Analysis of AI Models for Student Admissions: A Case Study

Kelly Van Busum
Department of Computer & Information Science
Indiana University Purdue University Indianapolis
kelvanbu@iupui.edu

Shiaofen Fang
Department of Computer & Information Science
Indiana University Purdue University Indianapolis
shfang@iupui.edu

ABSTRACT

This research uses machine learning-based AI models to predict admissions decisions at a large urban research university. Admissions data spanning five years was used to create an AI model to determine whether a given student would be directly admitted into the School of Science under various scenarios. During this time, submission of standardized test scores as part of a student's application became optional which led to interesting questions about the impact of standardized test scores on admission decisions. We first developed AI models and analyzed these models to understand which variables are important in admissions decisions, and how the decision to exclude test scores affects the demographics of the students who are admitted. We then evaluated the predictive models to detect and analyze biases these models may carry with respect to three variables chosen to represent sensitive populations: gender, race, and whether a student was the first in his family to attend college.

CCS CONCEPTS

• **Applied Computing** → Education

KEYWORDS

machine learning, bias, predictive model, test-optional, college admissions

ACM Reference Format:

Kelly Van Busum and Shiaofen Fang. 2023. Analysis of AI Models for Student Admissions: A Case Study. In *The 38th ACM/SIGAPP Symposium on Applied Computing (SAC'23),* March 27–31, 2023. DOI: 10.1145/3555776.3577743

1 INTRODUCTION

Artificial Intelligence (AI) has started to play an increasingly important role in higher education. AI-based predictive models built from existing datasets such as admissions data provide insight into the policies, practices, and biases in U.S. higher education system. At the same time, AI models have their own limitations and potential for harm such as biases towards certain populations. It is therefore important to understand these limitations and biases to truly integrate AI tools in future higher education applications. This paper describes an initial attempt, as a case study, to build, apply, and analyze AI models using an existing student admission dataset.

Admissions data from the School of Science at a large urban research university was used to create machine learning-based AI models. These models predict whether a student would be directly admitted into the School of Science, or not, under a variety of scenarios. The dataset spans five years, and over this time, the admissions policy of the university changed from requiring students to submit standardized test scores as part of their application, to making test scores optional. Such fundamental changes in admission policies can significantly impact our higher education system for different populations, so it is important to have a strong understanding of the implications of such policy changes.

When existing data is used to create predictive models, the behaviors and features of these models provide a powerful opportunity to better understand both the potential impact of policies on different populations, and the limitations of the models themselves. As a case study, this paper attempts to address both issues using a dataset collected from past admission data from the School of Science. Our analysis will first try to better understand what variables are important in admissions decisions, and how the decision to exclude test scores may affect the demographics of the students who are admitted. The predictive model contains a variety of demographic variables, and three: gender, race, and whether a student was the first in his family to attend college ("First-Generation"), were chosen to represent sensitive populations ("sensitive variables"). We then carefully evaluate the AI models for presence of potential biases with respect to performance relative to these three variables. The result of this analysis provides some initial evidence that AI algorithms can be harmful when used as part of the admission decision-making process if bias is not effectively mitigated.

Students who attend large urban universities have some unique challenges. They may be more likely than students who attend other types of universities to face the competing demands of work, family, and school leading to poorer academic experiences. Attending a university in a city with a high cost-of-living and the continuing effects of COVID-19 are both factors that can increase

SAC'23, March 27 – March 31, 2023, Tallinn, Estonia

DOI: 10.1145/3555776.3577743

the financial demands on students, also leading to increased stress and poorer academic outcomes. Although urban universities tend to be more diverse, minority students still report struggling with a sense of belonging, which can negatively affect academic performance. To best support all students, care must be taken to understand the impact of policy on equity.

Test-optional policies were designed to address the concern that standardized test scores are biased metrics for predicting student success and to increase equity in admissions procedures. But more research needs to be done to fully understand how these policies might change the demographics of the students admitted to the university and other impacts. Identifying the important factors in admissions decisions and how test-optional policies might change admitted student demographics is the first goal of this project.

