Statement of Purpose

I am Abdus Salam Azad. My research interests broadly span the field of Machine Learning (ML). In my undergraduate and Master's thesis, I have worked on Memetic Algorithms. I have also attempted to explore the domain of ML further and its application in relevant fields through several projects, courses, and research collaborations. In my Ph.D, I am interested in developing ML algorithms for solving real-world learning problems arising in different domains including Natural Language Understanding.

I had my first major research experience during my undergraduate thesis. I worked on Genetic Algorithms(GA) to solve MDPVRP—a lesser studied variant of the well-known Vehicle Routing Problem (VRP), which extends VRP with multiple depots and periods. I was supervised by Prof. Md. Monirul Islam, who has been working on GAs for the past 20 years. For GAs to perform well, maintaining the population diversity is very crucial. To keep the population diverse, the existing GA approaches for VRPs incorporate a diversity measure with the solutions' fitness, which can be computationally expensive. Our proposed method aimed at maintaining the population diversity solely by the use of selection operators. We also proposed a new formulation for MDPVRP which allows interdependent operations among depots to provide cheaper solutions at the cost of a bigger search space. Our work was acknowledged as the winner in the yearly thesis poster competition organized by CSE, BUET (1st out of 57 submissions).

In my Master's thesis, I continued my work with Prof. Islam on our proposed MDPVRP formulation. This time, we developed a Memetic Algorithm (MA)—a hybrid GA with a local improvement component. The existing MA methods focus extensively on greediness, which typically leads them to a premature convergence and require additional techniques such as population restart for further progress. Our proposed method introduces a stochastic local improvement component to address this problem. The component focuses simultaneously on both greediness and randomness to maintain the balance between exploration and exploitation, which consequently helps to avoid a premature convergence. We also proposed a heuristic, partly greedy and partly stochastic, to construct the initial solutions. Extensive experiments on the benchmark problems revealed significant improvements over the state-of-the-art methods. This work has been accepted in the IEEE Transactions on Cybernetics.

I developed a decent understanding of search techniques and constrained combinatorial optimization during my thesis. To get a greater overview and deeper understanding of the topics of AI & ML, I have taken a number of related courses during my undergrad and Master's, including AI, ML, Pattern Recognition, and Data Mining. I have also participated in MOOCs on ML (Coursera) and Deep Learning (Udacity). My interest towards Natural Language Understanding piqued during one of my Master's projects. I surveyed the literature of bidirectional image-sentence search, searching images with sentence descriptions (and vice versa), and analyzed three of the state-of-the-art methods. I also proposed a two-stage approach that decouples object detection within the images from the inference of their inherent semantic relations.

I am currently working on the machine/reading comprehension problem—answering questions based on passages, we have modeled the problem as a path-finding game in a passage

graph, where an agent traverses the graph to locate the answer. The graph is constructed from the passage utilizing word embeddings, parse trees, and coreference resolution. I am training the agent using reinforcement learning. The motivation behind using reinforcement learning is that the learned model will be able to show clear paths on the passage graph which it follows to generate the answer.

Since my undergraduate studies, I have attempted to explore AI, ML, and relevant fields through my research, courses, and different projects. In my Ph.D., I am interested in designing learning algorithms which can gather knowledge from data, especially from texts and/or images, to solve meaningful problems by utilizing their inherent semantics. There has been a significant advancement in this field, particularly with deep learning approaches. However, there is room for further improvement. The current state-of-the-art supervised deep learning methods need lots of labelled training examples. I would like to design algorithms which can utilize unlabeled or partially labelled data. I am also interested in developing algorithms that solve problem of a particular domain (e.g, text) by utilizing data from another relevant domain (e.g., image) as auxiliary supervision.

I consider the Department of Computer Science at the Purdue University one of the most suitable places to pursue my Ph.D. I am particularly interested to work with Prof. Dan Goldwasser. His works in NLP is very intriguing and matches with my interest. I also find the works of Prof. Elias Bareinboim on data fusion very exciting. Prof Yexiang Xue's work on machine learning and probabilistic reasoning intrigue me too. I am also open to working with others who have interest in developing machine learning algorithms to solve real-life socio-economic problems or, problems arising in different branches of science. I believe an opportunity to pursue my Ph.D. at the Purdue University will enable me to conduct impactful research and help me to advance towards a research-oriented career in academia.