Bridging Theory and Practice: Building an Inclusive Undergraduate Data-Science Program

Dr. Mehmet Ergezer, Wentworth Institute of Technology

Mehmet Ergezer holds a Doctor of Engineering degree from the Department of Electrical and Computer Engineering at Cleveland State University, Cleveland, OH. Currently serving as an Associate Professor of Computing and Data Science at Wentworth Institute of Technology in Boston, MA, Dr. Ergezer's expertise lies at the intersection of embedded systems and computational intelligence. He has co-authored publications on artificial intelligence and computer science education and holds consumer electronics patents in the USA and Europe.

Beyond his technical pursuits, Dr. Ergezer is deeply committed to fostering intercultural competence among undergraduate students. He has spearheaded initiatives to broaden participation in artificial intelligence, including co-founding and co-chairing the AAAI Undergraduate Consortium. Additionally, he has led transformative study abroad programs, such as a semester-long immersion experience in Berlin, Germany, providing students with invaluable global perspectives.

Dr. Mark Mixer, Wentworth Institute of Technology

Mark Mixer is an associate professor in the School of Computing and Data Science at Wentworth Institute of Technology. After completing a B.A. from Dartmouth College, he received his M.S. and Ph.D in Mathematics from Northeastern University. His research is in the fields of group theory, graph theory, combinatorics, and discrete geometry.

Dr. Weijie Pang, Wentworth Institute of Technology

Weijie Pang is an tenure-track assistant professor in School of Computing and Data Science at Wentworth Institute of Technology. After received B.S. and M.S. from Beijing University of Technology and Ph.D. from Worcester Polytechnic Institute, Dr. Pang finished the Post-doc program in Department of Mathematics and Statistics at McMaster University. Her research area is data science, financial mathematics, pandemic modeling and stochastic analysis.

Bridging Theory and Practice: Building an Inclusive Undergraduate Data Science Program

Mehmet Ergezer, Mark Mixer, Weijie Pang
Wentworth Institute of Technology
Boston MA, 02115 USA
{ergezerm, mixerm, pangw}@wit.edu

Abstract

As the field of Data Science (DS) continues to evolve, institutions of higher education face the challenge of developing curricula that prepare students for the industry's rapidly changing landscape. In this paper, we will present a case study of the development and implementation of the undergraduate Bachelor of Science in Data Science (BSDS) program at the Wentworth Institute of Technology. This new degree was developed by an interdisciplinary committee, including faculty members from computer science, humanities, management, mathematics, and sciences to underscore the importance of collaborative expertise in the field of Data Science. We will discuss the curricular development as well as our efforts required to successfully launch the new program. We will provide insights into the decision-making process for aligning the program with dynamic industry requirements.

A focus point for this program lies in fostering diversity and inclusivity, with a keen aim to amplify the presence of underrepresented and marginalized groups in computing, data analysis, and artificial intelligence. Our Data Science program offers a pathway for community college graduates to complete the program in a short time window. In particular, we are developing a "$2 + 2$" option for students, where 2-year associate degrees from various local community colleges transfer effectively, only leaving 2 years left for completion of the BSDS degree for the students at our university. This initiative is not just about accessibility but is a deliberate strategy to welcome individuals from diverse educational backgrounds, thereby enriching the learning environment with a multiplicity of perspectives. Additionally, we focus on diversity and inclusion at the course design level, implementing pedagogical tools with this ideal in mind. These tools are tailored to create an environment that respects and values diversity, ensuring that our educational offerings are not only academically rigorous but also considerate of the varied experiences and perspectives that each student brings to the table. Through this holistic approach, we aim to cultivate a learning community where every individual, irrespective of their background, feels seen, heard, and empowered to contribute meaningfully to the field of Data Science.

In this paper, we present the challenges and triumphs encountered during the program's launch, shedding light on proactive measures addressing potential barriers to entry and participation. Our work can be of value to others who are interested in designing a program that combines theoretical depth in both mathematics and computer science with practical applicability along with a focus on diversity and inclusivity.

I. Introduction

The field of computing is developing at a rapid pace, and in particular, the field of Data Science has been gaining popularity due to its importance in the current era of artificial intelligence and big data [1, 2]. The growth in the amount of available data, computational resources, and improved algorithms has resulted in the blossoming of Data Science in many domains [3, 4].