As an additional step toward understanding and mitigating bias in admissions policies, college admissions officers are increasingly turning to the use of machine learning-based AI algorithms. However, when AI models are trained using existing datasets, the models can introduce new bias and fairness issues with the potential for harm. This shows a need to better understand how bias might manifest as AI algorithms become more widely used in higher education. Identifying and measuring bias in the models built to predict admission decisions is another important goal of this project.

This work makes the following contributions:

1. College admissions data was used to generate AI predictive models, evaluated on their overall accuracy and effectiveness. These models were used to better understand the primary factors in admissions decisions and their variations within different cohorts.
2. The performance of the AI predictive models was evaluated relative to various sensitive groups, and the results show evidence of biases. This serves as a warning and contributes to an understanding of how bias can be identified, measured, and eventually mitigated.

The rest of this paper is organized as follows. In Section 2, we provide an overview of the literature related to this work and its impact. In Section 3, we discuss the methods used to analyze the admissions dataset, the details of the AI prediction model, and metrics related to fairness and bias. In Section 4, we describe the results of our analysis and identify the bias found in the AI predictive model. In Section 5 we provide some concluding remarks and directions for future research.

2 RELATED WORK

Students at urban universities face some unique challenges. For example, they are likely to experience the conflicting demands of work, family, and school, which can negatively affect their satisfaction with educational experiences [1]. Students, especially minority students, may face barriers that undermine academic achievement, reduce their sense of belonging, and interfere with degree completion [2]. Those who attend a university located in a city with a lack of affordable housing may struggle to find appropriate housing which can negatively affect their academic lives, health, and well-being [3]. Data collected about the impact of COVID-19 on urban college students showed that decreases in student earnings and household incomes led to significant disruptions in students' lives, and that these disruptions had an especially negative impact on first-generation students [4].

Recently, many universities have started to experiment with an admissions policy that allows students to apply without submitting standardized test scores. The hope is that these policies will address concerns of bias in the use of test scores as a metric to predict student success. One survey suggests that test-optional policies are changing enrollment demographics, especially with respect to underrepresented minorities, as Black and Hispanic students are 24% and 21% respectively more likely to apply to a school with a test-optional policy [5]. A second study found that test-optional admission increased the first-time enrollment by 10-12% of underrepresented minorities, and 6-8% by women [6]. Another study of liberal arts schools found that although test-optional policies enhance the perceived selectivity of a school, they did not increase the diversity [7]. Researchers are also exploring nuances of the policy. For example, what are the implications of giving students the option to submit test scores, when they are not required? Does it matter that some students choose not to submit scores because they are too low, while others choose not to take the test at all [8]?

Artificial intelligence tools, particularly machine learning algorithms, are increasingly used in higher education applications. Using AI to assist college admission process can be more objective and efficient [9]. Machine learning algorithms can also help universities better understand admission criteria and their impact in the admissions process [10] as well as admissions yield [11]. AI can also be used to predict the likelihood of admission for individual students [12-13] in some situations and can sometimes provide evidence for bias in human-centered admissions processes [14].

While AI tools bring many benefits to higher education applications, they can also be biased towards sensitive populations due to the intrinsic bias in existing datasets and in the algorithms themselves [15]. There are a variety of frameworks and taxonomies for classifying bias [16] and although some bias can be neutral/unobjectionable, many biases are problematic and require a response [17]. Identifying and categorizing bias in AI is a first step toward creating methods for mitigating bias, an area of active research [18-19].

3 METHODS and DATASET

This research was conducted using admissions data from the School of Science at a large urban research university. Students in the dataset were applying for admission from 2017-2021. The admissions process became test-optional during this period; first-year students applying for admission from Spring 2021 forward could choose not to submit standardized test scores. We built various machine learning-based models using different training sets and feature sets to predict which students were admitted directly into the School of Science ("Direct Admits"). Students who are categorized as "Not Direct Admits" are those who were admitted into the university but not into the School of Science, and those who were not admitted to the university at all.

