Josette Riep and Annu Prabhakar

8 Toward Bias-free Artificial Intelligence for Student Success in Higher Education

Abstract: Artificial Intelligence (AI) has continued to increase its footprint in (HCI). Systems that span every aspect of daily life have become increasingly reliant on algorithms to identify products, promote opportunities, and guide strategy and operations. The field of higher education has seen a dramatic increase in the use of AI to drive recruitment and retention decisions. Persistence predictors and risk factors for example have garnered broad use across institutions in some cases without a thorough assessment of the impact on underrepresented groups in areas such as science technology, engineering, and mathematics (STEM). Tangible examples of bias that exist within algorithms are too often developed by teams that lack inclusive representation or an inclusive approach through training and design. The study reported on in this chapter analysed US educational data and focused on leveraging AI to remove bias and inform the design of meaningful solutions to facilitate innovative pathways towards STEM graduation attainment for an increasingly diverse student body. The background and details of the study are outlined in the chapter, together with explanations of the methods chosen and an evaluation of the models chosen to minimize bias.

Keywords: Artificial intelligence; Big data; Machine learning; Neural networks (computer science); Bias-free language; Cultural pluralism; STEM

Introduction

Artificial intelligence (AI) and deep learning are transforming the way everyone engages, interacts and makes decisions. By executing system driven intelligent tasks, AI has a growing impact in multiple sectors including entertainment, marketing, the world of work, healthcare, and education.

The importance of science, technology, engineering and medicine (STEM) in society and education has been emphasized by many, including the Smithsonian Science Education Center.

> Four billion people on the planet use a mobile phone, while 3.5 billion people use a toothbrush. In the past two years, 90% of all of the world's data has been generated. NASA plans to set foot on Mars in the next 20 years, and driverless cars are already being tested in Europe. The future is here, and it requires a citizenry fluent in science, technology, engineering, and math (STEM) (Smithsonian Science Education Center 2024).

https://doi.org/10.1515/9783111336435-009

Yet in the years 2011 to 2017 there was an alarming overall decline in enrolments in STEM by multiple groups, with particularly stark declines for Blacks and African Americans (National Center for Educational Statistics 2023). The gap is real but nuanced by the complexity of today's environment. An increasing demand for professionals in technology and engineering is surpassed only by an inability to fill the demand. Companies are in dire need of professionals in fields like data science, software development and cybersecurity (Goupil et al. 2022). The pipeline challenge is real and spans many demographic areas. Data from the US National Center for Educational Statistics (NCES) longitudinal study from 2011 – 2017 provided in Table 8.1 shows that almost half of all students (48.99%) who pursued a degree in science, engineering and math attained the degree (Bryan et al. 2019). The numbers are significantly worse for Black/African Americans where only 30.53% attained the degree. Furthermore, 36.27% of Black/African American students left with no degree compared to 22.57% across all races. The gap highlights the need to determine the factors contributing to the disparity and to explore ways in which bias and AI might impact student success.

Table 8.1: NCES STEM Degree Attainment Longitudinal Study Data (Bryan et al. 2019)

	Attained bachelor, degree (%)	**Attained associate degree** (%)	**Attained certificate** (%)	**No degree Still enrolled** (%)	**No degree l Left without return** (%)	**Total** %
Science, engineering, and math (all)	48.9958	10.4094	4.1211	13.8993	22.5743	100
Science, engineering, and math (Black / African American)	30.5329	8.8300	5.9101	18.4473	36.2797	100

In addition to challenges demonstrated by the large number of students who fail to graduate with a STEM degree, there is a surprising trend in the overall year to year growth across most demographic areas. Figure 8.1 plots the growth from 2013 to 2021 in STEM related programs by race and ethnicity. There are only two demographics, Hispanic and Asian, that are seeing sustained growth in degree attainment. African Americans show overall low numbers. Perhaps most shocking is the dramatic drop in STEM degree attainment by whites, perhaps partially explained by Covid. In examining how outcomes for underrepresented groups in STEM are improved, the broader question should be asked about how to improve outcomes in STEM generally.

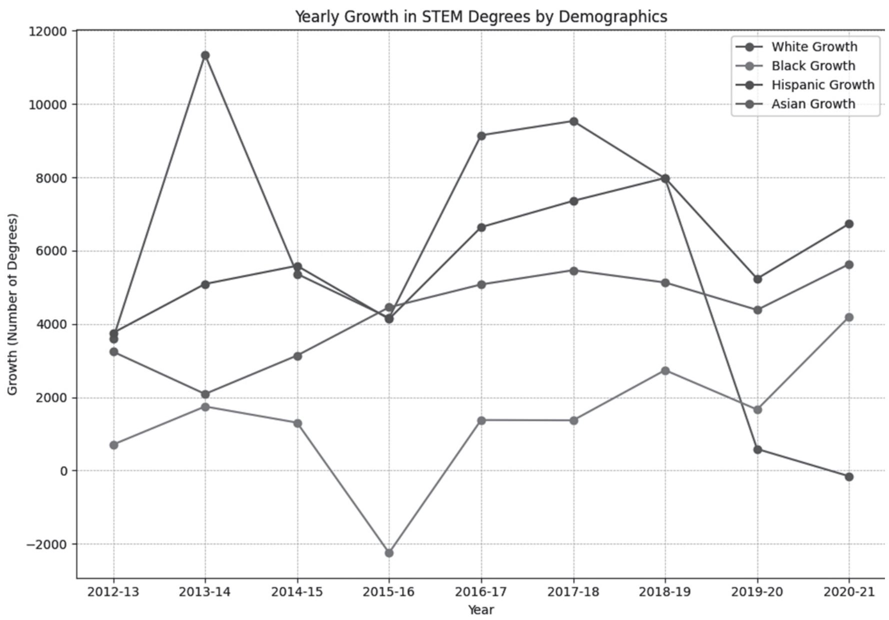

Figure 8.1: STEM growth by demographic (Data drawn from NCES 2023)

In addition, the track record on the use of AI and machine learning with African Americans has not been good. There have been many documented cases of AI bias. AI engines have disproportionately identified African Americans as animals or criminals. Incredibly, examples include an AI driven folder named gorillas being used to organize photos of Black people, which makes the disgraceful reality of inbuilt bias abundantly clear (Metz 2021).

This chapter reports on a pilot study conducted to assess potential bias when leveraging AI to support student success. Recruitment and retention are two factors but not the only evaluated in higher education in the US. By conducting the study on the use of AI and machine learning and the identification of impact factors, it was hoped to identify the potential ways and methods that bias is introduced and subsequently what can be done to raise awareness of the issues and provide solutions so that everyone can benefit from the great potential AI in higher education promises to deliver. The study was conducted leveraging available longitudinal data from the National Center for Education Statistics (NCES) within the Department of Education in the US. The research undertaken relied heavily on the ability to create a predictive model or set of models derived from the processing and analysis of data spanning multiple years. Given the vast amount of information available through NCES, it served as a foundational source for data analysis and training.

The Use of Artificial Intelligence in Higher Education

Enhanced personalization, dynamic content delivery, gaming, and predictive analytics are all driven by increased adoption of AI and machine learning (ML). AI and machine learning have transformed higher education through a wide range of applications. Student recruitment and retention initiatives in particular have garnered increased attention as educational institutions work to leverage new technologies including AI to improve student success. Systems such as Civitas Learning tout the ability to leverage machine learning to impact student attainment. Scientists and experts who work at Civitas Data and elsewhere have highlighted specifically the need to transition student success from risk factor assessment to impact factor analysis. They emphasize the need to conduct data mining and analysis within an ecosystem that encompasses the full student life cycle and promote what they describe as a caring culture that leverages a rich knowledge base of information that ties success to institutional activities (Kil, Baldasare, and Milliron 2020).

Researchers undertaking a systematic review of artificial intelligence in education in 2019 highlighted hundreds of higher education-focused AI implementations that spanned several core areas including intelligent tutoring systems (ITS); adaptive group formations for collaboration; intelligent virtual reality; and recruitment and retention strategies (Zawacki-Richter et al. 2019). The following content highlights some AI applications in higher education and refers to research undertaken into their effectiveness.

Intelligent Tutoring Systems

The systematic literature review by Zawacki-Richter et al. identified twenty-nine studies focused on intelligent tutoring systems including gap diagnosis, strength identification, learning material curation and collaboration (Zawacki-Richter et al. 2019). There are several examples of implementations in the area of STEM including tools such as ALEKS and Metatutor. ALEKS is used in multiple institutions to provide personalized online dynamic learning in the areas of mathematics and science. A 2021 study highlighted the effectiveness of ALEKS by studying 9,238 students from more than 50 institutions. The study found that ALEKS was most effective as a supplement to traditional instruction. The results of this study emphasized the ability to leverage ITS as a supplemental tool and not a replacement for instructor led courses. Stratification by populations based on race, gender or ethnicity were not the primary focus.

Additional research leveraging ALEKS has focused on analyzing drop-out and graduation rates and the impact of procrastination. The research attempted to make a correlation between time spent on a task and successful completion by analyzing student cohorts from 2014-2019 who were required to complete bridge course material in ALEKS. By analyzing student activity and comparing the amount of activity along with the timeliness of participation, the research sought to apply an empirical assessment that derived poor academic performance as a function of procrastination (Harati et al 2021). The approach provided an example of where bias might play a key role in the categorization applied to students with very little inclusion of other variables that could be contributing factors.

