Suxin Ji

1035 Barnett Shoals Rd Apt 721, Athens, GA 30605 | + 1 (650) 507-9557 | suxinji@gmail.com | LinkedIn

Technical strengths include systems programming, machine learning, and data engineering.
Open to full-time opportunities starting now.

EDUCATION

UNIVERSITY OF GEORGIA — Athens, GA
Ph.D. Student in Computer Science — Aug 2024 – Present

- **GPA:** 4.0/4.0
- **Research Focus:** Dynamic Binary Translation Optimizations

UNIVERSITY OF PENNSYLVANIA — Philadelphia, PA
Master of Science in Engineering in Data Science — Jan 2023 – Aug 2025

- **GPA:** 4.0/4.0
- **Core Courses:** Big Data Analytics, Machine Learning and Statistics, Computer Systems Programming, Algorithms and Computation, Natural Language Processing, GPU Computing, Artificial Intelligence, Medical Image Analysis

SOUTHERN UTAH UNIVERSITY — Cedar City, UT
Bachelor of Science (B.S.) Management Information Systems — Jan 2018 - Aug 2021

- **GPA:** 3.9/4.0
- **Core Courses:** Data Structure, Statistics, Discrete Structures, Algebra, Database, Algorithm, OOP in Java
- **Awards:** Honored with Summa Cum Laude and selected into the Dean's List

PROFESSIONAL EXPERIENCE

DATAWITCHING — Roseville, CA
Solutions Engineer — Mar 2023 – Jul 2024

- Solo developer of a production-grade Incorta data application and Java-based connector for Oracle JD Edwards, adopted by multiple clients across several versions. The solution outperformed legacy alternatives by increasing report coverage by 20% and reducing data pipeline latency by 50%.
- Led integration of AI/ML modules and enhanced analytics capabilities across platforms including BigQuery and Workday Adaptive Planning, driving automation and boosting product value.
- Created and maintained the company's official website and YouTube tutorials. One tutorial garnered over 3K views, was widely shared among partner companies and clients, and directly contributed to 20+ potential client leads.
- Presented technical solutions to stakeholders and clients, incorporating feedback and demonstrating strong ownership, problem-solving, and communication skills.

INCORTA — San Mateo, CA
Data Engineer — Nov 2021 – Nov 2022

- Engineered and deployed machine learning solutions to optimize business processes, utilizing Python, R, and PySpark for data modeling and analysis, improving client data processing speed by 30% compared to legacy solutions.
- Built the Incorta DataPrep API from design through documentation to support efficient data manipulation.
- Conducted rigorous integration testing and delivered critical feedback, leading to a 15% increase in product reliability and fewer customer-reported issues.
- Developed widely-used YouTube tutorials that helped customers onboard quickly and solve real-world use cases, accelerating product adoption and improving the overall user experience.
- Ranked among the top 3 contributors in the Incorta Community by helping users troubleshoot and guiding them to effective solutions, boosting customer satisfaction and platform adoption.

INCORTA — San Mateo, CA
Technical Intern — Jun 2019 – Apr 2021

- Performed A/B testing on pre-release Incorta Data API using Jupyter Notebook to support product development.
- Developed time series models using ARIMA and Prophet to support Incorta's one-click forecasting feature.
- Built visual dashboards in Incorta to demonstrate product features and job analytics.
- Documented ML pipeline instructions through Blogger postings, daily view 100+, increasing exposure for technical team.
- Refactored Python-based log parsers to unify data extraction tools across multiple Incorta versions.

TEACHING EXPERIENCE

UNIVERSITY OF GEORGIA Athens, GA
Graduate Teaching Assistant Jan 2025 – Present

- CSCI 1302: Software Development
  Assisted in teaching Java-based object-oriented programming, modular software design, and data structures. Provided in-person office hours and project-based support.
- CSCI 3360E: Data Science
  Supported instruction in Python programming, machine learning, dimensionality reduction, and data visualization. Answered student questions, clarified technical concepts, and graded assignments with timely feedback.

RESEARCH EXPERIENCE

UNIVERSITY OF GEORGIA Athens, GA
Research Assistant Aug 2024 – Present

- Conducted research on system call emulation correctness in QEMU under the supervision of Prof. Wenwen Wang, focusing on Dynamic Binary Translation optimizations.
- Designed and implemented an independent syscall testing framework to compare native and QEMU-emulated behavior, with automated test generation, execution, and result analysis.
- Identified syscall emulation inconsistencies and limitations in existing testing tools, and developed mechanisms to uncover previously undetected QEMU bugs.

PROJECT EXPERIENCE

KIWI – Database-Driven Web App Apr 2024
Database & Information Systems, University of Pennsylvania

- Designed and implemented a full-stack web app with SQL-backed data management and custom query functionalities.
- Developed both client and server modules for user interaction, data retrieval, and persistent storage.

Redfin Housing Market Prediction Dec 2023
Big Data Analytics, University of Pennsylvania

- Preprocessed large-scale housing market data using KNN imputation, outlier removal, and PCA.
- Built and evaluated predictive models including linear regression (Spark & scikit-learn), random forest, gradient boosting, and neural networks.

KAGGLE COMPETITION - COMMONLIT READABILITY PRIZE Jun 2021
Goal: Rating the complexity of reading passages for grade K 12 classroom use, optimizing reading selections, and benefiting reading skills

- Led in a team of four in Kaggle competition in PyTorch on Google Colab platform.
- Implemented a BERT model for NLP regression tasks in pretraining the passages and predicting the readability.
- Achieved top 1% scores in the leaderboard and won the silver medal.

CAPSTONE PROJECT - SCHOOL LEARNING MANAGEMENT SYSTEM Jan 2021
Goal: Improving the student and teacher performance with unstructured big data analysis

- Analyzed the requirements and functionality of the Canvas LMS using Canvas Docker image deployed in Azure Cloud.
- Developed ETL process for data manipulation and warehousing through PostgreSQL database.
- Built dashboards and reports for various dimensions for teachers, students, and administrators using Incorta.

Technical SKILLS

- **Languages:** Python, Java, C/C++, SQL, R
- **Frameworks & Tools:** Spark, PyTorch, CUDA, NumPy, Pandas, Scikit-learn, Docker, Git, Linux, Bash, MySQL, PostgreSQL, Tableau, Excel, Incorta
- **Cloud Platforms:** Azure, GCP, Alibaba Cloud