Statement of Purpose

Archiki Prasad (User ID: archikiprasad@gmail.com) *Ph.D. Applicant*

I am interested in problems in Machine Learning applied to Natural Language Processing (NLP), and Speech Processing. Exploring these fields through various courses, research projects, and internships has ushered me into pursuing a career in research. As a Ph.D. student, I hope to develop as an independent researcher, further my knowledge, and through my contributions, advance the research in state-of-the-art systems.

Research Experience: In my junior year, I interned at Adobe Research, India under the supervision of Dr. Shiv Kumar Saini where we worked on the problem of time-series forecasting for cold start products - with extremely scarce or no prior data. While LSTM-based architectures are widely used for time-series tasks, the extent of cross-learning across training instances within them is unclear. This poses a challenge for cold-start products, for which our goal is to maximize the learning from data-rich "similar" products. Thus, we proposed an extension to key-value memory networks [1] wherein the key-network learns common latent representations and reinforces cross-learning. In our modifications, the value-network is regression-based and dynamic − incorporating feedback for error-correction. Our proposed framework achieved at least 25% improvement on all metrics over the LSTM baseline. We were also interested in identifying products/time-series that were "similar" with respect to the network by analyzing the attention weights and similarity in embeddings. However, the limitations of these approaches sparked my interest in the field of interpretability. To showcase our research, we prepared an interactive demo for users to evaluate the strengths/weaknesses of the system for themselves. The culmination of this project resulted in a patent filed at the USPTO and a poster presented at WWW 2020.

To delve deeper into the interpretability of neural networks, I pursued a research project under Prof. Preethi Jyothi, with a focus on understanding what happens within black-box end-to-end automatic speech recognition (ASR) models. We aimed to expound upon the confounding effect of accents on systems trained on typical US-accented data (eg. [2]). By employing techniques such as gradient attributions, information-theoretic measures, and probing classifiers, we shed light on "how" the effect of accents manifests within an ASR system. The intuition behind using different kinds of techniques was to observe trends that varied across techniques (capturing a specific behavior), and trends that remained consistent. For most parts, we did not have a direct reference in the speech modality, so we directly built upon interpretability literature from NLP, through which I became familiar with a lot of downstream NLP tasks. However, extending these techniques to a new modality was not very straightforward. Unlike classification-based tasks like sentiment analysis and natural language inference (NLI), we wanted to explain the prediction of a sequence of characters, which can result in diffused gradients at the input frame level. We tackled this by meaningfully aggregating output tokens and then, calculating the respective attribution efficiently. Eventually, I published this work as a first-author long paper at ACL 2020.

Having dabbled with some ideas in NLP research, I was keen on continuing work on NLP. Under my advisor Prof. Preethi Jyothi, I am now working towards adapting multilingual BERT (mBERT) for code-switched (CoS) languages, wherein, more than one language is used in the same text. Prior works have achieved good performance on token-level CoS tasks (like POS tagging) but fallen short on other tasks like NLI requiring broader understanding of CoS sentences. Therefore, in our work we are focussing on NLI (dialogue-style) and sentiment analysis tasks of the GLUECoS [3] (English-Hindi) evaluation benchmark. Due to several reasons, CoS datasets are often low-resource and so, many tackle this by augmenting with simulated or real CoS data [3]. However, this may not be economical and at times not as effective. Thus, my approach involves leveraging the cross-lingual ability of mBERT [4] and pre-training on intermediate tasks (using single and multi-task frameworks) in English and Hindi.

In parallel with the above project, my ongoing thesis with Prof. Preethi Jyothi and Prof. Rajbabu Velmurugan involves making end-to-end ASR systems more robust to low-resource noisy and accented out-of-domain speech. Currently, we have employed ML techniques such as multi-task learning, and adversarial training for noise adaptation and compared them against state-of-the-art front-end speech enhancement (SE) techniques. We have shown that our simple ML techniques can outperform the best SE models for a large variety of noise types and have submitted this work to ICASSP 2021. One key-challenge during this project was to balance the trade-offs be-

tween in-distribution accuracy and out-of-distribution generalization in various methods. Moving forward, we aim to tackle the effect of accents in speech under the presence of background noise. There have been some listener studies under the same setting, but such a joint adaptation has not been explored for ML-based systems.

