

Goals Statement

of Muhtasim Noor Alif (PhD applicant for Fall—2024)

My primary research interest lies in the fields of Data Mining, Machine Learning, Deep Learning and Human Computer Interaction. My long-term career goal is to be actively involved in research as a faculty member or as a researcher. Pursuing an advanced degree is the initial step toward achieving my long-term goals. Through my research, I want to find solutions to real-life socio-economic problems or problems arising in different branches of science, using data analysis and machine/deep learning methods. At the same time, I want to explore ways to make such solutions useful and most effective to ordinary people. I believe my academic and research journey has enabled me to dive deep into a concept, and visualize its effects and challenges. I can quickly learn new tools and frameworks, and use them to give my solutions a concrete form. I am also accustomed to learning through trial and error, which I believe is one of the key components of research. I am confident my educational and professional pursuits have prepared me to pursue a career in research.

My interest in Machine Learning started before I was even aware of the field. I remember the time when I was new to using Facebook, and I was trying to upload a group photo with my friends. Suddenly, I noticed that the platform automatically detected the faces of my friends and asked me if I wanted to tag them. This face detection mechanism was very new in the early 2010s, and it blew me away. Later in my junior year of university, I came to know about ML and began to understand the high-level workings of several ML algorithms. I was quite surprised to discover that the fundamental mathematics behind these algorithms, such as regression and classification, were concepts I had learned back in high school, unaware of their extensive applications. I realized that in the realm of science, even the most seemingly straightforward things can have limitless potential. And thus my love for ML was born.

In my undergraduate years, I went through a range of theory and lab courses, but I enjoyed the courses most which involved a lot of algorithms and mathematics. During the COVID pandemic, I had some free time to explore topics of personal interest. I completed the deep learning specialization course on Coursera offered by Dr. Andrew Ng during that period. It was my first hand-on experience with ML and DL, and I vividly enjoyed the assignments. During the later years of my undergraduate studies, I had the opportunity to study in details in the courses on 'Artificial Intelligence' and 'Machine Learning'. The courses had various interesting assignments. In the AI lab, I got to make an AI agent for a two player game using adversarial search, predict the position of a ghost in a grid using hidden markov model, and perform many other tasks. In the ML course, I had to dive deep into various fundamental ML theories like linear/logistic regressions, support vector machines, clustering algorithms, etc. I also learned about some of the cutting-edge deep learning models like GANs, Transformers, etc. I hands-on implemented Logistic Regression, the Expectation-Maximization algorithm, a CNN from scratch, and numerous other key ML concepts. I did a project as well where I developed a Bangla news classifier using BERT. All these coursework and assignments made me even more interested in the field of ML, and I decided that ML should be the focus of my attention.

I grew a knack for research while I was working on my undergraduate thesis under the supervision of [Dr. Tanzima Hashem](#). In my thesis, I worked on conducting a comprehensive analysis and comparison on the state-of-the-art deep learning based crime prediction models. Inspired by the lack of experiments in a unified setting and the need to assess the applicability in different real-life scenarios, we designed various systematic evaluations to be conducted on the models. We picked the Chicago crime dataset available on the city portal as the input for our models after careful deliberation. During dataset analysis, I identified notable variations in crime density among regions, strongly correlated with external factors like traffic patterns, point of interests, etc. In addition, the models employed varied architectures, integrating components like diverse Graph Neural Networks, Attention Mechanisms, Convolution and Diffusion Modules, Recurrent Modules, etc., for the task of space-time aware crime prediction. As a result, I conducted numerous experiments on the models, altering the data sparsity, prediction timesteps, ablation study of external features, and different model components. We have submitted our manuscript in Neurocomputing, and it is currently under review. In the course of this research, I was introduced to the extensive domain of Data

Mining, further fueling my interest for ML. I came to understand that with systematic analysis and effective modeling, we can reveal novel insights and predict a range of phenomena that were once hard to understand. I firmly believe that there is substantial potential for research in the field of Data Mining.

After my graduation, I joined a research project under the guidance of [Dr. A. B. M. Alim Al Islam](#), and I am currently working on it. Our goal is to assess and enhance the applicability of GPT-based large language models(ChatGPT and GPT-4) in the context of English as a Foreign Language (EFL) education among young learners. We aim to build a system designed exclusively for students in school, employing a personalized mobile app. Through the app, we provide the students English language learning assignments, record their responses, and give them ChatGPT feedback so they can rectify their mistakes. We plan to gather, analyze, and refine the system based on the collected data from the students. At present, we have finished developing the app. I have had experiences of developing web-based systems before, but this was my first experience developing a mobile app. Thanks to my undergraduate studies, it has prepared me to quickly learn new things and apply the knowledge efficiently within a short amount of time. As I incorporated ChatGPT into the system for providing the students with feedback, I discovered several issues with enabling direct student interaction with ChatGPT. The interactions may lack relevance to the problem, and there is a need to filter for potential toxicity. These issues led me to engineer the prompts for ChatGPT carefully, so that any unpleasant results can be avoided and high-quality feedback can be ensured. Doing this research, I realized that while modeling and building intelligent systems is needed, the interactions that ordinary individuals have with these systems are equally significant. To get the best results from these advanced technologies, we need to carefully assess their impact on people's lives. Therefore, I see significant research opportunities in this particular field.

Motivated by my ambition for an academic career, I currently serve as a lecturer in the Department of Computer Science and Engineering at the University of Asia Pacific. While teaching, I actively look for ways to ensure students attain a solid understanding of the subject, craft problems that allow them to apply their acquired knowledge effectively, and enhance the overall enjoyment of the learning process. I have a genuine passion for teaching, and I like to think of it as a two-way experience; as I share my knowledge, I am also continuously learning along the way. I believe these teaching experiences will help me achieve my dream of becoming a faculty member in the future.

I consider the Department of Computer Science at Virginia Tech one of the most suitable places to pursue my Ph.D., considering its consistent and unique contributions to the fields of my interest. There are many esteemed faculty members and researchers here who are consistently making impressive contributions, and I want to be a part of that community as well. I am particularly enthusiastic about the prospect of collaborating with [Dr. Debswapna Bhattacharya](#) at the intersection of Computational Biology and Machine Learning. I would also like to collaborate with [Dr. Anuj Karpatne](#) in the field of Scientific Knowledge-guided Machine Learning. I am very interested in the work of Machine Learning Laboratory and would like to work with [Dr. Hoda Eldardiry](#)'s group as well. I am equally open to working with other researchers who share a similar interest with me. I believe an opportunity to pursue my Ph.D. at Virginia Tech will enable me to conduct impactful research and help me to advance towards a research-oriented career in academia.