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

Statement of Purpose

"This data is key in detecting down-syndrome early enough to terminate pregnancy" my manager explained on the first day at my recent internship. As someone who was misdiagnosed with down-syndrome, issues like these drove my desire to improve technological shortcomings through data. I chose data science because my strengths lie in the harmonious intersection of computer science, creativity, and their applications to real world problems like preventing misdiagnosis. Ultimately, I am applying to Berkeley's MEng EECS Program because the curriculum, resources and research opportunities, and emphasis on innovation and leadership align with my work experiences, leadership, and 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 querying and visualization tool to help researchers identify significant genes which contribute to virus (Ebola, COVID variants, Dengue, etc) transmission. Consequently, my contributions helped researchers better understand how dangerous viruses infect humans. Through this I learned about contributing directly to ongoing research with individuals from innovative institutions like Berkeley, Stanford, and UCSF and how I can positively impact society through data science. Understanding virus infections is invaluable, and using data to learn about these interactions inspires me to pursue a masters.

Building upon this, in my most recent internship as a data science and bioinformatics intern at Bio-Rad, I ensured that the limited amount of maternal data was protected and its integrity stayed intact. This was significant to ensuring families have enough time to decide to terminate pregnancies and also preventing misdiagnoses similar to mine. Similar to my previous 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 initiative of impactful data science. Through these experiences, I realized that my skills are most valuable when I collaborate with domain experts and utilize data science to positively shape domains and tackle problems similar to my experiences

Leadership. Collectively, these industry experiences have helped me share the value of data science. As an instructional assistant for data science courses serving over 1200 students, I've contributed to my department through empowering students to learn data science. From creating exam and homework questions, holding office hours, and creating helpful guides on Pandas, I've provided a sandbox for students to make mistakes and witnessed their growth over time. Relating to my misdiagnosis, I want to foster students to accurately represent data and represent as many perspectives and backgrounds in the industry. Another aspect that excites me is the Fung Institute's attention to detail on leadership and career development; specifically the intertwining of leadership courses in the curriculum will help me develop as I share my engineering experiences with others. Overall, I believe leadership is a reflection of my postgraduate goal to lead and ensure data science and ML are used for societal impact.

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

Future Plans & Fit. To prepare for a graduate curriculum, I cross enrolled and excelled in graduate classes such as Deep Learning, Recommender Systems, and Computer Vision. At Berkeley, I would be most excited to pursue coursework that strengthens my knowledge such as Cloud Computing & Infrastructure, Data Engineering, and healthcare ML applications. This curriculum will expand my technical skills for impactful data science towards healthcare; however, it is only one aspect as to why Berkeley's environment provides me a foundation to achieve my 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 integral to 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 my initiative to make ML/AI more reliable.

Relatably, Berkeley's focus on innovation perfectly aligns with my engineering experience in data science for healthcare/biotechnology. My future research interests are increasing accessibility and reliability and reducing algorithmic bias in healthcare ML/AI. Berkeley captivates me because there are many labs already working on these problems of societal benefit. 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 healthcare ML unbiased and equitable (Chen Lab). Having access to these opportunities is important because they currently tackle issues similar to projects where I influenced societal improvement during my internships. Knowing faculty like Professor Chen are investigating problems of similar impact to what I contributed to previously is a strong reason alone as to why Berkeley intrigues me intellectually. This institution provides the curriculum, faculty, unparalleled resources, and freedom for multidisciplinary research problems like these - the perfect setting to achieve as a graduate student and future data scientist.

A masters in the field is an intermediate goal. 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 research opportunities where I can continue to mature and learn as I step towards my goal of applying data science towards societal impact in my postgraduate career. Reflecting on my contributions to virus transmission research and prenatal down-syndrome tests 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.