

# Afiya Ayman

Phone: (346)-7756249, [afiya.ayman17@gmail.com](mailto:afiya.ayman17@gmail.com), [Google Scholar](#) | [LinkedIn](#) | [Web](#)

---

## SUMMARY

- **Six+ years of experience as an academic researcher**, developing **scalable deep learning and statistical models**, **optimizing** machine learning training and deployment processes for efficiency and impact
- **PhD** in Information Science & Technology, specializing in **Automated Multi-Task Machine Learning**
- **Strong publication record**, featuring contributions to prestigious ML conferences and journals
- **Research Focus:** Automated Machine Learning, Multi-task Learning, Data Science, Artificial Intelligence for Social Good

## EDUCATION

- **PhD in Information Sciences & Technology**, **Pennsylvania State University**, University Park, PA **Aug 2022 - Aug 2025**
- **MS in Computer Science**, **University of Houston**, Houston, TX **Jan 2019 - May 2022**
- **BSc in Computer Science & Engineering**, **CUET**, Bangladesh **Mar 2011 - Oct 2015**

## RESEARCH EXPERIENCE

### **Pennsylvania State University – University Park, PA**

Research Assistant - [Applied Artificial Intelligence Lab](#), Advisor: Dr Aron Laszka **Aug 2022 - Aug 2025**

- Advanced **multi-task learning (MTL)** research by developing **task-affinity-driven grouping strategies**, improving gain prediction accuracy, and **outperforming state-of-the-art methods by ~7%** across computer vision, tabular, time-series, and transportation benchmarks.
- Designed a **two-stage ensemble model** leveraging gradient dynamics to predict MTL performance, enabling cost-effective MTL optimization.
- Applied automated MTL frameworks to **real-world public transit forecasting**, enabling data-driven decision support for high-impact transportation systems.

### **University of Houston – Houston, TX**

Research Assistant - **Resilient Networks and Systems Lab**, Advisor: Dr Aron Laszka **Jan 2019 - Jul 2022**

- Led DOE- and NSF-funded ML research in transportation, energy, and security, producing **multiple peer-reviewed publications** and influencing transit agency operations.
- Built energy prediction models for electric and diesel transit fleets that reduced MSE by up to **33% compared to baselines**, informing fleet scheduling and routing.
- Developed **neural architecture search (NAS)** methods for transit ridership prediction, jointly optimizing prediction error and model complexity per task, achieving **9% prediction error reduction** over generally optimized models.
- Conducted large-scale NLP and statistical analyses on cybersecurity datasets, uncovering behavioral and security patterns in smart contract development and bug bounty programs.

## TEACHING EXPERIENCE

Teaching Assistant, Pennsylvania State University, University Park, PA

**Jan 2025 - May 2025**

- **Object-Oriented Programming** — Spring 2025: Assisted in instruction, grading, and student support for undergraduate programming course.

Lecturer, School of Science, Engineering & Technology, East Delta University, Bangladesh

**Sep 2016 - Dec 2018**

- Taught foundational CS and engineering courses, including Programming, Data Structures, Discrete Math, AI, and OS.
- Designed course materials and mentored students on research and coding competitions.

## SELECTED PUBLICATIONS

- **Ayman, A.** et al. (2022). Neural Architecture and Feature Search for Predicting the Ridership of Public Transportation Routes. In *8th IEEE International Conference on Smart Computing*.
- **Ayman, A.** et al. (2021). Data-driven Prediction and Optimization of Energy Use for Transit Fleets of Electric and ICE Vehicles. In *ACM Transactions on Internet Technology*.
- Sivagnanam, A., **Ayman A.** et al. (2021). Minimizing energy use of mixed-fleet public transit for fixed-route service. In *35th AAAI Conference on Artificial Intelligence (AAAI)*.
- **Ayman, A.** et al. (2020). Smart Contract Development from the Perspective of Developers: Topics and Issues Discussed on Social Media. In *Financial Cryptography and Data Security: FC 2020 International Workshop, Revised Selected Papers*, pp. 405–422. Springer.

