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 a misdiagnosis. Ultimately, I am applying to Harvard's MS Data Science because the curriculum, research opportunities, and emphasis on innovation and 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 sought 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 viral transmission was invaluable, and using data to investigate important problems like these inspired me to pursue a masters.

Building upon this, at my most recent internship at Bio-Rad, I built systems to protect the integrity of a limited amount of maternal data, ensuring that False Positives/Negatives for down-syndrome and other conditions were properly recorded. Maintaining these misclassifications was vital as each data point was also an affected family. A single error equated to potentially terminating a healthy pregnancy and vice versa. I live knowing I could have been a terminated pregnancy and therefore feel strongly for data driven technology like Bio-Rad's. Having completed my internship, my influence will be folds greater when my work eventually culminates in a working aneuploidy test for future mothers and families. Reflecting on this perspective, I am committed to applying data science to problems of great downstream magnitude like my misdiagnosis, and I plan to pursue graduate school to expand my skills for this purpose.

Academic Service. Collectively, these industry experiences have helped me share the value of data science. As an instructional assistant for data science courses serving over 1500 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 programming, I've provided a sandbox for students to make mistakes and witnessed their growth over time. Relating to my misdiagnosis, I wanted to foster students to accurately represent and utilize ML to ensure data driven tools are reliable. I would like to learn how to become a leader in the field through embracing a challenging, diverse curriculum, and I believe Harvard's close knit culture fits my academic background perfectly.

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

Future Plans & Fit. To prepare for a graduate curriculum, I cross enrolled in graduate classes such as Deep Learning, Recommender Systems, and Computer Vision. At Harvard, I hope to take a deeper dive in ML methodology to further develop my knowledge for impactful data science towards healthcare; however, it is only one aspect as to why Harvard's environment provides me a foundation to achieve my career goals. I noticed the faculty places an importance on ethical 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, this fall I am involved in Professor Zhiting Hu's group centered around LLM reasoning to continue my initiative to make ML/AI more reliable.

Similarly, Harvard's focus on innovation perfectly aligns with my engineering experience in data science for healthcare/biotechnology. My future research interests are increasing accessibility, reliability, and equity in healthcare ML/AI. Harvard captivates me because there are many labs working on problems that resonate with my misdiagnosis. For example, I would like to engage in research similar to Dr. Hormoz's "Deep learning-based 3D cell segmentation framework for future image-based disease detection" (Hormoz Lab) and especially Professor Zitnik's initiative on equitable AI for individualized diagnosis and treatment (Zitnik Lab). In Zitnik's paper, "DeepPurpose: a deep learning library for drug–target interaction prediction" she approached questions such as "how can we make ML/AI tools accessible for everyone in healthcare?" and "what post model-deployment considerations should we make in the context of a user with limited DL experience?" Having access to these opportunities is important because they resonate with my goal of making ML more reliable and equitable for societal impact. Knowing faculty like Professor Zitnik are investigating problems of similar nature to what I previously contributed to is a strong reason alone as to why Harvard intrigues me intellectually, and I plan to pursue the Master's Thesis track if admitted. This institution provides the curriculum, resources, and unparalleled faculty for multidisciplinary research problems like these, and will allow me the freedom to learn from minds like Zitnik to achieve as a graduate student and future data scientist.

Ultimately, my goals lie beyond a masters in the field. In the future I envision myself leading an R&D team to solve problems through data. This program would provide an environment where I can continue to mature, learn as I take courses at the cusp of innovation, and allow access to research opportunities that resonate with problems similar to my misdiagnosis. Reflecting on my contributions to virus transmission research, prenatal down-syndrome tests, and my future career goals, I am thrilled to pursue a graduate education at Harvard to continue growing as a student, data scientist, and leader.