Usman Khan

usmankhan.dev | usman@usmankhan.dev | **US Citizen** | linkedin.com/in/khanu | github.com/ukhan1219

Education

University of Central Florida — Orlando, Florida
B.S. in Computer Science — *Expected Graduation: December 2025*

Relevant Coursework: Algorithms in Machine Learning, Artificial Intelligence/Machine Learning, Robot Vision, Computer Vision

Work Experience

Software Engineering Intern — Aug 2024 – Jun 2025
Vcom3D — *Python, TensorFlow, OpenCV, Raspberry Pi 5, Meta Quest 3, BioGears (UW), C++, XML* — *Orlando, Florida*

- Built pose tracking models using TensorFlow on Raspberry Pi, boosting accuracy & reducing latency by **30%**
- Merged BioGears (University of Washington) for injury simulation, boosting training realism by **98%** across modules
- Created AR/VR apps on Meta Quest to support simulations ran by BioGears in a distributed system architecture
- Refined system integration across components via cross-functional collaboration, slashing errors & streamlining updates

Machine Learning/AI Undergraduate Research Assistant — Apr 2024 – Apr 2025
University of Central Florida — *Python, TensorFlow, Neo4J, NumPy, SKLearn, NetworkX, Pandas* — *Orlando, Florida*

- Enforced automated distributed data mining algorithms using AI/ML via Neo4J for enhanced predictive analytics
- Generated data mining methods for RandomForestRegressor on a DARPA dataset (**6.8M+** nodes) to detect illicit activity
- Devised scalable distributed data pipelines boosting entity tracking accuracy and speed by **30%** across datasets
- Deployed statistical methods for performance optimization, reducing processing time by **40%** for high-volume pipelines

Projects

PyChess | *Python, PyTorch, Hugging Face Transformers/TRL, Accelerate, python-chess, Stockfish, TensorBoard, DistilGPT-2*

- Engineered an end-to-end chess AI post-training pipeline that automates data generation and training with robust tracking
- Processed 90M positions; curated 1M supervised samples and 500k preference pairs using strict quality filters
- Reduced data generation time by **97%** via multithreading and optimized I/O; eliminated memory-related training failures

Mantle | *SwiftUI, Python, PyTorch, Core ML, Transformers, Hugging Face, Metal (MPS), Amazon Web Services EC2*

- Converted Transformer models (Mistral, Llama) from PyTorch to Core ML utilizing AWS EC2 instances
- Applied Core ML compression (quantization, pruning, palettization) shrinking models by **75%** while retaining accuracy
- Accelerated inference **25%** leveraging Metal Performance Shaders (MPS) optimization on for On-Device inference
- Developed privacy-first SwiftUI app (**iOS 18+**) for On-Device ML inference, enabling offline AI chatbot functionality

Glance | *SwiftUI, Golang, Firestore, Firebase Auth, Plaid API, Google Cloud Platform, Figma, XCTest*

- Architected a budgeting app using SwiftUI and a Go backend, achieving seamless Plaid API integration
- Implemented secure authentication via Firebase Auth & managed sessions, supporting **100+** concurrent users reliably
- Enhanced data retrieval speeds by **40%** through strategic caching & optimized Firestore queries in the Go backend

DUI | *Rust, Crates.io (Cargo), Homebrew, clap, rustyline, crossterm, tui-rs, serde*

- Reimagined a Docker CLI with **100%** command parity, interactive mode, and real-time charts/dashboard
- Published on Cargo as dui-cli and Homebrew tap; reached **1000+** downloads
- Improved UX with tab completion, contextual help, smart suggestions; optimized release builds

Whooga | *PERN: PostgreSQL, pgvector, Express.js, React, Node.js, Python, TypeScript, AWS RDS*

- Managed full stack project with 4-person team building specialized collections marketplace using Jira
- Architected vector database with pgvector for semantic search; engineered ML pipeline for item matching with BERT
- Designed computer vision system (YOLOv8, SAM, DINOv2) for semantic image matching across photography conditions
- Implemented real-time messaging system with WebSockets enabling seamless buyer-seller communication for transactions

Technical Skills

Languages: Python, Rust, TypeScript/JavaScript, C, Java, SQL, NoSQL
Frameworks: PyTorch, Keras, TensorFlow, NumPy, Pandas, SKLearn, Next.js, Node.js, Express.js, React, Tailwind
Tools: Git, Github, Docker, Linux, LaTeX, Prisma, Neo4J, Figma, Amazon Web Services, Google Cloud Platform