Student Success

The use of predictive analytics and systems that focus on data driven success models have continued to expand in use. Tools like Civitas Learning are used in multiple institutions in an effort to drive student success across a broad and diverse population of students. Since 2017 Civitas has participated in a collaborative initiative focused on promoting equity and degree completion. *Driving Toward a Degree* is a research collaborative for increasing student success across the higher education landscape. Data is collected and analyzed to offer insights to institutions for enhancing student support and success, retention, and completion. The survey responses shed light on perceptions across multiple institutions who have acknowledged the same challenges since 2017. The 2021 study evaluated responses from 2800 respondents over 1300 institutions (Shaw 2021). Questions on racial equity help gauge strategic direction and institutional priorities across institutions although there is no detail on the role that predictive analytics play in achieving equity and enhancing student success. Many companies offer proprietary algorithms that on the surface seem promising but lack detail to determine if the results are effective. Some highlight the importance of analyzing student success programs by effectively measuring impact factors (Carmean, Baer, and Kil 2021). The challenge and opportunity reside in ensuring that the impact factors being examined are holistic and representative for all groups. How for example are underrepresented groups impacted when the algorithms used are based on training data sets that lack representation? Is it an accurate predictor for all or does unintended bias play a key role?

Use of Intelligent Virtual Assistants

Neural networks or the practice of automating and processing data regarding pathways to support everything from autonomous vehicles to automated photo tagging to medical diagnoses has become one of the world's fastest growing industries. In the US alone, it was projected in 2019 that revenues would increase from 9.5 billion in 2018 to 118.6 billion in 2025 (Heinonen 2019). Page and Gelbach undertook a progressive experiment in 2016 at Georgia State University (GSU) to determine whether AI was capable of replacing human decision-making processes in addressing an individual's personal needs. They delved into the use of neural networks to facilitate student pathways to college by applying the analogy of autonomous cars to facilitate active learning and machine-driven response to external inputs. The student journey from high school to college over the summer was investigated in an effort to reduce summer melt or failure to proceed by implementing a new AI system between April and August of 2016 to determine if machine-oriented activities could automate human decision-making activities. Conversational technologies and AI were used to support students as they worked through the pre-enrollment, enrollment and orientation phases required for successful matriculation and university entrance.

Key matriculation and entrance activities including the financial aid application process, submission of transcripts, required health records, loans and tuition were included, and a virtual assistant focused on three key areas: required pre-enrolment tasks; reliable data regarding student accomplishments; and initial responses to frequently asked questions from students related to those tasks. The overall findings of the research study were encouraging. GSU was able to maintain a high level of engagement with only 6.6% of students opting out of the program Through the use of artificially intelligent virtual assistants, The hypothesis that AI driven intervention would increase successful matriculation was proven as students who participated in the program had an increased enrolment of 3.3 percent over the group that did not participate in the program resulting in a 21% reduction in failed summer matriculation. The project provided a glimpse into the new possibilities that AI could bring as well as the potential for improved outcomes and expanded impact (Page and Gelbach 2017).The student outreach program continues to operate at GSU through the Panther Experience which:

> is designed to maximize your success both inside and outside the classroom. Being a truly engaged Panther means having a diverse holistic experience built upon the POUNCE Pillars… the foundation upon which Georgia State Panthers build the necessary skills and talents to achieve all of your personal and professional life goals. Each pillar aligns with

specific programs, activities, and services available to all students: Promote Panther Pride; Own Your Academic Journey; Understand Ethics and Integrity; Nurture Healthy Habits; Connect College to Career; Embrace Campus, Cultures and Community (Georgia State University n.d.)

Bias Associated with Artificial Intelligence

Damaging examples of machine learning bias or algorithm bias have been well publicized (IBM 2023). Beyond the obvious are the subtler examples that go undetected. There are situations where the impact of AI models on staff and student recruitment, promotion, retention, educational opportunities, and educational support to name a few is not immediately known or identifiable. What happens when AI is used to identify students likely to fail particular programs or those more likely to succeed? Do models disproportionately steer African American students towards vocational opportunities and away from advanced technology fields such as cybersecurity? There are very real challenges that must be understood and addressed.

There are some leaders in this space who are working to remove bias from AI. Timnit Gebru is one. An AI expert and leader for more inclusive AI practices, Gebru has been an advocate for a strong ethical approach to AI that includes fairness, transparency, and accountability. Through a rich focus on ethics, Gebru has highlighted the lack of diversity in AI and the need to increase the number of African Americans and other underrepresented groups developing AI models (Metz 2021). Gebru and others postulate that models are in part biased because the creators of the models tend to be a monolithic group of primarily Caucasian men. Although including people of different experiences and backgrounds may help, there is also an acknowledgement that standards and guidelines must be established (Schiffer 2021). In addition, there is an acknowledgement that goes beyond the model creators, to the data used to generate the models. Models trained by data that is not also analyzed by subpopulations, for example, African Americans, may present outcomes that are not reflective of the needs of identified subgroups.

The AI industry has many opportunities for improvement. And although some glimmers of desire for improved eradication of bias have come from companies such as Microsoft and Google, there is undeniably the reality that those who speak out against bias are at risk as evidenced by the treatment of champions such as Timnit Gebru and Margaret Mitchell. Gebru and Mitchell were fired from Google after highlighting biases within Google's search engine (Schiffer 2021). Since being fired, Gebru has been subjected to accusations and racist comments with some saying that the efforts to oust prejudice were self-aggrandizing and others stating that Gebru, a Stanford computer science graduate, had no business in the AI field. Comments telling Gebru to "go back to Africa" highlight both indirect and overt forms of oppression that plague the technology space.

The Study of Bias in Artificial Intelligence through Analysis of Data from the National Center for Education Statistics

This study, reported in this chapter, was conducted against the backdrop of growing developments in the use of AI in higher education – particularly in student recruitment and retention – and the evident biases associated with these applications. The study is based on an analysis of data from the National Center for Education Statistics in the US.

The Data Available from the National Center for Education Statistics

NCES has at least two longitudinal studies pertinent to the research undertaken. Beginning Post-secondary Students and Baccalaureate and Beyond both look at data that spans multiple years

> The Beginning Postsecondary Students Longitudinal Study (BPS) currently surveys cohorts of first-time, beginning students at three points in time: at the end of their first year, and then three and six years after first starting in postsecondary education (NCES n.d.a)

> The Baccalaureate and Beyond Longitudinal Study (B&B) examines students' education and work experiences after they complete a bachelor's degree, with a special emphasis on the experiences of new elementary and secondary teachers (NCES n.d.b)

In the case of Beginning Post-secondary, the data was collected between 2011 and 2017. With approximately 22,500 respondents, the repository contains a wealth of information on students spanning their acceptance into college to the closure of their collegiate career either through degree completion, or an egress from post-secondary education.

The data analyzed includes both traditional and non-traditional students, covering over 80 categories of information. While it encompasses traditional measures such as financial status, first-generation status, and grade point average (GPA), the NCES data also captures elements that have historically been considered intangible, such as peer support, faculty-student relationships, campus climate, confidence, and other factors impacting student success. These elements provide a broader understanding of the challenges students face, including those stemming from an unfriendly or non-inclusive campus environment. Beyond measuring academic participation, the data offers insights into social interactions, experiences with faculty, and the support structures provided by peers or family. While this data alone helps narrate a comprehensive story, the application of AI could reveal success patterns that remain hidden with traditional analytical methods. Additionally, this approach could open new avenues for research into previously unexplored pathways.

The Methodology

The study used quantitative analysis to examine the educational data provided by NCES from 2011–2017. Data was first prepared, then explored, visualized, classified, and analyzed (Figure 8.2).

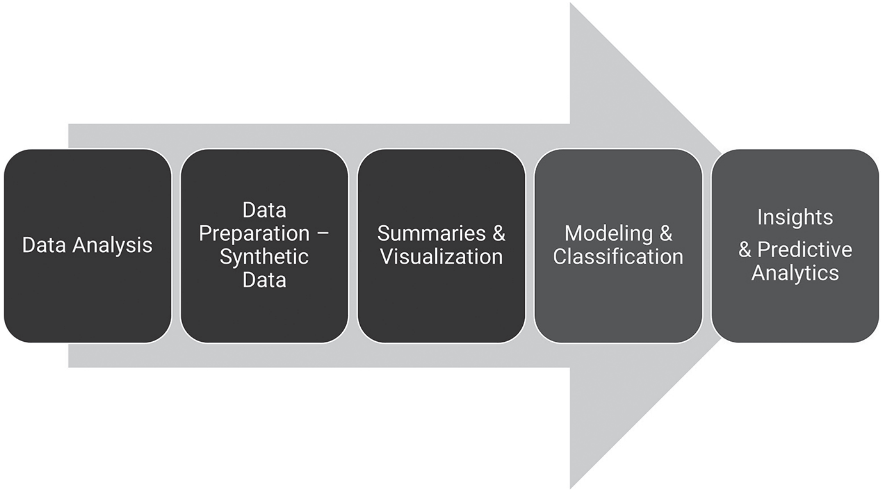

Figure 8.2: Research methodology

Data Preparation

The data was reviewed, scrubbed, prepared, and processed through multiple machine learning algorithms. Analysis for validity included the assessment of statistical pattern recognition over large data sets following three fundamental components:

- Training, using representative data from NCES's beginning post-secondary data
- Validation, using model analysis and fine tuning, and
- Testing to confirm of output accuracy.

NCES provides a detailed dictionary that summarizes available data elements along with source details including category, subject, description, source, time of collection and descriptive statistics for simple to complex data sets. The first step in data preparation included a review of all available data elements to determine which elements would be most beneficial to the development of the model. Part of the challenge in determining which elements to include was the fundamental acknowledgement that any decision to **e**xclude data may as a matter of practice introduce a certain level of bias in the form of assumptions. To alleviate any concerns, the study included most available data elements with the exception of elements that were identified to have a high level of correlation with other existing data. The following steps were taken to initially prepare the NCES data:

- High level review of beginning post-secondary (BPS) categories of data via the NCES codebook
- Creation of NCES DataLab reports filtered by STEM categories and stratified by degree attainment, race, and gender
- Execution of percentage distribution reports spanning 40 categories by subject leveraging the NCES DataLab
- Export of generated reports for post review and processing
- Generation of synthetic data, and
- Import of data into a relational database for further processing.

Although NCES provides institutions with access to raw data for analysis, raw data was not leveraged for this study. In order to address any concerns, synthetic data generation was adopted to reverse engineer aggregate report data into individual unidentifiable records.