In healthcare, Data Science will continue to improve the quality and efficiency of healthcare delivery, as well as enable personalized medicine and disease prevention. For instance, Data Science can help diagnose diseases [5], predict outcomes [6], and recommend treatments based on patient data [7, 8]. The use of big data and analytics in healthcare could save up to $100 billion annually in the US [9].

In education, Data Science can enhance the learning experience [10] and outcomes for students and educators [11, 12], as well as provide insights into the effectiveness of educational policies and programs [13, 14]. Furthermore, Data Science can help design adaptive learning systems, assess student performance, and identify learning gaps [15, 16]. e-Learning is predicted to reach over $300 billion by 2025 in the global market, making it one of the most favorable businesses [17].

Data Science has quickly entered the business world where it can be used to optimize operations, enhance customer experiences, and generate new revenue streams [18]. A modern marketing framework that leverages big data for business analytics is proposed by [19]. In the next decade, the demand for data scientists is expected to grow by 35%, compared to the average growth of 3% for all occupations[20].

With the expected growth in Data Science jobs, there will be a need for well-trained practitioners of Data Science. However, with all the clear benefits of the growth in Data Science, there is still a lack of clarity on what precisely Data Science is [21]. It is thus timely to develop a program that addresses this upcoming need while helping clarify what skills are needed for Data Science students going forward. We believe Wentworth is in a strong position for this task. In particular, at Wentworth Data Science, Computer Science, Statistics, and Applied Mathematics all are housed in one school (the School of Computing and Data Science). While many institutes have siloed programs in statistics, mathematics, and computing, Wentworth's structure allows for the interdisciplinary nature of Data Science to work nicely with all of the other programs in the school.

The development and implementation of the undergraduate Bachelor of Science in Data Science (BSDS) program at a teaching-focused university is a case study that highlights the importance of collaborative expertise in the field of Data Science and underscores the need for curricula that prepare students for the industry's rapidly changing landscape while fostering diversity and inclusivity. Our work can be of value to others who are interested in designing a program that combines theoretical depth in both mathematics and computer science with practical applicability along with a focus on diversity and inclusivity.

The following summarizes the immediate need to develop thoughtful Data Science programs [22]. We affirm the surge in data has led to the emergence of Data Science as a discipline, prompting an increasing number of universities to introduce undergraduate programs in the field [23, 24]. While many of these programs originate from computer science [25] or statistics [26] departments, the authors delineate a distinctive Data Science program created collaboratively by computer science and statistics departments. Unlike its counterparts, their program, like ours, strives for a balanced curriculum, delivering comprehensive training in both computer science and statistics, coupled with unique data-related skills rarely covered in these disciplines. The overarching objectives encompass producing students adept in both computer science and statistics, endowing them with distinctive data-focused skills, and facilitating the application of their expertise to specific domains. In recounting experiences in establishing this program, the authors underscore its growth, positive student feedback, and success in preparing students for internships.

The paper is organized as follows. Section II breaks down the curriculum development on a term-by-term basis; Section III provides some insight into our program and what it took to establish it; and Section IV presents how to establish an inclusive educational atmosphere, fostering diversity, equity, and inclusion (DEI) awareness among the students and inclusive curriculum design.

II. Curriculum

The development of the BSDS curriculum at Wentworth was a collaborative effort led by an interdisciplinary committee comprising faculty members from Computing, Mathematics, Sciences, and Humanities. This inclusive approach, drawing from diverse academic backgrounds, enriches the curriculum with varied perspectives on essential core ideas. Moreover, the curriculum is designed to align with the competencies outlined in the ACM Data Science Task Force's "Computing Competencies for Undergraduate Data Science Curricula" [27]. Grounded in the joint ACM and IEEE taskforces on Computer Science [28] and the American Statistical Association guidelines in statistical sciences [29], this work adheres to established undergraduate curriculum guidelines[30], ensuring a robust and comprehensive learning experience for BSDS students.

In this context, the BSDS program, akin to our Applied Math and Computer Science, encompasses 120 credits. Students engage with 28 four-credit courses and four two-credit Calculus classes, shaping their educational journey over four years. While this is the envisioned completion timeline, the program offers an alternative three-year trajectory. This accelerated path, although challenging, appeals to the aspirations of more ambitious students aiming for early graduation. Beyond the immediate benefits of

Table I
First-year Curricula for BS in Applied Math, Computer Science, and Data Science.

