

## RESEARCH SUMMARY

- Proficiency in Machine Learning, Time-series (spatio-temporal process, multivariate sensors), Image processing (material CT images), and graph (Pubmed, molecular structure, road network) data.
- Skilled in implementing deep learning frameworks: transformers, convolutional, generative models, sequence models.
- Experience in deploying deep learning frameworks to handle data paucity ( $\sim 50MB$ ) as well as large-scale data ( $\sim TB$  size).
- Experience working with huggingface deep learning models (image segmentation and transformers).
- Experience training deep learning models on supercomputers like SUMMIT and FRONTIER.
- 3 years of experience in national lab leading to open source and doe licensed projects and publication at venues like NeurIPS, AAAI, SigKDD.
- Highly adept at collaborating with interdisciplinary domains and cross-cultural teams.

## KEYWORDS

**Programming:** Python, Matlab, Pytorch, Pytorch-geometric, C++

**Deep Learning:** CNN, ResNet, UNet, variants of LSTM, Transformer, Vision Transformers, Image segmentation (SAM, UNeT, Mask-RCNN), NBeats, GAN, GNN, VAE, Diffusion models.

**Deep Learning Visualization libraries:** Tensorflow, Loss-landscape, pyHessian

**Physics informed libraries:** PySINDY, SimPY, PySNR, NVIDIA-Modulus

**Python Data Analysis:** Scikit-Learn, OpenCV, Pillow, pyWavelets, Numpy, SciPy, Pandas, Optimization, Clustering, Classification, Gradient Boosting, XGBoost, SVM.

**Large-scale Training:** Multiprocessing, PyPy, Pytorch DDP

**Database:** PostgreSQL, SQLite, Pandas.

**Tools:** VSCode, Jupyter Notebook, Google Colab.

**Opensource repositories:** GitHub, Huggingface.

**Supercomputers:** ORNL Summit, Frontier.

## EDUCATION

**Doctor of Philosophy** in Computer Science  
 Virginia Tech

October 2021  
 Advisor: Prof. B. Aditya Prakash

**Professional Certificate** in Urban Computing  
 Virginia Tech

May 2020

**Bachelor of Science** in Computer Science and Engineering  
 Bangladesh University of Engineering and Technology, Bangladesh

March 2016

## SELECTED PROJECTS

### Foundational Vision Transformer for Science (Oct 2024- Present)

- Developing to pretrain a large vision transformer wo work for different scientific datasets, like climate, pathology, XCT reconstruction, bio-medical imaging, etc.
- *Funding source:* Lab Director Research and Development (LDRD), ORNL
- *Experience:* Mask AutoEncoder, Model parallelization, Vision transformer, HuggingFace Timm, XCT reconstruction data

### Modeling Battery Degradation (Oct 2024- Present)

- Integrate real-time Machine-Learning training and inference with data from cloud (shared across DoE labs).
- *Funding source:* Rapid Operation Validation Initiative, ORNL
- *Experience:* Fourier Neural Operator Network, Time-series Transformers, ONNX

### Generative Turbulence Modeling (Oct 2024- Present)

- Developing a Physics-informed generative model for turbulence.
- *Funding source:* Director's Discretionary Time, ORNL
- *Experience:* Vision transformer, Diffusion Model, John Hopkins Turbulence Data

### Foundational Graph Neural Network for Material Science (March 2024- Sep 2024)

- Lab Directors Research Fund (LDRD) to train large-scale graph neural network with material datasets and fine-tune for material specific applications.
- *Collaborator:* ORNL CSMD, CSMD, CNMS
- *Experience:* Hierarchical graph neural network training using Python OoP and finetune training using ORNL framework HydraGNN on SMILES and Quantum chemical datasets (QM9, QM7, OGB, Moleculenet)

### LLM Influenced Knowledge Graph Neural Network (*September 2023- present*)

- Whole Nineyards Project to automatically generate knowledge graph for scientific data.
- *Collaborator: National Center for Computational Sciences (NCCS), ORNL, Microsoft Deepspeed*
- *Experience: LLM (GPT-3), Open Graph Benchmark (OGB) models , GNN.*

### Finetuning Image Segmentation for Material Microscopy (*July 2023- Sep 2024*)

- Develop a foundational image segmentation model to generalize for material microscopy images.
- *Collaborator: Material Science Division, ORNL*
- *Techniques: Foundational models, e.g., Meta Segment-Anything-Model, Physics-based GAN*

