSTRACT

ADM systems can be used to perform a task as inconsequential as recommending a song on Spotify, to making a decision that is instrumental to someone's life, such as determining their candidacy for college. If an algorithm is trained on biased data, it can propagate prejudice. Thus, it is pertinent to find methods to decrease ADM bias. This paper presents a way to potentially mitigate ADM bias by teaching high school students a intersectional data analysis activity that incorporates the second pillar of the liberatory computing framework, critical consciousness. This activity is designed to enable high school students to understand the bias and history behind the college admission process, which allows students to develop a critical consciousness. Establishing a critical consciousness will diversify the computing field and the data incorporated into ADM systems by encouraging minoritized high school students to get a degree in computer science. The National Institute of Standards and Technology (NIST) suggests that diversifying the computing field has the potential to reduce bias in ADM systems. Thus, the activity is focused on students developing a critical consciousness. This paper discusses the preliminary findings from teaching a two-day computing activity to high school students.

CCS CONCEPTS

• **Social and professional topics** → *Computing literacy.*

KEYWORDS

algorithmic-decision making, affirmative action, critical consciousness, data analysis, data science, liberatory computing framework

ACM Reference Format:
Olivia Dias, Raechel Walker, and Cynthia Breazeal. 2023. Teaching an Intersectional Data Analysis on Affirmative Action. In *Proceedings of the 54th ACM Technical Symposium on Computing Science Education V. 2 (SIGCSE 2023), March 15–18, 2023, Toronto, ON, Canada.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3545947.3573294

1 PROBLEM/MOTIVATION

Automated decision-making (ADM) systems have been used in the college admissions process to decrease the number of possible candidates that college admission officers have to sift through. The data used to train these ADM systems has the potential to propagate the bias present in the admission process. Due to the COVID-19 pandemic, there has been an increase in the number of colleges utilizing ADM systems during their college admission process [7, 11]. However, the use of these ADM systems has been shown to occasionally exacerbate the bias reflected in admission officers [11]. For instance, the University of Texas at Austin implemented a type of ADM system in their admission process to evaluate an applicant's application by producing "scores based on the likelihood of admission by a review committee" [11]. Ultimately, the ADM was discarded because the college admissions officers came to the conclusion that the system had a high likelihood to "replicate superficial biases in the scoring- scoring up some applications not because they were good, but because they looked like the kinds of applications that had been approved in the past" [11]. Biased ADM systems hinder prospective students from receiving a college education, limiting the diversity in computing and perpetuating the cycle of marginalization against the people that are already disenfranchised. The paucity of diversity in the computing field motivated the curation of an Intersectional Data Analysis activity. The Intersectional Data Analysis activity was designed with the objective of engaging minoritized high school students in computer science to encourage them to obtain a degree in computing. In order to convey the predicament with biased ADM systems to the high school student without explicitly telling them, the students created data visualizations to aid in conceptualizing the lack of diversity in colleges from states that have banned race-based affirmative action. An intersectional data analysis analyzes the "multiplicative effects of different but interdependent categories and factors" in the data [1]. Students developed a critical consciousness about the importance of diversity in colleges, after they analyzed the data visualizations. The goal of this activity was to help the students to understand the shortcomings of the present day admissions process and ultimately encourage them to pursue a career in computing.

2 BACKGROUND AND RELATED WORK

Current computing pedagogy for minoritized students does not teach advanced computing or data science concepts in conjunction with social justice topics [2–5, 8–10, 12, 13]. My research is novel since it compels students to not look at data objectively, but to use data science to analyze the bias that might be present in the data. Moreover, a vast amount of computing pedagogy does not provide context about the data it is trained on and thus ignores how the algorithm can impact minoritized groups [7]. The students learned about the history of affirmative action and the data analysis process which is an approach to computing pedagogy that developed the student's critical consciousness. Developing a critical consciousness enabled the students to perceive the historical bias that might be present in the data. The college admittance activity focused on

SIGCSE 2023, March 15–18, 2023, Toronto, ON, Canada

ACM ISBN 978-1-4503-9433-8/23/03.
https://doi.org/10.1145/3545947.3573294

the historical context behind the curation and evolution of affirmative action throughout the years. Students analyzed the effects of affirmative action, which is when an "organization [proactively] devotes resources (including time and money) to making sure that people are not discriminated against on the basis of their gender or their ethnic group" [6].

