

Haris Rehman

AI/ML Engineer — Technical Product Management

☎ Call | @ harisrehmanchughtai@gmail.com | 🔗 LinkedIn | 🐙 GitHub | 📁 Portfolio | 📍 Alberta, Canada

PROFESSIONAL SUMMARY

AI/ML Engineer and Technical Product Manager with hands-on experience in Computer Vision, Deep Learning, and MLOps. Proven ability to bridge engineering and product by translating user needs and data insights into scalable AI solutions deployed on cloud and on-prem environments.

EDUCATION

National University of Science and Technology (NUST)

Islamabad, Pakistan

B.Sc. in Software Engineering

WORK EXPERIENCE

University of Alberta

Edmonton, Canada

AI Researcher - On-Site

June 2026 – Present

- Developing multimodal generative and foundational models integrating **12-lead ECG** waveforms with longitudinal **EHR** data (diagnoses, procedures, labs, medications, hospitalization timelines) for population-scale cardiovascular modeling.
- Building large-scale patient embedding models that learn temporal disease trajectories from aligned ECG and administrative health record sequences.
- Forecasting clinical outcomes including **rehospitalization**, **MACE**, & **mortality** using multimodal time-series inputs.

VyroAI - Chatly

Islamabad, Pakistan

Associate Product Manager - On-Site

October 2025 – May 2026

- Scaled overall product from **200K** to **1.2M+** MAUs (**6x** growth) by driving feature prioritization, growth experiments, and cross-functional execution.
- Grew AI Chat from **7K** to **70K+** users (**10x**) and launched AI Docs from **0** to **2K+** users within **30 days**, achieving rapid product-market validation.
- Increased user engagement with **35–50%** uplift in DAUs through onboarding improvements, retention strategies, and iterative UX enhancements.
- Contributed to monetization efforts, supporting revenue growth to the **\$6.5M** range via pricing experiments, funnel optimization, and Stripe billing improvements.
- Drove data-informed decisions by implementing end-to-end event tracking (GTM → Mixpanel), building real-time dashboards, and automating feedback reporting across teams.
- Collaborated cross-functionally with Engineering, Design, SEO, Branding, Support, and QA to identify pain points, prioritize roadmap, and improve release velocity by **20%**.
- Scaled team operations from **5** to **35+** members, improving cross-functional coordination, delivery efficiency, and execution bandwidth.
- Built **5+** internal tools to automate workflows, significantly reducing manual effort and improving team productivity.
- Improved platform reliability by contributing to QA automation workflows, reducing production issues by **50%** and increasing release confidence.
- Synthesized **3,000+** user feedback points/month into actionable insights to inform product strategy and feature development.

Cowlar Design Studio (YC W-17)

Islamabad, Pakistan

Machine Learning Engineer - Hybrid

April 2024 – April 2025

- Assisted with automation efforts for a factory in North America, boosting production capacity by **40x** for a key Amazon-partnered Fiber Optics supplier. Developed a system leveraging ML and CV to align fibers with **5-micrometer** precision, scaling monthly output from **240** to **9,600** cables.
- Enhanced the Fiber Organizer Robot, integrating ML and CV to sort **250-micron** fibers with **96%** accuracy, increasing AWS cable production by **500%** (from **1** to **6** cables per hour).
- Pushed AI solutions to production by setting up CI/CD pipelines with Docker whilst load balancing deployment servers using techniques like Nginx.
- Developed and deployed real-time AI solutions for customer-centric applications, including an Action Recognition system for Smart Carts achieving **95%** accuracy and an Azure OpenAI chatbot to revolutionize online retail customer experiences.
- Optimized model inference speeds using NVIDIA-enabled dockerization, code optimization, and dimensionality reduction, reducing inference times from **200ms** to less than **100ms** per request.
- Used Docker Swarm orchestration to set up distributed GPU compute across nodes for training models **3–4x** faster.

PROJECTS

Fiber Organizer & Termination Line Robot | ML, CV, Automation, Patented Tech | [viaPhoton](#) Oct 2024 – Dec 2024

- Contributed to the development of the patented (**US 12001073 B2**) machine, currently deployed in Aurora, IL, streamlining operations and reducing manual intervention.
- Contributed to the automation of a factory for an Amazon-partnered Fiber Optics supplier, increasing production capacity by **40x** (from **240** to **9,600** cables per month).

- Assisted the development of the Fiber Organizer Robot, a patented system utilizing CV and ML to sort **250-micron** fibers, increasing production capacity by **500%** and reducing sorting time from **3 minutes** to **8 seconds**.
- Integrated multiple fiber termination processes into an automated system, achieving **5-micrometer** precision and a **96%** success rate, producing **50** AWS cables per day.