Math Fall	Math Spring
Discrete Mathematics	Foundations of Applied Mathematics
Calculus 1A	Calculus 2A
Calculus 1B	Calculus 2B
English Sequence	English Sequence
Science Elective	Science Elective
CS Fall	**CS Spring**
Discrete Mathematics	Computer Organization
Calculus 1A	Calculus 2A
Calculus 1B	Calculus 2B
English Sequence	English Sequence
Computer Science 1	Computer Science 2
DS Fall	**DS Spring**
Discrete Mathematics	Foundations of Applied Mathematics
Calculus 1A	Calculus 2A
Calculus 1B	Calculus 2B
English Sequence	English Sequence
Computer Science 1	Computer Science 2

financial savings, this option facilitates swifter access to advanced degrees, reflecting the program's commitment to flexibility, excellence, and academic prowess.

A. First-Year Curriculum

The first-year curriculum for our Data Science program was designed to closely match both the first years of our current Computer Science and Applied Mathematics programs. At Wentworth, students choose their major before matriculation. This synchronization of the programs allows students the flexibility to switch between programs seamlessly if they discover that their original intent does not match their desires after completing the first year. Similarly, it opens the doors for a variety of students to transfer into the Data Science program after the first year.

Table I presents the first-year curricula for Applied Math, Computer Science (CS), and Data Science (DS). There are two main differences: Applied Math students take their science electives in the first year, while DS and CS students take Computer Science 1 and 2; additionally, CS students take Computer Organization instead of Foundations of Applied Math. Despite these differences, it is effortless for students to switch between programs, allowing for a diverse set of students to choose the DS program.

B. Second and Third Year Curriculum

By the second year, the DS program comprises classes from both the Applied Math and the Computer Science programs. Namely, DS students take Probability and Statistics, Linear Algebra, and Advanced Statistics from the Applied Math program as well as Data Structures and Algorithms from the Computer Science program. Additionally, there is a required course named Data Science Fundamentals for second-year DS students; this course aims to provide the fundamental knowledge and skills commonly required to solve data-driven problems.

Similarly in the third year, DS students are taking a mixture of Mathematics courses (Multivariable Calculus) and Computer Science Courses (Databases), as well as starting with courses that are designed for students in the Data Science program, or for students obtaining a Data Science Minor. In the Spring of the third year, our DS students are required to obtain their first Co-Op Job Experience. Instead of taking classes, they must be working full-time in a job related to Data Science for at least 12 weeks. This Co-op experience is a transformation for our students; when they arrive back on campus they are more motivated than ever to obtain the remaining skills needed to succeed in industry. This strategic integration of real-world exposure ensures that our graduates are not only well-versed in academic theory but also possess the practical skills and industry awareness necessary for thriving in the professional realm of Data Science.

C. Fourth-Year Curriculum

The culmination of the DS program in the fourth year encompasses a compulsory Machine Learning class, complemented by a diverse array of electives within the Data Science program. Further enriching their skill set, students embark on their second Co-op work experience during the Spring semester, gaining hands-on exposure to real-world Data Science applications. The program concludes in the Summer with a combination of additional electives and a mandatory Senior Design class. In this concluding phase, students collaborate on a group project throughout the semester, applying their accumulated knowledge and skills to address complex challenges in the field of Data Science.

D. Educational Objectives

In this program design, we envision that integrating Computer Science courses, such as Analysis of Algorithms, will not only enrich the curriculum but also pave the way for students to explore software engineering career paths. Simultaneously, Math courses, including Multivariable Calculus, are strategically incorporated to enhance quantitative reasoning skills and provide a solid foundation for advanced Data Science concepts. Furthermore, students delve into domain-specific courses, such as Machine Learning, and See it and Say it with Data Viz, to acquire specialized knowledge that aligns with the evolving landscape of Data Science applications. This multifaceted approach reflects our commitment to equipping BSDS students with a diverse skill set, ensuring they are well-prepared for the intricacies of the field and fostering their success in various professional trajectories.

E. Master of Science in Data Science and 4 + 1

The $4+1$ program is an accelerated academic track designed to provide our undergraduate students pursuing relevant Bachelor's degrees in Applied Mathematics, Computer Science, Cybersecurity, Data Science, and similar with the opportunity to seamlessly transition into a Master of Science in Data Science (MSDS) program. This program follows the launch of our graduate programs, which initially introduced a two-year MSDS for students with STEM degrees.