High Level Analysis

Although the data available from NCES is not solely focused on technology and engineering, a deeper look at the data available from the Baccalaureate and Beyond longitudinal study identifies many factors that may be beneficial to understanding the higher education landscape for students pursuing a career in engineering and technology. Before delving into these high-level assessments, it is important to understand the foundation upon which the Beginning Post-Secondary repository rests.

Dataset Overview

The beginning post-secondary longitudinal study (NCES n.d.a) consists of a series of survey assessments spanning 6 years. The comprehensive study includes:

- Over 22,500 US participants
- Participants interviewed biannually over the 6-year period, and.
- Data results spanning over 1900 institutions across the United States.

In addition to survey data, NCES collected institutional data including transcripts, financial aid information and pre-enrollment data including first generation status. A demographic breakdown of assessments for the largest populations by race is contained in Figures 8.3 and 8.4.

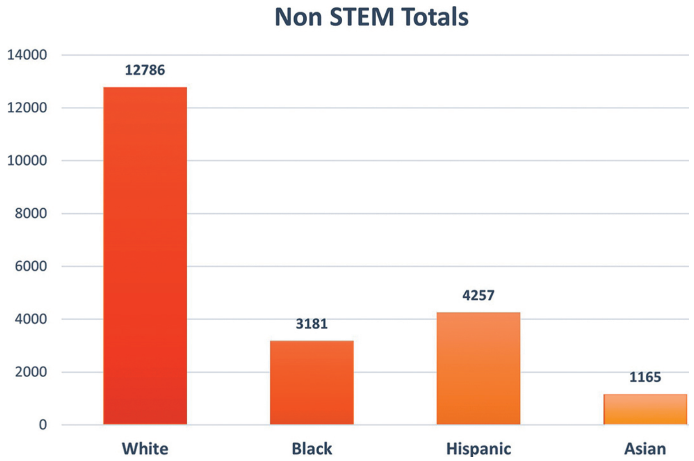

Figure 8.3: Breakdown by race

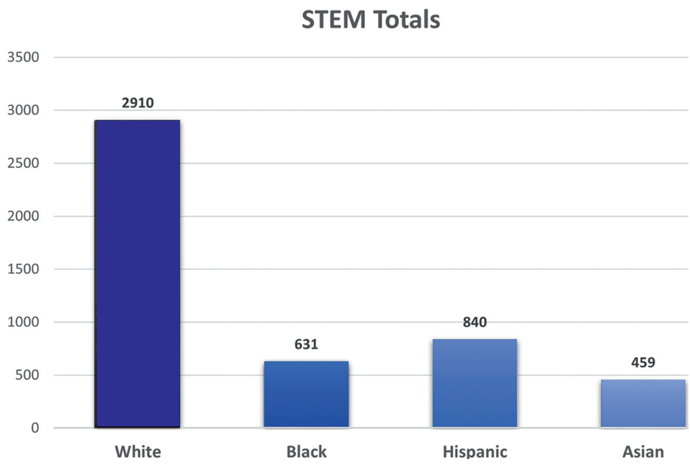

Figure 8.4: STEM Breakdown by race

Synthetic Data

Synthetic data is data generated for use in various forms of analysis including predictive analytics. There are multiple reasons researchers use synthetic data including:
- Anonymity or maintaining the privacy of study participants
- Lack of available data, and
- Algorithm training.

Synthetic data is useful only if the generated data mimics the characteristics of real data sets. The process used to generate synthetic data is key to a successful implementation. Synthetic data can be generated by different techniques (Dilmegani 2024):
- Distribution which is useful when the data is restricted or unavailable as individual records but the distribution of the data is well known across key criteria
- Variational autoencoder, which is an unsupervised model, and
- Generative adversarial network based on iterative model training.

Whether the distribution model is adopted using the Monte Carlo algorithm for example or a form of deep learning adopted for data generation, the use of synthetic data is still somewhat controversial. Concerns and debate over the validity of generated data persist. There are however many examples of proven uses for synthetic data.

Examples of Synthetic Data Use

Several tangible examples of synthetic data use in multiple areas exist including marketing, finance, systems development, quality assurance, security, clinical trials, and other forms of research (Chen et al 2019). Fraud detection methods for example is a growing area within synthetic data use. The ability to create data that provides testing and effectiveness evaluation prior to the implementation of new security measures can prove to be invaluable as it provides a means of validating controls without waiting for real data in the form of fraudulent activities to occur.

As the use of synthetic data continues to expand the processes to generate synthetic data and its overall effectiveness continue to grow. In a 2020 study, Tucker, Wang, Rotalinti and Myles set out to create more predictable outcomes and by using resampling, latent variable identification and outlier analysis, the researchers produced results that more closely tracked real world data (Tucker et al. 2020). A final example demonstrates the capabilities of synthetic data generation by introducing dependency modeling. SynC was created by researchers to move beyond the reconstruction of data from aggregate source information to modeling of dependencies across available parameters. Researchers used Canadian census data from 65,000 respondents to conduct two experiments. The process followed several key steps including synthetic reconstruction, outlier removal, dependency analysis, and predictive modeling. In addition, data sets used for analysis included variables such as personal and family history, which may have some level of impact when predicting outcomes. Researchers found that they were able to closely model results found in original data with a high level of accuracy (Li, Zhao, and Fu 2020).

Process for Synthetic Data Analysis

Although rich and plentiful access to aggregate data existed from NCES, the lack of access to the raw data warranted a novel approach that provided a means of taking the research forward. A considerable amount of time was committed to the analysis and creation of synthetic data leveraging aggregate data from NCES as the primary source. Key steps included data review and core dependency analysis, aggregate report generation, aggregate translation, and data generation.

NCES provides data tools in its DataLab for online review and analysis of source information. Powerstats in particular provides the ability to generate reports across all available subject areas. Because this study focused on predicting outcomes related to degree attainment, a foundational set of control and outcome variables was identified and included highest degree attained, race and gender and major field of study. Other data elements were reviewed and identified for detailed analysis. The list included elements that spanned institutional characteristics, education, social experiences, financial background, faculty and student interactions, confidence factors and other criteria that might have had indirect or indirect impact on the desired outcome. Ninety-seven different characteristics with multiple permutations were sourced. Each of the 97 characteristics was analyzed and a report generated from NCES DataLab. Each report was restricted to National Science Foundation STEM programs in the areas of computers, engineering technologies, mathematics, medicine, health, and science. Behavioral sciences were excluded for the purpose of this study. Each report provided a breakdown in percentages for a given characteristic stratified by a combination of race and gender.

To translate the individually generated reports from percentages into raw values or student counts, calculations were used to convert base percentages to overall counts first by the overall characteristics breakdown by race followed by conversions by both race and degree attainment type. Floor values were used for data analysis and model generation. Generating synthetic records for all parameters within each of the 97 characteristics was completed using a cursor driven implementation. Due to the nature of the data and variations of the parameters within each characteristic, cursors provided an effective means of ensuring that the accurate number of person records was created. Approximately 22,500 records were generated with counts per characteristic adhering to the identified values and boundaries by race and degree attainment. Once all data was imported into a small relational database, a simple query of the more than 90 elements as a flat file with 22,500 rows was used to create a data file or efficient importing and analysis against multiple model types.

Model Analysis

As part of the process to identify the model(s) best suited for the study, research was conducted on the most prevalent predictive modeling practices. Some form of supervised learning was the best approach for what was a classification problem, although there least one study argues that supervised learning introduces bias into the process (Andaur Navarro et al. 2021). Research had indicated that logistic regression, gradient boosting, and random forest might be strong candidates for the predictive modelling required. The criteria used for assessment were:

- Type of outcome expected
- Method for identifying and excluding characteristics with a high correlation
- Process for identifying and excluding outliers
- Model complexity with this study involving a large number of characteristics and parameters
- Accuracy both overall and by one of the three key groups: Black, Black, and White, All Races, and
- Weight analysis.

Figures 8.5 and 8.6 show the results of an initial run of models to determine the best candidates, with Figure 8.5 examining error rates and Figure 8.6 indicating weight comparisons. When examining accuracy alone, many of the models appear to be ideal candidates. Fast Large Margin for example appeared to be ideal in relation to both accuracy and classification errors (Figure 8.5). However, Fast Large Margin seemed to be the most unpredictable when making small and large changes to the weighting of characteristics in the data set. Changes in race, GPA, and other key characteristics had no impact or significantly opposite impacts when comparing its output to the output of other models. Random Forest on the other hand seemed to present more predictable and consistent outputs based on the weighting of selected parameters (Figure 8.6).

