

# Daniel Santiago Fleming

[Website](#) | [Github](#) | [LinkedIn](#) | [daniel.fleming@alum.mit.edu](mailto:daniel.fleming@alum.mit.edu) | (956) 252-5720

## EDUCATION

**Massachusetts Institute of Technology** | *BS in Computer Science and Engineering*

Aug 2020 - May 2024

Relevant Courses: Machine Learning, Computer Vision, Data Science, Discrete Math, Software Construction

Society of Hispanic Professional Engineers, TAMID Group, Chess Club, Poker Club, Sigma Nu Fraternity

## SKILLS

*Programming Languages:* Python, TypeScript, C++, Assembly, Java, R, Scala

*Software / Frameworks:* Next.js, Tailwind CSS, Flask, React, Vue, Django REST, NumPy, OpenCV, Tensorflow, PyTorch

*Tools / Tech:* SQL, PostgreSQL, Supabase, Prisma, Git, Bash, Google Cloud, IBM Cloud, Docker, CI/CD

*Languages:* Spanish (ILR Level 4)

## RESEARCH EXPERIENCE

**MIT Sports Lab** | *Undergraduate Researcher*

Sep 2023 - May 2024

- Improved data utilization, as measured by an 83% increase, by translating complex data into clear instructions
- Enhanced NBA injury risk prediction by a 14% accuracy improvement by analyzing NBA skeletal tracking data

**MIT Digital Humanities Lab** | *Undergraduate Researcher*

June 2021 - July 2022

- Accelerated the formatting of over 100,000 images by 50% through implementing a multithreaded architecture

**Summer Science Program for Astrophysics** | *Researcher*

June 2019 - July 2019

- Determined the orbital elements of asteroid (2001 XT30) with 99.6% accuracy via Gauss's method, Monte Carlo analysis, and Differential Correction; Published asteroid data to the Minor Planet Center by submitting verified observations

## TECHNICAL EXPERIENCE

**Semantic Retrieval Engine** | *TypeScript, Next.js, Vector Databases*

October 2024

- Built a Retrieval Augmented Generation pipeline using dense embeddings and OpenAI API
- Scraped over 3,000 recipe webpages and embedded the data into a Pinecone vector database for efficient retrieval

**Quarterly Report Parser** | *TypeScript, Next.js, Supabase*

October 2024

- Developed a full-stack web application using Next.js for user authentication and Supabase PostgreSQL for storage
- Implemented uploads for 10-Q Reports, utilizing LLM-based unstructured data parsing to automate financial data extraction into CSV format

**AI Chess Bot** | *Python, Tensorflow, Google Cloud Platform, Flask*

Aug 2024 - Sep 2024

- Engineered a full-stack chess web app with a CNN-based engine trained on 1 Million+ Lichess games and fine-tuned to my opening preferences, mimicking my play style
- Developed the backend with Flask and leveraged GCP with an NVIDIA L4 GPU for efficient training

**Enhancing Stable Diffusion with ControlNet** | *Python, OpenCV*

Apr 2024 - May 2024

- Fine-tuned Stable Diffusion by integrating ControlNet to incorporate color information by conditioning the model on colored edge maps alongside text prompts to improve image fidelity
- Improved image quality by 11% using Color Mean Squared Error (MSE) to compare an enhanced model against a baseline Stable Diffusion model

**AI Mammogram Analysis** | *Python, OpenCV, PyTorch*

Mar 2024 - Apr 2024

- Improved breast cancer detection to 93.4% validation accuracy by training a Mask R-CNN model for classifying and segmenting mammogram images
- Evaluated model performance, achieving recall of 1.0 with confusion matrices and precision-recall curves

**Computational Analysis of Hi-C and RNA-seq Datasets** | *Python, NumPy, Sklearn, Pandas*

Mar 2024

- Identified interchromosomal interactions in Hi-C data by detecting 442 high-interaction regions ( $p < 0.01$ ) by developing a greedy search algorithm
- Identified key genomic markers for cell clusters using PCA and regularized logistic regression on single-cell RNA-seq data

## HACKATHONS

**Health Lens** | *TypeScript, Next.js, Node.js, Flask, Supabase, IBM Cloud*

Aug 2024 - Oct 2024

- Created a free skin and eye diagnosing service to improve healthcare access and health literacy in underserved areas
- Developed a melanoma and cataract classifier using a CNN trained on over 10,000 images for skin condition diagnosis

## PUBLICATIONS

**New Observations of Near-Earth Asteroid 153842 (2001 XT30)**

June 2019 - July 2019

- Used astrometry and photometry to analyze data collected from the asteroid before calculating its trajectory. Analysis of its orbit and information from JPL Horizons shows that its a Mars-crossing asteroid with unlikely potential to hit Earth soon
- Minor Planet Center, July 27, 2019; Rolfness, J., Liang, L., Fleming, D., et al. 2019 "New Observations of Near-Earth Asteroid 153842 (2001 XT30)," M.P.S. 906049, M.P.C 11064