Prior to test-optional admissions, Direct Admit decisions were based on grade point average (GPA) and standardized test scores. When the admissions process became test-optional, Direct Admit decisions for students who opted to exclude standardized test scores from their application were based on GPA and an assessment of "math readiness" based on performance in high-school math courses. The dataset used in this research contains approximately 6000 students who were required to submit test scores, and approximately 1700 students who were applying under the new test-optional policy.

The features used in the predictive model were: whether the student is a beginning student, the number of campuses to which the student applied, age, gender, race/ethnicity, whether the student is a first-generation student, whether the student is an in-state resident, GPA, and standardized test scores. Standardized test scores include both ACT and SAT scores but were normalized to SAT scores. In this dataset, ethnicity (Hispanic/Latino) was part of the race variable. Through exploratory analysis, three variables were chosen as variables to represent sensitive populations for further analysis: gender, race/ethnicity, and first-generation students.

This study consisted of an analysis of three groups of data:

- Group 1: Data over all five years.
- Group 2: Test-required cohort (Fall 2017 – Fall 2020).
- Group 3: Test-optional cohort (Spring 2021 – Fall 2021).

For each of these groups, we conducted three analyses:

- GPA was included, but standardized test scores were excluded.
- Standardized test scores were included, but GPA was excluded.
- Both GPA and test scores were included.

Finally, each analysis was repeated three times: once where the sensitive variable was gender, once where the sensitive variable was race, and once where the sensitive variable was first-generation students. The goal was to understand how changing the value of a single sensitive variable affected the accuracy of the models. In addition, we note that the test-optional cohort contains only one year of data, and that recent data is confounded by the pandemic.

Demographic data with respect to the sensitive variables is summarized in Table 1:

Table 1: Sensitive Variables, Full Dataset

Gender	Race	First-Generation
Male (36%)	White (60%)	No (72%)
Female (64%)	Hispanic (13%)	Yes (28%)
	Black (10%)	
	Asian (8%)	
	Other (9%)	

The predictive models were built using a linear support vector machine. Other machine learning models can also be applied here and may lead to better prediction accuracies, but that was beyond the scope of this research. We randomly selected two-thirds of the data for training each time, using proportionate stratified sampling with respect to the sensitive variable. All three of the sensitive variables were included in each model, but only one was treated as a sensitive variable for sampling in each experiment. Each model was then validated using five-fold cross-validation.

The resulting model was evaluated for potential bias, and to do this a bias metric needed to be identified for bias detection. Google's What-If Tool contains multiple well-known metrics for fairness, including demographic parity, equal opportunity, and equal accuracy [20]. Demographic parity is where each subgroup of the sensitive variable must have a similar percentage for positive classification. Equal opportunity is where each subgroup of the sensitive variable must have a similar percentage of correct predictions for positive classification. Equal accuracy is where each subgroup must have a similar percentage for correct prediction. While useful, these metrics are primarily used for determining fairness which is a broader concept than bias. For example, admission policies and practices can be unfair by these definitions, but that does not imply bias in the AI algorithm itself.

As our main goal in this work was to identify bias in the AI models (as opposed to evaluating the fairness of admissions policies), we evaluate each model based on overall accuracy, as well as accuracies in each subgroup of the sensitive variables: Male/Female, White/Non-White, and Non-First-Generation/First-Generation. A difference in accuracy over a threshold of 5% between sub-groups was considered biased. We also apply the same accuracy threshold to evaluate and compare the specificities and sensitivities of the model for different sub-groups. Biases detected in specificity and sensitivity often lead to real fairness issues in the admission process.

Note that the threshold value used here was selected somewhat subjectively. The goal here was to experiment with an approach to helping people evaluate AI models trained from admission data, rather than determining if this particular dataset carries bias or not. Different threshold values may be used for different applications or scenarios.

4 RESULTS AND DISCUSSION

The results presented here are based on the accuracy of the AI predictive models in various scenarios. An exploratory analysis of the data and corresponding AI models found that:

1. Standardized test scores play a major role in admissions decisions.
2. Gender and First-Generation are two other variables that play prominent roles in admissions decisions.
3. Many students who would not be admitted under test-required policies would be admitted under test-optional policies, including more students from sensitive populations.

