Andrew Tan

(312) 868-1341 | andrewt8101@gmail.com | linkedin.com/in/andrewt8101 | github.com/andrewt1018

Education

Purdue University **West Lafayette, IN**
B.S. in Computer Science, B.S. in Mathematics | **GPA 4.0/4.0** *Aug. 2022 – May 2026*
Minor in Electrical Engineering & Philosophy

- Distinctions: CS Corporate Partners Scholarship (2023), CS Endowment Scholarship (2024), Dean's List
- Coursework: Algorithms, Data Structures, Software Testing, Abstract Algebra, Real Analysis

Experience

John Deere | Java, React, Scala, Apache Spark **2024 - 2025**
Software Engineering Intern *Waterloo, IA*

- Conducted an Agile spike into Creoson, a transaction server for automating Deere's 3D CAD Software (Creo)
- Integrated Creoson's APIs into current tech stack, streamlining user interaction with a custom launch tool
- Modified Java source code for Creoson, improving internal tools and functions for Creo
- Developed a React-based note taking tool that communicates with Creo via HTTP requests to Creoson

Data Engineering Intern *Des Moines, IA*

- Designed and implemented a table schema in Spark Scala to record **30+** known issues in Deere's Engineering Data Lake with automated insertions and SQL query generation for downstream data analysts and scientists
- Saved an estimated **$500,000** in annual implicit costs by implementing a data quality cleaning pipeline
- Implemented rigorous unit, integration, and acceptance tests and set up CI/CD pipeline using SBT and GH Actions
- Reported the differences between RAPIDS accelerated GPU clusters versus Photon accelerated CPU clusters

I-GUIDE Datamine | Python, GCP **Aug. 2024 – Jun. 2025**
Research Team Lead *Hybrid - West Lafayette, IN*

- Lead a team of 4 undergraduate researchers in utilizing Apache Sedona to optimize workflows that analyzed the correlation between socially vulnerable populations and dam integrity across the nation
- Published results onto Google Cloud Platform's Dataproc and profiled a **28%** increase in speed
- Presented our research findings at the SRI2025 (Sustainability Research + Innovation) Conference in Chicago

Office of the Indiana State Chemist | Python **Dec. 2023 – Aug. 2024**
Student Developer *West Lafayette, IN*

- Trained an image recognition pipeline in Python on 43,251 test labels to retrieve "Guaranteed Analysis" information
- Experimented with various OCR (Optical Character Recognition) and NER (Named Entity Recognition) technologies

Tokio Marine Highland | Python, PostgreSQL **Jun. 2023 – Aug. 2023**
Data Engineering Intern *Chicago, IL*

- Wrote a reusable, sub-second API in Python that returned elevation data given a lat/long coordinate by leveraging Python-wrapped PostgreSQL to query raster files inside AWS s3 buckets
- Validated geocoding data returned by Bing's public API with Decision trees and Random Forests
- Improved the company's flood eligibility API's accuracy in catching faulty addresses by **26%**
- Compiled and formatted 37, 062 claims notes into easily readable PDF files through Python and PostgreSQL

Extracurriculars

Society of Asian Scientists and Engineers **Sept. 2022 – May. 2024**
Internal Vice President *West Lafayette, IN*

- Directed a new branch of events called "workshops" that fostered a productive learning environment for members
- Halved the sorting time and manpower required for mentor-mentee pairings with a reusable java algorithm

Skills & Interests

Technical Skills: AWS, Python, Scala, Java, C/C++, Apache Spark/Sedona, Bash, Git, GitHub, PostgreSQL, GCP
Certifications: AWS SAA-C03 Solutions Architect, IEEE ICFTIC 2021 Publication, Coursera - Machine Learning
Languages: English (Native), Mandarin Chinese (Native), Spanish (Basic), Japanese (Basic)
Interests: Volleyball, Drums, Traveling, AI Tech, Dogs (Border Collies), Card games (Poker), Food (Chinese Cuisine)