### Li-ion battery Temperature Modeling. (*October 2022- August 2023*)

- Develop a deep-learning sequential model to predict a temperature sequence under various conditions of Li-ion battery.
- *Collaborator: Stevens Institute of Technology*
- *Experience: LSTM, Transformer, Meta-learning*

### Causal Representation Learning for Understanding Neuronal Dynamics. (*April 2023 - September 2023*)

- Design a transformer-like architecture to model causality for understanding brain neural interaction dynamics collected from large-scale neural simulator STACS.
- *Collaborator: Sandia National Lab*
- *Techniques: GNN, Transformer*

### Data Reduction for Fusion Plasma Simulator. (*June 2022 - December 2022*)

- Develop a data reduction technique for identifying plasma particle distributions fusing XGC gyrokinetic simulation data.
- *Collaborator: Princeton Plasma Physics Lab*
- *Experience: ResNet, Hypernetworks, Vision Transformer*

### CDC Covid-19 Forecasting Challenge. (*March 2020 - September 2021*)

- Predict hospitalizations and mortality for Covid-19 with a data-driven deep-learning model (Placed 1st in Facebook COVID-19 Symptom Data Challenge).
- *Collaborator: Georgia Institute of Technology (Team: DEEP-COVID)*
- *Experience: Time-series sequential models, Explainable AI*

## EXPERIENCE

---

### Research Scientist at Oak Ridge National Lab (*October 2023- present*)

- Discrete Algorithms, Computations Science & Mathematics Division (CSMD)
- Contribute towards developing novel deep learning models for scientific problems arising from large-scale data.

### Postdoctoral Research Associate at Oak Ridge National Lab (*October 2021- October 2023*)

- Discrete Algorithms, Computations Science & Mathematics Division (CSMD)
- Supervisor: Dr. Ramakrishnan Kannan

### Research Internship at Oak Ridge National Laboratory (*May 2019-August 2019*)

- **Smart Neighborhood:** Build a machine learning framework to understand and leverage optimization algorithms for modeling energy usage in smart electric meters.

### Teaching Assistant at Virginia Tech (*August 2018-May 2019*)

- CS 2114: Software Design & Data Structure

### Teaching Assistant at Virginia Tech (*August 2017-May 2018*)

- CS 1114: Introduction to Software Design
- Summer Internship at CodeBoxr, Bangladesh (*2015*)

## GRANT WRITING EXPERIENCE

---

### PI/CO-PI

- Foundational Vision Transformer for Science. ORNL LDRD 2025 (Accepted).
- Large Scale Causality Learning for Neuromorphic Simulator. ORNL Seed 2025.
- Generative Material Design under Extreme Environments. ORNL LDRD 2024.
- Deep Hierarchical Generative Modeling Framework. ORNL LDRD 2024.
- Uncertainty quantification in simulation-based inference using latent features. DOE FAIR 2024.
- Multi-scale Multimodal Deep Learning Models for Scientific Data. ORNL LDRD 2023.
- Blackbox optimization for Material Science. ORNL Seed 2023.

## AWARDS

---

- Rising Stars in Computational & Data Sciences, University of Texas Austin, 2023.
- Outstanding Postdoc researcher award in Computer Science and Mathematics Division (CSMD), 2022.
- Facebook CDC COVID-19 Symptom Data Challenge, 1st prize (Team *DEEP OUTBREAK*), 2020.
- NSF Fellowship, 2019-2021.
- Travel award CIKM, 2021.
- Travel award SIGKDD, 2019 & 2020.
- Undergraduate Best Poster award, Bangladesh University of Engineering & Technology, 2015.

## INVITED TALKS

---

- Rapid Operational Validation Initiative for Battery Analytics at ORNL Core-AI workshop, NC State, September 2024.
- Knowledge-Guided Machine Learning for Battery Design at KGML Summer Tutorial, ORNL, July 2024.
- Domain Generalizable Deep Learning for Multimodal Data on Scientific Domains at ORNL Staff Interview Seminar, August 2023.
- Success and Failure Analysis of Foundational and Few-shot Image Segmentation Models as a Case Study on Microstructure Characterization at Monterey Data Conference 2023, ORNL AI Expo 2023.
- Temperature Modeling through Invariance Learning Representation for ensuring Li-ion Battery Safety at ORNL AI Expo 2023.
- Data-driven and Knowledge-driven Deep learning Models for battery Safety Modeling at Mathematics in Computer Science seminar, ORNL, 2023.
- Deep Curriculum Learning for Multi-scale Battery Reconstruction at ORNL Postdoc Research Symposium (ORPA), 2022.
- Rule-based Quantification to Identify Crucial Power System Components for Mitigating Disaster Impact at IEEE BigData, 2022.
- Machine Learning Models for Critical Infrastructures at Women in Data Science (wIDS), Blacksburg, 2021.
- Connecting Critical Infrastructures through Explainable and Network-based Models at UrbComp Seminar, Virginia Tech, 2020.
- Urban-Net: A System to Understand and Analyze Critical Infrastructure Networks for Emergency Management at KDD 2019.

