My interest in graduate school is motivated by a combination of my fascination with machine learning and my passion for the research process more generally. An introductory course in machine learning during my sophomore year gave me my first taste of the discipline, and the opportunities that I have had to perform machine learning research since then have only grown my interest in the field. My projects as an undergraduate have included time series prediction with random forests of decision trees, generative modeling with generative adversarial networks, and image volume analysis with large 3D convolutional networks. These experiences have cultivated in me a desire to make machine learning research a more central component of my life.

I completed my first independent research project with Arvind Narayanan, assessing a proposed strategy for predicting movements in the price of Bitcoin and trading it profitably. While surveying the literature at the beginning of the semester, I encountered a paper by a professor at MIT describing a strategy that earned an impressive triple-digit return over the course of several months, and I decided that rigorously replicating and understanding these results would be an interesting project in and of itself. Ultimately, I found that although the researchers' result was valid, it took advantage of an assumption that the 'spread' between the bid and ask price of Bitcoin could be ignored. Thus the predictiveness of the features that they had extracted came entirely in the illiquidity of the market; when I implemented and re-ran their backtest but respected the bid/ask discrepancy (as any practical application of the strategy would have to), I found that their strategy netted almost exactly no profit or loss. As my first substantial academic research experience, I learned several important lessons from this project. Most importantly, I observed firsthand how important assumptions are when doing research, *especially* when that research is interdisciplinary; domain knowledge is crucial, and applying successful techniques across domains can sometimes yield results that are easy to misinterpret without familiarity with the specifics of the target field. Because machine learning is so broadly applicable, it is nearly always the case that research is interdisciplinary in some way, and thus this danger rears its head quite frequently. Second, I learned that goals change; my final project was only loosely related to my original idea. Finally, I learned that sometimes, no result is itself a result. This lesson might be the most enduring of all.

An independent project that really energized me about deep learning research, and specifically adversarial learning, was when I worked with Sandra Batista (now at USC) on a project to improve the stability of training generative adversarial networks, an exciting new training paradigm for neural networks with an enormous range of possible applications. We contrasted the efficacy of a novel technique, which we called adversarial minibatch selection, an alternative to minibatch discrimination, with several simpler techniques for stabilizing GAN training. We demonstrated this strategy can both improve sample quality and avoid low-entropy generators on MNIST; more investigation is necessary to determine if this technique generalizes to more complex domains. While cryptocurrencies and automated trading were (relatively) mature topics, GANs were (and still are) a very recently introduced idea when I performed this research, and the experience of performing research in an extremely quickly-changing field was invaluable. I specifically remember adjusting the course of the project during the semester due to a new result from USC addressing one of the project's key unknowns at the time regarding the effects of weight normalization. From this project, I also gained both familiarity with and an appreciation for the raw engineering skills that are necessary for carrying out research but often easy to overlook. In addition, while spending the summer at Google this year, I was able to talk to Ian Goodfellow, who invented GANs, about this project over lunch; his account of how much progress there is left to be made in the field of machine learning is deeply motivating to me.

I am now working on a year-long senior thesis addressing the challenge of aligning large volumes of anisotropic 3D biological imagery in Sebastian Seung's neuroscience lab. The lab's goal is to map every neuron and connection in the brain, or the connectome, and this ambitious project consists of a large analysis pipeline spanning roughly 25 lab members. An important stage in that pipeline is to adjust the slices of brain imagery so that consecutive slices are correctly aligned with each other. The lab currently uses a heuristic that performs elastic relaxation based on correspondence points identified between consecutive slices using template matching, but the solution that I am proposing in my thesis is an end-to-end differentiable neural network. This property is useful because the next step in the lab's analysis pipeline, segmentation of the biological entities such as neurons and axons present in each slice, is performed by a neural network; ideally, we would be able to combine this stage with our alignment network and train a single network for the task of alignment and and segmentation together. So far we have demonstrated effective alignment on simulated datasets with only affine deformations; next we hope to test on samples deformed by a richer class of transformations. By the end of the academic year, we hope to automate the entire alignment process, which currently occupies two lab members full-time.

The two months I have been working with Professor Seung's lab have already yielded very valuable learning experiences. Most memorably, the first nearly two months of the project yielded very unsatisfying results; attempts to replicate an old result from another lab member predicting the next slice of brain imagery were unsuccessful. I eventually shifted to a very different technique in order to achieve reasonable prediction. Even after substantial collaboration with the graduate student who had achieved the old result, we were never able to replicate it with his old architecture, which is still an unresolved mystery. In addition to managing this frustrating experience, unlike my previous research projects, there are more than ten graduate students and postdoctoral researchers working in the Seung Lab on components of the lab's pipeline related to mine; this means there is a huge repository of knowledge available to me through the other researchers in the lab. Using contextual knowledge from other lab members has contributed significantly to the richness of my research experience. Finally, working with Professor Seung has shown me an archetype for research leadership; he takes a surgical interest in the activities of every single member of his lab, and his ability to distill both the breadth and depth of this knowledge allows him to see the high-level direction and progress of his lab, and he can in turn help people subtly adjust their course.

During my PhD, I would like to pursue research with an over-arching focus on understanding the process of computation performed by the human brain. Of course, a fixation on modeling the brain's computation is ambitious; it's not clear to me that humanity will acquire any such meaningful understanding in my lifetime. Nonetheless I have been repeatedly fascinated by recent advances made in deep learning, which in at least some ways mirrors the brain's method of seeing the world. There are several technical problems within this subfield that I find especially troubling, and therefore interesting; in particular, I found the recent work done at Stanford by Dr. Fei Fei Li's group applying adversarial examples to policy learning to be very exciting progress in creating more human-like autonomous agents. Similarly, Professor Emma Brunskill's work in reinforcement learning in scenarios where stakes are high and justification for decision-making is crucial and Mark Woodward's recent paper on active one-shot learning are also very intriguing avenues of research that I would be excited to pursue as a graduate student.

I believe that the technical skills I have developed in my coursework and research experiences as an undergraduate would help me make a meaningful contribution to a research group in these fields at Stanford. I also believe that my time as a varsity athlete at Princeton have taught me valuable interpersonal skills that are crucial in a research environment where distillation and communication of complex ideas is essential. I hope that the experiences that I would gain as a graduate student at Stanford would not only prepare me for a career in machine learning research, but also enable me to better advocate for responsible applications of the technology in order to protect humanity's future. I feel especially strongly about this issue with respect to the defense industry, where the potential ramifications for the weaponization of AI are extremely serious. It is this thought that energizes me.