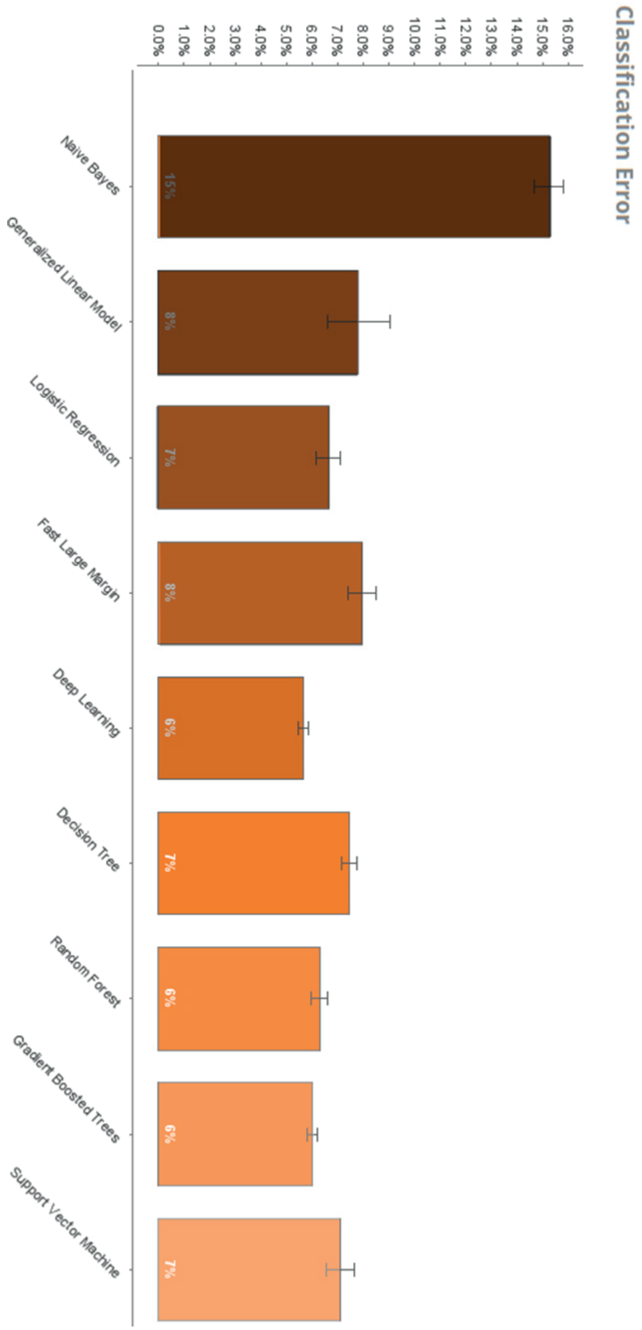

Figure 8.5: Models by Classification Error

Deep Learning		Logistic Regression		Random Forest		Gradient Boost	
Attribute	Weight	Attribute2	Weight3	Attribute4	Weight5	Attribute6	Weight7
LikelihoodExpectedEdu	0.264	LikelihoodExpectedEdu	0.2689	ProgramLastEnrolled17	0.170	LikelihoodExpectedEdu	0.2620
ProgramLastEnrolled17	0.164	Jobintensity1112	0.1133	healthservices12	0.166	intensity1314	0.1753
jobintensity1314	0.133	JobIntensity1213	0.1007	HighestDegreeExpected14	0.135	ProgramLastEnrolled17	0.1355
HighestDegreeExpected	0.106	NumberHSScienceCourses	0.0812	UndergradProgram1112	0.128	healthservices12	0.1181
health14	0.100	HighestDegreeExpected14	0.0761	LikelihoodExpectedEdu	0.126	intensity1112	0.0954
jobintensity1415	0.093	jobintensity1314	0.0747	jobintensity1314	0.122	LikelihoodEverExpected14	0.0899
NumberofHonors	0.085	StudentInteractions12	0.0736	NumberofHonors	0.100	SocialSatisfaction14	0.0795
NumberHSScienceCours	0.084	UndergradProgram1112	0.0632	StopsThrough17	0.096	FacultyInteractions14	0.0752
intensity1112	0.071	finaid14	0.0615	jobsoncampus17	0.095	jobintensity1314	0.0710
FamilyandFriendsSuppo	0.069	noservicesused12	0.0596	jobintensity1415	0.094	NumberHSScienceCourses	0.0676
RaceGender	0.068	intensity1112	0.0588	FacultyInteractions14	0.088	intensity2017	0.0673
intensity1516	0.066	Jobitensity1617	0.058	intensity1314	0.085	FamilyandFriendsSupport	0.0643
FacultyInteractions14	0.064	finaid12	0.0569	Jobintensity1112	0.084	academicsupport14	0.0596
LikelihoodEverExpected	0.064	LikelihoodEverExpected14	0.0567	intensity1415	0.083	StudiesSatisfaction12	0.0556
jobintensity1516	0.063	collegecreditsHS	0.0559	health14	0.077	finaid12	0.0546

Figure 8.6: Weight comparison across models

Modeling and Bias

Understanding the role that bias has in modelling can be examined from multiple perspectives. For the purpose of this study, bias was examined from two points of view: dominance of one population over another, and inherent bias introduced in modeling algorithms. A high-level view of the generated study data identified key factors to consider in recognizing and eradicating bias. The first was simply the breakdown of the population size by a core factor of race and gender. Figure 8.7 shows that the majority of college participants are white. Given the prevalence of white participants, it is highly likely that weights will be skewed and not fully consider characteristics that may vary based on multiple factors including race and gender.

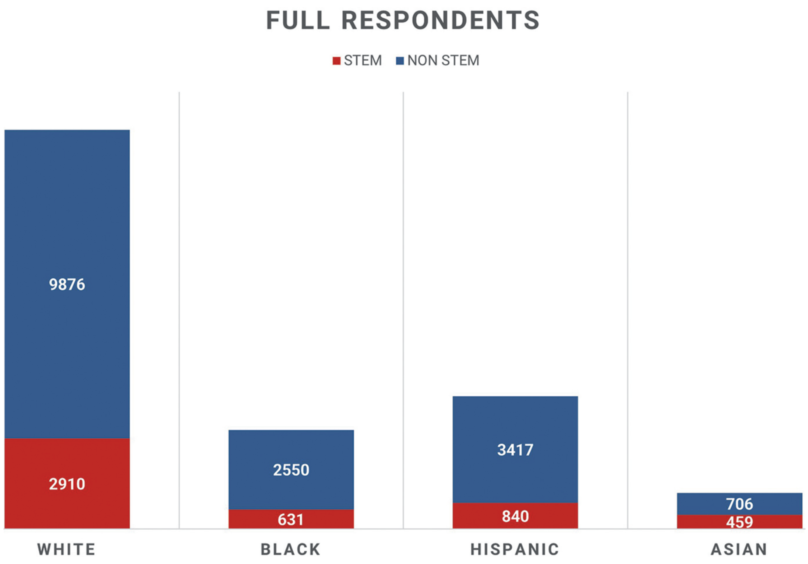

Figure 8.7: Beginning post secondary population breakdown by race

Determining the best method of addressing the challenges associated with dominance may not be as obvious as, for example, excluding the dominant race from the modeling analysis. Hanna, Denton, Smart and Smith-Loud rightfully argue that an examination of racial categories for algorithmic fairness must acknowledge that race is multifaceted and consists of both institutional and social constructs that are hard to quantify with a single attribute (Hanna et al. 2020).

Inherent Bias in Modeling

Benthall and Haynes argue that the use of supervised learning modeling introduces bias by falling toward a dependence on a single variable to quantify race. They recommend an adoption of unsupervised learning and a reliance on clustering to identify groupings based on multidimensional characteristics (Benthall and Haynes 2019). Although unsupervised learning provides for grouping based on multiple characteristics, it appears to be a first step that still includes some supervised learning modeling against clusters if not race. If it is a first step; questions regarding validity must be considered to ensure that the replacement of generated clusters does not supplant the importance of race and its potential value as part of the overall analysis (Benthall and Haynes 2019).

In this study, a model that focused on existing identifiable data by race and also addressed the need to consider the dimensions related to historical, institutional, and social constructs of race was deemed desirable and was part of the evaluation process. Based on an analysis of existing supervised learning models and a set of determining criteria including classification error, accuracy, weighting and identified key factors, the list of potential models was reduced to three for a final comparative analysis:

– Logistic Regression
– Gradient Boosting, and
– Random Forest.

Performance of the Models under Consideration

Logistic Regression

Although Logistic Regression is best used when the dependent variables are binary, it warranted inclusion in the top three models to be considered for multiple reasons:

– The overall performance of the model was one of best in the algorithms tested against bachelor's degree attainment
– The accuracy of assigned weightings in comparison to other models was statistically significant
– The identified factors that help improve the likelihood of obtaining a bachelor's degree meshed well with other high scoring models, and
– There appeared to be a reasonable amount of weight sensitivity and ability to adjust weights for multiple scenarios and identify outcome changes.

The overall performance of Logistic Regression is presented in Figure 8.8.

Performances

Criterion	Value	Standard Deviation
Accuracy	98.2%	± 0.1%
Classification Error	1.8%	± 0.1%

Confusion Matrix

	true Associates	true Bachelors	true Certificate	true NoDegree	class precision
pred. Associates	551	12	4	54	88.73%
pred. Bachelors	1	1787	1	0	99.89%
pred. Certificate	1	1	348	10	96.67%
pred. NoDegree	0	0	0	1791	100.00%
class recall	99.64%	99.28%	98.58%	96.55%	

Figure 8.8: Overall performance of the Logistic Regression model

Although the class precision was not as accurate for associate degrees and certificates, the model rightly predicted bachelor's degree attainment 99.8% of the time. The model appeared to work best when running the analysis against a population of Black and White Only (Figure 8.9) or Black Only. Although more visible in decision tree models, one can see that a lack of intensity or students who attend less than full time, students who lack confidence and students who are not living on campus tend to face greater challenges with degree attainment. Conversely, students whose parents have at least a high school degree and students who are well versed and know requirements for their desired degree tend to excel. Although the results seem plausible, they are not surprising and similar to other model results.

An interesting difference in use of the model emerged when run against a Black only population (Figure 8.10). Clear changes in important factors and weights were present. In particular, the model identified two additional criteria, mental health, and the number of breaks during college, which might positively impact degree attainment. African American students whose mental health remains constant are more likely to succeed in college. The area warrants additional study as it may reflect challenges related to race faced by African American students on and off campus.

