MATTHEW TORONTO

Oakland, CA 94607

508-439-1180 — matt.toronto97@gmail.com

https://github.com/mrtoronto/ // https://www.linkedin.com/in/matthewtoronto/

EXPERIENCE

Senior AI/ML Engineer, *66degrees*, Remote — Dec. 2023 to Current
- Led development of enterprise LLM automation suite for print media design
  - Scaled from Proof of Concept to largest AI project in company history ($3M+ revenue)
  - Built end-to-end pipelines for natural language edits, SVG generation, and automated IDML population
- Developed custom LLM-powered recommendation solutions for one of the largest US food distributors and others
- Created ML-driven dynamic pricing system for major US soccer franchise, optimizing for revenue and sellouts
- Fixed Two-Tower Recommendation System performance: 3x faster training, 2x inference speed, +60% accuracy
- Built enterprise LLM application for automated Technical Design Document generation
- Served as primary technical advisor for AI initiatives across multiple enterprise clients
- Delivered technical workshops and demonstrations to C-level executives and engineering teams
- Promoted from AI/ML Engineer to Senior AI/ML Engineer in March 2025

Machine Learning Engineer, *Evaluate.market*, Boston, MA — Jan. 2021 to Sept. 2023
- Engineered agentic AI system using GPT-4 for automated digital asset trading
  - Incorporating real-time pricing, inventory management, and multi-step negotiations
- Developed ML-based user matching algorithm analyzing portfolio similarities across 100K+ users
- Built real-time asset valuation model with high accuracy across asset segments using 1M+ historical transactions
- Created automated content generation pipeline using GPT-3 and Stable Diffusion
  - Produced 10,000+ unique digital assets with spritesheets and associated metadata

Data Scientist, *Supportiv*, Berkeley, CA — Apr. 2020 to Jun. 2022
- Trained, deployed and maintained critical BERT models for real-time crisis detection, achieving 95%+ recall
  - Model got so good it achieved superhuman performance and identified weaknesses in the moderators
- Built automated ML pipeline for continuous model training and deployment via Google Cloud
- Improved user similarity and content recommendation systems supporting real-time matching and resource delivery
- Re-designed content ETL process adding audio transcription, link validation, and enhanced monitoring
- Created end-to-end analytics infrastructure for real-time monitoring of user engagement and model performance
- Developed automated reporting system for key stakeholders, directly supporting CEO and CTO

Data Engineer Co-op, *Iora Health*, Boston, MA — Jan. 2019 to May 2020
- Developed predictive XGBoost models to forecast patient admission risk and medical costs
- Engineered high-performance HL7 processing pipeline, improving efficiency by 99.9% (from 5 mins to 2 sec per file)
- Worked with clinical teams to optimize healthcare reporting, improving data accessibility for patient care decisions

Market Research Intern, *ioVita*, Boston, MA — Sept. 2018 to Dec. 2018

Research Assistant, *Gabuzda Lab, Harvard Medical School*, Boston, MA — Sept. 2017 to Dec. 2017

Microbiology Lab Assistant Co-op, *Evelo Biosciences*, Cambridge, MA — Jan. 2017 to Aug. 2017

EDUCATION

Northeastern University, Boston, MA
- M.S. Bioinformatics — *Jan. 2018 - Dec. 2020*
  - 3.9 GPA
  - Graduation date: Dec. 2020
- B.S. Biology — *Sept. 2015 - May 2020*
  - 3.7 GPA
  - Graduation date: May 2020
  - Magna cum Laude

SKILLS

Languages: Python (sci-kit learn, Pytorch, HuggingFace, OpenCV, PIL, Langchain, LlamaIndex, FAISS, matplotlib, selenium, Flask, FastAPI, Pandas, NumPy, seaborn, Keras), Vanilla Javascript, SQL, R, Solidity, Visual Basic, HTML, CSS, Bash

Cloud Providers: GCP, AWS

Other: Cursor, Copilot, LLMs, AI, ML, Neural Networks, Artificial Intelligence, Machine Learning, Prompting, LLM Tool Use, Claude, GPT, Mistral, Stable Diffusion, Imagen, Groq, Llama models, Graphs and Networks, Knowledge Graphs, Docker, NoSQL (Firebase), Linux, GitHub Workflows, Ubuntu, PostgresSQL, MySQL, Redshift, Looker, Tableau, HL7, PHI

Interests: Guitar, music production, cooking, skiing, gaming

SELECTED PROJECTS

AI Leads (v1, v2)

https://github.com/mrtoronto/ai-leads
https://github.com/mrtoronto/ai-leads-v2

- Built an AI Agent to collect B2B Leads from the internet
  - Agent generated queries, searched queries, filtered results, analyzed leads and recorded relevant information
- V1: Too focused on real-time, website-scale UI
  - Ran V1 as a business and got ~180 sign-ups but too little traction or repeated use to continue
- V2: Using Google Sheets as the database and the gspread API to query and update the "database"

Web3HFT

https://github.com/mrtoronto/web3_hft

- Wrote code to trade and evaluating trading strategies with the freqtrade library
- Setup logic to screen assets, run and evaluate parameter grid searches and deploy params to live bot
- Pushed all metrics and logs to spreadsheet for manual inspection (example sheet)
- Preceded by a NEAT ETH model where I tried to use a genetic algorithm to generate strategies
  - The genetic algo was too much and not enough at the same time

Cover Letter Generator

https://github.com/mrtoronto/cover-letter-generator
https://mrtoronto.github.io/cover-letter-generator/

- Uses an LLM to generate cover letters using relevant context
- Context includes user's resume, example cover letter(s) and the job description for a specific job
- Deployed as a Github page to make it easy to share with less technical friends

Podcast Semantic Search

https://github.com/mrtoronto/podcast-semantic-search

- Built system to ingest, transcribe and run semantic queries on podcast transcripts
- Used Huggingface Wav2Vec model to transcribe text, embedded with Deep Passage Retrieval model

Lycan Protocol Developer

https://github.com/mrtoronto/lycan-protocol-contract
https://etherscan.io/address/0xaed8f91df327d9db27e5b5e0f11a0a001c727ce0

- Only technical team member for launch of ERC-721A token
- Wrote, tested and deployed smart contracts for token
- Generated 6,600 unique images and organized associated trait-level metadata
- Setup and executed decentralized payment methods for team

Forms Discord Bot Engineer

https://github.com/mrtoronto/forms-discord-bot

- Created Wavey, a GPT-4 powered discord bot with different personalities, commands and modes
  - Sarcastic and borderline rude but always hilarious
  - Created greeting channels and greeted each user that joined with a VIP link
  - Occasionally went into roast mode for a few hours
- Bot was deployed to a digital ocean server and worked using discord.py and OpenAI API's

NFT Shazam

https://github.com/mrtoronto/nft-shazam

- Created a system to let users find relevant NFT metadata from an image or screenshot
- System embedded tens of thousands of NFT images using alexnet and stored embeds in a FAISS index with metadata
- Queries sent to the system were embedded the same way and most similar stored assets were returned