For students interested in research-related roles, the $4+1$ program allows them to complete both their Bachelor's and Master's degrees in a shorter time frame, saving both time and money. Specifically, students enrolled in the $4+1$ track will need to complete two advanced four-credit undergraduate electives during their BS: one focused on computer science and the other on mathematics. The specified courses are Statistical Thinking in Math and either Big Data Programming or Applied Large Language Models in Computer Science. To be eligible for these advanced courses, students must have completed the prerequisites in Mathematics (Advanced Statistics) and Computer Science (Data Science Fundamentals).

It's important to note that the $4+1$ program, with a total of 32 credits, is slightly more condensed than the traditional two-year MSDS curriculum, which consists of 33 credits. The accelerated program replaces three three-credit graduate courses from the traditional curriculum (Applied statistics for research, Data science foundations, and one DS Elective) with the aforementioned advanced undergraduate courses, streamlining the academic path for students interested in a quicker progression toward a Master's in Data Science.

As in the two-year program, $4+1$ students undertake a comprehensive capstone project spread across two semesters, during which they engage in extensive research, write a detailed treatise, and present their work, showcasing their mastery of Data Science concepts and methodologies.

F. Minor and Related Classes

In addition to the undergraduate and graduate programs, we also have developed a Data Science minor. The Data Science minor provides students with the necessary analytical skills to gather and analyze data and introduces principal components of machine learning to infer from processed data and share results in a meaningful way. To earn the Data Science minor, students must successfully complete 20 credits (5 courses, 4 credits each). To enroll in the minor, students must complete two prerequisite courses in Computer Science and Statistics. The core courses in the minor are Data Science Fundamentals and

Linear Algebra. After completing the core courses, students must choose three electives from three different academic units. They can choose classes from Applied Mathematics, Computer Science, Applied Sciences, Humanities and Social Sciences, or Management. The minor mimics the interdisciplinary approach that was so important in our development of the full programs. The minor has been popular with Applied Math and Computer Science students. It is based on this demand that the faculty-led the charge to build the BSDS program.

For students who may not be pursuing a minor in Data Science, we offer a dedicated course in Data Analytics. This standalone course is accessible to students across all programs, boasting an inclusive approach with no prerequisites. Recognizing the significance of Data Analysis skills and concepts in various life domains, we designed this course to be open to any student eager to delve into the intricacies of data. As part of our vision, we aim to position Data Analytics as an alternative to our algebra-based statistics course. This strategic offering allows numerous majors to seamlessly integrate it into their programs with minimal overhead, equipping students with the essential skills to thrive in our increasingly data-rich world.

G. A Vision for $2 + 2$ Articulation Agreements

To foster a more diverse pathway and student body, we are actively working on articulation agreements with community colleges. Our vision is a $2 + 2$ model, which involves collaborating with community colleges to create seamless pathways for students to transition from their community college programs into our institution. This not only enhances accessibility to education but also contributes to a more diverse and inclusive student body at WIT, aligning with our commitment to providing quality education for all.

At WIT, our ability to support our students largely relies on the tuition they contribute. However, we are currently facing the challenge of declining enrollments across undergraduate programs in the U.S., and this trend is expected to continue for the next decade or longer [31]. This decline can be attributed to various factors, including changes in demographics that show a decrease in the number of college-aged students seeking higher education [32]. Additionally, the increasing costs of education, including tuition fees and operational expenses, have created financial barriers that may discourage prospective students from pursuing a college degree. In light of these challenges, we must take a proactive and sometimes intuitive approach to ensure the ongoing success and sustainability of our institution.

In addressing the expected decline in enrollment, we understand that tackling this challenge involves making judgment calls, as there are no clear metrics on how to navigate it. Our efforts to create programs with new pathways and approaches in Data Science and AI, accompanied by an intensified online campaign and targeted ads, embody an approach that seeks to find a balance between analytical decisions and intuitive choices. We aim to help our community adapt and thrive in a dynamic educational landscape. As educators and leaders in the academic realm, we have recognized the growing importance of Data Science in various industries. Feeling a sense of urgency, we are developing academic programs to meet the rising demand for skilled data scientists among our students.

