# Asadullah Hill Galib

asadgalib19@gmail.com +1-414-242-0786 31 Middlevale Rd, Apt 1578F, East Lansing, MI 48823
linkedin.com/in/galib19 galib19.github.io github.com/galib19

As a Ph.D. in Computer Science (ML), I offer a strong background in Software Engineering coupled with a proven track record of multi-disciplinary research collaboration. (Link: Explore the avenues of Machine Learning and AI I have worked on)

## Education

**Ph.D.** in **Computer Science** *August 2020 - July 2024*
Michigan State University, MI, USA
Advisor: Pang-Ning Tan, Domain: Predictive and Generative Modeling, Time Series/Spatiotemporal ML, Representation Learning

**M.Sc.** in **Software Engineering** *January 2019 - December 2020*
University of Dhaka, Dhaka, Bangladesh
Thesis: Significant Features Analysis For Android Malware Detection Using Machine Learning Techniques [Manuscript][Code]

**B.Sc.** in **Software Engineering** *November 2014 - December 2018*
University of Dhaka, Dhaka, Bangladesh

## Technical Skills (* Proficient)

**AI/ML Skills*:** Generative AI, Representation Learning, Forecasting, Timeseries/Spatiotemporal ML, Data Mining, Adversarial ML, Interpretable/Explainable AI, Large Language Models (LLMs), NLP, Stable Diffusion, OpenAI API, Reinforcement Learning

**AI/ML Tools:** PyTorch*, Lightning*, Captum*, SK-learn*, Pandas*, NumPy*, Matplotlib*, Anaconda*, MATLAB, Keras

**CS Skills:** Python*, Java, C*, PHP, JavaScript, Android*, GCP*, MySQL, SQLite, React Native, Laravel, Selenium, Agile, SRS*

## Experience

**Researcher (Internship)**, *Frontier Development Lab (FDL) 2022 by NASA and the SETI Institute* *June 2022 - August 2022*
- Carried out a statistical analysis that shows promising links between major earthquakes and ionospheric perturbations.
- Created the first machine learning-ready dataset and statistical tool comprising spatiotemporally varying seismic precursors.
- Built machine learning models for forecasting and detecting earthquakes from heterogeneous multivariate time series data.
- Designed a probabilistic model to learn the spatial variability of ionospheric observations around seismic locations.
- Research outcomes: 3 AGU abstracts, 2 papers, 1 technical memo, and 1 NASA NTR.

**Graduate Research Assistant**, *Michigan State University (CSE)* *January 2022 - Present*
- Developing novel deep learning algorithms addressing extreme events within spatio-temporal and time series data.

**Graduate Teaching Assistant**, *Michigan State University (CSE)* *August 2020 - December 2021*
- Lead classes and labs of 260+ students in *CSE 102: Algorithmic Thinking and Programming (Python)*

**Software Engineer & Executive Assistant (Internship)**, *Brain Station 23* *January 2018 - June 2018*
- Developed from scratch and maintained a web application and a mobile application, using Laravel Framework, PHP, MySQL, React-Native, Redux-Saga, Android Studio, Postman, and proper version-controlling (Git, SourceTree).

## Selected Publications(* Equal contribution) (Full list)

Authored **10** peer-reviewed publications, leading as the first author on **6** of them. Presented findings at prestigious international conferences including **KDD (twice), NeurIPS, IJCAI (twice), ICDM** and **AGU**. Selected publications:

- **DiffusionCF (KDD 2024)**: An explainable AI framework using counterfactual explanations and conditional diffusion to generate informative, realistic, and close counterfactual instances for explaining forecasts of time series extremes.
- **SimEXT (ICDM 2023)**: A representation learning framework for time series extremes that enhances representation learning performance by 1.1%-8.2% and improves the downstream prediction performance by **1.7%-11.6%**. [Manuscript]
- **Self-Recover (IJCAI 2023)**: A novel self-supervised learning framework for data fusion and imputation in time series data, boosting forecasting performance by **2.5%-10.5%**. [Manuscript]
- **DeepExtrema( IJCAI 2022)**: A novel framework for forecasting time series extremes with uncertainty estimations that integrates extreme value theory with deep learning techniques, significantly enhancing forecasting performance by **6.5%-16%**. [Manuscript]
- Wilson, T., McDonald, A., **Galib, A. H.**, Luo, L., & Tan, P. N. (2022, Aug.). Beyond Point Prediction: Capturing Zero-Inflated & Heavy-Tailed Spatiotemporal Data with Deep Extreme Mixture Models. In Proceedings of the 28th ACM **SIGKDD 2022** Conference on Knowledge Discovery and Data Mining (pp. 2020-2028).
- Cullen*, L., **Galib*, A. H.**, Smith*, A. W., Varshney*, D., Brown, E., Chi, P., ... & Svoboda, F. (2022, Dec.). Can We Forecast And Detect Earthquakes From Heterogeneous Multivariate Time Series Data? In I Can't Believe It's Not Better Workshop: Understanding Deep Learning Through Empirical Falsification. (**ICBINB@ NeurIPS 2022**).

