Riyasat OHIB

Ph.D. Candidate | Georgia Tech

riyasatohib.com | linkedin.com/in/riyasatohib | github.com/riohib | riyasat.ohib@gatech.edu | Atlanta, GA

EDUCATION

Present – Aug 2021 | **Georgia Institute of Technology, Ph.D. in ECE (Concentration in AI), Atlanta, GA**

- Research in efficient AI and sparsity in deep learning.
- Applications of efficient AI in federated, multi-task and multi-modal learning.
- Research interests : Efficiency in large models, reasoning, LLM modularization and mixture of experts.
- Supervised by Dr. Vince Calhoun and Dr. Sergey Plis.
- CGPA : 4.0/4.0

Aug 2021 – Aug 2019 | **Georgia Institute of Technology, Master's in ECE Program, Atlanta, GA**

- Research in Sparse Neural Networks and Neural Network Pruning.
- CGPA : 4.0/4.0

WORK EXPERIENCE

Dec 2024 – Sep 2024 | **Cohere : LLM Efficiency Research**

Incoming Research Intern, Atlanta, GA

- Focusing on activation sparsity and inference efficiency for LLMs.
- Exploring inference/test time compute.

Efficiency in AI | LLMs | Research

Aug 2024 – May 2024 | **Dolby Laboratories : Experience Delivery Lab, Advanced Technologies Group (ATG)**

Ph.D. Research Intern, Atlanta, GA

- Worked on efficient fine-tuning of LLMs.
- Work under internal patent review and conference review.

Efficiency in AI | LLMs | Computer Vision | Research | Vision Language Models | Multimodal

Aug 2022 – May 2022 | **FAIR at Meta AI : Fundamental (previously Facebook) AI Research**

Research Scientist Intern, Menlo Park, CA

- Designed & implemented a git-like library for version control & model compression called weigit.
- Weigit was integrated as part of the open-source facebookresearch/fairscale library.
- Research on extreme sparsity in deep learning models using signal processing based techniques (e.g. FFT and DCT) during training.

Sparse Neural Networks | Model Compression | Model Pruning | Efficient AI | Signal Processing | Research

Present – Aug 2019 | **TReNDS Center at Georgia Tech**

Graduate Research Assistant, Atlanta, GA

- Working on sparse deep learning, efficient AI and its applications in federated, reinforcement, multi-task and multimodal learning.
- Designed a new sparse projection algorithm :TMLR, ICLR-HAET.
- Developed a novel sparse offline-RL method : NeurIPS-offlineRL.
- Designed a novel communication efficient federated learning method : arXiv.

Sparse deep learning | Model Pruning | Efficient AI | Signal Processing | Research | pytorch

April 2018 – Oct 2017 | **BAT Bangladesh**

Team Leader, Full Time, Dhaka, Bangladesh

- One of the 4 Team Leaders in the Manufacturing Department of Bangladesh's largest production facility.
- Learned project management and data analysis in a large-scale multinational corporation by leading a group of over 80 Engineers, Technicians and Staffs.

Project Managemet | Team Leader | Data Driven Decision Making

Selected Research Projects

May 2023 – Aug 2020 | **Efficient AI, Sparsity and Compression**
TReNDS Center, Atlanta, GA

- Developed a novel Group Sparse Projection algorithm for training sparse deep models. published in TMLR, initial work at ICLR HAET workshop.
- Developed a communication efficient method for Federated learning (FL) in the non-IID data setup. Preliminary work published at ICLR Sparse Neural Network Workshop and full work on arXiv.

Model Compression | Sparse Deep Learning | Computer Vision | Neural Network Pruning | PyTorch | NumPy | Distributed Training

Present – May 2021 | **Sparsity in Reinforcement Learning and efficient multi-task Learning in RL**
TReNDS Center, collaboration with MILA, Montreal, CA, Atlanta, GA

- Working on network pruning for offline and online RL tasks before training. Preliminary work accepted at NeurIPS workshop
- Full work accepted at **NeurIPS 2024**.
- Work done in collaboration with Dr. Doina Precup's group at Montreal Institute for Learning Algorithms (MILA).

Reinforcement Learning | Network Pruning | Sparsity | Python | PyTorch | NumPy

Mar 2016 – Sep 2015 | **Predicting Location of Audio Recordings**
IEEE Signal Processing Cup : Team and Programming Lead IUT, Dhaka, BD

- Predicted the location of recording of audio files, exploiting embedded background power signatures from nearby electrical power lines via machine learning techniques.
- Led the Islamic University of Technology (IUT) Signal Processing Cup team to 11th rank worldwide and an Honorable Mention in IEEE Signal Processing Cup, 2016.

