leo@alexeyev.org
(727) 776-5537

Leonid Alexeyev

linkedin.com/in/leonid-alexeyev/
www.alexeyev.org

Education

Georgia Institute of Technology

- M.S. Computer Science, Concentration: Machine Learning — **Expected May 2026 | GPA: 3.67**
- B.S. Computer Science, Concentration: Intelligence and Devices — **May 2025 | GPA: 3.78**
- Teaching Assistant for CS 3630 (Perception and Robotics) and CS 2200 (Systems and Networks)

Skills

- **Languages:** Python, C++, C, C#, Java, JavaScript, Typescript, CSS, SQL, MATLAB, Assembly
- **Tools:** PyTorch, HuggingFace, React, React Native, Node.js, .NET, Firebase, PostgreSQL, MySQL, Git, CAD

Experience

Software Development Intern — **QGenda** — **May 2025 - August 2025**

- Developed and maintained full stack features for on-call and location-based scheduling using **React** and **.NET**, supporting a healthcare workforce management SaaS platform used by hundreds of thousands of clinicians across multiple hospital systems
- Fixed 21 production bugs improving system reliability and reducing support escalations
- Built a custom OpenSearch connection snippet and wrote a guide to fix a bug now used across the engineering team
- Designed, optimized, and troubleshot **PostgreSQL** and **MySQL** queries for performance-critical scheduling and reporting features

Pavement Quality Researcher — **Georgia Institute of Technology** — **January 2024 - May 2025**

- Investigated Full Reference Image Quality Assessment (IQA) and Point Cloud Quality Assessment (PCQA) indicators to develop a standardized validation metric of 3D scanners used for pavement crack detection, identified MS-SSIM as strong candidate for cracking inference
- Wrote **Python** scripts to validate identified IQAs and analyzed performance metrics using provided range data
- Pre-processed point clouds using **Matlab** and implemented candidate PCQAs on provided scans of pavement
- Developed new annotation procedure and software to address the issue of phantom cracks, improved manual annotation accuracy by 15%

Projects

AutoLLMTuner — **October 2025 - December 2025**

- Built an automated per-layer quantization system for LLM's that optimizes latency, memory, and output quality compared to standard per-model quantization
- Created components for model loading, layer-wise quantization, benchmarking, and visualizations using **Python**
- Designed and implemented an automated benchmarking pipeline to evaluate accuracy, latency, and VRAM usage across thousands of quantization configurations
- Represented model configurations as precision vectors and used an evolutionary algorithm to find a pareto frontier
- Ran large-scale experiments on Qwen-7B, Mistral-7B, and Trinity-Mini across 8× H200 GPUs, discovering that non-uniform precision vectors consistently outperform uniform quantization

Dolphin Chat Mask — **September 2025 - November 2025**

- Designed a head-mounted real-time spectrogram system for live underwater visualizations of dolphin vocalizations
- Built a signal-processing pipeline on a Teensy 4.0, using **C++** to compute 256-point FFTs in real time
- Engineered a ring-buffered display pipeline with waterfall spectrogram rendering for zero-drop frame rendering
- Reverse-engineered a SCUBAPRO Galileo HUD for power, display, and rotary-knob inputs

Liftr — **January 2025 - April 2025**

- Developed a mobile social fitness app using a **React Native** framework, **Node.js** backend, and a **Firestore** NoSQL database, allowing users to log workouts, track progress, and connect with others
- Designed an optimized follower feed system using caching, significantly improving cost efficiency and scalability
- Implemented a lazy-loading design for user activity fetching, reducing unnecessary Firestore reads by 10x

EMS Scheduler — **February 2025 - March 2025**

- Built a role-aware shift scheduling web application for volunteer EMS teams using **React** and **Firebase** with dynamic calendar views and live availability filtering to prevent coverage gaps and role conflicts
- Implemented admin-controlled user provisioning and shift management for access control