

Shaunak Wagh

U.S. Permanent Resident

shaunakwagh@gmail.com | 908-432-6883 | linkedin.com/in/shaunakwagh | github.com/shaunakwagh

EDUCATION

University of Colorado, Boulder

Master of Science in Computer Science

Expected December 2026

GPA: 4.0/4.0

Relevant Coursework: Advanced Data Structures, RSA and Quantum Algorithms, Dynamic Programming, Greedy Algorithms, Approximation Algorithms, Linear Programming

Massachusetts Institute of Technology MicroMasters Coursework

2023 - 2024

Machine Learning with Python-From Linear Models to Deep Learning, Data Analysis: Statistical Modeling and Computation in Applications, Fundamentals of Statistics and Probability

The Pennsylvania State University, University Park, PA

2018 - 2022

Bachelor of Science in Computer Science

Relevant Coursework: Object-Oriented Programming, Data Structures and Algorithms, Systems Programming, Theory of Computation, Operating Systems Design, Discrete Mathematics

WORK EXPERIENCE

Open Source, Volunteer Developer - PolicyEngine.

April 2024 – July 2024

- Worked on the React frontend of PolicyEngine to calculate the effects of public policy changes.
- Optimized GitHub Actions CI/CD pipelines, improving deployment reliability by 30%.
- Collaborated with remote contributors to fix production issues and ensure UI consistency.

Capstone Researcher – Penn State

Aug 2022 – Dec 2022

- Developed iOS app in Swift with CoreML analyzing 5,000+ ECG data points to assess heart disease risk.
- Tested with real ECG scans, validating accuracy using Xcode, Heroku Cloud, and MySQL data pipeline.
- Automated data pipeline with Python scripts to normalize and label ECG data, reducing pre-processing time by 60%.
- Deployed initial beta version to TestFlight with custom onboarding screens for trial participants
- Led 5-member team, coordinating development, testing, and documentation collaboratively.

PROJECTS

AI Document Retrieval System | Python, LangChain, FAISS, HuggingFace, Google Gemini[LLM], Gradio

- Developed an AI-powered web tool for municipal bond insurers that generates automated, source-cited due diligence PDF reports from uploaded official statements, using Google Gemini LLM, LangChain, and Gradio.
- Streamlined bond underwriting analysis by enabling instant extraction of key risks, covenants, and financial details from complex bond documents, reducing manual review time by over 80%.
- Enhanced transparency and auditability for bond insurers by including exact source page citations in all generated reports, supporting data-driven insurance decisions.
- Designed a modular frontend in Gradio with intuitive upload features, result rendering, and error handling, ensuring a production-ready UX for insurance analysts.

Text-Based Game Reinforcement Learning (RL) Agent | Python, PyTorch, NumPy

- Developed autonomous agents to complete quests in a multi-room text adventure game using tabular Q-learning, linear function approximation, and deep Q-networks.
- Designed NLP feature pipelines (BoW, TF-IDF) for state representation and enabled agents to learn optimal strategies under reward shaping.

Neural Network Digit Classifier | Python, PyTorch, NumPy

- Constructed multi-stage CNN layers with convolution, pooling, and fully connected classifiers, applying data augmentation and dropout regularization for performance robustness.
- Tuned hyperparameters and added batch normalization to optimize inference speed and accuracy for grayscale image input, enabling batch-level evaluation with scalability.

Virtual Memory Management System | C, Linux/Unix, Shell, Bash

- Built a software simulation of cloud storage systems like OneDrive and Dropbox in C. Developed code for file management, system startup/shutdown, and moving data using serialization and deserialization.
- Designed a system to split data into blocks and spread them across multiple nodes, with a caching mechanism that achieved 80% hit rate. Created a device driver to connect the application to virtual hardware smoothly.

Full-Stack Music Store Platform | Java, Spring Boot, H2, HTML/JS

- Built a web-based store with Spring Boot and H2 database, enabling secure CRUD for albums and responsive UI features.
- Designed a responsive frontend with role-based login and session handling, and implemented an admin dashboard that supports item addition, updates, and deletion with seamless interaction.

CERTIFICATIONS

- **AWS Solutions Architect: Associate** (Acquired March, 2025)
- HackerRank SQL(Advanced)
- JPMC Forage Software Engineer Simulation
- Jira Service Management Fundamentals
- IBM SkillsBuild Artificial Intelligence Fundamentals

TECHNICAL SKILLS

- **Languages:** Python, C/C++, Java, JavaScript, HTML5/CSS
- **Databases:** MySQL, PostgreSQL, MSSQL, Oracle 21
- **Frameworks:** LangChain, React, Gradio, Node.js, Express Tailwind/Bootstrap, SpringBoot
- **Technologies & Tools:** AWS(EC2, Lambda, SageMaker, S3), Shell Scripting, CodeceptJS, Linux/Unix, Git, Github, JIRA, CI/CD, Agile, Microsoft Office (Word, PowerPoint, Excel), Google (Docs, Slides, Sheets), LaTeX