## SELECTED PUBLICATIONS

---

### DISSERTATION

- Explainable and Network-based Approaches for Decision-making in Emergency Management. Anika Tabassum. PhD Dissertation, Virginia Tech 2021.

### SELECTED PEER REVIEWED JOURNAL & CONFERENCES

- Modeling Thermal Runaway Countering Data Paucity through Adversarial Invariant Encoding. **Anika Tabassum**, Srikanth Allu, Ramakrishnan Kannan, Nikhil Muralidhar. IEEE BigData, Washington DC, USA, 2024.
- Narayan Bhushal, **Anika Tabassum**, Sangkeun Matthew Lee, Supriya Chinthavali, Thomaz Carvalhaes, Nasir Ahmad, Nils Stenvig, and Teja Kuruganti. Historical Power Outages of the United States and the Social Vulnerability Index. IEEE Power Energy Society General Meeting (PESGM), Seattle, WA, USA, 2024
- Bharat Srikishan, **Anika Tabassum**, Ramakrishnan Kannan, Srikanth Allu, Nikhil Muralidhar. Reinforcement Learning Prediction Cascades: A Case Study for Image Segmentation. AAAI 2024.
- Sangkeun Lee, Supriya Chinthavali, Narayan Bhushal, Nils Stenvig, **Anika Tabassum**, Teja Kuruganti. Quantifying the Power System Resilience of the US Power Grid Through Weather and Power Outage Data Mapping. IEEE Access Journal of Power Systems and Energy 2024 (to appear).
- **Anika Tabassum**, Nikhil Muralidhar, Ramakrishna Kannan, Srikanth Allu. MatPhase: Material Phase Prediction for Li-ion Battery Reconstruction using Curriculum Learning. IEEE BigData 2022.
- **Anika Tabassum**, Supriya Chinthavali, Sangkeun Lee, Bill Kay, Nils Stenvig, and B. Aditya Prakash. Efficient Contingency Analysis in Power Systems via Network Trigger Nodes. IEEE BigData 2021.
- **Anika Tabassum**, Supriya Chinthavali, Varisara Tansakul, and B. Aditya Prakash. Actionable Insights in Urban Multivariate Time-series. ACM CIKM 2021.
- Alexander Rodriguez, **Anika Tabassum**, Jiaming Cui, Jiajia Xie, Javen Ho, Pulak Agarwal, Bijaya Adhikary, and B. Aditya Prakash. DeepCOVID: An Operational DL-driven Framework for Explainable Real-time COVID-19 Forecasting. Annual Conference on Innovative Applications of Artificial Intelligence (IAAI) 2021.
- Alexander Rodriguez, Nikhil Muralidhar, Bijaya Adhikary, **Anika Tabassum**, Naren Ramakrishnan, B. Aditya Prakash. CALINET: Steering a Historical Disease Forecasting Model Under a Pandemic. AAAI 2021.
- Nikhil Muralidhar, **Anika Tabassum**, Liangzhe Chen, Supriya Chinthavali, Naren Ramakrishnan, and B. Aditya Prakash. Cut-n-Reveal: Timeseries segmentations with explanations. ACM Transactions on Intelligent Systems and Technology (TIST) May 2020.

- Sorour E. Amiri, **Anika Tabassum**, E. Thomas Ewing, and B. Aditya Prakash. Tracking and analyzing dynamics of news-cycles during global pandemics: a historical perspective. ACM SIGKDD Explorations Vol. 21 Issue 2 December 2019.
- **Anika Tabassum**, Supriya Chinthavali, Sangkeun Lee, Liangzhe Chen, B. Aditya Prakash. Urban-Net: A System to Understand and Analyze Critical Infrastructure Networks for Emergency Management. ACM SIGKDD 2019.
- **Anika Tabassum**, Sukarna Barua, Tanzima Hashem and Tasmin Chowdhury. Dynamic Group Trip Planning Queries in Spatial Databases. SSDMB 2017.

