

# Erim Ozben Evren

425-435-9074 | [eevren@uw.edu](mailto:eevren@uw.edu) | [linkedin.com/in/erim-evren](https://www.linkedin.com/in/erim-evren)

## Profile

---

Electrical & Computer Engineering student focused on machine learning for high-performance systems and optimization under real-world constraints. Experience building ML-driven pipelines, training models from scratch, and designing low-latency systems in Python and C++, with interest in applying ML to hardware and large-scale system optimization.

## Experience

---

**GoDaddy** **June 2025 - September 2025**

Software Engineering Intern Tempe, AZ

- Built an LLM-based system to analyze large-scale production pipelines, identifying upgrade risks across complex dependency graphs
- Designed a RAG pipeline with vector search over code, configs, and system dependencies for structured querying and reasoning
- Developed data representations and evaluation workflows to assess model output quality and system reliability
- Reduced manual debugging effort by surfacing failure modes pre-deployment through automated analysis
- Owned end-to-end system design, implementation, and validation in a production environment

**Husky Satellite Lab** **December 2024 - Present**

CubeSat Software Engineer Seattle, WA

- Developed telemetry and onboard data handling systems in C++ under strict memory and bandwidth constraints
- Designed and implemented custom serialization/deserialization logic for efficient binary communication between subsystems
- Built validation frameworks to ensure deterministic behavior in reliability-critical spacecraft software
- Independently owned implementation of core data handling components, from protocol design to testing and debugging

**Washington Aerial Robotics** **June 2025 - Present**

Computer Vision, Mapping, and Autonomy Lead Seattle, WA

- Built a real-time computer vision pipeline for autonomous drone navigation under strict latency constraints
- Optimized neural network inference for real-time performance and stable throughput
- Measured system latency, frame throughput, and failure modes under varying environmental conditions
- Integrated perception outputs into control systems for multi-agent coordination

## Projects

---

**ML-Guided Design Space Optimization (Python, PyTorch)**

- Built a learned cost model to predict system performance across a high-dimensional design space (design space exploration)
- Modeled tradeoffs between competing objectives (e.g., performance, power, and area)
- Used learned model predictions to guide search toward optimal configurations under constraints
- Compared learned optimization against brute-force and heuristic search methods
- Developed automated evaluation pipeline for benchmarking and iterative model improvement

**Transformer Language Model (From Scratch, PyTorch)**

- Implemented a transformer architecture from first principles (attention, positional encoding, autoregressive decoding)
- Trained on 6B+ tokens and evaluated using cross-entropy loss and perplexity
- Ran controlled experiments on model depth, context length, and hidden size to analyze scaling behavior and convergence
- Optimized training for iteration speed and stability across configurations
- Explored tradeoffs between model capacity, compute cost, and performance

**Learned System Performance Predictor (Python, PyTorch)**

- Built a machine learning model to predict system latency and throughput from structured input features
- Engineered features representing system configurations and workload characteristics
- Trained and evaluated regression models to approximate performance metrics under varying conditions
- Analyzed prediction error and model generalization across unseen configurations
- Used model outputs to identify performance bottlenecks and guide system-level improvements

**Low-Latency Order Book & Matching Engine (C++)**

- Designed a deterministic, event-driven matching engine optimized for low-latency processing.
- Implemented cache-efficient data structures and memory layouts to maximize throughput under high-frequency workloads.
- Instrumented p50/p99 latency and throughput metrics to evaluate performance under load.

## Education

---

**University of Washington** **September 2024 - June 2028**

Electrical and Computer Engineering Seattle, WA

- UW Academy (early admission after 10th grade with merit based scholarship), Honors

## Skills

---

**Languages:** Python, C++, SQL

**ML & Systems:** PyTorch, transformers, RAG, vector databases, low-latency systems, real-time systems

**Foundations:** probability, statistics, optimization