(509)336 − 9938 mhmethun@gmail.com mhmethun Folsom, CA, USA

MD (METHUN) KAMRUZZAMAN

PROFESSIONAL SUMMARY

More than five years of experience developing advanced machine learning models and digital twins using GPUs. Competent in GenAI, LLMs, and deep learning, with a proven ability to improve predictive accuracy in complex datasets. Excited about utilizing cutting-edge technologies to promote cross-domain research through applied AI/ML, fostering innovation and efficiency. Skilled in high-performance software development.

PROFESSIONAL EXPERIENCES

Sandia National Laboratories **May** 2023 **– present**
Postdoctoral Researcher (Applied Machine learning) *Livermore, CA*

- Spearheaded SFT and PEFT with LoRA on Huggingface/dnaBERT2 LLM model using 10M 5'UTR samples in a distributed cluster of 10 nodes with A100 GPU each, gained $4x$ performance on time.
- Trained a Huggingface/molgpt LLM model using $300k$ SMILEs string of small molecules, fine-tuned using $1,000$ candidate molecules, generated $10k$ new candidate-like molecules, and improved 95% lab efficiency.
- Spearheaded ensemble-based predictive model and rule-based recommendation system, generated two industry-manufacturable novel candidate molecules from $10k$ new molecules, saved years of research cost and time.
- Developed a Drug repurposing recommendation system using human-virus PPI, transcriptomic, and FDA-approved viral gene-drug interaction data.
- Led multilayered anomaly detection on wearable health data, reduced the execution time by $5x$ by utilizing NVIDIA Hopper GPU.
- Advanced anomaly detection through digital twin development (a small town with a population of $45,000$), utilizing generative AI (VAE, WGAN, and Diffusion), significantly reduced model uncertainty.
- Spearheaded a generative machine learning model for lattice QCD matrix elements at short distances and large momentum.

University of Virginia **Oct** 2020 **– May** 2023
Postdoctoral Researcher (Applied Machine learning) *Charlottesville, VA*

- Developed an ETL process to extract data daily from two hospitals, transform data into a common data model, and prepare data for model training, testing, and validation using CI/CD.
- Constructed a $10GB$ temporal network considering the last 10 years of contact information between patients and providers to extract latent features (stored in mangoBD for optimized query) that improve the existing model AUC by 15%.
- Designed and implemented a novel algorithm (hierarchical model) to address heterogeneity in the unbalanced HIPAA compliance EHR datasets, improved model accuracy by 7%.
- Designed and implemented machine learning solutions for healthcare applications, integrating privacy-preserving frameworks to enhance the Google Health Studies app collaboration with the Google research team.
- Developed an agent-based model (HPC cluster using SLURM) that helps to identify the source of infection and missing infections using claim and prescription data.

Washington State University **Aug** 2015 **– Sep** 2020
Research Assistant (Data science and visual analytics) *Pullman, WA*

- Developed data analysis framework (named Hyppo-X) in C++, integrating Topological Data Analysis (TDA) algorithms with interactive visualization tools to uncover complex patterns in high-dimensional datasets.
- Engineered graph-based algorithms to extract scientific hypotheses from structural properties, enabling breakthrough discoveries in maize phenomics research.
- Implemented statistical modeling techniques to analyze genotypic diversity under environmental stress, delivering actionable insights for agricultural research.
- Optimized graph-based data analysis algorithms to streamline scientific hypothesis generation, reducing processing time and enhancing research efficiency.
- Orchestrated integration of NodeJS and D3.js technologies for advanced data visualization, revealing crucial patterns in hospital patient trajectories.

KBGroupUK & NRG **May** 2009 **– Jul** 2015
Software Development Engineer *Dhaka, Bangladesh*

- Developed full-stack e-commerce websites from scratch using .NET Framework, HTML/CSS, and JavaScript, and increased ticket sales by more than 13%.
- Led a 6-person development team (2013–2015). Company revenue increased 5% YoY.
- An iOS app embedded with localized and optimized marketing content to market globally.
- Cost efficiency, as well as the use of creativity, and reduced time to market.

INTERNSHIP EXPERIENCES

Amazon **Jun** 2019 **– Sep** 2019
Applied scientist intern *Seattle, WA*

- Optimized machine learning models for predictive analytics, leveraging Python frameworks and SQL to enhance data lake architecture and processing workflows.
- Implemented machine learning models using ARIMA and LSTM, collaborating with the data science team to enhance sales predictions by 3% and optimize database architecture.
- Streamlined data lake performance through SQL analysis, resolving at least ten critical database issues while maintaining system reliability and data integrity.

Pacific Northwest National Laboratories **Jul** 2018 **– Aug** 2018
Research intern *Richland, WA*

- Engineered high-performance C++ solutions utilizing the SPDLOG framework, delivering optimized system benchmarks and improved processing efficiency.
- Diagnosed and resolved asynchronous processing bottlenecks in Python applications, implementing data-driven solutions with the Pandas framework.

EDUCATION

Washington State University **Aug** 2015 **– Aug** 2020
Ph.D. in Computer Science *Pullman, WA*

Bangladesh University of Engineering and Technology **Mar** 2004 **– Mar** 2009
B.Sc. in Computer Science and Engineering *Dhaka, Bangladesh*

SKILLS

- GANs, VAE, DIFFUSION, LLM, LSTM, TensorFlow, PyTorch, SFT, PEFT with LoRA
- Python, R, C++, CUDA, NodeJS, .NET, JavaScript, HTML/CSS, SQL, D3js, Java.
- Collaborative, team player, bias for action, deliver results, curious to learn, and take challenges.