Academic Statement of Purpose

Saige Rutherford
Computer Science and Engineering PhD
UMID:

The world is highly complex and nonlinear, and so is the path which led me here. While the beginning of my education and research career was somewhat serendipitous, I am proud to say that my decision to pursue a PhD in Computer Science to study **the integration of artificial intelligence and application of machine learning to improving health,** has been carefully and fully thought out.

During my undergraduate education as a student of physics, I learned critical ideas and life lessons. First and foremost, studying physics taught me how to make mistakes and still persevere, which in retrospect, is a most useful skill for life. I am not intimidated when confronting difficult problems in science because I have built strong problem-solving skills and I have learned that it is a strength to ask for help sometimes. Five years of research experience (2.5 years full time) has motivated me to pursue a PhD, allowed me to learn how to perform rigorous research and exposed me to a range of topics such that I have a clear picture of the type of PhD I want to pursue.

My current research is in Psychiatry at the University of Michigan where I work as a research computer specialist under the supervision of Chandra Sripada, MD (Psychiatry), PhD (Philosophy). Our research focuses on improving multi-modal neuroimaging methodology. I work closely with a team of analysts who provide computational and statistical support to graduate students, post-docs, and faculty within Psychiatry and Psychology. We use supervised and unsupervised machine learning models to find patterns in brain images that predict cognitive and clinical phenotypes. The project I am most proud of is my work developing the Brain Basis Set, a parsimonious feature set for quantifying inter-individual differences in the brain, that improves interpretation and prediction compared to baseline brain-behavior predictive models. In this project I was involved in the development of the approach, the programing (in python) of aspects of the algorithm, and in the analysis of a very large data set for training, testing, and validating the model. In the Brain Basis Set project, we were among the first to use leave one ***site*** out cross-validation and properly control for a range of confounding variables. To date, this basis set has produced some of the best predictive models in the field, for a range of behavioral phenotypic measures. Highlights of other projects I have worked on include the use of dynamic causal modeling to studying the visual system in schizophrenia, quantifying social dysfunction in autistic patients, developing graph theory approaches to examine neural abnormalities in children who stutter, investigating brain connectivity changes pre/post mindfulness training in PTSD patients and changes pre/post cognitive behavioral therapy in OCD patients.

As an analyst in Psychiatry, I also work closely with collaborators in Computer Science (**Jenna Wiens, Danai Koutra**) and Statistics (Liza Levina) to develop better methodology for identifying subtle patterns in brain images. In 2018 I started working with Jenna Wiens, on a range of machine learning projects: a domain guided 3D convolutional neural network for predicting age from structural MRI images, a CNN to segment the fetal brain from functional MRI images, and currently a reinforcement learning paradigm using the Michigan Genomics Initiative data to discover heterogeneity in opioid prescription patterns aimed at minimizing opioid prescriptions. In working with Jenna in CS, I have also been exposed to projects such as **Emily Provost Mower**'s work on context specific algorithms to detect emotional shifts in speech in bipolar patients and **Rada Mihalcea's** recent precision health grant, "Precision Counselor: Natural Language Processing for Enhanced Behavior Counseling". These projects have also deeply inspired me. There is no single source of data that can represent human health. To that end, work is needed to build models of human health that can integrate diverse sources of data. There is exponential growth in neuroimaging, genetics, mining electronic medical records, and learning speech patterns from recorded clinical interviews. Research that integrates these disparate data sources is currently lacking. I am drawn to computer science, and specifically machine learning, because of its potential to draw from these different data sources in order to advance patient care. **Computer science as the foundation of my graduate school education will allow me to develop a deeper understanding of machine learning so I can effectively use these complex data sources to improve psychiatric diagnosis and treatment.**

During my PhD, I envision improving federated learning and differential privacy algorithms, to assuage data sharing concerns and build on the wealth of neuroimaging and clinical data that is available. Multi-task and transfer learning algorithms are ideal to probe the dimensions of RDoC, a research framework created by the National Institute of Health for investigating mental disorders that shifts from a categorical to a continuous perspective for describing mental health and illness. Causality in machine learning, as discussed in "The Book of Why" by Judea Pearl, is a topic that fascinates and inspires me. Neuroimaging methods for functionally mapping the brain are dominated by correlations. When we no longer work with the observational distribution, but a distribution where certain variables or mechanisms have changed, (and potentially perturb these systems, say by trans magnetic stimulation), we can make casual inferences from brain data, and I find this topic appealing. There are also open problems associated with using multi-site imaging data to train deep learning models, specifically tackling the cross-site confounds need to be addressed.

My ultimate career goal is to develop applications of machine learning (with a specific interest in AI-adjacent methods) in psychiatric settings to improve treatment plans through prediction of treatment outcomes. The best way to accomplish this goal, is to become a computer scientist. I love linear algebra, statistics, electronics, programming, and developing new ways to analyze data. Within the existing strong interdisciplinary collaboration between medical and engineering disciplines at the University of Michigan, **I seek to strengthen the connections between psychiatry and computer science research and envision myself one day joining the faculty of a major university as a professor of Computer Science and Psychiatry.**

I bring a unique perspective to my work in computer science. The fields of computer science, machine learning, and data analysis are increasingly intertwined, and novel algorithm development needs to be designed with "the end in mind". The "end" here represents not only the application of these algorithms, but specifically the platforms they will be run on, the people who will be using them, and the impact on clinical outcomes. All too often consideration of outcomes is not sufficiently addressed when new analysis tools are created. As computer science becomes an integral part of every researcher's toolkit, the need for this mindset is crucial for successful adoption of these tools. **Trained as a physicist, I have come to this perspective working as a computer scientist in psychiatry, and I believe this is an important strength. I am well positioned to think about the novel development and application of machine learning algorithms from both perspectives: the scientist and the physician.**

My goals are ambitious, but realistic. Machine learning may never be able to predict suicide, overdose, or a manic episode the morning before it occurs. But I believe it will one day help us detect periods of increased risk for such adverse psychiatric events – the central aim of my work will be to bring this future closer.