III. Establishing the Program

In this section, we delve into the steps taken to establish the Bachelor of Science in Data Science program at Wentworth Institute of Technology. We provide an overview of the program's inception and early developments including events aimed at promoting the program to market insights and preliminary enrollment statistics.

A. Events to Promote the Program

In our commitment to empowering the next generation, we co-orchestrated the organization of the Girls STEM Summit at the School of Computing and Data Science. This initiative, aimed at fostering gender diversity and inclusion in STEM fields, not only inspired young minds but also underscored WIT's commitment to promoting equity in education and opportunities. The tangible impact of this event is evidenced by the participation and engagement shown in Figure 1.

Figure 1. Students and faculty at the Girls STEM Summit, Summer 2023

We extended similar activities to other recruitment events, such as the welcome days for the accepted students day and the Wentworth preview days. Faculty volunteered at these events, often during the weekends, and their involvement allowed them to personally connect with prospective students, showcasing the joy of problem-solving through engaging activities and puzzles. This hands-on approach not only made our events more vibrant and interactive but also served as a testament to the dedication of our faculty in inspiring the next generation of STEM enthusiasts.

B. Market Insight

EAB conducted a feasibility study of a Bachelor of Science in Data Science at Wentworth. Their report revealed promising opportunities for program graduates based on strong employer demand trends and positive employment projections. However, EAB also cautioned about the challenging competitive landscape, suggesting potential difficulty for the proposed program to capture student demand. To address this, the report recommended offering accelerated degree options, such as a 4 + 1 program, and incorporating experiential learning opportunities into the curriculum. Both of these suggestions were incorporated into our curriculum.

C. Program Enrollment Statistics

Our first class of Data Science undergraduate students commenced in the fall of 2023. Table II describes the number of applications, admitted students, and students who enrolled in the program. The number of students who will enroll in 2024 remains officially undetermined.

Table II
Number of students for the undergraduate Data Science Program

Year	Applications	Accepted Students	Students Enrolled
2023	77	30	18
2024	68	42	TBA

IV. Diversity, Equity, and Inclusivity

In higher education, the importance of diversity and inclusivity in teaching cannot be overstated [33, 34]. Numerous educators actively structure their courses and labs to enhance diversity, equity, and inclusion (DEI) throughout their teaching practices[35, 36]. Within the realm of Data Science, effective teaching methodologies serve to ensure a comprehensive understanding of diverse perspectives, foster innovation, and equip students with the skills necessary to adeptly navigate real-world challenges within the complex,

interdisciplinary landscape of this field, spanning across diverse industries and contexts [37]. To establish an inclusive educational atmosphere in our Data Science program and prepare students for the diverse working environments of the future, we have undertaken numerous initiatives in collaboration with professors and officers from various departments. The details of these actions are outlined as follows.

A. Faculty Training on Diversity, Equity, and Inclusivity

To cultivate a supportive learning environment for both students and faculty, faculty members must comprehend the significance of such an environment and possess a clear understanding of bias [38, 39]. In alignment with this objective, the Division of Diversity, Equity, and Inclusion (DEI) at Wentworth has instituted a comprehensive series of workshops. The workshop series comprises nine sessions, addressing critical topics including the data-informed approach to identify equity gaps, considerations of identity, inclusive communication strategies, course accessibility, gender bias, the GROW framework, stereotype threats, and various other pertinent subjects. The symposium serves as a platform for faculty to engage in collaborative dialogue, enhancing their understanding of diversity, equity, and inclusion in education. With a wealth of knowledge gained from these initiatives, faculty members make informed adjustments to their teaching methodologies, thereby fostering an open and supportive learning environment for students pursuing majors in Data Science.

B. Establishing a Inclusive Educational Atmosphere

Given the interdisciplinary nature of Data Science, the faculty designs the curriculum to accommodate students with diverse backgrounds and distinct expectations [37, 40]. In our university, the curriculum includes the course "Data Science Fundamentals" to provide the foundational knowledge and skills commonly required to solve data-insightful problems. The course introduces computational and inferential approaches using set off cross-disciplinary skills. This foundational course guides students starting with the basics, such as logging into the collaborative JupyterHub environment and executing notebooks. It introduces core topics such as computational programming, data manipulation and visualization, and introduction to machine learning.

