Dylan Lewis

dylanrl97@gmail.com | +1 210-818-8787 | dyllew.github.io | LinkedIn: /dyllew | GitHub: /dyllew
New York, NY

Education

Massachusetts Institute of Technology — Cambridge, MA
***Master of Engineering in Electrical Engineering and Computer Science**; GPA: 5.0/5.0* — *May 2022*
Thesis: Towards Automated Assessment of Crowdsourced Crisis Reporting for Enhanced Crisis Awareness and Response — *Concentration: AI*

***Bachelor of Science in Electrical Engineering and Computer Science**; GPA: 4.3/5.0* — *May 2020*

Relevant Coursework: Software Studio **(TA)**; Software Construction; Data Structures & Algorithms; Computer Systems Engineering; Statistics, Computation, & Applications; Machine Learning; Advanced Natural Language Processing; Machine Learning & Data Science in Politics

Experience

AreaHub — New York, NY
Backend Software Engineer with Data Science & Frontend Contributions — *March 2023 - Present*

- Transformed a manual Python-based pipeline into an automated, dockerized Kafka consumer microservice, enabling reliable and scalable generation of a core company product upon API-triggered requests. Significantly reduced time and processing pipeline for bringing the product to clients.
- Built a cloud-native chatbot application powered by Retrieval-Augmented Generation (RAG) using LLMs, deployed on AWS using Lambda and Bedrock, and orchestrated via Serverless and Terraform
- Provisioned an AWS MSK Kafka cluster and NestJS Kafka Consumer microservice to offload conversation analytics to a dedicated Kafka consumer, enabling parallel processing and preserving chat latency.
- Developed an internal full-stack web application to streamline client onboarding and product generation workflows, leveraging NestJS for the backend and HTMX + Alpine.js for responsive, lightweight frontend interactions.
- Enhanced observability of several core applications using Datadog SDKs in TypeScript and Python to forward application logging and enable application performance monitoring (APM)
- Researched and designed a novel algorithm (patent-pending) to address data robustness and data availability challenges in algorithms dependent on real-time sensor data. Conducted extensive analysis of existing solutions to inform a proprietary design, which was submitted for a USPTO patent for its technical innovation and commercial potential across multiple sectors and applications.

MIT Urban Risk Lab — Cambridge, MA
Research Assistant — *September 2020 - July 2022*

- Utilized insights from crisis managers in the US and Fukuchiyama (FC), Japan to develop the Human Risk/No Human Risk classification task and determined a performance metric that strongly aligned with their priorities; achieved a 92.8% F2 score on a test set of past flood event Japanese text data in FC with a SVM model and pretrained BERT embeddings
- Built pipeline and visualization tool for experimenting with various featurizations of Japanese crisis text data, dimensionality reduction techniques, and clustering algorithms, in order to yield human-interpreted labels from the documents found in each cluster; uncovered 9 human-interpreted labels from past flood event data in FC
- Developed an open-source Python package for training, testing, and predicting with pretrained CNNs for classifying crowdsourced crisis image data; developed an open-source Python package for featurizing crisis text data, training and testing with a variety of classification machine learning models, and visualizing clusters of featurized text data

Publications & Open-source Projects

System and method of multi-modal multi-task environmental quality forecasting — AreaHub
Named Inventor on U.S. Patent Application No. 18/792,269 — *February 2025*

Urban Risk Lab Image Analysis Module / Text Analysis Module — MIT Urban Risk Lab
Creator of open-source Python packages created from Masters Thesis Research — *July 2022*

Towards Automated Assessment of Crowdsourced Crisis Reporting for Enhanced Crisis Awareness and Response — MIT Urban Risk Lab
Author of MIT Master of Engineering in Electrical Engineer and Computer Science Thesis — *May 2022*

Skills

Programming Languages
Expert: {TypeScript, Python}; Proficient: {SQL}

Technologies
PostgreSQL, Kafka, AWS, LLMs, Terraform, Datadog, Docker, Fly.io

Libraries & Frameworks (Full-stack)
NestJS, TypeORM, htmx, Alpine.js, React

Libraries & Frameworks (Data Science)
PyTorch, Hugging Face, pandas, scikit-learn