Smart Shopping Cart | *Tensorflow, MoViNet, MQTT, MLflow, NGINX, Docker*

Nov 2024 – Mar 2025

- Contributed to the development and improvement of Cowlar Design Studio’s smartcarts for Al-Meera Supermarkets in Qatar.
- Optimized and improved the Action Recognition architecture for predicting customer interactions with smartcarts, achieving **95%** accuracy in real-time on edge devices.
- Continuously cleaned and improved the data pipeline through rigorous validation and testing to debug and better the model performance.

FleadrAI | [LinkedIn](#)

- Building an AI-driven lead generation system that automates **70%+** of the lifecycle by integrating LLMs with a dynamic “Lead Memory Vault” for personalized outreach. Developing web crawlers using headless browsers to populate a multi-zone, replicated DB, optimized through read replicas and indexing.
- Integrating DSPy for prompt optimization and ACP for inter-agent communication between help and action modules. React frontend with MCP server, CI/CD pipelines, containerization, GitHub workflows, and version control.

gptree-cli | [GitHub](#)

- Go-based CLI to export and summarize codebases into LLM-friendly prompts with directory trees, markdown, and token-aware chunking. It supports prompt generation for Claude/GPT, README generation, ‘.deb’ builds, and intelligent context formatting. Lets you filter the directories and files before passing to downstream tasks.

gitmind | [GitHub](#)

- GitMind is a powerful Go-based CLI tool that revolutionizes your Git workflow by providing AI-powered commit message generation, comprehensive multi-language security analysis, automated test generation, and intelligent multi-commit splitting. Works entirely offline with local LLMs for maximum privacy and performance.

Web Page Monitoring | [GitHub](#)

- Dockerized full-stack app that monitors webpages for changes and sends real-time email alerts. Supports account creation, project management, and multi-recipient alerts. Built using React, FastAPI, Node.js, MongoDB.
- Deployed via CI/CD to Render; designed for reliability and usability.

AI Enabled Gym | [GitHub](#)

- Built a real-time gym assistant using MediaPipe, OpenCV, and custom CNNs to monitor posture and provide corrective feedback. Classifies exercise types and prevents injuries through visual analysis. Also counts the reps for the user.

Open-Source Contributions

- Contributed open source to Ultralytics. I needed to enable distributed parallel computing for on-premises architecture for YOLO training. The pipeline was scattered. I connected the dots and got a new version released with the new support available for everyone.

SKILLS

Machine Learning & AI: Data Science, Deep Learning, Generative AI, Natural Language Processing (NLP), Computer Vision, Reinforcement Learning, Collaborative Filtering, MLOPS

Libraries & Frameworks: Scikit-Learn, StatsModels, Matplotlib, NumPy, Pandas, Hugging Face Transformers, TensorFlow, PyTorch, MySQL, React, JSX

Cloud & Deployment: Azure AI Studio, Azure Machine Learning Studio, Docker, Kubernetes (Docker Swarm), CI/CD, MLflow, TensorBoard, TorchServe, TensorFlow Serving, FASTAPI, Uvicorn, Flask, Nginx Load Balancing

Big Data & Infrastructure: Hadoop HDFS, Apache Spark, Hive, Zookeeper, MQTT, Prometheus, AlertManager, Redis with MongoDB, PostgreSQL, MySQL, MongoDB, Edge Model Deployment, GPU Optimization, Distributed Training, Remote Docker Image Registry Management (GitLab, DockerHub)

Development Tools: Postman, Insomnia, Git, Conda, virtual envs, poetry, UV, RedHat tmux

Computer Vision: YOLO, OpenCV, MediaPipe, Image and Video Classification, Object Detection, and Action Recognition

Soft Skills: Team Collaboration, Task Management, Critical Thinking, Project Management, Problem-Solving

Languages: English (Professional), Urdu (Native), Punjabi (Native), German (Novice), Japanese (Starter)

CERTIFICATIONS (CLICK TO VERIFY)

AWS Certified Machine Learning - Specialty [MLS-C01]	July 2025
AWS Certified Solutions Architect - Associate [SAA-C03]	August 2025
Certified Generative Adversarial Networks (GANs) Specialist	August 2023
IBM Certified Data Scientist	July 2023
DeepLearning.AI Certified TensorFlow Developer	June 2023
DeepLearning.AI Certified Machine Learning Engineer	April 2023

AWARDS & ACHIEVEMENTS

Cambridge Outstanding Learner Award: Ranked **2nd** in **5** different subjects in A Levels among all the students of Punjab district.

National Science Talent Contest (NSTC): Stood in the **top 50** selected candidates in mathematics from all over the country who then competed for a chance to represent the country in the international math Olympiad.