Md Abdur **Rahaman**

1087 Westshire PL NW Apt 12, Atlanta, GA 30318, USA

☐ 505-318-5391 | ✉ aabdur.rahaman007@gmail.com | ☗ https://a-rahaman.github.io | ▣ md-abdur-rahaman-a43b2a153

Education

Ph.D. Computational Science and Engineering | Expected August 2024 | Advisor: Dr. Vince Calhoun
Georgia Institute of Technology, GA, USA

M.Sc. Computer Science, July 2019
University of New Mexico, NM, USA

Research Interests

Machine Learning, Big Data, Computer Vision, Generative AI, Reinforcement Learning, Multi-modal Fusion, Pattern Mining & summarization, Biclustering, NLP, Statistical Modeling, Computational Neuroscience

Experience

Georgia Institute of Technology   *Atlanta, GA*
Graduate Research Associate   *Jun 2019 - present*

- Designing AI/ML framework for classification, biclustering, multi-modal fusion, pattern recognition, summarization, etc.
- Reinforcement learning from human (domain expert) feedback (RLHF) to improve model's decision.
- Generative models for learning high-dimensional image reconstruction, segmentation, and prediction.
- Leveraging neuro-computational paradigms to enhance SOTA AI (Neuromorphic Computing).

○ **Selected 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
  - A novel edge embedding enriched with graph semantics. Can capture submodules of a dynamic graph.
  - A cluster-guided attention for learning edge (token) inspired by ViT with computational complexity $O(kn)$
- **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% with a rational fusion of neuroimaging and genomics data.
- **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 their frequency subspace.
  - Devised a probabilistic framework for selecting the summary shapes with maximum prevalence and diversity.
- **SpaDE: Semantic locality preserved clustering & segmentation**
  - Joint optimization of sample-feature distributions using an Auto-encoder architecture for instance-feature co-clustering
  - Differentiable heuristic for sparsity & semantic locality to enhance 2D segmentation.
  - RLHF fine-tuning with a reward model on experts' subgrouping preference vs. clustering results from the base model.
- **Generative modeling with Concept activation vector (CAV) interpretability**
  - A latent diffusion model for 3D brain image generation & disease prediction.
  - Introspect the trained model 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**
  - Model biological networks by instantiating their components as nodes and interaction/causal inferences as the edges.
  - Multi-headed self-attention to learn node embedding and an orthonormal readout for graph-level representation.
  - Improve performance on downstream tasks (e.g., segmentation, classification) with insights into the underlying system.

NOKIA BELL LABS   *Murray Hill, NJ*
Data Science Research Intern   *Sep 2021 - Dec 2021*

○ **Project Title: Routing failure tickets to the corresponding service team**
- 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 user-provided ticket descriptions.
- A multi-modal framework to combine system logs and user error descriptions to route the failure alert with high precision.

University of New Mexico   *Albuquerque, NM*
RESEARCH ASSISTANT   *June 2017 - Apr 2019*

- Worked on data mining models for biological multi-variate time series.
- Developed a time series motif (recurrent pattern) detection and a probabilistic summarization framework.
- Designed exhaustive-search based biclustering and tri-clustering algorithms with model order(k) relaxation.

  ○ **Selected Projects**

- **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 to quantify feature compression.

Department of Computer Science, University of New Mexico   *Albuquerque, NM*
TEACHING ASSISTANT   *Aug 2016 - May 2017*

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

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, OpenCV, Ray, Stanford CoreNLP, Pandas, NLTK, Scikit-Learn, PySpark, Hugging Face
Tools	MATLAB, R, MLOps, FMOps, XGBoost, Docker, DeepSpeed, Slurm, SPM, Git, Heroku

Selected Publications

- **Md Abdur Rahaman**, Zening Fu, Armin Iraji and V. D. Calhoun, 2024, "A Deep Biclustering Framework for Brain Network Analysis". Accepted. In ***CVPR*** *Workshop on Domain adaptation, Explainability, Fairness in AI for Medical Image Analysis.*

- **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)***.

- **Rahaman, Md Abdur**, Jiayu Chen, Zening Fu, Noah Lewis, Armin Iraji, Theo GM van Erp, and Vince D Calhoun. 2023. "Deep multimodal predictome for studying mental disorders". ***Human Brain Mapping***.

- **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***.