Statement Of Purpose

Kameswari Devi Ayyagari
akdevi.92@gmail.com

I am applying to the PhD program in Computer Science for the Fall 2021 semester. Through this program, I seek to study machine learning algorithms and develop new techniques to explore and expand their applicability to different applied science fields.

I have been working as a research engineer for the last three years applying machine learning algorithms to biomedical data. I got interested in Machine Learning while pursuing my Masters degree in Computer Science at the University of Utah. I explored my interest in ML by undertaking academic projects. As part of one such project, I developed a prediction algorithm that predicted stock prices of different airlines based on on-time performance and twitter sentiment analysis. I was also interested in understanding the theoretical mathematical foundations and my independent study with Dr. Aditya Bhaskara on approximation guarantees for subspace clustering in the spirit of algorithms for k-means helped me explore that. Intrigued by ethical implications of machine learning algorithms, I also studied with Dr. Suresh Venkatasubramanian on Convex Framework for Fair Regression Proposing a convex family of fairness regularizers of varying strength that spans the spectrum from group to individual fairness. These projects have helped me gain a well rounded view of both the theoretical and practical aspects of machine learning.

Upon graduation in 2017, I joined nView as a medical imaging scientist where I developed a neural network infrastructure that significantly reduced reconstruction artifacts in the tomosynthesis process. The approach uses Convolutional Neural Networks(CNN) as regularizers during different stages in the iterative reconstruction process. The early stage CNNs add data in the tomosynthesis axis and late stage CNNs refine features and limit biases (potentially introduced by the early stage CNNs) to improve the quality of CBT-fluoroscopy. This has facilitated reduction in radiation dosage to the patient while producing consistently accurate images resulting in safer and faster surgical procedures. This work is published and available in SPIE [1].

More recently, I developed and validated a deep learning pipeline that classifies whole slide images(WSI) of HE-stained skin biopsies into diagnostically-relevant classes (Basaloid, Squamous, Melanocytic and Other), increasing diagnostic accuracy and reducing turnaround time. Additionally, the pipeline was designed to return a measure of confidence in the assessment. By filtering out predictions that are unlikely to be correct, this confidence based classification allowed for the practical use of machine learning in the challenging and nuanced domain of dermatopathology. This was the first real-world-validated pathology deep learning system and it was published in Nature Scientific Reports [2]. I performed this work as a research engineer at Proscia Inc. for the past two years.

In addition to classifying the WSIs, I was curious to explain the decisions made by the classifier and correlate classification decisions with the spatial locality of tumors on the

WSIs. Since the classification was achieved by aggregating instance-wise decisions from a WSI into a single slide-wise decision using Multiple Instance Learning(MIL) and attention mechanism, I initially used the instance-wise attention weights generated from the attention mechanism to create heatmaps. I later proceeded to experiment with model-agnostic interpretability techniques like Local Interpretable Model-Agnostic Explanations(LIME) for developing more localised and robust explanations. Identifying specific regions of interest in this manner helped build trust in the deep learning system with medical practitioners which, I think, is the cornerstone for widespread and sustainable integration of deep learning in the next generation of healthcare applications.

These experiences in my professional career have reinforced my belief that efficient machine learning, statistical and data analysis tools can reveal invaluable insights into large and complex datasets while delivering faster, more accurate results on vast scales. In my work I also realized that different datasets with their varied characteristics respond differently to various ML algorithms. This highlighted a need for a more nuanced education in different ML algorithms, their fundamental mathematical makeup and predisposition towards certain dataset characteristics. This education, I believe, is the key to successfully applying ML to solve practical problems in various domains. This has served as a driving force for me to pursue a PhD as the next step in my research career.

My decision to apply to the University of Toronto is deeply motivated by the outstanding contributions from the research groups and professors here. I am particularly inspired by the Computational Biomedicine Group which is currently working on interesting development of novel computational approaches for analysis and modeling of medical and biological data. My interests in leveraging deep learning techniques to advance biology, medicine and healthcare align well with those of Dr. Brandon Frey. It would be a privilege to study in the CS department of the University of Toronto under the guidance of its remarkable faculty. Thank you for considering my application and I look forward to hearing from you.

References

[1] Devi Ayyagari et al. "Image reconstruction using priors from deep learning". In: Medical Imaging 2018: Image Processing. Ed. by Elsa D. Angelini and Bennett A. Landman. Vol. 10574. International Society for Optics and Photonics. SPIE, 2018, pp. 104–110. DOI: 10.1117/12.2293766. URL: https://doi.org/10.1117/12.2293766.

[2] Julianna D. Ianni et al. "Tailored for Real-World: A Whole Slide Image Classification System Validated on Uncurated Multi-Site Data Emulating the Prospective Pathology Workload". In: Scientific Reports 10.1 (Feb. 2020), p. 3217. ISSN: 2045-2322. DOI: 10.1038/s41598-020-59985-2. URL: https://doi.org/10.1038/s41598-020-59985-2.