"Sometimes innovation is only old ideas reappearing in new guises . . . [b]ut the new costumes are better made, of better materials, as well as more becoming: so research is not so much going round in circles as ascending a spiral."[1]

In the ever-evolving landscape of artificial intelligence, I am captivated by the notion that innovation often breathes life into age-old ideas, donning them in new guises that are not only more aesthetically pleasing but also remarkably robust. My venture into intelligent systems has been a purposeful exploration of this concept. Over time, I have come to realize that true progress lies not merely in chasing novel concepts but in harnessing the potential embedded in the foundations of our knowledge. In one experience, during a self-directed project, I used ideas from statistics and machine learning to create a novel consumer targeting system by modeling the rich data from Augmented Reality systems. I worked with multiple tools and methods, including statistical modeling (to identify augmented visuals influencing the consumer purchase), an exemplar part-based 2D-3D alignment method to find the best matching 3D models of furniture present in the user's preferred purchase viewpoint in an AR app and used a combination of 3D style compatibility and color compatibility algorithms to recommend relevant furniture. Having published this work at ICCVW 2019, I learned how to use ideas from data science and ML to develop useful solutions for systems such as AR. The realization that these 'ideas' can be refined and fortified with modern techniques and materials resonates deeply with my academic and professional pursuits.

Looking back, my interest in engineering stems from my father, who, as a civil engineer, ingrained in me the benefits of engineering, mathematics principles, and advanced software tools in assisting him with his work. Therefore, deciding to pursue Electrical Engineering and Computer Science during my undergraduate studies was a natural decision. Over the course of the program, I was exposed to a vast range of diverse domains, of which I particularly enjoyed the courses I took in multivariable calculus (Mathematics-I), linear algebra (Mathematics-II), probability and stochastic processes, artificial intelligence, data analytics, big data processing, programming languages, and blockchain. My curiosity towards learning as much as possible about these domains helped me score near the top of my class.

During college, I also undertook several internships in an effort to practically apply the theoretical knowledge I was gaining. As a remote research collaborator with the Sustainability and AI Lab at Stanford University, I helped build and validate machine learning models with remote sensing data for socioeconomic, computational sustainability, and computer vision tasks using tools from object detection, reinforcement learning, noise contrastive estimation, transfer learning, and statistics. The experience of working on these projects kindled my interest in research and was one of the main factors that motivated me to learn math, statistics, computing, and machine learning further.

To supplement my knowledge, I also undertook multiple online courses on machine learning, neural networks, and natural language processing and implemented several algorithms from scratch, like logistic regression and Naive Bayes, for sentiment analysis. This exposure fascinated me with the evolving landscape of language understanding, tracing advancements from word counting in the 1960s and latent semantic analysis in the 1990s to automated algorithms in the 2010s and Large Language Models (LLMs) in the 2020s. It is surreal to know that 'transformers,' the most common model architecture for LLMs, is essentially a massive system of nonlinear regression equations and has resulted in unprecedented performance in various NLP tasks.

After graduation, I joined Microsoft Turing Research as a Research Fellow, where I got the opportunity to dabble with large-scale models in language (and vision). Having developed an avid interest in language understanding, I subsequently forayed into this area and personally undertook several initiatives that were aimed at developing evaluation frameworks for various LLMs. In one of the projects, I actively sought inspiration from historical ethical frameworks and moral psychology and developed a psychometric assessment tool to measure the ethical reasoning capabilities of LLMs based on Kohlberg's Cognitive Moral Development model and Defining Issues Test. Additionally, I established a comprehensive framework to facilitate the infusion of ethical policies for moral alignment in LLMs by leveraging in-context learning to address complex social dilemmas characterized by conflicting values (value pluralism). This project made me recognize that understanding the ethical implications of LLMs requires a nuanced comprehension of human behavior and moral understanding. It provided valuable insights into the development of ethical frameworks, principles, guidelines, methodologies, and tools essential for the responsible and ethical design, evaluation, and deployment of LLMs. I learned a lot from these projects as I overcame initial failed attempts by iteratively improving for 8 months until successfully publishing as a joint first-author in top-tier avenues like EMNLP, EACL, and Neurips Workshop, along with a submission in LREC-COLING.

In another project, I performed a literary survey of state-of-the-art methods in visual language understanding and co-led the development of DUBLIN, a large-scale transformer-based encoder-decoder model adept at interpreting diverse document types. I successfully trained DUBLIN to excel in a wide array of tasks by using techniques like multitask pre-training, curriculum learning, template-based multimodal instruction tuning, and a novel and optimized method for variable input resolution

[1] Karen Sparck Jones. 1994. Natural Language Processing: A Historical Review, pages 3–16. Springer Netherlands, Dordrecht.

training. I also realized the importance of thorough data analysis and visualization in gaining valuable intuitions about the problem, allowing me to come up with novel solutions. From implementing basic methods in statistics and machine learning from scratch during undergraduate school to developing these complex systems, this project served as a crucible for the development of my technical acumen. It provided me with a profound understanding of the intricacies involved in crafting large-scale machine learning systems and further catalyzed my interest in weaving together existing methods with modern technological advancements. This work also resulted in multiple patents and a publication in EMNLP 2023, where I was the joint first-author.

