

Research Interests: As a mathematics major, I am overjoyed when I find out why a specific method works, particularly in Machine Learning (ML). I am particularly interested in Computational Learning Theory (CoLT) and the Foundations of Deep Learning (DL). Learning algorithms are frequently created to produce and optimize bounds on the relevant values. These bounds offer assurances and provide insight into black-box machine learning systems which I find exciting as it gives me a deeper understanding of the algorithm, and it's challenging yet fun to prove these bounds and correctness. Despite their high capacity, numerical instability, abrupt minima, and non-robustness, DL models generalize well in practice, which appears to be a paradox. I would like to work in this field to better understand the paradoxical nature DL models exhibit and also explore the field of Deep Reinforcement Learning, Meta-Learning, Neural Operators and contribute to AI4Science projects. Caltech is my top choice for my graduate research because I believe my research objectives and interests can continue flourishing intellectually through the Ph.D. program, working under the supervision of Professor Anima Anandkumar.

Research Experiences - Theoretical: My scholastic journey started during my second year of undergraduate studies when I joined Microsoft as a Data Science Intern, collaborating with Microsoft Research Redmond and China. I laid the foundational work for the team by investigating these particular clusters in Azure that don't completely get used and hence cost the company a lot of money and resources get wasted. I performed data mining and worked on building Interpretable and explainable classifiers on why these nodes do not get filled up (Like the Bin Packing problem; to pack the clusters efficiently). I enjoyed my experience in the industry for the autonomy and ownership over my work as well as the collaborative nature and support system in place. My project had significant positive improvements on the corporation, savings of over \$200 million and the community. Nevertheless, I had some burning questions regarding my research interests going forward.

I subsequently joined the Reinforcement Learning and Artificial Intelligence lab and the Alberta Machine Intelligence Institute under professors Martha White and Adam White (DeepMind), where I worked on two projects. The first project looked at understanding the critical differences between Kernelizing Bayesian Linear Regression (BLR) versus using Gaussian Processes Regression (GPR). They generalize differently because GPR uses the kernel implicitly, and in BLR, we compute the kernel explicitly; hence we are trying to reason through the different representability powers of these two methods using functional analysis and understanding the RKHSs associated with them. One of the primary theorems we established was that there is no regularizer of a function in the reproducing kernel Hilbert space associated with the explicit kernel that yields a close approximation of the implicit kernel space's function. I worked towards implementing the entire codebase, studying various theoretical properties and the paper's write-up, which is in progress to be submitted to TMLR. In addition to learning effective habits of theoretical collaboration over many hours of discussion, I found my passion for ML theory through this project. The second project was the optimization of the MinAtar codebase (which is inspired by the Arcade Learning Environment but simplifies the games to make experimentation with the environments more accessible and efficient) and producing standard benchmarks comparing various algorithms, such as comparing Actor-Critic Algorithms, Deep - Q Networks using a new proposed hyperparameter tuning approach. I was the sole researcher on the project. I produced results where the training time of the algorithms decreased by almost 50% across all environments by including Just in Time compilation and increased the efficiency of algorithms and other PyTorch optimizations. The codebase has been released as open source and is now a testbed for researchers to test their RL and AI agents.

To further pursue research in Theoretical Machine Learning, I recently joined Professor Anima Anandkumar's (Nvidia) lab at Caltech as a research intern. One of the fundamental questions we hope to tackle is "Why are Neural Networks so good at generalizing data?" from the perspective of the Fourier space, which includes understanding the relationship between the Implicit Bias of gradient descent and the Spectral Bias that Neural Networks exhibit. Fourier Neural Operators are used to solve PDEs and have been widely successful due to them being discretization invariant. However, it is still a challenge to select an appropriate number of frequency modes and training resolutions for different PDEs. Too few frequency modes and low-resolution data hurt generalization, while too many frequency modes and high-resolution data are computationally expensive and lead to over-fitting. We finally propose the Incremental Fourier Neural Operator, which augments both the frequency modes and data resolution incrementally during training and has been accepted in the NeurIPS 2022 AI for Science Workshop. I ran all the experiments, increased the efficiency of all these models, provided the results and studied other theoretical properties of FNO exhibits. Finally, we are now trying to expand these results to more general PDE solvers and other DL models to understand the relationship much better, and the work is still in progress.

Research Experiences - Applied: Apart from being interested in fundamental research, I am passionate about projects that would impact the community positively and where Neural Networks are commonplace to gain practical insight and apply improvements rooted in theory. I have had the chance to work on projects involving Climate Change, Healthcare and PDEs.