## SELECTED PROJECTS

**Machine Learning Research:** Led multiple NSF- and DOE-funded research projects in collaboration with city agencies and transit operators, designed machine learning frameworks across diverse domains for solving real-world problems, significantly enhancing the accuracy and efficiency of computational tasks.

### **Data-driven Energy Optimization for Multi-Modal Transit Agencies (Project Summary):**

- Designed and developed a framework for predicting energy consumption for various transit vehicle types using multi-month sensor data, outperforming classic learning algorithms (decision trees and linear regression) by ~ **33% in MSE reduction**.
- Achieved **<5% prediction error** for 6-hour trips, by aggregating sample-level predictions across time-series segments.
- Built a decision tree-based **map-matching** module linking noisy GPS to road geometry with 90% accuracy, enabling elevation and distance features.
- Models trained on multi-month, multi-vehicle telemetry data (6 vehicle types across 8 months); prediction outcomes informed energy-aware routing and scheduling strategies for transit agencies in Tennessee.
- Integrated into a distributed ML pipeline for real-time energy prediction

### **Neural Architecture & Feature Search for Transit Ridership Prediction (Project Summary):**

- Developed a neural architecture and feature search framework for route-specific ridership prediction using Automatic Passenger Count and weather data - jointly optimizing prediction error and model complexity
- Evaluated on real-world transit ridership data across ten routes, showing that route-specific neural network architectures and features outperform generally optimized models in prediction accuracy (**9% lower error**).
- Architectures optimized per route-task consistently yielded the best results compared to hand-designed baselines, demonstrating the benefit of customizing both model complexity and feature set to task-specific patterns in real-world transit systems.

### **Automated Multi-Task Machine Learning for Ridership Prediction of Public Transportation Routes:**

- Proposed an efficient affinity-driven MTL framework that pre-selects task groups to maximize MTL performance gains.
- Achieved **15% lower prediction error** than single-task models and **~7% MSE improvement** over baseline MTL grouping methods. Validated on real-world public transit ridership data, demonstrating consistent performance gains across task groups.

**Relevant Skills:** *Python, Pytorch, TensorFlow, Sci-kit Learn, Deep Learning, Transformers, Computer Vision*

**Data Analysis Studies:** Collected, integrated, and analyzed data from multiple sources to conduct exploratory data analysis (EDA), topic modeling, and statistical studies.

### **Impact of COVID-19 on Public Transit Accessibility and Ridership (Project Summary):**

- Analyzed **3.3M+ transit boarding events** from Nashville and Chattanooga by integrating farebox, GPS, and telemetry data; performed temporal, spatial, and socio-economic analyses to assess ridership declines across demographics, locations, and time-of-day patterns.
- Identified **persistent COVID-19 impacts** on transit accessibility, informing transit agencies' strategies for equitable service restoration; findings published in *Transportation Research Record*.
- Findings informed adaptive strategies for transit agencies to prioritize equitable service restoration and plan for future disruptions.

**Smart Contract Security Awareness Analysis:** Investigated security concerns and awareness in the smart contract developer community by analyzing Q&A discussions, blog posts, and associated source code from multiple platforms.

**Bug Bounty Ecosystem Analysis:** Examined Chromium bug bounty program data, including activity logs and rules descriptions, to characterize participant incentives, behaviors, and vulnerability reporting processes.

**Relevant Skills:** *Python, NLTK, Matplotlib, Numpy, Pandas, Sci-Kit Learn*

## TECHNICAL SKILLS

- **Languages:** Python, Java (basic), C++ (basic), SQL, Bash
- **ML Libraries:** PyTorch, TensorFlow, Keras, Scikit-learn, NLTK, SciPy
- **Tools:** Git, Docker, Jupyter, VSCode, AWS (basic)
- **Methods:** Deep Learning, Multi-task Learning, AutoML, NAS, EDA, NLP, Statistical Modeling, Transformers

## HONORS/AWARDS

- **UH NSM Alumni Scholarship** (2021-2022)
- **GHC Scholar** awarded Student Scholarship for attending **Grace Hopper Celebration (GHC'20)**.
- **Rising Star 2020** at **CRCS Workshop on AI for Social Impact, Harvard**