Md Abdur Rahaman

https://a-rahaman.github.io/; aabdur.rahaman007@gmail.com; +15053185391

Education

Georgia Institute of Technology, GA, USA
PhD in Computational Science and Engineering | Expected August 2024

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

Experience

Graduate Research Associate – June 2019 to present.
Center for Translational Research in Neuroimaging and Data Science (TReNDS) at Georgia Tech

- Research Interests: Machine Learning, Deep Learning, Computer Vision, Pattern Mining, Computational Neuroscience
- Developing computational frameworks for learning discriminative and semantically meaningful patterns from big data
- Implementing robust AI for multi-dimensional clustering, classification, pattern recognition, and multi-modal fusion.
- Leveraging neurocomputational paradigms to enhance SOTA AI (Neuromorphic Computing).

Data Science Research Intern – September 2021 to December 2021
Nokia Bell Labs

- Worked with the log analytics group to develop a log summarizer to compress the machine logs (billions of lines).
- Applied BERT models for learning log representation to detect a system failure.
- A multi-modal framework for combining system logs and user's error descriptions to route the failure alert.

Graduate Research Assistant – May 2017 to April 2019
Mind Research Network at the University of New Mexico

- Designed exhaustive biclustering and tri-clustering algorithms by relaxing the specification of the model order (k).

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 - 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'.
- The model's architecture and processing are powered by neuromorphic computing.
- Combines Neuroimages (fMRI, sMRI) and genomics to classify mental disorders.

SpaDE: Semantic locality preserving Auto-decoder for deep biclustering

- Auto encoder-based feature learning with a novel bi-clustering regularization – uncovering data point's true manifold.
- Formulated the regularization terms for semantic locality preservation (increases biological relevance) and sparsity.
- Designed a latent space-based meta-heuristic for two-dimensional cluster assignment of samples and features.

Statelet: a data summarization framework for time series data

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

BrainGraph: a graph neural network (GNN) for modeling brain's functional connectivity

- G (V, E). The regions of interest (ROI) are nodes (V) and the functional connection (correlation) between them are edges 'E'.
- Used self-attention to learn enhanced graph embedding through a novel readout.

IBRNN: Information-theoretic introspection method for Recurrent Neural Networks (RNNs)

- CBOW for word2vec embedding of the text corpus and bi-LSTM for the downstream task.
- Inspired by information Bottleneck theory, compute MI around labels, features, and layers and quantifies feature compression.

mriCAV: Concept Activation Vector (CAV) for model interpretability

- Introspect the fully trained deep models by finding active concepts - orthogonal vectors towards learned features.
- Allows testing model's inclination towards pre-defined concepts e.g., neuro-vision concepts ROI, brain networks, activations.

N-BiC: greedy biclustering Algorithm without specific numer of biclusters (k).

- Constraint depth-first search (DFS) - based algorithm to semi-exhaustively explore all possible combinations of instances.
- Doesn't require the specification of the number of clusters (k). Continously optimizes for a list of intrinsic biclusters.

Skills

Languages: C++, Python, C#, Java, JavaScript, CUDA, Jquery, SQL **Tools:** Anaconda**,** Panda, FSL, SPM, Git, MATLAB
Cloud Technologies: AWS, Google Cloud, Slurm, Spark **Libraries:** PyTorch, TensorFlow, OpenCV, Stanford CoreNLP, NLTK, Scikit Learn