This increased the students' confidence in pursuing a career in computing and to mitigate the potential bias ingrained in ADM systems [14]. This approach to computing pedagogy comes from the liberatory computing framework, which combines general computing concepts with the five pillars of the liberation framework- "1) sound racial identity, (2) critical consciousness, (3) liberation-centered academic/achievement identity, (4) collective obligation, and (5) activism skills" [15]. In the curriculum, students developed their critical consciousness about affirmative action and recognized the lack of minoritized groups represented in their data visualizations they developed as a part of the workshop.

3 APPROACH AND UNIQUENESS

3.1 Participation

For this research, twelve high school students were recruited to do the college admittance activity. Each student received an hourly wage to participate in this research study. The students participated in a six week long program that taught the data activism curriculum. The college admittance activity was taught over the course of two days, for four hours each day. Seven students consented to participate in the research study; five were girls and two were boys. Six of the students were African American and one student was Asian American.

3.2 Activities

Students learned about some of the historical biases present in the college admissions process and generated a critical consciousness about affirmative action. Then students will create and analyze data visualizations, which will demonstrate the lack of minoritized groups admitted into college. Finally, the students will discuss how minoritized groups are disproportionately affected by biased college admittance algorithms and the lack of diversity in colleges.

3.2.1 Affirmative Action Section. For this activity, I taught students about ideologies and practices that influenced the college institution. Students discussed the effects of various injustices, such as the legacy of slavery and poverty on a person's ability to enter college and obtain financial stability. Next, the students learned about affirmative action and how it has changed over the years due to the numerous lawsuits against it. Then, the students discussed and considered the bias in college admissions ADM systems.

3.2.2 Intersectional Data Analysis Activity. This activity is focused on performing an intersectional analysis on the gender and race demographics of colleges in different states that did and did not ban race-based affirmative action. I guided the students through a data analysis activity on Deepnote, an online data science notebook where data can be cleaned, manipulated, and visualized. In the Deepnote portion of the activity, the students learned computing concepts related to Python, a programming language, and Pandas, a Python library that is used for data analysis. The students worked with a dataset that consisted of aggregated data related to standardized exam scores, the state of the college, race and gender demographics of the student population, the percentage of the student population that received financial aid, and various other metrics from around 1500 colleges in the US.

In this specific activity, students learned how to do Boolean indexing. Boolean indexing is a programming technique that allows a programmer to clean data and limit the number of rows in the dataset that is analyzed. Also, the students learned how to use helper functions, which are used in data science to clean and manipulate the data so that the data can be visualized. Before this activity, none of the students knew how to do boolean indexing nor helper functions.

4 RESULTS AND CONTRIBUTIONS

Every student demonstrated that they developed a critical consciousness about affirmative action. Additionally, they all showcased the data science skills they gained by visualizing the lack of representation of minoritized groups in the college enrollment dataset. Specifically, all students incorporated Boolean indexing or helper functions to clean, manipulate, and visualize their data. I performed a thematic analysis to assess the impact of the intersectional analysis of college admittance activity.

After completing the college admittance activity, students utilized their skills do a final project. The final project gave students the ability to conduct their own data analysis on data related to a social justice topic of their choice. Six students utilized boolean indexing and two students used helper functions during their final project. Even though this curriculum focused on racial disparities in college admittance, several students were interested in analyzing racial disparities in different fields. Three students found this topic engaging because they continued to analyze racial disparities, but in different topics, such as food insecurity, advanced placement classes, and healthcare.

Student 1 chose to research the historical issues that have led to the lack of diversity in students that take Advanced Placement (AP) classes in high school (Fig 1). The student found that "resource inequities caused by funding gaps in school districts" and "inequitable access to quality early childhood opportunities contributes to differences in the childs' quality of education".

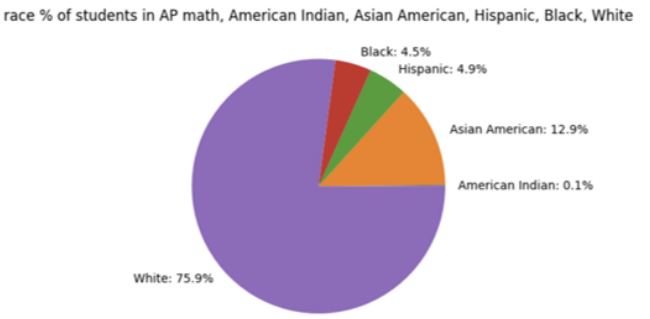

Figure 1: An example of Student 1 visualizing the race demographics of students that take AP classes

Student 1 utilized boolean indexing to aggregate the data so that they could create data visualizations and analyze the data. Their ability to think critically and understand the bias present in datasets can mitigate the adverse effect biased ADM systems have on marginalized communities.

