Joseph Williams

(251) 888-0787 | joseph.williams747@proton.me | github.com/jwilliams2023 | linkedin.com/in/jwilliams2023 | Personal Website

EDUCATION

Auburn University

Accelerated Master of Science in Computer Science, GPA: ***4.00 / 4.00*** **Aug 2025 – May 2027**

Bachelor of Science in Computer Science, Minor in Statistics, GPA: ***3.85 / 4.00*** **Aug 2023 – Dec 2025**

- **Coursework:** Machine Learning, Deep Learning, Artificial Intelligence, Data Science, Data Mining, Probability & Statistics, Adv. Linear Algebra, Databases, OS, Networks, Data Structures & Algorithms, Software Engineering

TECHNICAL SKILLS

Languages: Python, SQL, Java, JavaScript, R, C++, HTML/CSS
Tools & Frameworks: scikit-learn, PyTorch, pandas, NumPy, Databricks, Hugging Face, Matplotlib, React, Git, Anaconda, Docker
Concepts: Machine Learning, NLP, LLMs, ETL, EDA, Data Wrangling, Pipelines, APIs, Testing, Scraping, Agile, CI/CD, Azure Cloud
Certifications: Databricks Fundamentals, Databricks Generative AI, AWS Educate Introduction to Cloud 101

EXPERIENCE

Protective Life Insurance **June 2025 – Aug 2025**

Data Science Intern | Python, SQL, pandas, NumPy, Git, Azure DevOps, Hugging Face, Databricks *Birmingham, AL*

- Modernized and rebuilt end-to-end legacy SAS pipeline in Python using pandas, NumPy, and SQL to process underwriting data; investigated transformation mismatches, validated table outputs, **improved speed by 70%**, and **saved $50K+ annually** in licensing costs
- Developed **SQL queries** with Python to extract and clean reporting data for forecasting, underwriting types, and channel performance; removed duplicates and missing/invalid values, calculated metrics, and formatted outputs for dashboards
- Implemented an **LLM-powered tool** to convert SAS and R to Python by integrating **3 LLMs** via Hugging Face Transformers and Databricks Mosaic AI Model Serving; validated outputs in pandas to select the most accurate **model conversion** and support **scalable future migrations**

Auburn University, Dr. Akond Rahman **Oct 2024 – May 2025**

Software Security Engineering Research Assistant | Python, docker, OpenAI API, Ansible, Git, GitHub, Linux *Auburn, AL*

- Analyzed **300+ malicious Python packages** for threats like obfuscation, shell access, and data exfiltration; manually reviewed code and leveraged GPT via OpenAI API to generate structured **JSON reports**
- **Simulated 4 attack scenarios** in Ansible to test for bypasses, code execution, and file exposure, using docker and Linux

PROJECTS

Covid Classifier | GitHub | *Python, PyTorch, pandas, NumPy, Matplotlib, Anaconda, CUDA, Git*

- Processed **1M+ COVID patient records** via row deletion for missing data; feature engineered 13 predictive features using label encoding of 12 conditions (1 = present, 0 = absent), date-to-binary mortality flag, and Min-Max age normalization
- Designed 4-layer PyTorch model with ReLU, Adam optimizer, and binary cross-entropy loss, achieving **~80% recall**; prioritized recall for sensitivity to reduce life-critical false negatives in a highly imbalanced mortality dataset

Credit Card Rewards Tracker | Live Site | GitHub | *React, SQL, Tailwind CSS, Node, Netlify, Git*

- Developed full stack rewards tracker with **React + Tailwind CSS** frontend and central dashboard for accounts and bonuses
- Deployed with GitHub and Netlify CI/CD; integrated storage and backend logic via **Supabase with PostgreSQL**

GrubHub Price Match AI Automation | GitHub | *Python, Tesseract OCR, Selenium, Anaconda, Git*

- Designed OCR pipeline with Tesseract LSTM and image preprocessing, achieving **99% text extraction accuracy** on receipts
- Automated browser workflows with Python and Selenium to streamline GrubHub's price match claim process**, reducing manual effort by 90%**

LEADERSHIP & ACHIEVEMENTS

Association for Computing Machinery (ACM) **Aug 2023 – Present**

- **Web Dev Club**, *Vice President (2025)* – **Led weekly lectures and co-developed annual club app with 20+ students**
- **Competitive Programming Team,** *Competitor* – **12th out of 110**, ICPC 2024 Southeast Division 2 Regional