Md Abdur **Rahaman**

📱 +1 505-318-5391 | ✉ aabdur.rahaman007@gmail.com | 🏠 https://a-rahaman.github.io | in md-abdur-rahaman-a43b2a153/

Education

Georgia Institute of Technology, GA, USA

Ph.D. in Computational Science and Engineering, Expected August 2024

University of New Mexico, NM, USA

M.S. in Computer Science, July 2019

Research Interests

Machine Learning, Deep Learning, Computer Vision, Generative AI, Reinforcement Learning, Multi-modal Fusion, Biclustering, Big Data, Natural Language Processing, LLM, Statistical Modeling, Computational Neuroscience

Experience

Center for Translational Research in Neuroimaging and Data Science (TReNDS) at Georgia Tech — *Atlanta, GA*

GRADUATE RESEARCH ASSOCIATE — *Jun 2019 - present*

- Developing computational models to extract knowledge from large scale data
- Designing AI/ML framework for exploratory data analysis: clustering, multi-modal fusion, pattern recognition and summarization
- Generative models for learning high-dimensional image reconstruction, segmentation, and classification.
- Reinforcement learning from human (domain expert) feedback (RLHF) to improve the AI model's performance.
- Leveraging neurocomputational paradigms to enhance SOTA AI (Neuromorphic Computing).

NOKIA BELL LABS — *Murray Hill, NJ*

DATA SCIENCE RESEARCH INTERN — *Sep 2021 - Dec 2021*

- **Routing failure tickets to the corresponding department**
- Collaborated with log analytics group and developed a log summarizer to compress the machine logs (billions of lines) into a manageable & learnable representation.
- Employed BERT models for learning ticket descriptions written by the user about the error occurred.
- A multi-modal framework to combine system logs and user error descriptions to route the failure alert with high precision.

Mind Research Network — *Albuquerque, NM*

GRADUATE RESEARCH ASSITANT — *May 2017 - Apr 2019*

- Designed novel exhaustive biclustering and tri-clustering algorithms with model order(k) relaxation.
- Developed a time series motif detection and a probabilistic summarization framework.

University of New Mexico — *Albuquerque, NM*

GRADUATE TEACHING ASSITANT — *Aug 2016 - Apr 2017*

- Assisted with Linear algebra, Declarative Programming, Computer Algorithms courses designed for CS undergrad students.
- Roles: TA Office hour, Grading, Conducting tutorials on Haskell, Scheme, GNU Emacs

Selected Research Projects

Bi-clusformer: a Transformer based end-to-end biclustering framework.

- Leveraged transformer's self-attention across feature and sample dimensions to generate coherent submatrices - biclusters.
- A transformer-powered GNN for learning intricate submodules in graph-structured data.
- Computational complexity reduction $[\mathcal{O}(kn)]$ with cluster-guided self attention for ViT.

mBAM: deep multi-modal fusion with neuromorphic design

- A multi-modal latent space fusion using spatial and modality-wise attention inspired by the 'Bottleneck Attention Module'.
- Model architecture with neuromorphic design - introducing a feedback loop from multi-modal neurons to unimodal.
- Improved disease classification accuracy by 20-30% by successful integration of neuroimaging and genomics data.

SpaDE: Semantic locality preserving Auto-decoder for deep biclustering

- Novel VAE adaptation for bi-clustering applied to recommender systems, biological systems, document terms modeling
- Regularizers for sparsity & semantic locality in the biclusters to enhance interpretability.
- RLHF fine-tuning and reward model on experts' subgrouping preference vs. biclustering results from the base model.

Statelet: a summarization framework for time series data

- Discovers a set of 'k' most dominant and explanatory motifs from an extensive collection of time series.
- Novel implementation of Earth Mover Distance (EMD) for motifs comparison and Kernel Density Estimator (KDE) for smoothing the frequency subspace is applied on the brain's functional connectivity (FC) time course.
- Developed a module for selecting the summary shapes with maximum prevalence and diversity.

Generative modeling with Concept activation vector (CAV) interpretability

- Utilized a generative model Wasserstein GAN (WGAN) for disease label prediction.
- Introspect the fully trained deep models by finding active concepts - orthogonal vectors towards learned features.
- Allows testing against pre-defined concepts. Compute CAV (inclination) towards neuro-vision concepts: brain's region of interest (ROI), activation, and connectivity patterns.