Machine Learning | Signal Processing | Fourier Analysis | FFT | Short Time Fourier Transform | Audio Data | Matlab

Technical Strengths

- Deep Learning, Machine Learning, Computer Vision, Efficient AI.
- Python, C++, Matlab.
- PyTorch, Numpy, Pandas.
- Linux, slurm, cluster computing, bash scripting.

Relevant Coursework

Statistical Machine Learning | Convex Optimization | Linear Algebra | Advanced DSP | Fourier Analysis | Advanced Programming Techniques | Real Analysis | Information processing in Neural Systems

Publications and Pre-prints

2024 Samin Yeasar, **Riyasat Ohib**, Sergey Plis, Amy Zhang, Alessandro Sordoni, and Doina Precup. *Efficient Reinforcement Learning by Discovering Neural Pathways.* **NeurIPS, 2024** *(poster)*.

2024 **Riyasat Ohib**, Bishal Thapaliya, Gintare Karolina Dziugaite, Jingyu Liu, Vince Calhoun and Sergey Plis. *Unmasking Efficiency : Learning Salient Sparse Models in Non-IID Federated Learning*. [arXiv]

2024 **Riyasat Ohib**, Bishal Thapaliya, Jingyu Liu, Vince Calhoun and Sergey Plis. *Efficient Federated Learning on distributed NeuroImaging Data.* **Frontiers in Neuroinformatics**. webpage

2023 **Riyasat Ohib**, Bishal Thapaliya, Jingyu Liu, Vince Calhoun and Sergey Plis. *Decentralized Sparse Federated Learning for Efficient Training on Distributed NeuroImaging Data*. **Neurips Medical Imaging Workshop, 2023**

2023 **Riyasat Ohib**, Bishal Thapaliya, Pratyush Reddy, Jingyu Liu, Vince Calhoun and Sergey Plis. *SalientGrads : Sparse Models for Communication Efficient and data aware Distributed Federated Training.* **ICLR Sparsity in Neural Networks workshop (SNN), 2023.** *webpage* .

2022 **Riyasat Ohib**, Nicolas Gillis, Niccolo Dalmasso, Vamsi Potluru and Sergey Plis. *Explicit Group Sparse Projection with applications to Deep Learning and NMF*. **Transactions on Machine Learning Research (TMLR), 2022.** webpage

2021 Samin Yeasar, **Riyasat Ohib**, Sergey Plis and Doina Precup. *Single-Shot Pruning for Offline Reinforcement Learning*. **NeurIPS Offline Reinforcement Learning workshop, 2021**. paper webpage

2021 **Riyasat Ohib**, Nicolas Gillis, Sameena Shah, Vamsi Potluru, Sergey Plis. *Grouped Sparse Projection for Deep Learning.* **ICLR Hardware Aware Efficient Training workshop, 2021**. paper webpage

2018 **Riyasat Ohib**, Samin Arnob, Muhtady Muhaisin, Riazul Arefin, Taslim Reza and MR. Amin. *ENF Based Machine Learning Classification for origin of Media Signals : Novel Features from Fourier Transform Profile.* **Accepted at ICEECS 2018** presented on Nov 13-14, 2018.

2017 Samin Yeasar, **Riyasat Ohib**, and Muhtady Muhaisin. *Power file extraction process from Bangladesh grid and exploring ENF based classification accuracy using machine learning.* **IEEE R10HTC Conference, 2017**. paper

2016 **Riyasat Ohib**, Samin Yeasar Arnob, Md Sayem Ali, Rakibul Hasan Sagor, and Md Ruhul Amin. *Metal nanoparticle enhanced light absorption in Ga-As thin-film solar cell.* **IEEE Asia-Pacific Conference on Applied Electromagnetics**, pages 89–93, 2016. paper

Projects and Open Source Contributions

WeiGit : A git-like neural network model-weight tracking library — 2022

github.com/facebookresearch/fairscale

- Open source contribution, project was added as part of the open source fairscale library maintained by Meta AI FAIR.
- Designed & implemented a git-like model weight tracking library for tracking the changes of model weights during training.

Software Engineering | Open Source Contribution | SW Design | library implementation | Compression

Drone Simulation using OpenGL and OpenMPI — 2019

github.com/riohib/UAV-Simulation-OpenGL-OpenMPI

- A C++ implementation of flight simulation for a pack of drones following physics mechanics equations.
- Graphics was rendered using OpenGL on C++.
- Each drone physics was handled by a separate compute node and all drones were coordinated among nodes using OpenMPI.

C++ | OpenGL | OpenMPI | Physics Simulation | Graphics

ENF Data Acquisition and Analysis : — 2016

github.com/riohib/IEEE-SP-Cup-2016

- Collected 10 hours of Electric Network Frequency (ENF) data from the Bangladesh Power Grid.
- Analyzed data using Fourier Analysis and classified with Support Vector Machines.

Machine Learning | Fourier Analysis | Support Vector Machines | Matlab