Figure 8.9: Logistic regression model key factors: Black and White

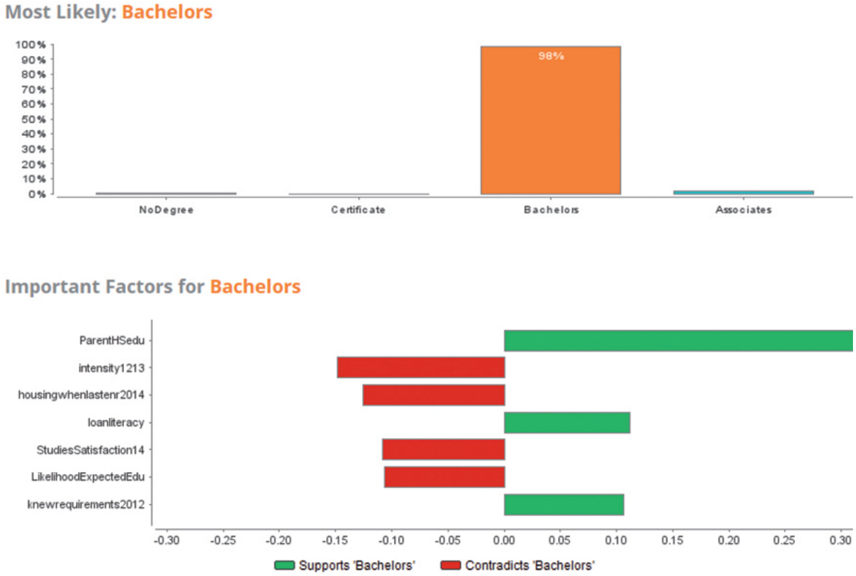

Figure 8.10: Logistic regression model key factors: Black only

Gradient Boosting

Gradient boosting is one of two decision tree models analyzed against the study data. Gradient boosting strengthens the accuracy of decision trees by creating a series of weaker decision trees that build upon one another with the goal being that each successive tree becomes a stronger predictor of outcomes (Chong 2021). From a strictly performance perspective, Gradient Boosting trees performed exceptionally well against the study data. The model accurately predicted bachelor's degree attainment with 100% accuracy (Figure 8.11).

Performances

Criterion	Value	Standard Deviation
Accuracy	99.9%	± 0.1%
Classification Error	0.1%	± 0.1%

Confusion Matrix

	true Associates	true Bachelors	true Certificate	true NoDegree	class precision
pred. Associates	555	2	1	1	99.28%
pred. Bachelors	0	1795	0	0	100.00%
pred. Certificate	0	1	354	1	99.44%
pred. NoDegree	0	0	0	1852	100.00%
class recall	100.00%	99.83%	99.72%	99.89%	

Figure 8.11: Gradient Boosting accuracy and classification

The key to Gradient Boosting trees is identifying the number of successive trees required to obtain optimization. In this study, 90 trees were created with a depth of four.

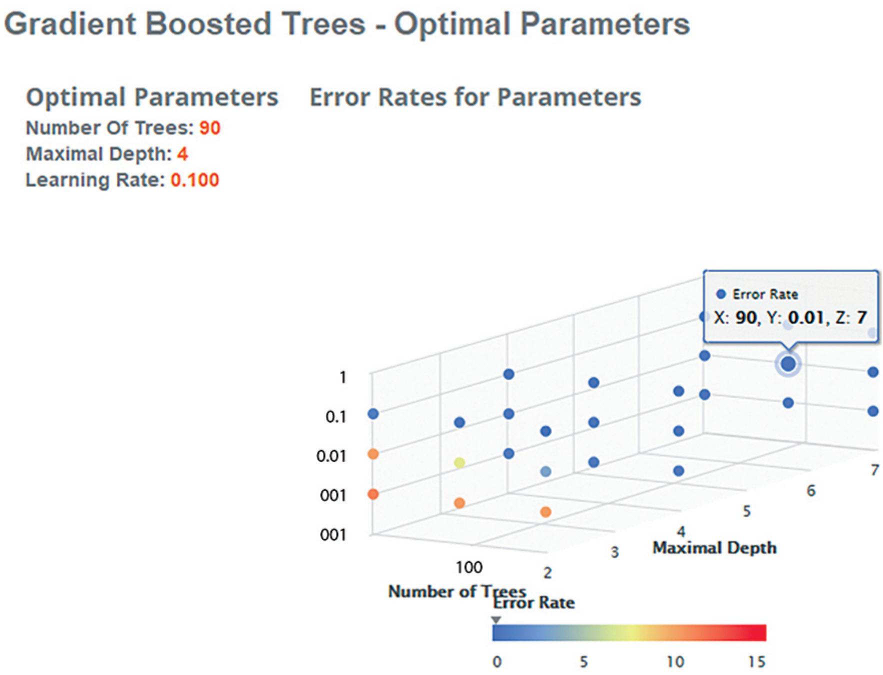

Figure 8.12: Error rate parameters

In addition to the low error rate short learning rate and optimized tree depth (Figure 8.12), the Gradient Boosting tree model also identified key factors that mirrored factors found in both the logistic regression model and random forest model. The gradient model added additional factors of health and faculty interactions (Figure 8.13).

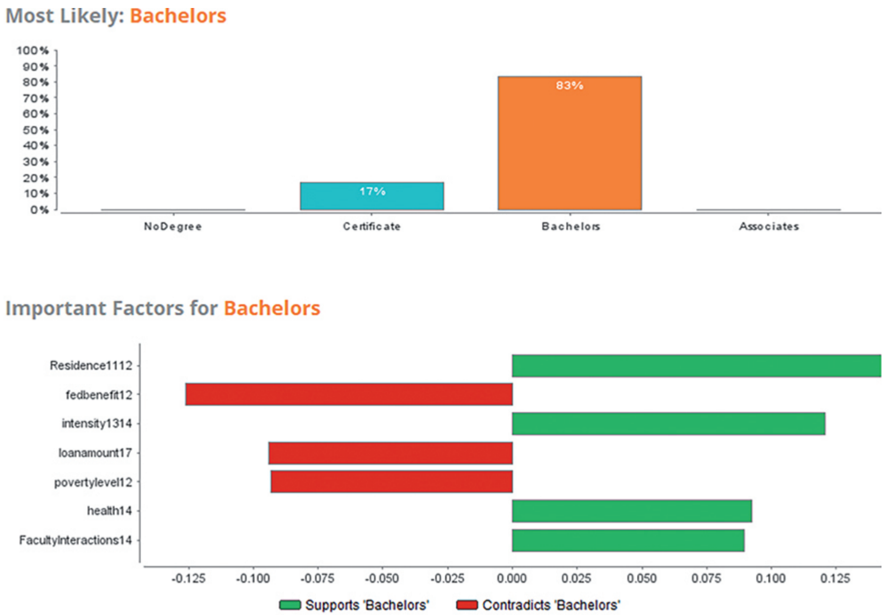

Figure 8.13: Gradient Boosting trees key factors: Black and White

Gradient boosting trees were a strong candidate as the optimal model for the study although the actual tree definitions appeared limited for practical application. As an example, one of the generated trees (Figure 8.14) depicts groupings that appear to work better for binary decisions, that is, in a group or not. The multiple dimensions of the actual study data might not fit well within this dynamic.

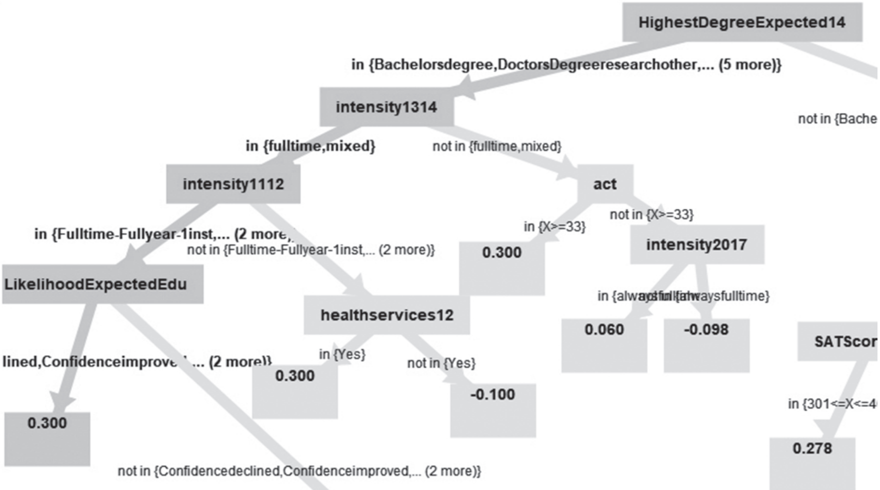

Figure 8.14: Gradient Tree: Black and White

Random Forest

Random Forest is like Gradient Boosting in the sense that both comprise an ensemble of trees. One of the main differences between the two is bagging. In Random Forest trees, a collection of independent trees is created using varying subsets of training data (Yiu 2019). The independence of trees is key as they have no knowledge of each other and therefore the outcomes of each are not influenced by one another. Random Forest in this case might provide additional flexibility in analyzing parameters that are not simply binary. The overall performance of Random Forest was comparable to Gradient Boosting trees. The Random Forest model was able to predict bachelor's degree attainment 100% of the time (Figure 8.15). In addition, the Random Forest model's optimal tree generation is less than Gradient Boosting with 60 trees 7 levels deep (Figure 8.16).

Performances

Criterion	Value	Standard Deviation
Accuracy	99.8%	± 0.1%
Classification Error	0.2%	± 0.1%

Confusion Matrix

	true Associates	true Bachelors	true Certificate	true NoDegree	class precision
pred. Associates	550	1	0	1	99.64%
pred. Bachelors	0	1797	0	0	100.00%
pred. Certificate	3	1	354	2	98.33%
pred. NoDegree	0	0	0	1852	100.00%
class recall	99.46%	99.89%	100.00%	99.84%	

Figure 8.15: Random Forest model key factors: Black only

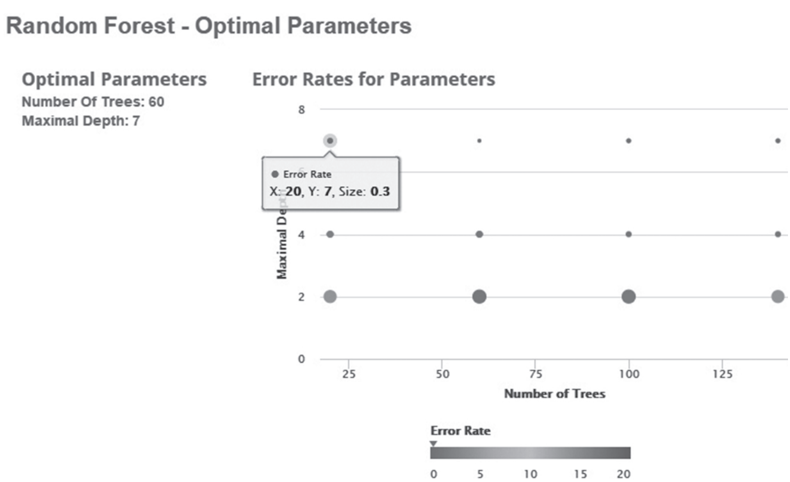

Figure 8.16: Random Forest Tree: Black only

Like Logistic Regression and Gradient Boosting tree models, Random Forest was also able to identify comparable key factors for successful completion of a bachelor's degree (Figure 8.17).

