Machine Learning: Predicting Graduation Rates of Virginia High Schools

Algorithmic Bias in Education

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
Lillian Cochrane

May 3rd 2022

Technical Team Members: Matthew Gerace, Matt Koehler

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

Caitlin D. Wylie, Department of Engineering and Society

Kathryn A. Neeley, Department of Engineering and Society

Daniel G. Graham, Department of Computer Science

Introduction

My technical research topic is going to be focusing on a group project that I did for a machine learning class. The goal was to train a program that predicted the on-time graduation rate of Virginia high schools, and therefore create a tool that educators could use to understand how certain factors could contribute to overall student performance. Based on the Census Bureau's 2017 Current Population Survey, the median income of high school dropouts was $6,000 dollars less than those who graduated high school but did not enter higher education (Trends in High School Dropout and Completion Rates in the United States, n.d). By focusing on improving graduation rates, schools are setting students up for more financial stability in their careers. Understanding what factors might contribute to graduation rates can help with understanding how to help students. My technical research paper will go over the models and methods used to create this program.

For my STS research topic, I want to focus on algorithmic bias, specifically in the field of education. While machine learning is becoming more ubiquitous, the ethical issue of how models can perpetuate bias has gained more attention (Kirkpatrick, 2016). If and when artificial intelligence becomes more pervasive in the classroom, it will have the potential to make decisions that can end up affecting student and teacher performance. So, there should be an understanding of the effects of algorithmic bias in education in particular, and how it can be avoided or mitigated. I want to analyze case studies to illustrate some of the effects of algorithmic bias, and also analyze techniques recommended by experts using the ethics of care framework. This is to hopefully understand best practices to apply when creating these technologies, rather than advising not to create them at all.

Technical Topic

My team was interested in providing educators with an understanding as to how certain factors might affect the performance of their school's student population in general. We wanted to help schools who might not know how they might be able to improve their school's overall performance. We used data such as free and reduced lunch eligibility rate and number of students enrolled in advanced programs as features in our predictive models. As shown by the following examples, machine learning is increasingly being used by schools. Schools have already begun using predictive algorithms as tools, such as how some colleges have systems that indicate when a student may need extra help, or that help students plan their degrees (Ekowo & Palmer, 2017). Another example are the advances being made in precision learning, which seek to understand student's unique learning needs in order to provide more individualized instruction (Luan & Tsai, 2021).

I worked with two teammates to train and test different types of regression models. We were guided by a teaching assistant who answered questions we had about the project. Our goal was to train and test newer and older models, and compare them to identify the best performing model. We used root mean squared error (RMSE) as our performance metric, with the lowest RMSE indicating the lowest error. We chose to test and train linear regressors, decision trees, random forest from the scikit learn library, and neural networks from the keras library, as well as to combine some models in a voting regressor. We hoped to find a best performing model and lower its RMSE, as well as being able to analyze models to understand how some features contributed to decision making. Next steps that involve the life of the project past the due date would involve improving the chosen model to allow schools to predict their graduation rates by inputting their own data, as well as see how the predicted rate changes when altering their own data points. This would give school an idea of what student needs could be. However, the caveat is that correlation does not mean causation. The model would not be foolproof, as there are many factors that could influence and be influenced by graduation rate.

The earliest research paper that I could find that discussed using machine learning to predict graduation outcomes was published in 2017 (Pang et al.). In the paper, the authors propose and evaluate their predictive model, which is designed to be used by higher education institutes. The proposed model uses support vector machines to classify whether or not a student is expected to graduate on time. The model combines the predictions of several support vector machines through ensemble learning. The most recent research paper that I could find that involved using machine learning techniques to predict graduation outcome was published in 2022 (Demeter et al.). In this paper, the authors describe their machine learning model that was designed to predict whether or not students who are pursuing a degree in higher education for the first time will graduate on time. The model that the authors eventually chose is a two-level model, consisting of two random forest classifiers. The first level predicts whether or not a student will graduate. The second classifier predicts whether or not a student will graduate on time. Comparing the models that we tested to the similar models above, our models predict the graduation outcomes of an entire high school, rather than individual students. Additionally, we trained regression models, while the two previously described models are classifiers. Our project was hoping to help schools (rather than individual student) identify and understand contributing factors to their graduation rate, so they might understand where they need to focus effort.

STS Topic

For my STS research topic, I want to address the problem of algorithmic bias in education. I plan to apply the ethics of care framework, which focuses on understanding the needs of people in a system, and considers everyone involved to be responsible for others in the system (Taylor, 2020). This topic is important to discuss because there have already been instances of algorithms used determining student success, as well as accompanying public backlash for the bias inherent in these technologies. For example, in 2020, A-level students in the UK received grades that were calculated by an algorithm (Smith, 2020). On twitter, some students shared how they received lower grades than they usually expected. While the Office of Qualifications and Examinations Regulation released a statement claiming that there was no evidence of introduced bias, it was found that the algorithm unfairly lowered the grades of disadvantaged students. This is mostly due to how the algorithm would adjust a student's score to match past score distributions from their school (Coughlan, 2020). This unfairly capped the scores of students who attend schools that had lower past performance, which in itself could be due to a number of systemic factors. Due to public pressure, the Scottish Government decided that grades would be reassigned based on teacher judgement. The UK government decided to follow suit. Scholars are aware of how students are using tools that are built using data that represents historical and systemic bias, and it is considered one of the larger ethical challenges involved with introducing AI to the classroom. It can be seen when using google translate, with how certain jobs become gendered, and with tools that are used to introduce personalized learning and predictive data (Akgun & Greenhow, 2021).

