Andrew Hao Tan

(312) 868-1341 | andrewt8101@gmail.com | [LinkedIn: Andrew Tan](#) | [GitHub: Andrew Tan](#)

EDUCATION

Purdue University – West Lafayette   **West Lafayette, IN**
Bachelors in Computer Science, Bachelors in Mathematics, Minor in ECE, Minor in Philosophy | **GPA:** 4.0/4.0   *Expected May 2026*
- **Distinctions:** CS Endowment Scholarship (2024), CS Corporate Partners Scholarship (2023), Continued Dean's List

WORK EXPERIENCES & PROJECTS

John Deere | *Scala, Apache Spark, SQL*   **Des Moines, IA**
Data Engineering Intern   *May 2024 – Aug. 2024*
- 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 through the implementation of a data quality cleaning pipeline
- Implemented rigorous unit tests, integration tests, and acceptance tests and set up CI/CD pipeline using SBT and Github Action
- Researched the differences between RAPIDS accelerated GPU clusters versus Photon accelerated CPU clusters on Databricks

Office of the Indiana State Chemist | *Python, NLP, OCR*   **West Lafayette, IN**
Part-time SWE Developer   *Dec. 2023 – Aug. 2024*
- Trained an Optical Character Recognition pipeline in Python on **43,251** test labels to retrieve Guaranteed Analysis information
- Practiced information security and setting up containerized applications via the use of *Docker*

Tokio Marine Highland | *Python, PostgreSQL*   **Chicago, IL**
Data Analytics Intern   *Jun. 2023 – Aug. 2023*
- Validated geocoding data returned by Bing's public API with supervised ML models (Decision trees and Random Forests)
- Improved the company's flood eligibility API's accuracy in catching faulty addresses by **26%**
- 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
- Compiled and formatted **37, 062** claims notes into easily readable and accessible PDF files through Python and PostgreSQL

MoneyMango A.I. Initiative | *Python, Sentiment Analysis*   **West Lafayette, IN**
Founder   *Nov. 2022 – May 2023*
- Directed a project with a team of 3 that personalized an investment plan based on risk tolerance and investor personality
- Data scraped **100+** Reddit and Twitter pages and performed sentiment analysis to predict market optimism for **15+** stocks
- Oversaw the product's aesthetics, pitch deck, and various marketing features

LEADERSHIP & EXTRACURRICULAR EXPERIENCES

Society of Asian Scientists and Engineers   **West Lafayette, IN**
Internal Vice President   *Sept. 2022 – May 2023*
- Planned events for a club of **200+** active members and **doubled** member engagement
- **Halved** the sorting time and manpower required for mentor-mentee pairings with a reusable java algorithm
- Directed a new branch of events called "workshops" that fostered a carefree environment for members to learn new skills in
- Lead a team of 10 junior board members and successfully organized an End-of-Year banquet for 150 attendees

IDEAS Lab | *Professor Aniket Bera*   **West Lafayette, IN**
Researcher   *Jan. 2023 – Present*
- Collaborated with Ph.D student Phu Pham and conducted research on fine-tuning text-to-image conversions using Stochastic Diffusion Models (SDMs) and Denoising Diffusion Probabilistic Models (DDPMs)
- Attended weekly meetings discussing ongoing research about augmented intelligence with Ph.D students

ML@Purdue   **West Lafayette, IN**
Project Leader & Collaborator   *Sept. 2022 – Present*
- Organized meetings as team leader to compete in IMC Prosperity and Optiver Algo Trading competitions
- Studied and implemented various trading algorithms as an active member in the ML team of the Algo Trading group
- Attended weekly reading group sessions that discussed pioneering research in recent AI and ML topics

SKILLS & INTERESTS

Technical Skills: Python, Scala, Java, C/C++, Apache Spark, Bash, Git, PostgreSQL, Postman, AWS s3, QGIS, MuseScore
Certifications: [IEEE ICFTIC 2021 Publication](#), Machine Learning (Certification from Coursera)
Soft Skills: Leadership, Time Management, Adaptability, Work Ethic, Teamwork, Logic, Efficiency, Consistency, Empathy
Languages: English (Native), Mandarin Chinese (Native), Spanish (Basic), Russian (Basic)
Interests: Volleyball, Drums, Traveling, AI Tech, Dogs (Border Collies), Card games (Poker), Food (Chinese Cuisine)