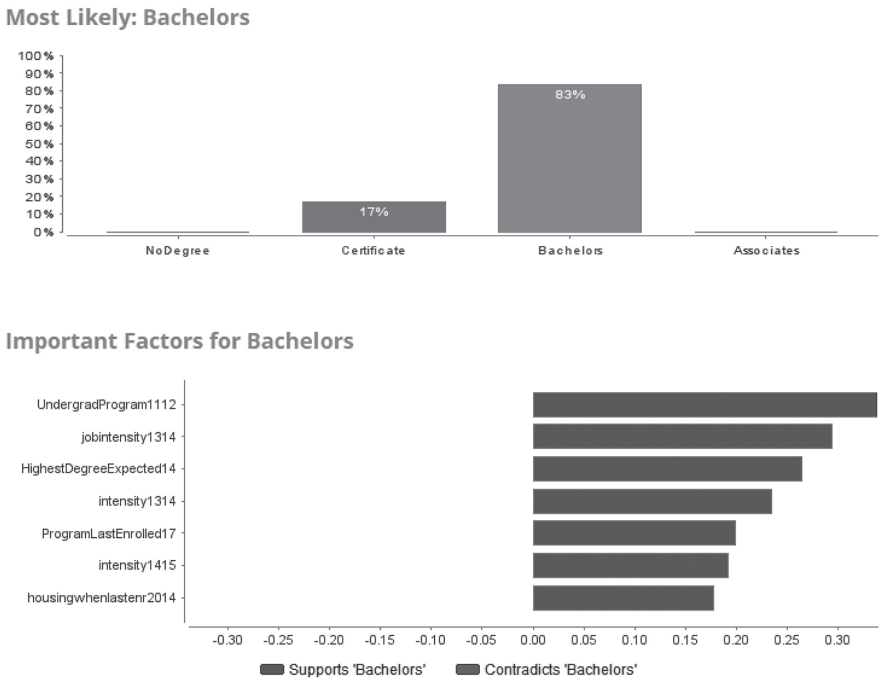

Figure 8.17: Random Forest key factors: Black only

The strength of the Random Forest model appears when examining the independent trees. As an example, the tree in Figure 8.18 shows pathways for attainment related to housing and other key factors. The tree shows multiple pathways toward successful attainment of a bachelor's degree while clearly identifying additional pathways with some probability of success.

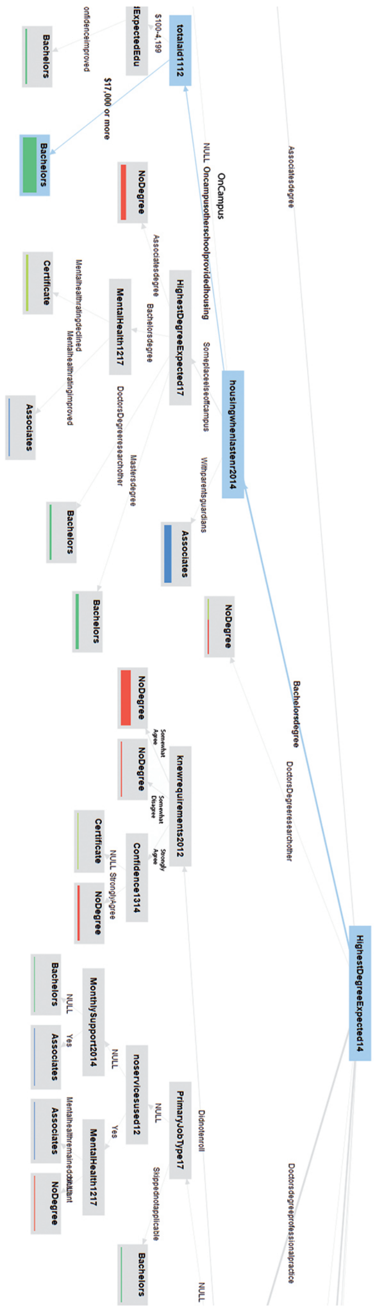

Figure 8.18: Logistic Regression model key factors: Black only

Selected Model: Supervised Learning Random Forest

Although each of the models had advantages and disadvantages, Random Forest was deemed to be the best for the study for the following reasons:

- Good balance between bias and variance with minimal difference between the expected and actual values
- Lack of dependence on other trees, unlike Gradient Boosting's dependence on pathways defined in previous trees
- Majority voting provides a means of returning optimal results as a prediction from the parallel and independently processed trees
- Overall performance, and
- Ease of analysis.

Some of the results from the survey are shown in the sample trees below generated from the Random Forest model.

The Results of the Survey

Faculty Interactions and Campus Climate

Across most models, faculty interaction was heavily weighted. Student experiences with faculty appeared to be a core factor in determining success. Both positive and negative experiences with faculty can either improve or diminish a student's ability to attain a bachelor's degree. The Random Forest tree (Figure 8.19) points out the importance of interactions with students and support from parents, family, and friends.

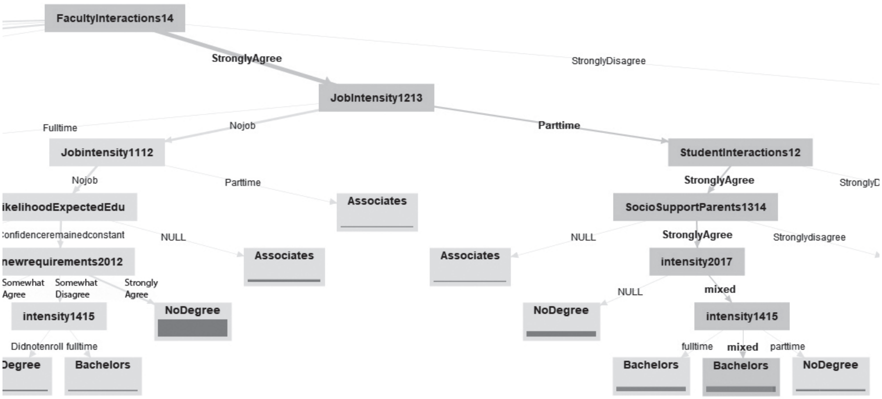

Figure 8.19: Faculty interaction, job intensity and student interactions

Financial Aid and Confidence

Overall and not surprisingly financial aid is a key factor for many African American students. However, even for those who receive financial aid it is not the only factor. A student's confidence level as well as the time spent on campus contributes significantly to success. For example, students who take breaks in their collegiate career or who are not able to attend classes full time are less likely to graduate (Figure 8.20).

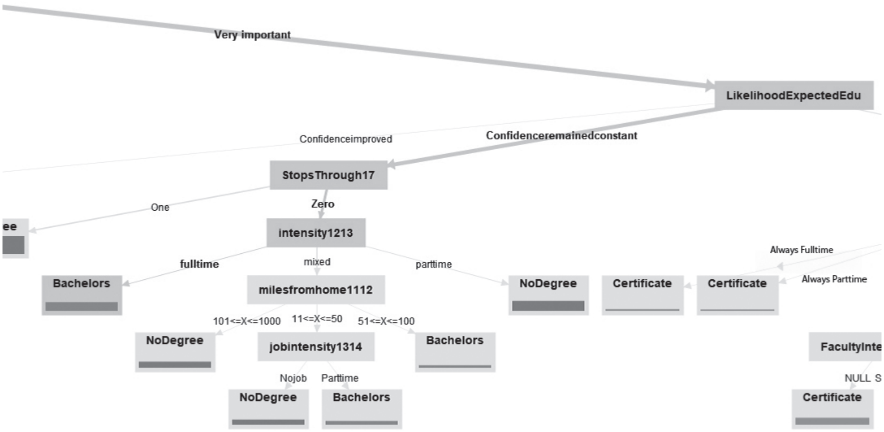

Figure 8.20: Financial aid, confidence, and the number of breaks

Intensity, Confidence, and Belonging

The advantages of full-time attendance are identified repeatedly over multiple models and multiple trees within the decision tree models. Intensity alone however is not enough. Figure 8.21 shows the importance of confidence and the importance of feeling a sense of belonging. Students who feel part of the institution or have a high level of confidence are more likely to succeed.

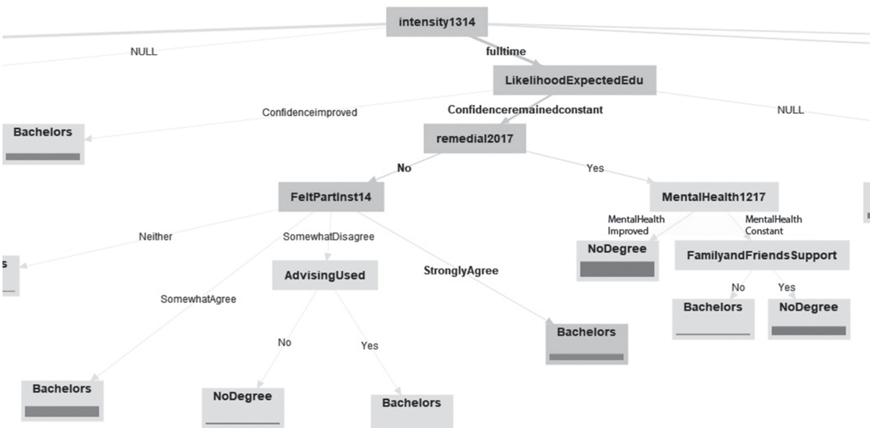

Figure 8.21: Academic intensity, confidence, remediation, and belonging

Parental Support, Positive Campus Climate and Housing

Students to whom parents have provided support from a socialization perspective are more likely to reach degree attainment. But parental support alone is not sufficient. Figure 8.22 indicates that a positive campus climate along with on-campus housing helps to increase the likelihood of graduating.