While the initial waves of these Large Models (including but not limited to LLMs) will be based on the inferencing capabilities and availability of cloud computing, I believe we should prepare for a future wave of "on-device large models." If done right and made to fit into the constraints of such devices, this will not only offer obvious "edge computing" advantages such as improved latency and offline availability but also democratize the reach of large models to emerging markets. In an effort towards this direction, I am exploring methods from matrix compression, quantization, early exit decoding, and uncertainty quantification to make these models amenable to single GPU deployment and faster inference. However, such "on-device" prospects also raise concerns about the security and privacy implications of running LLMs on mobile devices. In doing so, I am working on several security and privacy-preserving techniques that can be implemented in the design of on-device LLM-based features. One of the directions we are exploring is based on differentiating instructions from data by employing cryptographic delimiters to prevent man-in-the-middle threats.

These Large Models represent a new and rapidly evolving field, with more unknowns than "knowns." Over time, new insights, applications, and risks will emerge. I now desire to pursue further research in this area and apply tools from applied math, statistics, ethics, and machine learning towards the ethical development of socioculturally informed, reliable, efficient, and secure ML systems, focusing on large language (and vision) models.

While I have thoroughly enjoyed my stint at Microsoft so far and learned a lot, I feel the time is right to challenge myself and learn more about the areas of my interest. Hence, I want to pursue a Master's degree from Stanford ICME to explore areas such as applied math, statistics, computational data science, machine learning, and ethics in greater depth. I am fascinated by the prospects of creating intelligent systems using such integrations. I am eager to join Stanford ICME because of the relevance of the courses the program offers to my career trajectory and the strong alignment of my research interests with those of some professors. Courses such as Convex Optimization, Numerical Linear Algebra, Applied Statistics, Distributed Algorithms and Optimization, Parallel Computing, Human-Centered Design Methods in Data Science, Discrete Mathematics and Algorithms, Advanced Multi-Core Systems, Causal Inference, NLP, and Computer Vision have exciting content and would help me bridge the gap between my existing knowledge and my field of interest. I am also enthusiastic about the prospect of joining ICME's Computational Consulting group, where I look forward to offering guidance in applied math, statistics, data science, and ML to students of diverse backgrounds and levels, contributing to a collaborative learning environment.

I was thrilled to come across some exciting projects being pursued under Prof. Madeleine Udell. Her recent work, OptiMUS, caught my attention, especially in addressing heuristic problem-solving challenges through an LLM-based agent capable of formulating and solving MILP problems from natural language descriptions. Additionally, her research in low-rank approximations is fascinating. In an ongoing effort towards data- and compute-efficient multilingual learning, I am exploring a parameter-efficient continual pre-training framework using scaled low-rank adapters, aligning closely with her research interests. In another project, I am investigating methods from matrix compression, early exit decoding, etc., to make LLMs suitable for edge computing, aligning with her interests in faster and memory-efficient algorithms. Her research in interpretable and automated ML also fascinates me. In one of the projects focused on poverty estimation using remote sensing data, I explored model interpretability using SHAP. I am also drawn to the work led by Prof. Benjamin Van Roy, particularly his efforts in data-efficient reinforcement learning agents. Furthermore, I am intrigued by the research of Prof. Johan Ugander, especially his focus on understanding important social systems and processes. One of my research directions at Microsoft involves understanding ethical reasoning and moral value alignment in LLMs within the context of diverse languages and cultures, loosely aligning with his areas of interest. I am interested in collaborating with him on the following topic: Can we utilize community-specific language models trained on social-network data to probe and understand the divergent worldviews of partisan groups in the United States, thereby bridging political divides? The prospect of collaborating with these research groups greatly excites me. Having worked in AR and LLMs, I am also extremely excited by the prospect of being a part of Stanford's XR student organization and work towards designing, building, and evaluating new approaches at the intersection of AI, on-device LLMs and Extended Reality for education, learning, and healthcare.

In conclusion, I am embarking on this program with a focused and determined approach toward research backed by a strong background in computing, applied math, statistics, machine learning, language and visual understanding, along with an eagerness to learn more about advanced mathematical and statistical principles, as well as computational systems. I am well aware of the rigorous study program, strong and experienced research faculty, and state-of-the-art research facilities at Stanford, and I am convinced that it is the ideal place for me to attain a master's degree and lay a strong foundation for my future in industrial R&D.