# Carolyn Yang

carolynyang08@gmail.com carolynyang08 carolynyang08 carolynyang.dev

## Education

**Carnegie Mellon University** **May 2025**
*Bachelor of Science in Electrical and Computer Engineering* *Pittsburgh, PA*

- **Relevant Coursework:** Machine Learning, Statistical Computing, Linear Algebra, Reasoning with Data, Probability Theory, Data Structures & Algorithms, Computer Systems, Web Development

## Experience

**General Motors** **July 2025 – Present**
*Data Scientist* *Warren, MI*

- Built and deployed a Streamlit application on Databricks connected to a SQL Warehouse to generate ML-driven forecasts for a $\sim$ \$120M prototype and tooling budget, aiding vehicle managers in strategic resource allocation.
- Designed pipelines to clean and transform program data with SQL, pandas, creating interactive visualizations.
- Applied regression and clustering models using scikit-learn to estimate per-unit vehicle component costs, achieving within 85% accuracy when comparing predicted spend to historical vehicle program data.

**General Motors** **May 2024 – August 2024**
*Software Engineering Intern* *Milford, MI*

- Developed Python scripts using Selenium and Regex to automate manual software setup for 5+ teams, converting a multi-step initialization process into a repeatable routine later integrated into the CI/CD pipeline.
- Presented automation tools to 100+ engineers to facilitate knowledge sharing and drive adoption across teams.

**Hubbell Incorporated** **May 2023 – August 2023**
*Embedded Software Developer Intern* *Avon, CT*

- Implemented Rust firmware to collect and process real-time sensor data from inputs, ensuring reliable readings.
- Designed concurrent, interrupt-driven routines and optimized Rust code for real-time performance on a MCU.
- Validated system integrity via TDD and benchmarking, comparing Rust and C to optimize for maintainability.

## Projects

**Carnegie Mellon Capstone: Cyclify** | *Swift* **January 2025 – May 2025**

- Created an iOS application to analyze and detect poor cycling posture using pressure sensor and biometric data.
- Constructed time-series visualizations using Swift Charts to illustrate biking form performance through heatmaps and trend graphs, supported by a persistent SQLite-backed relational store via SwiftData.
- Recognized with 3rd place out of 50 teams by judges and faculty in Carnegie Mellon's ECE Capstone Showcase.

**Modular Neural Network Framework** | *Python, NumPy* **October 2024**

- Orchestrated a multi-layer neural network from scratch in NumPy by manually deriving and implementing forward and backpropagation logic for Linear, Sigmoid, and Softmax-Cross-Entropy layers.
- Developed a modular class-based system to manage forward and backward passes, utilizing vectorized NumPy operations to optimize matrix multiplications and improve training performance on high-dimensional data.

**Yelp Sentiment Analyzer** | *Python, NumPy, NLP* **September 2024**

- Engineered a text classification pipeline to categorize restaurant reviews by sentiment using Logistic Regression and GloVe word embeddings to map unstructured text into a semantic vector space.
- Implemented a custom SGD optimizer to manage gradient updates and optimize for model convergence.

**Journey Jotter** | *Python, JavaScript, AWS, Django, Gemini API* **March 2024 – April 2024**

- Designed a full-stack trip-planning app with Django backend, supporting real-time collaborative itinerary sharing.
- Integrated Google OAuth, Google Maps API, and Gemini API for AI-powered travel recommendations.
- Secured the app against XSS, CSRF, and SQL injection; utilized Django with MVC architecture deployed via AWS.

## Technical Skills

**Languages**: Python, SQL, R, JavaScript, HTML, CSS
**Frameworks & Technologies**: Pandas, NumPy, scikit-learn, PySpark, Databricks, PyTorch
**Additional Technologies**: AWS, Django, React, Node.js, Express.js, Flask, MongoDB