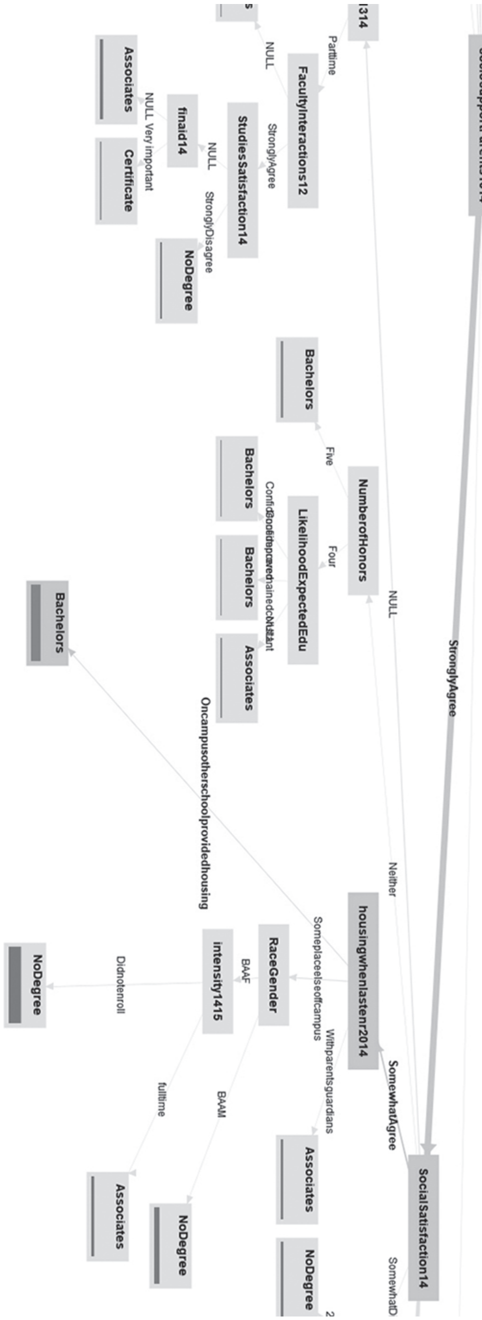

Figure 8.22: Parental support, social satisfaction, and housing

Confidence, Housing, and Intensity

Student confidence is critical to overall success. In addition, campus climate and campus living seem to be dominant factors as well (Figure 8.23).

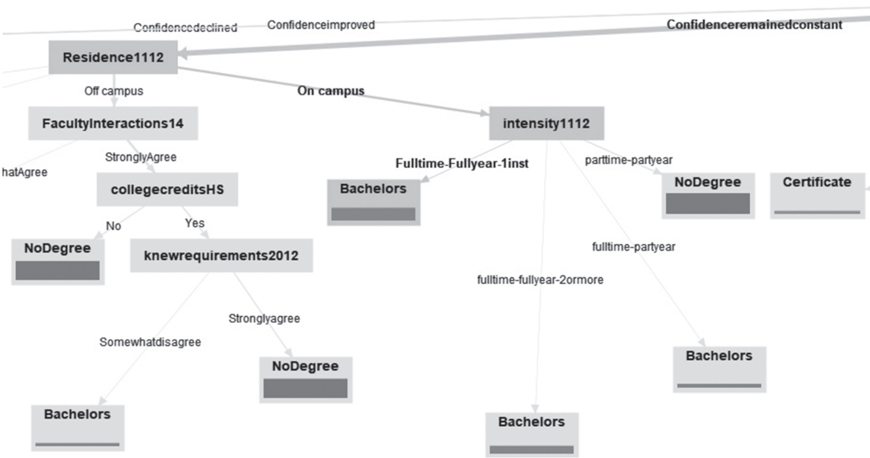

Figure 8.23: Housing

Years of Mathematics and Number of Breaks from Study

Data has historically demonstrated the importance of a strong mathematics background for students in STEM. Additional research in this study provides further validation of that experience. However, the study also emphasized that a strong mathematical background is not sufficient. The number of breaks during a student's collegiate career has significant impact. Figure 8.24 shows that students who took at least one break were much less likely to graduate regardless of their mathematical background.

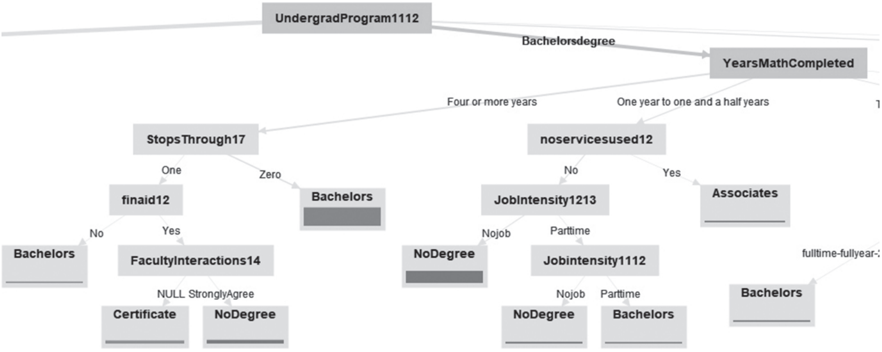

Figure 8.24: Years of mathematics and breaks from study

Mathematics, Support, and Social Satisfaction

Another example that shows the weight of a strong mathematics background along with underlying dependencies is found in Figure 8.25. One can see the impact campus climate can have on degree attainment. Students who experienced strong social satisfaction or satisfaction with relationships spanning fellow students, friends, and faculty were more likely to succeed than students with negative campus experiences.

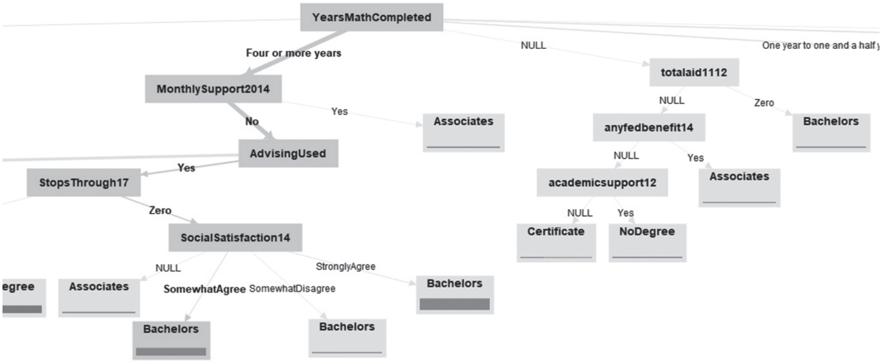

Figure 8.25: Years of advising, support, and social satisfaction

The Random Forest trees presented are but a few examples from 60 independent parallel trees within the model. Random Forest encapsulates the complexity and inter-dependency of a multitude of variables that have some level of impact on student success and present the results in a clear way.

Key Findings

The data was not only analyzed using multiple models but also through two key perspectives. Each model was run initially using four key groups by race: White, African American, Hispanic, and Asian. The data was then run through the various models using African Americans only. Consistently across models the results showed that factors identified as success or failure factors for African Americans differed when focusing on African Americans only as opposed to the entire group. The differing outcomes reinforced the potential for bias when one population is overshadowed by the dominant group.

Confidence

Across multiple models, confidence became a dominant factor for African Americans in higher education. Factors such as high school GPA continued to play a key role; however, in many cases confidence alone was a key differentiator between students who drop out and students who successfully complete a bachelor's degree.

Faculty Interactions

Faculty interactions regardless of race appeared to play a key role as well. African American students who documented poor faculty interactions more frequently than not failed to graduate. White peers however had significantly fewer poor faculty interactions. Furthermore, white peers who had negative interactions were still more likely to complete their degree.

Mental Health

Mental health was a key success factor for all race groups. Students challenged with mental health concerns were at a significant disadvantage. In all groups students who measured their mental health as good or higher were more successful. The data also appears to show that many students in this study experienced a decline in their mental health. This area warrants more research.

Advising

Surprisingly for both White and African American students the use of Advising in the first year was seen as a detractor to successful graduation. More analysis is needed to understand the implications of the finding in this area.

General Findings

The project highlighted several instances where institutions leveraged AIML with varying degrees of success and exemplified common traits in terms of both successes and continued challenges. One challenge in particular is the continued lack of student success for underrepresented groups in areas such as STEM. "Black and Hispanic workers remain underrepresented in the science, technology, engineering and math (STEM) workforce compared with their share of all workers, including in computing jobs, which have seen considerable growth in recent years" (Fry, Kennedy, and Funk 2021). Even with the advent of AI and machine learning and the wide adoption of tools like Civitas Learning the overall numbers in areas like engineering have not improved. The Hechinger Report consistently reports on race and equity issues in American education and has drawn attention to the ongoing persistent and painful gaps in educational outcomes. The problems continue and the number of African American engineering and mathematics graduates declined in 2021 (Collins 2021; Newsome 2022).

Conclusion

The use of AI and machine learning in higher education opens up a world of possibilities. The thorough analysis of existing data collected over multiple years provides a glimpse into what is possible. By running analytical models for both the collective population and subpopulations by race, a glimpse of the significant outcome differences can be gained. The differences highlight the importance of ensuring subpopulations are adequately represented and weighted during the training and modelling phase. If done correctly, there is an opportunity to improve outcomes for not only African Americans but all students (Riep and Prabhakar 2021). If done incorrectly, Artificial intelligence and machine learning have the power to perpetuate existing bias and stereotypes that make it difficult for students to succeed. Although this study focused primarily on STEM there is every reason to believe that the same process could be used across multiple subject areas.