Our $4+1$ graduate students, on the other hand, are offered the introductory course "Applied Statistics for Research." This advanced course encompasses a broad spectrum of fundamental statistical topics as well as advanced topics, such as support vector machines, and unsupervised learning. $4+1$ graduate students with a strong statistical background have the option to waive these courses based on their prior knowledge. This flexibility allows students to tailor their study plans according to their specific requirements and backgrounds, ensuring a customized and effective educational experience in the realm of Data Science.

Moreover, instructors employ a multifaceted approach by providing diverse examples and projects to inspire students with varied backgrounds. Our faculty members, hailing from diverse disciplines such as Computer Science, Mathematics, Statistics, Engineering, and Communications, contribute to the richness of perspectives. This diversity enables instructors to share the theoretical underpinnings of Data Science, appealing to students from a wide array of backgrounds. Illustratively, when covering the topic of data classification, instructors delve into the connection between algebraic topology and data clusters. This approach fosters a profound comprehension of intricate patterns within complex datasets, particularly capturing the interest of students with a mathematical background.

In addition, the instructor specializing in data visualization, rooted in a communications background and housed in the School of Humanities and Social Sciences, incorporates examples and datasets tailored to engage students using their humanities expertise. For instance, if sharing a data visualization from a Congressional hearing, the professor elucidates the significance of maintaining consistency in units across different datasets and emphasizes the importance of a coherent legend.

Furthermore, we empower students to autonomously select project topics aligned with their interests and expectations. Since the inaugural course, "Data Science Fundamentals," we have afforded students the liberty to choose from a plethora of open-source datasets. This flexibility enables students to explore captivating topics, ranging from NFL sports data to social media analysis, and from NYC auto accident

data to global wireless network construction datasets. In a collaborative effort to support students, Data Science professors collaborate with officers from the Wentworth Library to curate an online resources list dedicated to open-source datasets, Data Science, and Machine Learning, accessible through the library website. Within the lectures of "Data Science Fundamentals", library officers deliver guest lectures to disseminate information about these resources and instruct students on leveraging the library to access educational videos and free books in the field. With this comprehensive support system, most students complete their projects productively and submit captivating reports. Some students even leverage their projects and outcomes to secure co-op internships.

Faculty collaborates with the Teaching and Learning Center at Wentworth to enhance the accessibility of our teaching materials. Our institution has implemented the Panorama accessibility tools within the BrightSpace learning management system to achieve this objective. When professors upload teaching materials to the BrightSpace website, students gain the ability to customize their viewing preferences.The Panorama tool facilitates the generation of accessibility reports for each course, enabling instructors to comprehensively review the accessibility features and challenges associated with their teaching materials. This valuable feedback equips professors to make informed decisions and provides practical suggestions for improving accessibility. For students, the Panorama tool generates alternative accessible formats, such as EPUBs, audio podcasts, language translations, and more. Collectively, these tools empower students to adapt teaching materials according to their individual preferences, thereby enhancing overall accessibility.

Through collaborative efforts and curriculum design, Wentworth enhances diversity and inclusivity in the instruction of Data Science.

C. Fostering Diversity, Equity, and Inclusivity Awareness among Data Science Students

In the context of most Data Science courses, students are tasked with the collection or selection of datasets, subsequent analysis, and the presentation of their findings through reports or presentations. To cultivate an awareness of unconscious bias, instructors employ case studies that involve the comparative analysis of different datasets and their corresponding conclusions. For instance, within our introductory Data Science course, we examine the trend of restaurant visitors based on a specific dataset. Following the presentation of our analysis results, students are prompted to identify any shortcomings or limitations within the dataset. Through collaborative discussion, students discern that the dataset exclusively encompasses data recorded between Thursday and Sunday, omitting information from Monday and Wednesday. Subsequently, the class engages in a discourse to evaluate the significance of this omission in the context of data analysis. The discussion unveils numerous restaurants offering deals or happy hour discounts during workdays, attracting substantial visitors during weekdays. Consequently, any conclusion drawn from an analysis involving this missing data is deemed biased and inherently unreliable. This illustrative example serves to underscore to students that data collection practices can serve as a source of unconscious bias, emphasizing the potential oversight of opinions from underrepresented groups.

Through our instructional methods, students gain valuable insights into the intricacies of unconscious bias in Data Science and the critical importance of comprehensive and representative data collection practices. This deliberate focus on diversity and inclusivity contributes to the establishment of an unbiased learning environment amongst our students, ensuring that they carry these principles into their future careers, thereby benefiting a broader audience upon graduation.