I think seeing research by experts, both in education and artificial intelligence, will be useful when writing my paper. Experts in artificial intelligence and machine learning have research on methods that could reduce bias during the development process, and can help provide an understanding as to how bias is introduced into these technologies. Educators have more first-hand experience using the technology in a classroom, and therefore might know more about the effects on students and teachers. They might also have mitigation techniques for when technology has already been released. Analyzing mitigation techniques through the lens of ethics of care will help gain a better understanding of how developers can design AI with the goal caring about users and their needs. Ethics of care allows for focus to be spent on caring for marginalized groups, who are most affected by algorithmic bias. Additionally, I will use case studies that provide real examples of algorithmic bias and its effects. Analyzing case studies through ethics of care can help better understand how people involved in the use of a technology can care for each other.

In the academic article "Algorithmic reparation", the authors argue that current corrective standards used in Machine Learning are insufficient because they focus on equality rather than equity (Davis et al., 2021). Social equity is taking into account differential needs, which is in contrast with social equality, that aims to treat everyone the same regardless of their circumstances. One method that the authors suggest to promote equity is distributed AI power, where developers and stakeholders engage with each other during the development process, with stakeholders teaching developers about their experiences. From the standpoint of ethics of care, understanding how to care for users would involve understanding their experiences. Distributed AI power promotes the same values that constitute ethics of care.

The book “Weapons of Math Destruction” includes several case studies that analyze algorithmic bias (O’Neil, 2017). One such case study involves teachers who were fired due to receiving poor results in their mathematically generated IMPACT evaluations. There were several issues with the system, such as using little data and not receiving any feedback as to if the analytics were “on track” or not. This is antithetical to the ethics of care. Teachers could have been cared for better if they understood how or why their evaluations were so low, as this provides a chance for improvement. Instead, the algorithm's evaluations were trusted implicitly, harming teachers.

Conclusion

The technical deliverable is a program that analyzes multiple machine learning models performance in predicting on time graduation rates of Virginia high schools. The STS deliverable is a better understanding of the effects of algorithmic bias in education, as well as analysis of recommendations from professionals, both using the lens of the ethics of care framework. By having more knowledge on the bias that algorithms used in education can perpetuate, it can be better understood how to create regulations for use and how to effectively mitigate bias in the technology.

Reference List

Akgun, S., & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*. https://doi.org/10.1007/s43681-021-00096-7

Baker, R. S., & Hawn, A. (2021). Algorithmic Bias in Education. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-021-00285-9

Coughlan, S. (2020, August 14). Why did the A-level algorithm say no? *BBC News*. Retrieved May 3, 2022, from https://www.bbc.com/news/education-53787203

Davis, J. L., Williams, A., & Yang, M. W. (2021). Algorithmic reparation. *Big Data & Society*, 8(2), 205395172110448. https://doi.org/10.1177/20539517211044808

Demeter, E., Dorodchi, M., Al-Hossami, E., Benedict, A., Slattery Walker, L., & Smail, J. (2022). Predicting first-time-in-college students' degree completion outcomes. *Higher Education*. https://doi.org/10.1007/s10734-021-00790-9

Ekowo, M., & Palmer, I. (2017). Predictive Analytics in Higher Education: Five Guiding Practices for Ethical Use. *New America*, 3. https://reportcenter.highered.texas.gov/sites/docid/cache/file/AC2E2644-9DF0- 2DF2467C1FD0F2571918.pdf

Kirkpatrick, K. (2016). Battling algorithmic bias. *Communications of the ACM*, 59(10), 16–17. https://doi.org/10.1145/2983270

Luan, H., & Tsai, C.-C. (2021). A Review of Using Machine Learning Approaches for Precision Education. Educational Technology & Society, 24(1), 250–266. https://www.jstor.org/stable/26977871

O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown.

Pang, Y., Judd, N., O'Brien, J., & Ben-Avie, M. (2017). Predicting students' graduation outcomes through support vector machines. *2017 IEEE Frontiers in Education Conference (FIE)*. https://doi.org/10.1109/fie.2017.8190666

Smith, H. (2020). Algorithmic bias: should students pay the price? *AI & Society*, 35(4), 1077– 1078. https://doi.org/10.1007/s00146-020-01054-3

Taylor, L. (2020). The price of certainty: How the politics of pandemic data demand an ethics of care. *Big Data & Society*. https://doi.org/10.1177/2053951720942539

Trends in High School Dropout and Completion Rates in the United States. (n.d.). National Center For Education Statistics. https://nces.ed.gov/programs/dropout/intro.asp#ref2