We also evaluated the models for bias with respect to sensitive variables. The results were:

1. **Gender**: The model incorrectly predicted admission for women more often than for men in the test-optional cohort.

2. **First-Generation**: The model incorrectly predicted rejection for first-generation students more often than for non-first-generation students in both the test-required and test-optional cohorts. Both models also incorrectly predicted admission for non-first-generation students more often than for first-generation students.
3. **Race/Ethnicity**: The model incorrectly predicted admission for white students more than non-white students in both cohorts. The model also incorrectly predicted rejection for non-white students more than white-students in the test-optional cohort.

4.1 General Analysis

The general analysis focused on two scenarios: the test-required cohort where the predictive model included both GPA and standardized test scores, and the test-optional cohort where the predictive model included GPA but excluded test scores. Table 2 shows the permutation importance of the top ten variables included in the experiment.

Table 2: Permutation Importance

Test-Required	Test-Optional
Test Scores .278	GPA .246
Female .070	Male .082
First-Gen .061	First-Gen .081
Not First-Gen .041	Female .062
Male .031	Not First-Gen .051
GPA .028	Out-Of-State Res .020
Out-Of-State Res .010	In-State Res .017
In-State Res .008	Black .016
Hispanic/Latino .005	Asian .007
Age .005	2+ Races .007
Cross Val: .88 (+/- .03)	.90 (+/- .03)

From this table, it is clear that Test Scores are the dominant variable for predicting Direct Admits in the test-required cohort and are well above the importance of GPA. For the test-optional cohort, GPA is the most important variable by far, as expected. Gender and First-Generation are the other important variables in both models. It is somewhat surprising that, for the test-required cohort, test scores played a pivotal role in admissions, and GPA only played a minor role. It means that the decision of making test scores optional is perhaps more significant than many people believed, particularly for major urban public universities.

To better understand how the test-optional policy might change which students are admitted, we plotted in Figure 1 the values of GPA and Test Scores of all students in the test-required cohort who were not Direct Admits. Under the test-required policy, students with Test Scores under 1080 or a GPA under 3.0 would not be Direct Admits; these are the students in all but the upper-right quadrant. The test-optional policy requires a higher GPA of 3.3 to be a Direct Admit, but many of those in the lower right quadrant clear that threshold and would become Direct Admits. These are the students whose GPAs are high enough under the test-optional policy to be Direct Admits, but whose Test Scores would exclude them under the test-required policy.

The significance of the admission policy change is evident in Figure 1 where the majority of the students who would not be directly admitted under a test-required policy may be admitted under a test-optional policy. This result is not a statement about which policy is better. It simply shows that a change from test-required to test-optional can fundamentally change the admissions decisions for a large number of students.

Table 3 shows how students in sensitive populations would be affected by the change from a test-required to a test-optional policy. More women, non-white students, and first-generation students would meet admissions thresholds under a test-optional policy. This data suggests that not only may there be more students admitted overall, but many of those students are also students who are part of a population more likely to report negative academic experiences. More research needs to be done to understand how to better prepare and support incoming students so that they might be successful academically.

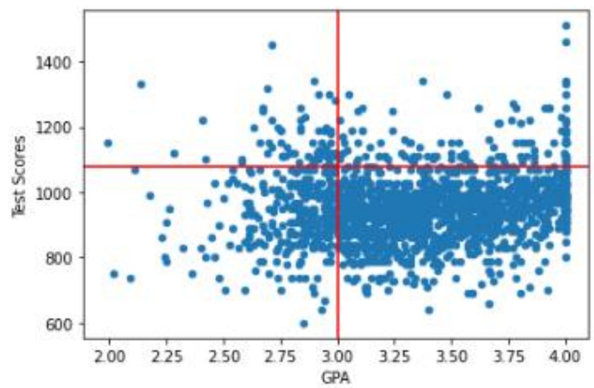

Figure 1: GPA and Test scores for Test-Required Cohort who were NOT Direct Admits

Table 3: Demographics of Students in Test-Required Cohort Who Meet Admission Thresholds

