Statement of Purpose
Vaishnavi Shrivastava, MS Computer Science

We live in an exciting age where Artificial Intelligence (AI) is radically transforming our lives. Today, we can find information using search engines, have conversations with digital assistants, use translation services to comprehend most languages, and have email clients compose our emails. This promise of AI integrated into our lives, automating mundane tasks and empowering us to do much more, is what drew me to Machine Learning (ML) and Natural Language Processing (NLP). It has been exciting to see the recent breakthroughs in NLP, with large language models (LMs) like GPT-3 single-handedly writing code, composing narratives, and summarizing any text. Despite these advances, NLP techniques still struggle with factual correctness, lack common-sense reasoning abilities, often propagate societal biases, are less performant on low-resource languages, and have large carbon footprints. These problems limit the potential of NLP, and tackling them will require new leaps in thinking. I wish to pursue a Master's in Computer Science (CS) at UCI to gain the broad knowledge and in-depth research experience necessary to address these challenges for the wider adoption of AI. I believe that my expertise in building real-world NLP systems has given me the skill-set to excel in graduate studies and has equipped me with a novel perspective to bring to the incoming class of future technological leaders at UCI.

Working as an Applied Scientist at Microsoft helped me experience how the challenges faced by real-world applications of NLP limit the adoption of cutting-edge research. Working on the Suggested Replies team, providing users with automated responses to emails, I noticed that our system curbed training and inference times by using a biLSTM-based model instead of reaping the advantages of transfer learning from large pre-trained models. Motivated to leverage pre-trained models, I experimented with compression approaches to bring down their fine-tuning and inference times and found low-cost techniques to be surprisingly effective at reducing these times while still bringing in the benefits of transfer learning. This work helped us deploy a fine-tuned pre-trained model, leading to gains in user engagement. Further investigating how pre-trained models, domain adaptation, and dataset size affect model compression, I showed that the large volumes of data available in the industry could make low-cost compression approaches very competitive. These findings, published in Microsoft's AI Journal, demonstrated that industrial settings of NLP problems could have optimal solutions that are different from those in academia. I want to bring this perspective to grad school to bridge the gap between academic and industrial settings and explore robust solutions that can be seamlessly adopted in both areas.

Leveraging user information to provide personalized experiences is crucial for the large-scale adoption of NLP. With its strict privacy requirements and a large number of users, the industry has unique challenges and opportunities in developing personalization solutions. To personalize our GPT-2 based dialog system, I trained embeddings for users in our training set and used them to condition the GPT-2 model for personalized reply generation. Since this technique could only work for users seen during training, I added a projection network to generate embeddings for unseen users on the fly using n-grams from just a single one of their emails. To achieve this, I jointly trained the projection network with the GPT-2 model to project sparse n-gram features to the dense user embedding space. While this helped us personalize responses for all users, it remained expensive to periodically update embeddings for all users. Motivated by prompting of models like GPT-3, I replaced trained user embeddings with non-trainable user-specific prompts to induce personalized outputs and found these prompts to show superior performance. This work led to a paper submission at an upcoming NLP conference and taught me how to utilize the distinct constraints of the industry to research novel solutions. Despite these promising

results, developing personalized experiences robust to adversarial attacks on user privacy remains a challenge. During my Master's, I wish to use my understanding of the industry's needs to advance the theoretical underpinnings of privacy-preserving techniques to enable secure AI applications of the future. The expertise I've gained through this work in analyzing the weaknesses of cutting-edge techniques and developing innovative solutions will be invaluable in making impactful research contributions during my Master's.

While my current role as an Applied Scientist in a product team offers opportunities to innovate and gain experience with state-of-the-art technologies, the constraint of needing to fulfill an immediate business need through every innovation doesn't provide me the flexibility and environment that I need to grow as a leader in NLP. I strive to develop NLP solutions for users that significantly advance the field and are generalizable across applications. This will involve thinking beyond the immediate needs of a given product. Therefore, my long-term career objective is to be at a research lab in academia or industry, working on such visionary research investments. A Master's program will enable me to gain a broader understanding of the field, conduct in-depth research mentored by outstanding faculty, and be surrounded by brilliant, motivated peers, and is thus the best next step for me to move towards my ambitious goals.

My broad interests lie in Deep Learning for NLP, but two areas that particularly excite me are pre-trained language models (PLMs) and low-cost alternatives to fine-tuning these models, such as few-shot and zero-shot learning. Since increasing the size of PLMs will have unsustainable costs and has made these models less interpretable and more susceptible to adversarial attacks, I wish to study new pre-training approaches and novel architectures to further improve PLMs without incurring additional computational costs. Furthermore, although large PLMs are capable of zero-shot and few-shot learning, open challenges exist in using prompts to guide PLMs, including PLM performance varying greatly with small changes to prompts and prompting not aligning well with the training objectives of PLMs. My research interests align greatly with those of UCI's world-class faculty and AI2 Irvine's incredible researchers, and I hope to work closely with these pioneers in the field. I would love to work on improving PLM interpretability and robustness by building on Prof. Sameer Singh's recent work on the deficiencies of counterfactual explanations, interpreting compositional neural networks, and defending against data poisoning attacks. My interest in improving zero-shot and few-shot learning and experience in using prompting for user-personalization also aligns with Prof. Singh's research on contextual calibration of PLMs, automated prompt creation, and fine-tuning of PLMs in a few-shot setting. I would also be interested in expanding Prof. Xiaohui Xie's work on self-undermining knowledge distillation and on the knowledge distillation ticket technique to the NLP domain.

Beyond research, UCI's exceptional offering of courses like *Statistical Natural Language Processing* and *Projects in AI (in Adversarial ML)* will provide me with the intellectual depth and breadth that I seek. UCI will also bring together some of the best minds, and I am excited to collaborate with and learn from such excellent peers. I also dream of contributing to UCI's energizing community by taking on leadership roles in organizations such as Women in Information and Computer Sciences (WICS). With its rigorous CS curriculum, dynamic student body, and record of trailblazing research, UCI will be an ideal place for me to obtain my Master's and bring the power of NLP to people worldwide.