

## SUMMARY

Data Scientist with expertise in **machine learning, optimization, and operations research**. Skilled in designing, building, and deploying **end-to-end ML systems** using **Python, PyTorch, TensorFlow, and Scikit-learn**, scaling **training pipelines** with Ray and SLURM. Experienced in **translating research innovations** into **production-ready solutions**.

---

## TECHNICAL SKILLS

- **Programming Languages:** Python, C, C++, Java, C, JavaScript
  - **Databases:** Oracle DB, MySQL, MongoDB, SQLite
  - **Data Analytics & Visualization:** NumPy, pandas, matplotlib, Spark
  - **Cloud & Distributed Systems:** AWS (SageMaker, S3, Lambda), Docker, Slurm, Linux
  - **Machine Learning & AI:** PyTorch, TensorFlow, Scikit-learn, Transformers,
  - **Software Development & Practices:** Object-Oriented, Agile Practices, Behavior Driven Development (BDD), Git
  - **Optimization:** LP, MILP, CPLEX, Gurobi, Mosek, OR-Tools
- 

## EDUCATION

**Ph.D. Informatics**, (August 2025): **Pennsylvania State University**, University Park, PA

**M.S. Computer Science**, (August 2022): **University of Houston**, Houston, TX

**B.S. Computer Science and Engineering**, (January 2018): **University of Moratuwa**, Moratuwa, Sri Lanka

---

## EXPERIENCE

### Graduate Research Assistant – Penn State University

Aug 2022 – May 2025

- Analyzed the [LODES](#) dataset to uncover origin–destination travel patterns in U.S. cities with limited public transit, evaluating the potential effectiveness of multi-modal transportation solutions
- Identify the limitation in the state-of-the-art approaches with respect to implementation, and perform necessary **reorganization** (i.e., de-coupling components) and **optimization** to speed-up the computation

### Graduate Research Assistant – University of Houston

Sep 2019 – Aug 2022

- Study the **ridership patterns** of paratransit services in two U.S. cities in Tennessee, with respect to the aspect of before and after COVID-19
- Process and store Chromium Vulnerability dataset as SQLite database, and implement automated processing in Python using pandas to identify **identify** the original reporters (achieve an **accuracy of 98%**), **duplicates** issues, and time at which the issue got **patched** and **released** to public

### Software Engineer – LSEG Technology

Jan 2018 – Jul 2019

- Developed and maintained application software using **Java, Python, and C++** with **Object-Oriented principles**, implementing Agile practices (**Scrum**) to improve team productivity, system performance, and scalability
- Automated reporting, managed **database** and **back-end integrations**, and contributed to front-end development and deployment plans, streamlining operations, minimizing downtime, and improving overall system efficiency

### Software Engineering Intern – WS02

Jul 2016 – Dec 2016

- Built an end-to-end **HL7/FHIR monitoring** system using WS02 ESB, DAS, and BAM for both **real-time (Siddhi)** and **batch (Spark)** analytics of healthcare data
- Developed interactive dashboards (**JavaScript, jQuery, Leaflet.js**) with **geospatial** and **temporal** visualizations, supporting decision-making through data analytic