

NAVANEETH KRISHNAN

AI & ML Engineer

Navi Mumbai, India | navaneethks.official@gmail.com | navaneethks.site
linkedin.com/in/navaneethksofficial | github.com/navaneethkrishnansuresh | upwork.com/freelancers/navaneeth

SUMMARY

ML and inference engineer with 2+ years of production experience building model workflows, computer vision systems, OCR pipelines, agentic platforms, evaluation harnesses, and full-stack AI products for startups and international clients.

WORK EXPERIENCE

Core AI Engineer | BrainDrive LLC

Remote (US-based) | Jul 2025 - May 2026

- Built BrainDrive's full-stack hosted app for a user-owned AI system with billing, authentication, isolated VMs, and model-routing gateways.
- Used containerization and LiteLLM APIs to support model access, usage workflows, platform reliability, and security.
- Built ModelMatch, an evaluation harness for open-source models using LLM-as-judge scoring across finance, health, summarization, and therapy tasks.

Independent AI/ML Solutions Architect | B2B Consulting (Upwork Top Rated)

Remote | US / UK / EU clients | Jan 2024 - Jun 2026

- Delivered AI/ML systems for international clients across computer vision, generative AI, data workflows, and applied automation.
- Public profile proof: Top Rated, 100% Job Success, 5.0 rating, and 10 total jobs.
- Owned discovery, architecture, evaluation, delivery, and client handoff for business-facing AI workflows.

SELECTED PROJECTS

- **RailCompute, Agentic Fine-Tuning Platform:** Agentic platform that automates AI model fine-tuning across data prep, setup, and training with specialized agents - Cursor for AI training.
- **Geometric Volume Estimation Pipeline (UK client):** Engineered a deterministic inference pipeline using monocular depth estimation and geometric heuristics to calculate 3D volume from unconstrained 2D photos for an Upwork client.
- **Padel Vision Analytics (Serbia sports analytics company):** Computer vision system combining YOLOv8, TrackNet for ball tracking, DeepSORT for player ID, and MMPose HRNet for geometry-first stable identity tracking through occlusions, achieving 90%+ accuracy.
- **OneCollector AI OCR System (OneCollector, Germany):** Built an end-to-end AI system for iOS and Android for real-time TCG card scanning on a product with 50K+ users, combining YOLOv8 detection with custom OCR fine-tuned for small noisy ID regions and motion blur correction, plus data annotation, training, and mobile inference optimization.

OPEN SOURCE CONTRIBUTIONS

- **SGLang Project - SpecForge (Merged):** Fixed EAGLE3 loss-mask alignment during target preprocessing, restoring missing target supervision and removing stale boundary gradients; added regression coverage for mask edge cases.
- **BrainDrive ModelMatch (Merged | Code Owner | Repo Owner):** Built and maintained an LLM-as-judge evaluation harness for comparing open-source models across finance, health, summarization, and therapy tasks.

PUBLICATIONS & RESEARCH

IEEE Xplore: Micro-Expressions with Swin-T on CASME II | 2026

Published research on micro-expression recognition using Swin Transformer Tiny on CASME II, comparing spatial and temporal heads with gains from temporal cues and RAF-DB pretraining.

SKILLS

Programming: Python, C++, SQL

ML Frameworks: PyTorch, Transformers, OpenCV, RLHF, LoRA, QLoRA, LLM, VLM

Infrastructure & Cloud: Docker, GitHub Actions (CI/CD), PostgreSQL, AWS, GCP, RunPod, Kubernetes

Applied AI: RAG, AI memory, prompt engineering, evaluations, MCP, agentic systems, observability, LangChain

Math & Theory: Deep learning, neural networks, computer vision, natural language processing, LLM inference

EDUCATION

Level	Board / Program	Institution	Result
B.TECH	Artificial Intelligence and Data Science	NMIMS, Navi Mumbai	CGPA 3.39 / 4 2022 - 2026