PERSONAL STATEMENT FOR ACI FELLOWSHIP

```
"RuntimeError:  CUDA out of memory;consider increasing the CUDA memory limit."
```

Perhaps no other error that I have encountered in my Artificial Intelligence (AI) journey has made me realize that advancing from small-scale deep learning projects to real-world applications required more than just knowledge of AI models—it demands powerful, scalable computing resources, and the technical skills to develop frameworks to use such resources efficiently. Across the world, advanced cyberinfrastructure (ACI) is powering discoveries in nuclear, environmental, biological sciences, and many other areas at unimaginable scales, and I intend to use the ACI fellowship as an opportunity to dive deeper into how 64 GPU accelerated nodes can be leveraged to advance healthcare and medicine at Indian University.

My journey into ACI at IU began with the "Supercomputing for Everyone Series." The abundant knowledge base and workshops conducted by Dr. Scott Teige and Laura Huber were phenomenal to get me started and help me through my CSCI-P 556: Applied Machine Learning coursework in Fall 2023. Since then, I have transitioned from a student to a research assistant, thanks to Professor Xiaoqing Huang and Dr. Yijie Wang, and my overarching desire to be part of applied machine learning research in healthcare. As part of my research, I am focused on Alzheimer's disease (AD) prediction using a multi-modal approach, combining magnetic resonance imaging (MRI) scans and paired genetic data to advance the understanding of the underlying AD initiation and progression mechanisms. Particularly, we are working to identify regions of interest in the brain and corresponding genetic markers at the single nucleotide polymorphism (SNP) level. Our aim is to link driving biological genetic markers with brain morphological alternations, discovering new drivers and biomarkers that can assist in early diagnosis, disease progression tracking, and promote drug discovery in this challenging disease with very limited effective treatments. Ultimately, we aim to develop a scalable and generalizable multi-modal model that researchers and AI practitioners in broader disease and cancer areas can use to uncover novel mechanisms and promote precision medicine.

As of August 2024, to establish our research hypothesis we have developed a foundational model with over 75 million parameters in the imaging channel and 12 million parameters in the genetic channel, consuming close to 18 GB of VRAM just to initiate training. This model combines multi-layer architecture with 3 convolutional, fully connected layers and a single attention-based transformer block, to process a sample size of 11,000 2-D MRI images and over 250 patients' 250 x 610,000-dimensional genetic data. The model currently takes up to 8 minutes to run for every epoch on a single GPU partition node on the Big Red 200. By the end of September, we will scale to larger datasets—such as 3D and 4D imaging data (fMRI, dMRI, PET, CT) and genome-wide SNP data of larger cohort to train multiple convolution neural networks (CNNs) and benchmark them against state-of-the-art architectures like 3-D U-Net, Resnet-101, and vision transformers, enhancing feature extraction efficiency with increased attention blocks in the genetic channel. This would scale up our model by at least 15-20 times, exceeding the limits of our present resources and expertise. This is the point where we seek assistance from the ACI project consulting team, for overcoming this challenge by next month and advancing our work into a high-impact contribution to Alzheimer's disease research.

We are eager to make our training process as modular and efficient as possible, but we are sure transitioning to multi-GPU training would pose new challenges. We seek guidance on implementing effective high-performance computing (HPC) software design patterns, particularly in memory management and minimizing deadlocks. Additionally, support in GPU debugging to interpret job logs, identify errors and inefficiencies, and create SLURM scripts for enhanced fault tolerance and dynamic resource allocation would tremendously help us use the limited ACI resources effeciently and reach our goal faster.

We are currently looking into PyTorch's distributed data parallel (DDP) to distribute model parameters across nodes to reduce bottlenecks. As we progress to working with multidimensional data, large-scale data management and processing will become increasingly important. The memory-intensive nature of processing these data tasks requires help in parallelizing pre-processing pipelines and optimizing data loading to minimize overheads. I believe this will help us reduce the overall project timeline and would be crucial for

our success. Additionally, the ACI Fellowship timeline fits perfectly with our project schedule, as we plan to submit our first draft to Nature Methods by the end of December 2024. I would be honored to present our findings and progress with fellow team members and the ACI staff.

Another reason I am applying for the ACI Fellowship is to engage with the ACI staff and fellow students who share the same passion. With IU being home for talented researchers from around the world, I am eager to learn about their research, thought processes, and approaches, and to understand how they leverage ACI resources. The fellowship offers an invaluable opportunity to gain insights that extend beyond technical skills, enriching our perspective and our ability to contribute and lead advanced research projects in the future.

Lastly, with our lab's growing number of members and diverse, ongoing healthcare and medicine-related projects, the ACI Fellowship will provide me with the knowledge and skills needed to better support my team and fellow students. It will help me translate our ideas into actionable results, drive our research forward, and potentially lead to high-impact publications. I see this fellowship as a pivotal opportunity to deepen my expertise and make a meaningful contribution to our lab's success. Thank you for giving us students this incredible opportunity. I am very much looking forward to being part of the ACI community at IU through this fellowship and taking the next step to contribute to the advancement of research in medicine and healthcare at Indian University.