	Test-Required Threshold	Test-Optional Threshold
Gender		
Female:	59%	67%
Male:	41%	33%
Race		
Non-White:	31%	38%
White:	69%	62%
First-Gen		
First-Gen:	21%	27%
Non-First-Gen:	79%	73%

4.2 Bias in Predictive Models

In the process of identifying bias with respect to three sensitive variables in the predictive models, the focus was on two scenarios: the test-required cohort where the predictive model includes both GPA and test scores and the test-optional cohort where the predictive model excludes test-scores. Two of the sensitive variables are Gender and First-Generation, two variables with high permutation importance in both models. The third sensitive variable chosen was Race/Ethnicity, which was categorized as White and Non-White. The bias we tried to detect was a difference of 5% or more in overall prediction accuracy, or in specificity or sensitivity, between values of the sensitive variable. In both the test-required cohort and the test-optional cohort, we did not find any significant differences in overall prediction accuracies between the sensitive variables. But differences in specificity and/or sensitivity were present in both cohorts. In this application, specificity is a measure of students who were incorrectly admitted, and sensitivity is a measure of students who were incorrectly denied admission.

The results for Gender are summarized in Table 4. The predictive model for the test-required cohort where Gender is the sensitive variable does not show bias. The model for the test-optional cohort, however, shows bias in specificity. The model predicts that women will be admitted more often than they are in reality, more than it makes the same error for men.

Table 4: Model Bias when Gender is Sensitive Variable

	Overall	Male	Female
Test-Required Cohort			
Model Accuracy:	.882	.892	.876
Specificity:	.817	.825	.813
Sensitivity:	.921	.919	.922
Cross Val: .88 (+/- .03)			
Test-Optional Cohort			
Model Accuracy:	.900	.896	.902
Specificity*:	.690	.730	.667
Sensitivity:	.958	.949	.963
Cross Val: .90 (+/- .03)			

The results for First-Generation are shown in Table 5. The predictive models for both the test-required and the test-optional cohorts show bias. Both models predict that non-first-generation students are admitted more often than in reality, and both models predict that first generation students are rejected more often than in reality.

Table 5: Model Bias when First Gen is Sensitive Variable

	Overall	Non-First-Gen	First-Gen
Test-Required Cohort			
Model Accuracy:	.877	.883	.862
Specificity*:	.810	.783	.853
Sensitivity*:	.917	.931	.872
Cross Val: .88 (+/- .03)			
Test-Optional Cohort			
Model Accuracy:			
Specificity*:	.913	.923	.881
Sensitivity*:	.783	.753	.848
Cross Val: .90 (+/- .03)	.952	.968	.895

The results for Race are presented in Table 6. Both models show bias when Race is the sensitive variable. Both models predict that white students will be incorrectly admitted more than non-white students. The model for the test-optional cohort shows that non-white students will be incorrectly rejected more than white students.

Table 6: Model Bias when Race is Sensitive Variable

	Overall	Non-White	White
Test-Required Cohort			
Model Accuracy:	.892	.895	.891
Specificity*:	.847	.900	.790
Sensitivity:	.919	.890	.933
Cross Val: .88 (+/- .03)			
Test-Optional Cohort			
Model Accuracy:	.905	.896	.911
Specificity*:	.726	.804	.649
Sensitivity*:	.963	.934	.981
Cross Val: .90 (+/- .03)			

It ought to be noted that biases identified here are based on ground truth data. They are not indications of any bias in the actual admission policies or practices. Rather it is an evaluation of errors present in the machine learning algorithms and how these errors were distributed unevenly for different student populations. These biases need to be understood if the models are to be used in the future for an AI-assisted admissions process, or other similar applications.