Other Experiences: This summer I presented my work [5] at ACL 2020. It allowed me to get valuable feedback from my peers as well as the experts within the community. By attending other talks, tutorials, and birds-of-a-feather sessions, I not only became familiar with the research trends in the field but also engaged in thoughtful discussions with graduate students, and researchers, filling me with optimism about my research direction. I was also exposed to the work in conversational/dialogue systems. My interests in this field were further augmented while attending the Natural Language Understanding track of the Google Research India AI summer school.

Working on problems in diverse fields ranging from Time Series to NLP has allowed me to calibrate my interests as well as appreciate the inherently transferable nature of research. My mathematically-intensive background in Electrical Engineering and probability theory has informed my research perspective. However, working on problems in ML and NLP has required me to do additional coursework and pursue research projects outside my department. This instilled in me the importance of collaboration in productive research, as an example, my ongoing thesis is also a collaboration with professors from both EE and CS departments. I have thoroughly enjoyed my experience as a teaching assistant for five courses and I eagerly await these responsibilities during my Ph.D.

Interests: My interest broadly lies in building data-efficient models that are robust to distributional shifts and are interpretable or generate explainable predictions. I hope to contribute towards demystifying black-box models irrespective of their diverse end-tasks. To this end, I am fascinated by the different kinds of approaches ranging from probing to creating specific challenge sets. Working on code-switched NLI has brought my attention to work highlighting the unintended dependence of models on heuristics or spurious correlations in NLI. I am interested in designing techniques that allow fast adaptation to complex instances and improve robustness. However, detailed work on known biases may not be available for other tasks. Thus, I am also interested in developing training strategies and evaluation schemes that facilitate robustness to unknown biases and out-of-distribution data. In addition to these problems, I want to continue to work on language understanding tasks for low-resource languages (such as code-switched data) by facilitating cross-lingual learning and transfer. I understand that my interests are influenced by my research experience so far and may evolve as I work more closely in the field.

Looking Forward: The excellent faculty and peer-group at CMU have drawn me to pursue graduate studies there. I believe that the diverse and vibrant research happening at CMU's Language Technologies Institute matches my interests and background. I am interested in working with **Prof. Emma Strubbel** on robustness to out-of-domain data and improving the generalization of existing methods to low-resource languages and domains. I specifically wish to explore using multi-task learning and adversarial training for this task. I believe that my work on ASR systems matches **Prof. Shinji Watanabe**'s interests and line of work. I am keen to work under **Prof. Graham Neubig** on cross-lingual learning and multilingual systems as well as under **Prof. Alan W. Black** on code-switched languages and spoken language understanding.

My research experience thus far has been very fulfilling and rewarding. I am positive that I can work well in a research environment and with like-minded colleagues to achieve my goals. It is clear to me that pursuing a Ph.D. is the logical next step. With perseverance and a zeal to learn, I am confident that I can pick up the necessary skills required to tackle open problems in various fields and crystalize my interests. I look forward to joining and contributing to the esteemed CMU research community.

References

[1] A. Miller, A. Fisch, J. Dodge, A. H. Karimi, A. Bordes, and J. Weston, "Key-Value Memory Networks for Directly Reading Documents," in EMNLP 2016.
[2] A. Koenecke, A. Nam, E. Lake, J. Nudell, M. Quartey, Z. Mengesha, C. Toups, J. R. Rickford, D. Jurafsky, and S. Goel, "Racial disparities in automated speech recognition," in PNAS 2020.
[3] S. Khanuja, S. Dandapat, A. Srinivasan, S. Sitaram, and M. Choudhury, "GLUECoS: An Evaluation Benchmark for Code-Switched NLP," in ACL 2020.
[4] K. Karthikeyan, Z. Wang, S. Mayhew, and D. Roth, "Cross-lingual ability of multilingual BERT: An empirical study," in ICLR 2019.
[5] A. Prasad and P. Jyothi, "How Accents Confound: Probing for Accent Information in End-to-End Speech Recognition Systems," in ACL 2020.