## PEER REVIEWED WORKSHOPS

- Adapting Segment Anything Model (SAM) to Experimental Datasets via Fine-Tuning on GAN-based Simulation: A Case Study in Additive Manufacturing. **Anika Tabassum**, Amir Ziahari. NeuRIPS Foundational Model for Science Workshop. 2024.
- Explaining Neural Spike Activity for Simulated Bio-plausible Network through Deep Sequence Learning. **Anika Tabassum**, Shruti Kulkarni, Seung Hwan Lim, James B. Aimone, Fred Rothganger, Felix Wang, Brad Theilman. Neuro Inspiring Computational Elements (NICE) Workshop 2024.
- Attention for Causal Relationship Discovery from Biological Neural Dynamics. Ziyu Lu, **Anika Tabassum**, Shruti Kulkarni, Nathan Kutz, and Eric Shea Brown, Seung-Hwan Lim, Causal Representation Learning Wokshop, NeuRIPS 2023.
- **Anika Tabassum**, Nikhil Muralidhar, Ramakrishnan Kannan, and Srikanth Allu. Li-ion Battery Material phase prediction through Hierarchical Curriculum Learning. AI for Science Workshop, NeuRIPS 2022.
- Bill Kay, Hao Lu, Pravallika Devineni, **Anika Tabassum**, Supriya Chintavali, and Sangkeun Lee. Identification of Critical Infrastructure via PageRank. IEEE BigData (BTSD). 2021.
- Alexander Rodriguez, Nikhil Muralidhar, Bijaya Adhikary, **Anika Tabassum**, Naren Ramakrishnan, and B. Aditya Prakash. Steering a Historical Disease Forecasting Model Under a Pandemic: Case of Flu and COVID-19. NeuRIPS Workshop on Machine Learning in Public Health (MLPH), 2020.
- Pravallika Devineni, Bill Kay, Hao Lu, **Anika Tabassum**, Supriya Chintavali, and Sangkeun Lee. Towards Quantifying Vulnerabilities in Critical Infrastructure Systems. IEEE BigData Workshop on Big Data Tools, Methods, and Use Cases for Innovative Scientific Discovery (BTSD), 2020.
- Supriya Chinthavali, Varisara Tansakul, Sangkeun Lee, **Anika Tabassum**, JeffMunk, Jan Jakowski, Michael Starke, Teja Kuruganti, Heather Buckberry, JimLeverette. Quantification of Energy Cost Savings through Optimization and Control of Appliances within Smart Neighborhood Homes. ACM International Workshop on Urban Building Energy Sensing, (UrbSys), 2019.

## SURVEY ARTICLE

- Supriya Chinthavali, Varisara Tansakul, Sangkeun Lee, Matthew Whitehead, **Anika Tabassum**, and others. COVID-19 Pandemic Ramifications on Residential Smart Homes Energy Use Load Profiles. In Proc. of Energy and Buildings, Volume 259, pp 111847, (Elsevier) 2022. .
- Cramer, Estee Y., et al. Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the US. In Proc. of the National Academy of Sciences of U.S.A. (PNAS) 2022.
- **Anika Tabassum**, Supriya Chinthavali, Liangzhe Chen, and B. Aditya Prakash. Data Mining Critical Infrastructure Systems: Models and Tools. IEEE Intelligent Informatics Bulletin, 2018.

## SERVICES

### Program Committee (2017-2024)

- AAAI 2025, ICLR 2025, AAAI for Social Track 2025.
- AAAI 2024, ICLR 2024, SDM 2024, PKDD 2024, ICML 2024, BigData 2024, NeuRIPS 2024
- NeurIPS 2023, ICML Synergy ICML Synergy of Scientific ML and Modeling 2023
- ECML PKDD 2022, 2023
- IEEE Big Data 2022

### Workshop Organizer (2022)

- IEEE BigData Tools, Methods, and Use-cases for Innovative Scientific Discovery (BTSD) Workshop 2022, 2024.

### Professional Society (2023)

- ACM, SIAM, IEEE

### Session Chair (2021)

- ECML PKDD

### Reviewer (2017-present)

- ACM, SIAM, ICDM, KDD, WWW, SDM, TKDD, IEEE BigData, ECML PKDD, IEEE TPAMI.