- Cullen*, L., Smith*, A. W., **Galib*, A.H.**, Varshney*, D., Brown, E., Chi, P. J., ... & Svoboda, F. (2024, Jan.). A Global Analysis of Pre-Earthquake Ionospheric Anomalies. arXiv preprint arXiv:2401.01773.
- Cullen*, L., **Galib*, A.H.**, Smith*, A. W., Varshney*, D., Brown, E., Chi, P. J., ... & Svoboda, F. (2022, Dec.). Open-Source Data Pipelines and Statistical Tool for Studying Pre-Seismic and Post-Seismic Disturbances in the Ionosphere and Geomagnetic Field. In **AGU** Fall Meeting Abstracts (Vol. 2022, pp. IN25A-07).
- Cullen*, L., **Galib*, A.H.**, Smith, A. W., Varshney, D., Brown, E., Chi, P. J., ... & Svoboda, F. (2022, Dec.). Comprehensive Statistical Analysis of Ionospheric and Geomagnetic Signatures Before and After Earthquakes. In **AGU** Fall Meeting Abstracts (Vol. 2022, pp. NH13A-04).
- Varshney*, D., Cullen*, L., **Galib*, A.H.**, Smith, A. W., Brown, E., Chi, P. J., ... & Svoboda, F. (2022, Dec.). Multimodal Machine Learning for Earthquake Identification and Forecasting. In **AGU** Fall Meeting Abstracts (Vol. 2022, pp. INV44A-05).
- **Galib, A. H.**, & Bashyal, B. (2022, May.). On the Susceptibility and Robustness of Time Series Models through Adversarial Attack and Defense. arXiv preprint arXiv:2301.03703.
- Wilson, T., Tan P., Luo, L., & Galib, A. (2021, Dec.). Deep Learning With Extreme Value Theory for Modeling Precipitation Events. In **AGU** Fall Meeting Abstracts (Vol. 2021, pp. A15Q-07).
- **Galib, A. H.**, & Hossain, B. M. (2020, Jul.). Significant API Calls in Android Malware Detection (Using Feature Selection Techniques and Correlation-Based Feature Elimination). In Proceedings of the 32nd International Conference on Software Engineering Knowledge Engineering (**SEKE 2020**) (pp. 566-571).
- **Galib, A. H.**, & Hossain, B. M. (2019, Dec.). A Systematic Review on Hybrid Analysis using Machine Learning for Android Malware Detection. In 2019 2nd International Conference on Innovation in Engineering and Technology (**ICIET 2019**).
- **Galib, A. H.**, & Hossain, B. M. (2020, Jul.). A Review on Hybrid Analysis using Machine Learning for Android Malware Detection. In Dhaka University Journal of Applied Science and Engineering (DUJASE), Volume 5, Issue 1&2, pp. 49-55.
- Yasir, R. M., Asad, M., **Galib, A. H.**, Ganguly, K. K., & Siddik, M. S. (2019, May). GodExpo: an automated god structure detection tool for Golang. In Proceedings of the 3rd International Workshop on Refactoring (**IWOR 2019**) (pp. 47-50). IEEE.
- **Galib, A. H.**, Nahar, N., & Hossain, B. M. (2020). The Influences of Pre-birth Factors in Early Assessment of Child Mortality using Machine Learning Techniques. arXiv preprint arXiv:2011.09536.

## Academic and Research Projects (Details and full list)

- **On the Susceptibility and Robustness of Time Series Models through Adversarial Attack and Defense:** The vulnerability and robustness of several time series models are investigated through adversarial attacks and defense. [Manuscript][Code]
- **Image-to-Image Translation using Conditional GAN:** It generates colored images from sketches using a generative model - Conditional GAN. It incorporates the architecture and guidelines proposed by a CVPR 2017 study (Isola et al.). [Manuscript][Code]
- **Predicting GitHub Issues Lifetime using Machine Learning and Topic Modeling (LDA):** It outperforms the previous approach with a high precision and f1- measure. It extracts distinguishable and comprehensible topics from issues. [Manuscript].
- **Pre-birth Factors in the Early Assessment of Child Mortality using Machine Learning Techniques:** It achieves an AUC score of 0.947 which outperforms the clinical standards. Also, it assesses the relative importance of the factors. [Manuscript].
- **LifeBlood:** A GPS-based blood donor finder android app that searches and sorts nearer blood donors. [Technical Report][Code]
- **Analyzing co-authorship network: Centrality Measure, Link Prediction, and Community Detection:** It analyzes a network of co-authorship relations, predicts missing links and detects community using the network modularity algorithm. [Code]
- **AutoPilot-Web:** A web-based digital transformation of BTS (Base transceiver station) management. Its purpose is to optimize and automate the existing network management system. [Code]
- **AutoPilot-Mobile:** A mobile application (iOS and Android) for the digital transformation of BTS (Base transceiver station) management. [Code]
- **Optimizing Search Space in Code Smells Detection using a Novel Metric:** Significantly reducing search space (i.e., 93% to 21%) using a novel metric called - NCPC, while maintaining the performance of code smells detection. [Manuscript].
- **Heart Disease Prediction and Factors Analysis:** It predicts heart disease effectively in terms of performance and analyzes significant factors using machine learning techniques. [Manuscript] [Code]

## Leadership Activities

**Organizer**, First Software Industry-Academia Collaboration Session with 10 leading companies (2017), Boot Camp on technology for peace, Seminar on the fourth industrial revolution, Workshops on secured internet protocol and IT awareness for females.

**Vice President & Treasurer**, IIT Software Engineers' Community, University of Dhaka.
**Editor & Author**, Shoshikkha - A web-based knowledge platform in Bengali & English.