5 CONCLUSIONS

This research analyzed admissions data at a large urban research university using machine learning-based AI models to predict whether a given student was directly admitted into the School of Science. The dataset is interesting because the university’s admissions policies changed from test-required to test-optional during this period. This research contributes to a better understanding of the variables that play an important role in admissions decisions and shows that there is a significant change in admitted student body under a test-optional policy. In addition, the predictive models create new bias with respect to different populations under sensitive variables of gender, race, and first-generation. The conclusions we can draw from this study include:

1. Test-optional policies may significantly alter the demographics of the students who are admitted. The impacts of this need additional study.
2. AI models are increasingly used to create and understand policy in higher education, so it is critical to carefully analyze these models for errors and biases. Further, the bias

may not be apparent when examining the overall accuracy of the model, so more thorough evaluations must be done.

3. Although policies may reflect social bias, separate metrics for bias must be developed when evaluating AI models themselves.

One avenue for future work in this area is to analyze a more expansive dataset. This research focused on School of Science data, but are the results the same when applied to a university-wide dataset? In addition, the test-optional policy is recent, and this dataset only includes one-year of test-optional data. Continuing this analysis over the next few years could provide some interesting insight and more robust results. Similarly, test scores for the test-optional cohort were excluded entirely in this analysis. The overall accuracy of these models was high, so this was sufficient for the models, though not entirely reflective of reality. Under a test-optional policy, test-scores can be included in admissions decisions, if this is what a student prefers. A model that includes this nuance may lead to additional insights about the effects of test-optional policies.

A second area for future work is in bias mitigation. When bias in AI models is detected, how can it be corrected? Are there scenarios under which it should not be corrected? What is the interplay between social bias reflected in admissions policies, and bias in AI models?

As AI algorithms become more widely used, they can lead to greater accuracy, consistency, time-savings, and understanding in many domains. AI is a powerful tool and must be used with the understanding that as it is being used to solve problems, it can also introduce new problems. Techniques for identifying, measuring, and mitigating these problems are critical.

ACKNOWLEDGEMENTS

The authors want to thank the following people at IUPUI for their help in providing the dataset used in this study, for answering questions about the dataset format, and for explaining policies related to admissions: Jane Williams, Joe Thompson, Steve Graunke, Matt Moody, Norma Fewell, and Lori Hart.

REFERENCES

[1] Leslie B. Hammer, Tenora D. Grigsby, and Steven Woods. 1998. The Conflicting Demands of Work, Family, and School Among Students at an Urban University. *The Journal of Psychology* 132, 1 (1998), 220-226. DOI: https://doi.org/10.1080/00223989809599161

[2] Shaun R. Harper, Edward J. Smith, and Charles H. F. Davis, III. 2018. A Critical Race Case Analysis of Black Undergraduate Student Success at an Urban University. *Urban Education* 53, 1 (2018), 3-25. DOI: https://doi.org/10.1177/0042085916668956

[3] Luisa Sotomayor, Derya Tarhan, Marcelo Vieta, Shelagh McCartney, and Aida Mas. 2022. When students are house-poor: Urban universities, student marginality, and the hidden curriculum of student housing. *Cities* 124 (May 2022), 103572. DOI: https://doi.org/10.1016/j.cities.2022.103572

[4] Nuria Rodriquez-Planas. 2022. Hitting where it hurts most: COVID-19 and low-income urban college students. *Econ Ed Rev* (Apr. 2022), 102233. DOI: https://doi.org/10.1016/j.econedurev.2022.102233

[5] Jon Edelman. 2022. Survey: Test-Optional is Appealing to Minority Students. Retrieved Oct. 12, 2022 from https://www.diverseeducation.com/students/article/15292751/survey-finds-testoptional-policies-a-significant-motivator-for-minority-college-applicants

[6] Christopher T. Bennett. 2022. Untested Admissions: Examining Changes in Application Behavior and Student Demographics Under Test-Optional Policies. *American Educational Research Journal* 59, 1 (Feb. 2022), 180-216. DOI: https://doi.org/10.3102/00028312211003526

[7] Andrew S. Belasco, Kelly O. Rosinger, and James C. Hearn. 2015. The Test-Optional Movement at America's Selective Liberal Arts Colleges: A Boon for Equity or Something Else? *Educational Evaluation and Policy Analysis* 37, 2 (June 2015), 206-223. DOI: https://doi.org/10.3102/0162373714537350

