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

Statement of Purpose

"This data is key in detecting down-syndrome early in pregnancy" my manager explained on my first day at my recent internship. As someone who was misdiagnosed with down-syndrome, this unexpected moment confirmed that my reason to pursue data science had come full circle. I chose data science because my strengths lie in the harmonious intersection of computer science, creativity, and their applications to different domains and real world problems. During my undergraduate degree, I pursued this goal of applicable, impactful data science through multiple internships, extracurricular & research projects, and academic leadership positions. Ultimately, I am applying to Berkeley's MEng EECS Program because the curriculum, resources and research opportunities, and emphasis on leadership align with my career goal to become an industry leader in applying data science towards improving healthcare, biotechnology, and other domains.

Engineering Efforts. Beyond my coursework, I aimed to apply my knowledge to real world problems. During my first internship at Chan Zuckerberg Biohub, I developed a database visualization tool to help researchers identify significant genes which contribute to virus (Ebola, COVID variants, Dengue, etc) transmission. My contributions reflected my interest in applying data science to health related problems as they helped researchers better understand how dangerous viruses infect humans. Through this I learned about contributing directly to ongoing research with individuals from Berkeley, Stanford, and UCSF in an industry setting and how I can positively impact society through data science.

After working with graduates from these schools, I knew I was comfortable collaborating in this environment and considered the value graduate school would bring to my career goal. In my most recent internship as a data science and bioinformatics intern at Bio-Rad, I was involved in developing a test to predict down-syndrome during pregnancy so families have enough time to make a decision to terminate or continue pregnancy. Similar to my project around virus transmission, my impact will be folds greater when my work eventually culminates in a working aneuploidy test for future mothers across the world, further reinforcing my career interest and contributing to my decision to pursue graduate studies. Through these experiences, I realized that my skills as a data scientist are most valuable when I collaborate with domain experts and apply data science and ML to domains like healthcare and biotechnology.

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 empowering new students to learn data science and ML 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.

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

Future Plans & Fit. At Berkeley, I would be most excited to pursue coursework in deep learning architectures and applications, cloud computing & infrastructure, and more machine learning applications to healthcare; however, these courses are only one aspect as to why Berkeley's environment provides me a foundation to achieve my future career goals. To preface, I enrolled in CS188 during a summer internship to apply its topics to my project and familiarize myself with Berkeley's academic environment. Unexpectedly, I realized the faculty placed an importance on ML/AI for societal impact - a culture of integrity that perfectly aligns with why I pursued data science. My current research experience also aligns with this culture; to prepare for potential graduate research, I was guided by Professor Ilya Zaslavsky where I designed a regression model and robust accessibility metric to fairly score regions on how equitable public transit infrastructure is. Additionally, for my senior capstone, I am involved in Professor Zhiting Hu's group centered around LLM reasoning to continue pursuing AI applications to healthcare benchmarks and make LLMs more reliable.

In line with my experience in data science for healthcare/biotechnology, my future research interests focus on algorithmic bias in healthcare ML as well as making AI powered healthcare tools more accessible and reliable. Berkeley captivates me because there are many labs already working on these types of problems. For example, 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). Having access to these opportunities is important because they currently tackle problems similar to ones I found most impactful during my internships. Specifically, they aim to answer questions like "What kind of biases occur in supervised learning methods," and "How can we train future models to account for these potential biases?" Knowing faculty like Professor Chen are investigating problems of similar impact to what I have contributed to previously is enough reason alone as to why Berkeley intrigues me intellectually. This institution provides unparalleled resources and freedom for multidisciplinary research problems like these - the perfect setting for my intellectual interests to thrive.

To conclude, a masters in the field is an intermediate goal. Referencing my career goal, I hope to explore as many applications and current problems in data and ML to improve its impact on biotech and other related domains. I excel in an environment where I can collaborate and share my knowledge to others, and Berkeley would provide such an environment in coursework, faculty, and student body where I can continue to mature and learn. Reflecting on my contributions to virus transmission research and prenatal down-syndrome testing through data science, the imprint I left on my students as an instructional assistant, and my future career goals, I am thrilled to pursue a graduate education at *Berkeley* to continue growing as a student, data scientist, and leader as I transition to industry.