A robust graph neural network (GNN) for modeling Biological Networks

- Genrate a weighted graph G(V, E) from brain connectivity using the regions of interest (ROI) as nodes (V) and the functional connection (correlation) between them as edges 'E'.
- Use multi-headed self-attention to learn enhanced node embeddings and an orthonormal readout for graph-level representation.
- Improve disease classification accuracy by 15% with unique insights into brain functioning.

N-BiC: greedy biclustering algorithm without specific model order (k)

- Constraint depth-first search (DFS) - based algorithm to semi-exhaustively explore all possible combinations of instances.
- Discover biclusters of co-expressed samples and features in the dataset without specifying 'k'.

IBRNN: Information-theoretic introspection for Recurrent Neural Networks

- Explore the theoretical upper/lower bound of information compression in RNN layers.
- CBOW for word2vec embedding of the text corpus and bi-LSTM for the downstream task.
- Inspired by information Bottleneck theory, approximate mutual information around features, labels, and layers. Quantifies feature compression.

Skills

Programming	Python, CUDA, C/C++, Scala, JAVA, C#, JavaScript, Jquery
Cloud Technologies & DB	AWS, Google Cloud, Flask, Hadoop, Spark, GCP, Hive, MySQL, BigQuery, MongoDB, PostgreSQL
Libraries	PyTorch, TensorFlow, Keras, OpenCV, Stanford CoreNLP, Pandas, NLTK, Scikit-Learn, PySpark, Hugging Face
Tools	MATLAB, R, MLOps, FMOps, Docker, DeepSpeed, Slurm, SPM, Git, Heroku

Selected Publications

- **Rahaman, Md Abdur**, Yash Garg, Armin Iraj, Zening Fu, Jiayu Chen, and Vince Calhoun. 2022. "Two-Dimensional Attentive Fusion for Multi-Modal Learning of Neuroimaging and Genomics Data." In 2022 IEEE 32nd International Workshop on Machine Learning for Signal Processing **(MLSP)**.

- Baker, Bradley Thomas, Noah Lewis, Debratta Saha, **Md Abdur Rahaman**, Sergey Plis, and Vince Calhoun. "Information Bottleneck for Multi-Task LSTMs." In **NeurIPS** 2022 Workshop on Information-Theoretic Principles in Cognitive Systems.

- Dolci, G., **Rahaman, M. A.**, Galazzo, I. B., Cruciani, F., Abrol, A., Chen, J., ... & Calhoun, V. D. (2023, June). "Deep Generative Transfer Learning Predicts Conversion To Alzheimer's Disease From Neuroimaging Genomics Data". In 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops **(ICASSPW)**

- **M. A. Rahaman**, E. Damaraju, D. K. Saha, V. D. Calhoun and S. M. Plis, "Statelets: A Novel Multi-Dimensional State-Shape Representation Of Brain Functional Connectivity Dynamics". 2021 IEEE 18th International Symposium on Biomedical Imaging **(ISBI)**.

- **Rahaman, Md Abdur**, Jessica A Turner, Cota Navin Gupta, Srinivas Rachakonda, Jiayu Chen, Jingyu Liu, Theo GM Van Erp, Steven Potkin, Judith Ford, and Daniel Mathalon. 2019. " N-BiC: a novel biclustering algorithm for brain and behavioral features without specifying the model order". IEEE Transactions on Biomedical Engineering **(TBME)**.

- **M. A. Rahaman**, J. Chen, Z. Fu, N. Lewis, A. Iraji and V. D. Calhoun, "Multi-modal deep learning of functional and structural neuroimaging and genomic data to predict mental illness," 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society **(EMBC)**.

- Dolci, Giorgio, **Md Abdur Rahaman**, Jiayu Chen, Kuaikuai Duan, Zening Fu, Anees Abrol, Gloria Menegaz, and Vince D Calhoun. 2022. "A deep generative multimodal imaging genomics framework for Alzheimer's disease prediction." In 2022 IEEE 22nd International Conference on Bioinformatics and Bioengineering **(BIBE)**.

- Du, Yuhui, Zening Fu, Jing Sui, Shuang Gao, Ying Xing, Dongdong Lin, Mustafa Salman, Anees Abrol, **Md Abdur Rahaman**, and Jiayu Chen. 2020. "NeuroMark: An automated and adaptive ICA based pipeline to identify reproducible fMRI markers of brain disorders". **NeuroImage**

- **Md Abdur Rahaman**, Zening Fu, Armin Iraji and V. D. Calhoun, 2024, "SpaDE: a deep semantic locality preserving biclustering framework",(under review)