

KELLAN MCINTOSH

AI/ML Engineer | MLOps

Amherst, MA | 413-663-0343 | kellanmcintosh@gmail.com

[linkedin.com/in/kellan-mcintosh-202355399](https://www.linkedin.com/in/kellan-mcintosh-202355399) | github.com/kellanmcintosh | <https://kellanmcintosh.com/>

SUMMARY

AI/ML engineer who shipped a production agentic RAG system on AWS for a 6,300-employee organization and deployed a CNN end-to-end (92% test accuracy). Strong fundamentals: classical ML algorithms built from scratch and graduate coursework in Machine Learning and Simulation. Accelerated BS/MS candidate at UMass Amherst (4.0 GPA) seeking Summer 2027 AI/ML internships.

SKILLS

- **Machine Learning & AI:** Deep Learning (CNNs), Classical ML from scratch (Decision Trees, Random Forests, k-NN, Naive Bayes), RAG, Agentic AI, LLM Integration, Hybrid Retrieval (BM25 + dense), Model Deployment, MLOps
- **Languages & Libraries:** Python (PyTorch, TensorFlow/Keras, NumPy, Pandas, Scikit-learn), SQL, C, Java, TypeScript
- **Cloud & Data:** AWS (Bedrock, ECS Fargate, Amplify, Cognito, DynamoDB), Vector Databases (OpenSearch Serverless, S3 Vectors, pgvector), Docker, Google Cloud Run
- **Statistics & Simulation:** Monte Carlo & Discrete-Event Simulation, Bayesian Inference, Hypothesis Testing, Regression
- **Tools & Practices:** Git, Linux, CI/CD (GitHub Actions), Agentic coding workflows (Claude Code, Kiro)

EXPERIENCE

AI for the Commonwealth (AI4CW) UMass Amherst

Amherst, Massachusetts

AI Engineer Intern

01/2026 – 08/2026

- Led human-centered design discovery with HR stakeholders, translating ambiguous departmental pain points into a scoped technical solution for a RAG-based AI assistant.
- Architected and built a full-stack agentic RAG system (Next.js/FastAPI on AWS Bedrock) with complexity-based query routing and hybrid BM25 + dense retrieval, cutting mean policy lookup time from ~15 minutes to ~15 seconds (60× faster) with cited answers and 90%+ user satisfaction across a test group of 15 HR staff.
- Evaluated OpenSearch Serverless, S3 Vectors, and pgvector as vector storage backends; integrated OpenSearch Serverless to power production hybrid retrieval.
- Deployed the frontend on AWS Amplify with Cognito authentication and a GitHub Actions CI/CD pipeline, delivering a production tool for a 6,300-employee organization.

Massachusetts College of Liberal Arts

North Adams, Massachusetts

Tech Help Desk Consultant

07/2024 – 12/2025

- Resolved network, account, and hardware issues for students and faculty; supported campus-wide deployments and trained new staff.

PROJECTS

MRI Tumor Detection Model | [Code](#) | [Live Demo](#)

- Engineered a CNN with dual global pooling (average + max, concatenated) to classify brain MRI scans into four tumor types; used Grad-CAM analysis to diagnose a background-shortcut-learning bias and eliminated it via content-aware cropping and targeted data augmentation.
- Achieved 92% accuracy on a 1,600-image held-out test set; deployed via FastAPI and Docker on Google Cloud Run with CI/CD and Hugging Face Hub model versioning (pinned revisions for reproducible deploys).

Berkshire Bounty Data Report

- Cleaned and merged messy multi-source data from a biannual regional food-pantry funding survey; used hypothesis testing and regression to validate funding disparities, identifying Pittsfield and Springfield as the largest gaps, and presented findings in a data dashboard.

EDUCATION

University of Massachusetts Amherst

Amherst, Massachusetts

B.S. / M.S. Computer Science (Accelerated BS/MS) | GPA: 4.0

Expected B.S. May 2027, M.S. May 2028

Graduate Coursework: Machine Learning (COMPSCI 589), Introduction to Simulation (COMPSCI 550)

Relevant Coursework: Probability & Statistics, Algorithms, Systems Programming (C), Discrete Math & Theory of Computation

Honors: Dean's List (every semester), Alpha Lambda Delta Honor Society

Massachusetts College of Liberal Arts | CS coursework, GPA: 4.0 (transferred to UMass)

01/2024 – 12/2025