V. Conclusion

The development of the Bachelor of Science in Data Science program at Wentworth was a collaborative and interdisciplinary effort. It was designed to ensure a robust learning experience for our students. A key aspect of our approach was a focus on diversity and inclusivity within the teaching and learning environment. We incorporated these values into every stage of our program. By detailing the specifics of our curriculum, we hope to provide a useful blueprint for others. We are particularly optimistic about the program being offered in three or four years, both with mandatory co-ops, a $2+2$ pathway for community college students, and a $4+1$ option leading to a graduate degree in Data Science. This manuscript could help guide other institutions in developing programs that prepare Data Science students to successfully

navigate the diverse challenges they will encounter in both professional and educational environments.

References

[1] M. T. Özsu, "Data science—a systematic treatment," Commun. ACM, vol. 66, no. 7, p. 106–116, jun 2023. [Online]. Available: https://doi.org/10.1145/3582491

[2] F. Provost and T. Fawcett, "Data science and its relationship to big data and data-driven decision making," Big data, vol. 1, no. 1, pp. 51–59, 2013.

[3] Z. Sun, L. Sun, and K. Strang, "Big data analytics services for enhancing business intelligence," Journal of Computer Information Systems, vol. 58, no. 2, pp. 162–169, 2018.

[4] S. Negash, "Business intelligence," Communications of the association for information systems, vol. 13, no. 1, p. 15, 2004.

[5] S. Siuly and Y. Zhang, "Medical big data: neurological diseases diagnosis through medical data analysis," Data Science and Engineering, vol. 1, pp. 54–64, 2016.

[6] D. A. Jenkins, M. Sperrin, G. P. Martin, and N. Peek, "Dynamic models to predict health outcomes: current status and methodological challenges," Diagnostic and prognostic research, vol. 2, no. 1, pp. 1–9, 2018.

[7] J. Chen, K. Li, H. Rong, K. Bilal, N. Yang, and K. Li, "A disease diagnosis and treatment recommendation system based on big data mining and cloud computing," Information Sciences, vol. 435, pp. 124–149, 2018.

[8] L. Sun, C. Liu, C. Guo, H. Xiong, and Y. Xie, "Data-driven automatic treatment regimen development and recommendation," in Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 1865–1874.

[9] J. Cattell, S. Chilukuri, and M. Levy, "How big data can revolutionize pharmaceutical r&d," McKinsey Center for Government, vol. 9, 2013.

[10] T. P. Tran and D. Meacheam, "Enhancing learners' experience through extending learning systems," IEEE Transactions on Learning Technologies, vol. 13, no. 3, pp. 540–551, 2020.

[11] M. W. Kurniawati, S. Anitah, and S. Suharno, "Developing learning science teaching materials based on scientific to improve students learning outcomes in elementary school," European Journal of Education Studies, 2017.

[12] C. Romero and S. Ventura, "Educational data science in massive open online courses," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 7, no. 1, p. e1187, 2017.

[13] J. Scheerens, "Evidence based educational policy and practice: The case of applying the educational effectiveness knowledge base," Journal of Educational, Cultural and Psychological Studies (ECPS Journal), vol. 1, no. 9, pp. 83–99, 2014.

[14] M. Bienkowski, M. Feng, and B. Means, "Enhancing teaching and learning through educational data mining and learning analytics: An issue brief." Office of Educational Technology, US Department of Education, 2012.

[15] T. Kabudi, I. Pappas, and D. H. Olsen, "Ai-enabled adaptive learning systems: A systematic mapping of the literature," Computers and Education: Artificial Intelligence, vol. 2, p. 100017, 2021.

[16] P. Johanes and L. Lagerstrom, "Adaptive learning: The premise, promise, and pitfalls," in 2017 ASEE Annual Conference & Exposition, 2017.

[17] C. Richardson, "Council post: The next revolution in global e-learning," Jul 2018. [Online]. Available: https://www.forbes.com/sites/forbestechcouncil/2018/07/25/the-next-revolution-in-global-elearning/?sh=66f38d357c25

[18] Y. Niu, L. Ying, J. Yang, M. Bao, and C. Sivaparthipan, "Organizational business intelligence and decision making using big data analytics," Information Processing & Management, vol. 58, no. 6, p. 102725, 2021.

