Applicant: Nathaniel del Rosario | Website: natdosan.github.io | Email: nadelrosario@ucsd.edu

"The data you will be working with is vital in detecting down-syndrome as early as possible in pregnancy - great question," my manager answered on the first day of work as I learned more about my project at my most recent internship. As someone who was misdiagnosed with down-syndrome, such an unexpected moment confirmed that my reason to pursue data science had come full circle. My strengths lie in the harmonious intersection of data science, creativity, and their applications to different domains. Consequently, I have demonstrated my ability to thrive in this intersection through multiple internships, extracurricular & research projects, and academic leadership positions. Ultimately, I am applying to the Berkeley MEng Program because the curriculum, resources, and emphasis on leadership complement my desire to become an industry leader in data science applications towards improving healthcare & biotechnology, and I have an established history of steps towards this goal throughout my undergraduate studies.

Beyond my coursework, I have engaged in multiple internships, research projects, as well as leadership positions in and out of the classroom. During my first internship at Chan Zuckerberg Biohub, I developed a pipeline and dashboard that supports CRISPR screen queries, data transformations, and important visualization functionalities, helping researchers identify significant genes which contribute to virus transmission. Additionally I wrote the entire documentation to ensure that the project would be easily transferable to future employees. Though a beginner, I learned a lot about working in an industry setting contributing directly to ongoing research with individuals from Berkeley, Stanford, and UCSF, pair programming with my manager, and more importantly how a data scientist can contribute to the big picture solution in a large research project by considering the perspectives of those who benefit most from my work.

By working with graduates from these schools, I began considering graduate school because of the value it would bring to my transition to the industry setting, as well as knowing I felt comfortable learning in such an environment. Jumping to my most recent internship as a data science and bioinformatics intern, my contributions spanned the entire data science life-cycle, ranging from data collection, preprocessing, data integrity, writing tests to improve fault tolerance, explaining the architecture to my team and lab scientists, and contributing to analysis algorithms to predict down-syndrome early on during pregnancy so families have enough time to make a decision to terminate or continue pregnancy. In addition to building upon my previous applications of data science and coding to healthcare, I felt the most impact and desire to succeed with this project because its impact will be folds greater when my work eventually culminates in a working aneuploidy test for future mothers across the world, further contributing to my decision to pursue graduate studies.

Leadership has also been an integral part of my goals as a student and a data scientist. As an instructional assistant for 2 data science courses serving over 1200 students, I've contributed to my department through helping many new students become data and coding literate by guiding them through my past mistakes. From tangible efforts such as guides on how to manipulate data using Pandas, video demonstrations, grading and beta-testing exam, quiz, and homework questions for the course staff, I've provided a sandbox for students to make

mistakes and allow them to grow as the course went on. Overall, I believe leadership is a reflection of my career goal - to manage data science and machine learning projects and ensure they are used for improving our lives. Through the culmination of these diverse experiences in industry and academia, I realized that maturing as a data scientist happens when acknowledging our skills are most valuable when we collaborate and apply our technical skills to domains like healthcare and realize results only numbers and code could lead us to.

This layering of industry and leadership experience has helped me prepare for graduate school and a career beyond. At UC Berkeley, I would be most excited to pursue studies in deep learning architectures and applications (CS 282, 285), cloud computing & infrastructure (CS 267), and more machine learning applications to healthcare. Having worked in industry-research based environments, I know first hand what working with former Berkeley graduate students in a research environment is like and even took CS188 one summer to familiarize with Berkeley's rigor. Even when I was not working at these companies, I often found myself looking at Masters and PhD students' portfolios at BAIR, Sky, BEAMMO, and CTML. Specifically I would like to engage in research similar to Dr. Petersen's "Adaptive Designs for HIV Control" (CTML) and especially Professor Irene Chen's initiative on making machine learning in healthcare more robust and equitable (Chen Lab). Following the trend of my industry experience in ML for healthcare/biotech, my research interests focus on algorithmic bias in healthcare ML as well as making AI powered healthcare tools such as fMRI classifiers more accessible. I aim to answer questions such as "what kind of biases occur and how do they differ in different supervised learning methods", "why do these biases differ between on ethnic majority/minorities?", and "how can we train future models to account for these potential biases?" The resources Berkeley provides for multidisciplinary research problems like these are second to none and provide a perfect setting for my intellectual interests to thrive.

To conclude, a masters in the field is an intermediate goal. In the big picture, I hope to explore as many applications and current problems in data and ML to improve its impact on biotech and other related domains. As previously mentioned, my career goal is to direct a team that is dedicated to utilizing data science and ML to improve our lives. As demonstrated through my undergraduate experiences in industry, academia, and within the student body, I excel in an environment where I can collaborate and share my knowledge to others. Graduate school at Berkeley would provide such an environment in both coursework, faculty, and student body where I can continue to mature and learn. Reflecting on my industry contributions to virus transmission research and non-invasive prenatal down-syndrome testing through data science, the imprint I left on my students as an assistant in the data science department, and my future career goals, I am thrilled to pursue a graduate education to continue growing as a student, data scientist, and leader.