

Jonathan Duran-Ortiz

U.S. Citizen | 252-739-0978 | jonathan.duran8754@gmail.com | jbduran.com | [LinkedIn](#) | [GitHub](#) | [Hugging Face](#)

EDUCATION

North Carolina State University, College of Engineering

Expected May 2027

B.S. Computer Science, Minor in Business Entrepreneurship

GPA: 3.88

Pack Promise Scholar (full ride)

Coursework: Neural Networks, Automated Learning and Data Analysis, Web App Computing, Software Engineering, Data Structures & Algorithms, Operating Systems, DevOps, Neurobiology

Independent Study: Neural Networks: Zero to Hero (Andrej Karpathy)

TECHNICAL SKILLS

Languages: Python, Java, C, SQL, JavaScript, TypeScript, HTML/CSS, Kotlin, Swift, Bash

ML & Experimentation: PyTorch, scikit-learn, fastai, Hugging Face, Weights & Biases, TensorBoard, LLM Benchmarks, Pretraining, Data Cleaning

Frameworks & Data: React, Flask, Spring Boot, SwiftUI, SwiftData, PostgreSQL, Supabase

Cloud & Tools: AWS (Lambda, S3, EC2), Modal (serverless GPUs), Vast.ai, Docker, GitHub Actions, CI/CD, Git, Linux, Playwright

EXPERIENCE

Unbounded Labs | *Research Engineer* | [Technical Report](#)

May 2026 – Present

- Founded an independent research lab and co-developed Bartholomew: A Vintage LLM with a **three-person team**, pretraining a **2.82B-parameter** decoder-only transformer from scratch on **20.1B tokens** of pre-1931 English using nanocat and PyTorch. Model architecture: untied token/output embeddings, rotary positional embeddings (RoPE), QK normalization, ReLU² MLPs, grouped-query attention (GQA), and FlashAttention-3.
- Filtered Harvard's **242B-token** Institutional Books collection to build the pretraining corpus using OCR-quality, date, language, deduplication, and anachronism filters; implemented the team-designed ablation harness and Weights & Biases tracking.
- Configured Docker environments on rented Vast.ai H100s and ran architecture and context-length ablations; the team's final pretraining run sustained approximately **60% MFU** at **50K tokens/second** and ran for **5 days**.
- Created Vintage CORE's **20-task** evaluation suite and a historical-event leakage benchmark; created GSM8K-Vintage, performed QA on midtraining/SFT datasets, and ran supervised fine-tuning on our larger models.
- Adapted Karpathy's autoresearch to run **100 experiments** in **10 hours** on one H100, reducing best search-run validation bits per byte by **1.70%**; the team carried the Muon momentum finding into final training.
- Deployed inference on serverless GPUs from Modal with cached model weights: **13-second median cold start** across six cold containers; the model endpoint received **6,000 POST requests** in its first 24 hours after release.

Voice First | *Software Engineer*

Sep 2025 – May 2026

- Deployed Vapi-integrated voice agents for **three paying customers** and **two trial customers**, supporting approximately **50 calls** per company per week; built client knowledge bases and handled sales and onboarding.
- Built a React/TypeScript dashboard with Supabase PostgreSQL, Edge Functions, row-level security, Stripe subscriptions, and call analytics spanning recordings, transcripts, sentiment, and outcomes.
- Compared STT, TTS, LLM, and turn-detection configurations through Vapi telemetry and live-call transcripts; integrated tool calling, agent instructions, AWS functions, scheduled jobs, and WebSockets.

RDU Technology | *Software Engineer (Contract)*

Jun 2024 – Jun 2026

- Delivered Dreamline Glazing's production website, with responsive project showcases, service pages, and quote collection.
- Automated pricing for **1,000+ trading cards** using Python, Playwright, and Pandas, saving local stores **5+ hours weekly**.

PROJECTS

NeuroPawn Head Cursor | *Python, scikit-learn, BrainFlow* | [Source](#)

Apr 2026

- Built hands-free computer control using a NeuroPawn headset, combining IMU head tracking with electrode-based eyebrow-raise and jaw-clench detection and a personalized classifier trained on **400+ samples**. Implemented sensor streaming, calibration, live inference, and a diagnostic GUI.
- Extracted **16 time- and frequency-domain features** from electrode signals; compared Random Forest and Extra Trees classifiers using separate training and validation sessions.

3Define | *Trellis 2, Gemini, WebAR, iOS Quick Look* | [Demo](#)

Feb 2026

- Led a **four-person team** at Hack_NCState and built the image-to-3D pipeline in **under 24 hours**, converting product photos into 3D assets while delegating frontend development.
- Created the 3D visualizer and integrated WebAR and iOS Quick Look, enabling users to preview generated objects in their physical room.

LEADERSHIP

Thinking Machines, NCSU AI Club | *Co-President*

Jul 2026 – Present

- Founded and co-lead NC State's largest AI club with **100 members**; organize weekly meetings covering ML news, technical deep dives, and model-building competitions. Teach ML, lead LLM and AlphaGo projects, and organize hackathons and industry speakers.
- Designed a PyTorch Mandelbrot challenge with a 30-second Tesla T4 training limit, hidden MAE evaluation, and **217 leaderboard submissions**.

NCSU App Club | *iOS Lead*

Jan 2024 – May 2026

- Led **eight developers** through a **five-week** Swift/SwiftUI bootcamp. Led a **five-person team** to build Sequel, a SwiftUI social discovery app for film and television. Coordinated architecture, system design, wireframes, sprint planning, code reviews, and TMDB integration.