Jacob M. Million

Github | Jacobmmillion@gmail.com | Website

Education:

- Bachelors, Computer Science, *Columbia University, NY* — Expected May 2025
- AA Psychology, *Pierce College, WA* — 2018 - 2020

Experience:

- Internal Automation Engineer, **Vert Ventures** — January 2025 - Current
  - Developing and implementing internal automation tools to streamline processes, reduce manual tasks, and enhance productivity within the company.

- Data Science and LM Intern, **VytalSigns** — August 2024 - Dec 2024
  - Collaborated closely with the CTO and Head of Data to develop the backend of VytalSign's alpha platform.
  - Engineered robust data pipelines to aggregate, transform, and integrate diverse data sources into an LM knowledge graph using Python.
  - Facilitated communication between services using Azure Service Bus, allowing for automated and streamlined scoring of data after ingestion.
  - Designed and implemented automated verification methods to ensure the precision, safety, and reliability of LLM-generated outputs.

- LLM Code Production Training, **Scale AI** — April 2024 - August 2024
  - Worked with Scale AI and G2I to train large language models, enhancing their ability to generate high-quality code across complex domains by conducting in-depth reviews of downvoted LLM interactions, crafting optimized responses for training data.
  - Applied expertise in Python, Swift, Java, C, C++, Assembly, Lua, and Rust, and other languages, to address advanced topics such as object-oriented programming, RSA encryption, and iOS application development.

- AI Software Engineer Intern, **Radical AI** — August 2023 - Dec 2023
  - Developed an innovative AI-driven quiz generation tool leveraging Google Gemini API and VertexAI text embeddings. This tool consists of a pipeline for extracting and embedding content from user-provided PDF documents, dynamically generating quizzes with instant feedback and detailed explanations.
  - Delivered scalable solutions to enhance accessibility and retention of complex topics, promoting effective learning outcomes.

- Computer Science Instructor with **Coding4Youth** — May 2023 - August 2023
- Computer Science Instructor with **LearningWorld** — January 2025 - Current
  - Taught computer science fundamentals (e.g., conditionals, loops, OOP, algorithmic thinking) to students aged 6–18. Fostered technical confidence and problem-solving skills in students.
  - Developed lessons on AI and model training such as with scikit-learn, game development with Pygame, and website development with Flask for more advanced students.

Technical Projects and Applications:

- Stock and Crypto Price Prediction Application, Lead Developer, *Python and Shell* — 2025
  - Developed a Stock and Cryptocurrency Price Prediction Pipeline using historical data and machine learning. The pipeline ingests, cleans, and labels data, then trains a **Random Forest Classifier**

Jacob M. Million

jacobmmillion@gmail.com

optimized with **GridSearchCV**. The model predicts price movement with **58% precision** for cryptocurrency data, achieving an **F1-score of 0.63** on testing data sets. Deployed for real-time predictions based on market features such as price, market cap, and volume.

- HealthHub, Lead Developer, *Python and HTML* 2024
  - HealthHub is a web application designed to help users monitor diseases in specific states and engage in discussions about health-related topics. It is built on top of a **relational database**, and is implemented in Python using **Flask**.
- Menu-ly, Lead Developer, *Swift and Firebase backend* 2024
  - Menu-ly is an **iOS application** designed to encourage restaurant staff to reference and study their restaurant's menu descriptions. The app empowers management to update, edit, and remove menu items and descriptions, ensuring the information is always up to date. It uses **Firebase backend** for secure user authentication and data storage. The application is available on the IOS App Store.
- AI-Generated Quiz Tool, Lead Developer, *Python* 2023
  - A **web application** designed to generate quizzes based on input documents (PDFs) and topics provided by the user. It utilizes **machine learning** models for text embeddings and leverages Google's **Gemini** and **Vertex AI API** for document processing and quiz generation. **Streamlit** is used for the user interface.
- Oven Scheduler, Lead Developer, *C* 2025
  - Designed and implemented the Oven scheduling class, a **custom multi-core scheduler** optimized for runtime improvements over the Completely Fair Scheduler (CFS). Improved tail completion times by 14% for high-priority tasks and reduced average taskset completion time across various workloads. Developed a dual scheduling algorithm: round-robin for low-priority tasks and virtual time scheduling for high-priority tasks, with **dynamic load balancing** to distribute tasks efficiently across CPUs, minimizing execution bottlenecks.
  - Conducted **performance analysis** using taskset benchmarking, demonstrating reduced latency and enhanced throughput.

Other Activities:

- Columbia Journal of Science, Technology, Ethics, and Policy, *Assistant-Editor* 2022-2023
- Columbia University Arete Effective Altruism Fellowship, *Participant* 2023