[8] Zhi Liu and Nikhil Garg. 2021. Test-optional Policies: Overcoming Strategic Behavior and Informational Gaps. In *Proceedings of the ACM Equity and Access in Algorithms, Mechanisms, and Optimization (EEAAMO '21)*, Oct. 5-9, 2021, ACM Inc., New York, NY, 1-13. DOI: https://doi.org/10.1145/3465416.3483293

[9] Austin Waters and Risto Miikkulainen. 2014. GRADE: Machine Learning Support for Graduate Admissions. *AI Magazine* 35, 1 (Spring 2014), 64-75. DOI: https://doi.org/10.1609/aimag.v35i1.2504

[10] Nicholas T. Young and Marcos D. Caballero. 2019. Using machine learning to understand physics graduate school admissions. arXiv: 1907.01570. Retrieved from https://arxiv.org/abs/1907.01570

[11] Joseph Jamison. 2017. Applying Machine Learning to Predict Davidson College's Admissions Yield. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education (SIGCSE '17)*, Mar. 8-11, 2017, Seattle, Washington. ACM Inc., New York, NY, 765-766. DOI: https://doi.org/10.1145/3017680.3022468

[12] Mohan S. Acharya, Asfia Armaan, and Aneeta S. Antony. A Comparison of Regression Models for Prediction of Graduate Admissions. In *2019 International Conference on Computational Intelligence in Data Science (ICCIDS)*, Feb. 21-23, 2019, Chennai, India. IEEE, 1-5. DOI: 10.1109/ICCIDS.2019.8862140

[13] Ch. V. Raghavendran, Ch. Pavan Venkata Vamsi, T. Veerraju, and Ravi Kishore Veluri. 2021. Predicting Student Admissions Rate into University Using Machine Learning Models. In: D. Bhattacharyya and N. Thirupathi Rao, *Machine Intelligence and Soft Computing*. Advances in Intelligent Systems and Computing, vol 1280. Springer, Singapore. DOI: https://doi.org/10.1007/978-981-15-9516-5_13

[14] A.J. Alvero, Noah Arthurs, Anthony Lising Antonio, Benjamin W. Domingue, Ben Gebre-Medhin, Sonia Giebel, and Mitchell L. Stevens. 2020. AI and Holistic Review: Informing Human Reading in College Admissions. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '20)*, Feb 7-9, 2020, New York, NY. ACM Inc., New York, NY, 200-206. DOI: https://doi.org/10.1145/3375627.3375871

[15] Barbara Martinez Neda, Yue Zeng, and Sergio-Gago-Masague. 2021. Using Machine Learning in Admissions: Reducing Human and Algorithmic Bias in the Selection Process. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (SIGCSE '21)*, Mar. 13-20, 2021. Virtual. ACM Inc., New York, NY, 1323. DOI: https://doi.org/10.1145/3408877.3439664

[16] Brian d'Alessandro, Cathy O'Neil, and Tom LaGatta. 2017. Conscientious Classification: A Data Scientist's Guide to Discrimination-Aware Classification. *Big Data* 5, 2 (Jun 2017), 120-134. http://doi.org/10.1089/big.2016.0048

[17] David Danks and Alex John London. 2017. Algorithmic Bias in Autonomous Systems. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI '17)*, Aug. 19-25, 2017. Melbourne, Australia, 4691-4697. DOI: https://doi.org/10.24963/ijcai.2017/654

[18] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. 2019. A Survey on Bias and Fairness in Machine Learning. arXiv: 1908.09635. Retrieved from https://arxiv.org/abs/1908.09635

[19] Simon Caton and Christian Haas. 2020. Fairness in Machine Learning: A Survey. arXiv: 2010.04053. Retrieved from https://arxiv.org/abs/2010.04053

[20] James Wexler, Mahima Pushkarna, Tolga Bolukbasi, Martin Watterberg, Fernanda Viegas, and Jimbo Wilson. 2019. The What-If Tool: Interactive Probing of Machine Learning Models. arXiv: 1907.04135. Retrieved from https://arxiv.org/abs/1907.04135