DIEGO TORIBIO

917-743-2170 | toribio.diego23@gmail.com | github.com/toribiodiego | Queens, NY

EDUCATION

The Cooper Union for the Advancement of Science and Art — New York, NY

MEng, Electrical Engineering | GPA: 4.0/4.0 — 2025 – 2026

Courses: Deep Learning · Frequentist Machine Learning · Remote Sensing · Cloud Computing

BEng, Electrical Engineering – Track: Computer Engineering | GPA: 3.2/4.0

Courses: Operating Systems · Databases · Software Engineering · Stochastic Processes

Bard College — Annandale-on-Hudson, NY

Associate of Arts, Liberal Arts | GPA: 3.75/4.0

PROFESSIONAL EXPERIENCE

COLLECTIF Engineering — New York, NY

Electrical Engineering Intern — Summer 2024

- Parsed 9,000+ HVAC sensor records with Python to generate risk charts, cutting review from 4h to <1 min.
- Produced NEC-compliant load calculations, panel schedules, and CAD/Revit layouts for 12 projects, five managed solo.

PROJECTS

Multilabel Emotion Classification, Natural Language Processing — Fall 2024

- Processed 58,000 Reddit comments for multi-label emotion tagging across 27 categories.
- Fine-tuned various BERT-based transformer models with Hugging Face, comparison variants to maximize accuracy.
- Achieved 96% with RoBERTa-Large and 94% accuracy with a lightweight variant using half the compute.

Interactive Multimodal Agent, Generative Machine Learning — Spring 2025

- Streamed live webcam and microphone input through the Gemini Live API to create an interactive internet troll.
- Created a Gradio interface with YAML-configurable settings for audio, video, and agent system instructions.
- Ran the low-latency demo on consumer hardware at an exhibit, allowing attendees to chat live with the Agent, Agnus.

Checkers Agent, Reinforcement Learning — Fall 2024

- Trained a deep RL agent to play Checkers by playing against itself using Proximal Policy Optimization.
- Improved training stability by applying advanced initialization and normalization methods to prevent erratic learning, with progress logged in Weights & Biases.
- Achieved a ~60% win rate against random play and doubled average game length after 5,000 matches.

RESEARCH EXPERIENCE

Multimodal Alzheimer's Detection, Keene AI Labs — Spring 2025

- Isolated patient speech from 155 doctor-patient recordings and generated aligned Whisper transcripts.
- Extracted acoustic and language embeddings using pretrained transformer models.
- Combined those embeddings to train Random Forest & XGBoost classifiers for Alzheimer's detection.

TECHNICAL SKILLS

Languages: Python · JavaScript · C/C++ · Bash · SQL

Frameworks: TensorFlow · PyTorch · Ray · MLX

Libraries: Hugging Face· scikit-learn · OpenCV · Google Earth Engine

Tools: Git · Docker · Kubernetes · Jupyter · Weights & Biases · GCP (Compute Engine) · AWS (S3, EC2, SageMaker)

Interest: Machine Learning, Data Science, Deep Learning, Reinforcement Learning, Running, Cycling