Statement of Purpose

Prathmesh Savale
(Statistics-Computer Science Joint Masters, Fall '21)

The idea of using mathematics, statistics, and computation to analyze financial markets, create self-driving vehicles or synthesize vaccines for novel viruses, etc. has always intrigued me. This fascination motivated me to develop a skill set in Computational and Data Science. I chose Computer Engineering for my undergraduate studies, because of its heavy emphasis on mathematics and computation. I developed a strong understanding of linear algebra, calculus, and probability theory, but found the exposure to the specific courses in Data Science such as Bayesian methods, optimization, stochastic modeling, very limited. This gap in skill set is what motivates me to pursue an advanced degree in computing with a focus in statistics.

In 2015, I graduated from the University of Pune with a First Class with Distinction grade. I realized that the best testament for my academic knowledge would be an endorsement from real world industry experience. With my first job at Mu Sigma, one of India's largest pure play data analytics companies I focused on developing skill set in predictive and prescriptive analytics. Working with Apple and Tesco as part of client engagements in Mu Sigma helped me develop a core skill set in exploratory data analysis, data wrangling, feature engineering, and use of various statistical learning techniques in real-world scenarios. I also gained expertise in creating production ready CI/CD pipelines using distributed frameworks like spark and airflow to parallelize data preprocessing, parameter tuning and model building processes. In recognition for my work in these projects, I received the Spot and Impact Award from Mu Sigma, given only to the top 2 percent employees based on performance.

My most recent assignment with Kiewit on projects that involved natural language processing piqued my interest in using unstructured data to generate business insights. I was particularly drawn toward the potential of deep learning (DL) as a powerful enhancement to the traditional statistical modeling paradigm used in predictive analytics. Subsequently, I focused my attention to understand neural networks and some of the popular frameworks used to implement them. To continue working in a role that would help me develop expertise in DL and enable me to employ my skills as a developer, I moved to work with LTI. Subsequently, I also took the TensorFlow Developer Certification test to deepen my understanding of the framework. Currently, I am working on enhancing the DL capabilities of LTI's proprietary product called Leni. My work encompasses building FastAPI based microservices using TensorFlow and scikit-learn for forecasting and anomaly detection. The APIs are integrated into Leni as part of the TFX airflow pipelines and deployed in production.

My area of interest lies in developing automated systems to generate insights on unstructured data. I chart my career goals as a data scientist with leading institutions. After graduation, similar to my current role, I want to continue developing ML integrated product frameworks leveraging autoML for insight generation on structured and unstructured datasets. Five years down the line, I hope to develop my product suite that will help business owners make data-driven decisions with quicker turnarounds, eliminating the standard boilerplate required for prescriptive analytics.

After spending 5 years since my undergraduate degree, working on data science specific projects across the industry, I realized that the gaps in my understanding of applied statistics and mathematics necessary for Data Science can only be filled in an academic setting. Specifically, a program offering a dedicated curriculum in statistics along with computation. Because this skill set is also going to help me in my long term goal of developing an autoML based product suit for prescriptive analytics. The Statistics-Computer Science joint masters program at Purdue caters perfectly to my requirements since its courses are offered by both the Statistical and Computer Science departments.

Thank You for reviewing my application.