[19] S. Fan, R. Y. Lau, and J. L. Zhao, "Demystifying big data analytics for business intelligence through the lens of marketing mix," Big Data Research, vol. 2, no. 1, pp. 28–32, 2015.

[20] Sep 2023. [Online]. Available: https://www.bls.gov/ooh/math/data-scientists.htm

[21] M. A. Halwani, S. Y. Amirkiaee, N. Evangelopoulos, and V. Prybutok, "Job qualifications study for data science and big data professions," Information Technology & People, vol. 35, no. 2, pp. 510–525, 2022.

[22] J. C. Adams, "Creating a balanced data science program," in Proceedings of the 51st ACM technical symposium on computer science education, 2020, pp. 185–191.

[23] K. D. Schubert and M. D. Rossetti, "Creating a multi-college interdisciplinary bs data science program with concentrations," in 2021 ASEE Virtual Annual Conference Content Access, 2021.

[24] P. Anderson, J. Bowring, R. McCauley, G. Pothering, and C. Starr, "An undergraduate degree in data science: curriculum and a decade of implementation experience," in Proceedings of the 45th ACM technical symposium on Computer science education, 2014, pp. 145–150.

[25] P. T. Tosic and J. Beeston, "Designing undergraduate data science curricula: A computer science perspective," in 2018 ASEE Annual Conference & Exposition, 2018.

[26] D. Liu, "Prepare data science program student outcomes and curricula for abet accreditation," in 2022 ASEE Annual Conference & Exposition, 2022.

[27] A. D. S. T. Force, Computing Competencies for Undergraduate Data Science Curricula. New York, NY, USA: Association for Computing Machinery, 2021.

[28] A. f. C. M. A. Joint Task Force on Computing Curricula and I. C. Society, Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science. New York, NY, USA: Association for Computing Machinery, 2013.

[29] A. S. Association et al., "Curriculum guidelines for undergraduate programs in statistical science," Retrieved December, vol. 15, p. 2017, 2014.

[30] B. Cassel and H. Topi, "Strengthening data science education through collaboration," in Workshop on data science education workshop report, vol. 7, 2015, p. 27.

[31] M. I. Service, "Sector in depth: Higher education us closures while still rare will gradually increase," Jul 2023. [Online]. Available: https://www.moodys.com/research/Higher-Education-US-Closures-while-still-rare-will-gradually-increase-Sector-In-Depth–PBM$_1$172586?$cid = YJZ7YNGSROZ5414$

[32] M. Kline, "The looming higher ed enrollment cliff," 2019. [Online]. Available: https://www.cupahr.org/issue/feature/higher-ed-enrollment-cliff/

[33] R. Barnett, "Leading with meaning: Why diversity, equity, and inclusion matters in us higher education," Perspectives in Education, vol. 38, no. 2, pp. 20–35, 2020.

[34] C. C. Rogers and R. S. Valdez, "Designing for diversity, equity, and inclusion in systems engineering education," in 2021 ASEE Virtual Annual Conference Content Access, 2021.

[35] K. Miller-Roberts, "Designing for diversity, equity, and inclusion in engineering curricula," Ph.D. dissertation, 2022.

[36] A. Miguel, J. M. Sloughter, J. M. Dorsey, R. Hartley, F. J. Shih, and J. Crevier, "What's next? from analysis to action," in 2021 CoNECD, 2021.

[37] M. Dogucu, A. A. Johnson, and M. Ott, "Framework for accessible and inclusive teaching materials for statistics and data science courses," Journal of Statistics and Data Science Education, vol. 31, no. 2, pp. 144–150, 2023.

[38] K. Thomas, D. Satterfield, J. Sanders, A. Kirn, and K. Cross, "Call without response: Faculty perceptions about diversity, equity, and inclusion," in 2022 ASEE Annual Conference & Exposition, 2022.

[39] A. Jayakumar, L. Abrams, L. Sheppard, S. Siliman, and T. M. Calbert, "Increasing faculty participation in pedagogical diversity and inclusion activities," in 2021 ASEE Virtual Annual Conference Content Access, 2021.

[40] N. Hatanaka, "Inclusive environments through data science and ict," in Proceedings of the International Conference on ICT Application Research The 1st International Conference on ICT Application Research. International Institute of ICT Application Research, 2023, pp. 1–10.