References

Andaur Navarro, Constanza L., Johanna A. A. Damen, Toshihiko Takada, Steven W. J. Nijman, Paula Dhiman, Jie Ma, Gary S. Collins et al. 2021. "Risk of Bias in Studies on Prediction Models Developed Using Supervised Machine Learning Techniques: Systematic Review." *BMJ*: 375 n2281. https://doi.org/10.1136/bmj.n2281.

Benthall, Sebastian, and Bruce D. Haynes. 2019. "Racial Categories in Machine Learning." In *FAT* '19: Proceedings of the Conference on Fairness, Accountability, and Transparency*, *January 2019*, 289–298. New York: Association for Computing Machinery. https://doi.org/10.1145/3287560.3287575. Available at https://arxiv.org/abs/1811.11668v1.

Bryan, Michael, Darryl Cooney, Barbara Elliott, and David Richards. 2019. *2012/17 Beginning Postsecondary Students Longitudinal Study (BPS: 12/17); Data File Documentation.* US National Center for Education Statistics. Institute of Education Sciences. https://nces.ed.gov/pubs2020/2020522.pdf.

Carmean, Colleen, David Kil, and Linda Baer. 2021. "Why Data Matters for Student Success in a Post-Pandemic World." *Educause Review*, August 10, 2021. https://er.educause.edu/articles/2021/8/why-data-matters-for-student-success-in-a-post-pandemic-world. An adapted version is available on the Civitas Learning website at https://www.civitaslearning.com/blog/sustainable-outcomes/why-data-matter-for-student-success-post-pandemic/.

Chen, Junqiao, David Chun, Milesh Patel, Epson Chiang, and Jesse James. 2019. "The Validity of Synthetic Clinical Data: A Validation Study of a Leading Synthetic Data Generator (Synthea) Using Clinical Quality Measures." *BMC Medical Informatics and Decision Making* 19: 44, https://doi.org/10.1186/s12911-019-0793-0.

Chong, Jason. 2021. "Battle of the Ensemble - Random Forest Vs Gradient Boosting: Two of the Most Popular Algorithms in the World of Machine Learning, Who Will Win?" *Medium,* June 2, 2021. Published in *Towards Data Science*. https://towardsdatascience.com/battle-of-the-ensemble-random-forest-vs-gradient-boosting-6fbfed14cb7.

Collins, Michael. 2021. "Opinion: Why College Majors Are Another Form of Implicit Bias in Higher Education: Instead of Being a Great Equalizer, Higher Education is Leaving Too Many Students Behind." *The Hechinger Report,* December 20, 2021. https://hechingerreport.org/opinion-why-college-majors-are-another-form-of-implicit-bias-in-higher-education/

Dilmegani, Cem. 2024. "Synthetic Data in 2024: Techniques & Best Practices." *AIMultiple Research,* February 14, 2024. https://research.aimultiple.com/synthetic-data-generation/#what-are-the-techniques-of-synthetic-data-generation.

Edu.gov. "Federal Role in Education." (Jun, 2021). https://www2.ed.gov/about/overview/fed/role.html.

Franklin, Jeremy. 2016. "Racial Microaggressions, Racial Battle Fatigue, and Racism-Related Stress in Higher Education." *Journal of Student Affairs of the University of New York* 12, no. 1: 44–55. Available at https://static1.squarespace.com/static/50ccaa7ee4b00e9e60845daa/t/5768045be58c628d3598f632/1466434651823/Franklin.+J.+-+Racial+Microaggressions%2C+Racial+Battle+Fatigue%2C+and+Racism-Related+Stress+in+Higher+Education.pdf.

Fry, Richard, Brian Kennedy, and Cary Funk. 2021. "Stem Jobs See Uneven Progress in Increasing Gender, Racial and Ethnic Diversity." *Pew Research Center. Report,* April 1, 2021. https://www.pewresearch.org/science/2021/04/01/stem-jobs-see-uneven-progress-in-increasing-gender-racial-and-ethnic-diversity/.

Georgia State University. n.d. "The Panther Experience." Explore Student Life. Student Engagement. https://engagement.gsu.edu/the-panther-experience#pride.

Goupil, Francois, Pavel Laskov, Irdin Pekaric, Michael Felderer, Alexander Dürr, and Frederic Thiesse. 2022. "Towards Understanding the Skill Gap in Cybersecurity." In *ITiCSE '22: Proceedings of the 27th ACM Conference on Innovation and Technology in Computer Science Education*, vol. 1 July 2022, 477–483. https://doi.org/10.1145/3502718.3524807.

Hanna, Alex, Emily Denton, Andrew Smart, and Jamila Smith-Loud. 2020. "Towards a Critical Race Methodology in Algorithmic Fairness." In *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, January 2020*, 501–512. https://doi.org/10.1145/3351095.3372826.

Harati, Hoda, Laura Sujo-Montes, Chih-Hsiung Tu, Shadow Armfield, and Cherng-Jyh Yen. 2021. "Assessment and Learning in Knowledge Spaces (ALEKS) Adaptive System Impact on Students' Perception and Self-Regulated Learning Skills." *Education Sciences* 11, no. 10: 603. https://doi.org/10.3390/educsci11100603.

Heinonen, Henri. 2019. "Artificial Intelligence AI — Over 100 Billion in Revenues by 2025." *Medium*, December 3, 2019. Published in *Towards Data Science*. https://towardsdatascience.com/artificial-intelligence-ai-over-100-billion-in-revenues-by-2025-e737459a345e.

IBM. 2023. "Shedding Light on AI Bias with Real World Examples." by IBM Data and AI Team." *[Blog]*, October 16, 2023. https://www.ibm.com/blog/shedding-light-on-ai-bias-with-real-world-examples/.

Kil, David, Angela Baldasare, and Mark Milliron. 2021. "Catalyzing a Culture of Care and Innovation Through Prescriptive and Impact Analytics to Create Full-Cycle Learning." *Current Issues in Education* 22, no. 1 Special Issue. https://cie.asu.edu/ojs/index.php/cieatasu/article/view/1903/887.

Li, Zheng, Yue Zhao, and Jialin Fu. 2020. "SynC: A Copula Based Framework for Generating Synthetic Data from Aggregated Sources." In *International Conference on Data Mining Workshops (ICDMW), 17–20 November 2020. Proceedings,* 571–8. New York: IEEE. https://doi.org/10.1109/ICDMW51313.2020.00082. Available at https://www.andrew.cmu.edu/user/yuezhao2/papers/20-icdmw-sync.pdf

Metz, Cade. 2021. "Who is Making Sure the A.I. Machines Aren't Racist?" *The New York Times*, March 15, 2021, updated June 23, 2023. https://www.nytimes.com/2021/03/15/technology/artificial-intelligence-google-bias.html.

National Center for Education Statistics (US) (NCES). n.d.a. "Beginning Postsecondary Students (BPS)." https://nces.ed.gov/surveys/bps/.

National Center for Education Statistics (US) (NCES). n.d. "Baccalaureate and Beyond Longitudinal Study (B&B)." https://nces.ed.gov/surveys/b&b/.

National Center for Education Statistics (US) (NCES). 2019. "Status and Trends in the Education of Racial and Ethnic Groups 2018." NCES 2019-038. Washington DC: NCES Institute of Education Research. https://nces.ed.gov/pubs2019/2019038.pdf.

National Center for Education Statistics (US) (NCES). 2023. "Table 318.45.Number and Percentage Distribution of Science, Technology, Engineering, and Mathematics (STEM) Degrees/Certificates Conferred by Postsecondary Institutions, by Race/Ethnicity, Level of Degree/Certificate, and Sex of Student: Academic years 2012–13 through 2021–22”. *Digest of Education Statistics*. https://nces.ed.gov/programs/digest/d23/tables/dt23_318.45.asp.

National Center for Science and Engineering Statistics (US). 2023. “Diversity and STEM: Women, Minorities and Persons with Disabilities.” Alexandria, VA.: National Science Foundation. https://ncses.nsf.gov/pubs/nsf23315/.

Newsome, Melba. 2022. “Computer Science has a Racism Problem: These Researchers Want to Fix It.” *Nature. News Feature,* October 19, 2022. https://www.nature.com/articles/d41586-022-03251-0

Riep, Josette, and Annu Prabhakar. 2021. “Leveraging Artificial Intelligence to Increase STEM Graduates among Underrepresented Populations.” In SIGITE ’21: Proceedings of the 22nd Annual Conference on Information Technology Education October 2021, 73–74 https://doi.org/10.1145/3450329.3476848.

Schiffer, Zoe. 2021. “Timnit Gebru was Fired from Google – Then the Harassers Arrived.” *The Verge*, March 6, 2021. https://www.theverge.com/22309962/timnit-gebru-google-harassment-campaign-jeff-dean.

Shaw, Catherine. 2021. *Driving Toward a degree 2021: Caseload, Coordination, Student Engagement, and Equity*. Tyton Partners. https://tytonpartners.com/driving-toward-a-degree-2021-caseload-coordination-student-engagement-and-equity./

Smithsonian Science Education Center. 2024. “The STEM Imperative.” https://ssec.si.edu/stem-imperative.

Tucker, Alan, Zhenchen Wang, Ylenia Rotalinti, and Puja Myles. 2020. “Generating High-Fidelity Synthetic Patient Data for Assessing Machine Learning Healthcare Software” *npj Digital Medicine* 3: 147. https://doi.org/10.1038/s41746-020-00353-9.

Yiu, Tony. 2019. “Understanding Random Forest: How the Algorithm Works and Why It Is So Effective.” *Medium,* June 12, 2019. Published in *Towards Data Science.* https://towardsdatascience.com/understanding-random-forest-58381e0602d2.

Zawacki-Richter, Olaf, Victoria I. Marín, Melissa Bond, and Franziska Gouverneur. 2019. “Systematic Review of Research on Artificial Intelligence Applications in Higher Education – Where are the Educators?” *International Journal of Educational Technology in Higher Education* 16: Article no.39. https://doi.org/10.1186/s41239-019-0171-0.