

# Cyrus Amalan

cyrus.amalan@gmail.com | 831-236-0325 | <https://www.linkedin.com/in/cyrus-amalan/> | Portfolio

Machine Learning Engineer and M.S. Data Science candidate specializing in Deep Learning and Multimodal AI. Proven expertise in building sophisticated neural architectures, including CNNs and Reinforcement Learning (DQN) agents. Experienced in developing end-to-end ML pipelines, from large-scale data ingestion to model deployment on AWS. Passionate about applying Foundation AI to human physiology and biomarker data to extend healthspan.

## EDUCATION

New Jersey Institute of Technology, M.S. Data Science  
University of California, Santa Cruz, B.A. Computer Science

(Expected 2027)  
May 2025

## PROFESSIONAL EXPERIENCE

**Flash Bird** Santa Cruz, CA  
*Data Analyst Intern* Jun 2025 – Oct 2025

- Applied statistical methods including linear regression, classification, and hypothesis testing to evaluate program effectiveness for 100+ participants.
- Cleaned and analyzed complex datasets in Python and R, identifying trends to optimize curriculum outcomes.
- Developed interactive dashboards and automated reports to track engagement and performance metrics.
- Collaborated with cross-functional stakeholders to translate analytical findings into actionable recommendations.

**Flash Bird** Santa Cruz, CA  
*Data Analyst Intern (Applied ML)* June 2024 - Jan 2025

- Applied statistical modeling and classification methods to evaluate behavioral outcomes for 100+ participants, identifying key predictors of program success.
- Optimized data cleaning pipelines in Python/R to ensure high-fidelity inputs for downstream predictive modeling.
- Collaborated with cross-functional researchers to translate complex analytical findings into actionable product recommendations.

**Upwork, Remote** Berkeley, CA  
*Machine Learning & Data Infrastructure* June 2023 - Present

- Architected and deployed scalable ML workflows on AWS (Lambda/RDS) for high-volume client datasets, ensuring production-grade reliability and 99.9% uptime.
- Developed predictive models that improved forecasting accuracy by up to 25%, utilizing Python-based automated feature engineering and selection.
- Automated high-throughput ETL/ELT pipelines, reducing manual data preparation for model training by 60%.

## TECHNICAL PROJECTS

**Nutritional Tracker & Biomarker Dashboard (Multimodal Focus)** Monterey, CA  
*Full-Stack/Backend* Aug 2024 – Jan 2025

- Architected a full-stack Flask application with a PostgreSQL backend to manage and analyze user-reported nutritional and behavioral data.
- Developed a REST API to bridge user inputs with predictive modeling engines, improving retention by 12% through personalized feedback loops.
- Designed time-series data pipelines to ensure consistent reporting across thousands of user sessions.

**Spotify Scalable Data Pipeline** Monterey, CA  
*Data Analyst* Jun 2025 – Aug 2025

- Engineered a high-throughput pipeline to process 1M+ records using Python and SQL, optimizing for system performance and throughput.
- Built automated monitoring and alerting tools to ensure pipeline health, reducing downtime by 17% through proactive error handling.
- Engineered ML-ready datasets and built visualization tools that supported data-driven product decisions.
- Achieved a 17% improvement in forecast accuracy compared to baseline statistical models.

**Forecasting & Visualization Portfolio** Feb 2025 – May 2025

- Engineered R Shiny applications to visualize complex retention and marketing data across 10K+ entries.
- Automated data refresh pipelines, cutting report generation time by 40% and enabling real-time metric updates.
- Presented statistical findings to stakeholders, leading to an 8% improvement in targeting ROI through optimized modeling.

**StockVision AI – Reinforcement Learning Project** Jun 2025 – Dec 2025

- Engineered a Deep Q-Network (DQN) agent to automate decision-making in high-frequency environments, achieving 15% Alpha.
- Optimized agent convergence by 25% through reward shaping and hyperparameter tuning of the experience replay buffer.
- Integrated real-time API state representations, simulating the challenges of low-latency, real-world data streaming..

**Computer Vision Image Classifier (Deep Learning)** Jan 2025 – Mar 2025

- Developed a CNN using PyTorch to perform complex image classification, applying normalization and data augmentation to improve generalization.
- Optimized model performance by 15% through meticulous hyperparameter tuning and architecture refinement.
- Implemented interpretability tools, including feature map visualization and confusion matrices, to analyze model decision-making processes.

## SKILLS & INTERESTS

**Deep Learning:** PyTorch, TensorFlow, CNNs, Transformers, Reinforcement Learning (DQN), Self-Supervised Learning.

**Languages & Frameworks:** Python (Expert), SQL, R, FastAPI, NumPy, Pandas, Scikit-learn.

**MLOps & Cloud:** AWS (Lambda, RDS, S3), Docker, CI/CD for ML, Experiment Tracking (Tableau/Custom Dashboards).

**Data Science:** Multimodal Data Integration, Predictive Modeling, Hypothesis Testing, Signal Processing.

**Certifications:** Google Data Analytics Professional Certification, 2025 [Proof of Certificate](#)