Last winter semester, I joined the Artificial Intelligence Computational Team. My role consisted of developing algorithms to predict the weather of the Prairies (Western Canada), both short-term and long-term. This was a fun project, as I enjoyed researching and using various DL models to tackle climate change. Although my work was implemented and is part of the \$3 million project, I discovered firsthand the problem in applying deep learning for applications where dependability was critical because DL models infamously do not give performance guarantees, which reinforced why theory was vital.

Long Term Goals and the role of Caltech: I enjoy and excel at solving interdisciplinary problems and all my experiences collectively have shaped my research interests which in turn motivated me to pursue graduate school. My ultimate goal is to work in academia, researching and contributing to projects that have cutting-edge fundamental computer science research. Caltech being a pioneer in foundational research would cater well to my aspirations because of the diversity of its students and faculty, especially Professor Anima's Lab's past and current work indicates its members' unique strengths in this topic. I am particularly interested in doing research under Professor Anima Anandkumar as she has significant expertise in Deep Learning and has contributed to several AI4Science projects that have impacted the community positively. I am currently working with her and her graduate students too and have discussed my research interests with them and love the flexibility that her group offers and the supportive and collaborative environment that is built is somewhere where I would truly thrive.

Some of the research goals that I would love to contribute to are to contribute to a better understanding of these ML models, provide mathematical frameworks for designing new robust and stable ML algorithms and bridge the gap between theory and practice. I want to propose and implement improvements to DNN-based systems inspired by recent theoretical advances. Such as Neural Collapse, Double Descent, Grokking, Spectral Bias and Benign Overfitting (which I find fascinating as deep neural networks predict well, even with a perfect fit to noisy training data). Secondly, there is also a need for a clear definition of model "complexity" for deep neural networks to comprehend regularisation in deep learning. A recent study that inspired me is the topic of Geometric Complexity. I am curious to explore this topic further and see how it can be used to understand better the performance and generalization capabilities of Deep Learning models. Thirdly, I've been curious about what precisely representational learning is, why these DL models have a structured internal representation, and what this structure is. While neural networks can now do some abstraction tasks effectively (such as interpolation), the majority are still well above their current capabilities. So we must be able to test abstract thinking capacity both behaviorally and representationally to optimize for it. Lastly, I would also like to continue contributing to AI4Science projects such as using Neural Operators to solve PDE's and other applications of FNO's. I am also currently a research scholar at Google Brain, where I'm getting mentored by several researchers who have all guided me in my research interests and helped solidify them.

Turning lived experiences into value for Caltech's community, research and society: My other objective in graduate school is to provide equal access to educational resources for students from underprivileged communities. You can expect these additional commitments from me as a graduate student. Apart from my academic pursuits, I have a unique perspective, leadership, decision planning, and execution skill set that helps me push for impact in the real world beyond just academia. I ensured I achieved this goal by empowering students to grow their knowledge, pursue what they are passionate about in a peer-to-peer learning environment, and build solutions for local businesses and communities through workshops and events (40 in total; examples include Google Cloud, Flutter Workshops and a ML bootcamp). Leading these clubs allowed me to advance equal access to a higher level of education and help as many students thrive as possible.

Secondly, my experiences as an international student in Canada has given me diverse opinions and perspectives. I tried being more inclusive and emphatic to underrepresented students in the clubs I run, avoiding stereotypes and treating them with respect. While I've had to prioritize my academics, I feel that advocating for my classmates from underrepresented groups is still vital in CS and Mathematics, which is why I started running the ML theory group at Cohere For AI, where I could share my love of theory with individuals from all over the world. I intend to attain my goal by sharing my experiences and being more connected with Caltech's CS Graduate community. Finally, throughout my time at the University of Alberta, I served as a teaching assistant for courses such as Machine Learning, Theoretical Computer Science, and Real Analysis. This ingrained in me a desire to share information and highlighted the necessity of solid principles in learning.

I strongly believe that this program will fulfill my passion for applying mathematical theory to solving complex problems, but will also allow me to leverage my analytical skills and industry experience to contribute to the program's intellectual diversity. I'm also interested in continuing to be a TA and helping students understand foundational concepts. Lastly, furthering my education at Caltech, I strongly feel that my experiences, including leadership, work and research experience, commitment to education and giving back to the community, would be an asset to this program.