In this research, I saw three students developed a critical consciousness about racial disparities. All the students analyzed data with their critical consciousness in tandem with the computing skills they gained through the college admittance activity. This computing curricula can be used to encourage minoritized students to consider a computing career, and mitigate biases present in ADM systems.

In the future, I will teach my college admittance activity to more high school students. Additionally, I will monitor which students actually pursue a computer science career.

REFERENCES

[1] 2022. *Intersectional Approaches.* Retrieved October 1, 2022 from http://genderedinnovations.stanford.edu/methods/intersect.html#:~:text=Analyze%20data,but%20interdependent%20categories%20and%20factors

[2] L. Alvarez, I. Gransbury, V. Cateté, T. Barnes, Ákos Ledéczi, and S Grover. 2022. A Socially Relevant Focused AI Curriculum Designed for Female High School Students. *Proceedings of the AAAI Conference on Artificial Intelligence* 36, 11 (2022), 12698–12705. https://doi.org/10.1609/aaai.v36i11.21546

[3] M. A. Beasley and M. J. Fischer. 2012. Why they leave: The impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors. *Social Psychology of Education* 15, 4 (2012), 427–448.

[4] S. Erete, N. Thompson, M. Standberry-Wallace, B. Ju, D. Nacu, and N. Pinkard. 2021. Honoring black women's work: creating a parent and caring adult community to support youth STEAM engagement. *2021 Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology (RESPECT)* (2021), 1–5.

[5] Jayne Everson, F. Megumi Kivuva, and Amy J. Ko. 2022. "A Key to Reducing Inequities in Like, AI, is by Reducing Inequities Everywhere First": Emerging Critical Consciousness in a Co-Constructed Secondary CS Classroom. *In Proceedings of the 53rd ACM Technical Symposium on Computer Science Education* 1 (2022), 209–215. https://doi.org/10.1145/3478431.3499395

[6] Sirinda Sincharoen Faye J. Crosby, Aarti Iyer. 2006. Understanding Affirmative Action. *Annu. Rev. Psychol. 2006* (2006). https://doi.org/doi:10.1146/annurev.psych.57.102904.190029

[7] Christopher Starke Frank Marcinkowski, Kimon Kieslich and Marco Lunich. 2020. Implications of AI (un-)fairness in higher education admissions. *2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)* (2020). https://doi.org/10.1145/3351095.3372867

[8] Joseph Isaac and Christina Gardner-McCune. 2021. Developing Introductory Computer Science Pedagogy for Black Boys. In *2021 Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology (RESPECT).* 1–6. https://doi.org/10.1109/RESPECT51740.2021.9620642

[9] Irene Lee, Safinah Ali, Helen Zhang, Daniella DiPaola, and Cynthia Breazeal. 2021. Developing Middle School Students' AI Literacy. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (Virtual Event, USA) *(SIGCSE '21).* Association for Computing Machinery, New York, NY, USA, 191–197. https://doi.org/10.1145/3408877.3432513

[10] Jane Margolis, Jean J. Ryoo, Cueponcaxochitl D. M. Sandoval, Clifford Lee, Joanna Goode, and Gail Chapman. 2012. Beyond Access: Broadening Participation in High School Computer Science. *ACM Inroads* 3, 4 (dec 2012), 72–78. https://doi.org/10.1145/2381083.2381102

[11] Derek Newton. 2022. Artificial Intelligence grading your 'neuroticism'? Welcome to colleges' new frontier. (apr 2022). https://www.usatoday.com/story/news/education/2021/04/26/ai-infiltrating-college-admissions-teaching-grading/7348128002/

[12] N. Pinkard, S. Erete, C. K. Martin, and M. McKinney de Royston. 2017. Digital youth divas: Exploring narrative-driven curriculum to spark middle school girls' interest in computational activities. *Journal of the Learning Sciences* 26, 3 (2017), 477–516.

[13] M. A. Qazi, J. Gray, D. M. Shannon, M. Russell, and M. Thomas. 2020. A state-wide effort to provide access to authentic computer science education to underrepresented populations. *Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (2020), 241–246.

[14] Kristen Greene Lori Perine Andrew Burt Reva Schwartz, Apostol Vassilev and Patrick Hall. 2022. Towards a Standard for Identifying and Managing Bias in Artificial Intelligence. (March 2022). https://doi.org/10.6028/NIST.SP.1270

[15] Raechel Walker, Eman Sherif, and Cynthia Breazeal. 2022. Liberatory computing education for African American students. *2022 Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology (RESPECT)* (March 2022), 85–89. https://doi.org/10.6028/NIST.SP.1270

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009