Neo Mohsenvand

AI/ML Engineer and Product Designer

Agentic systems / Evaluation / Enterprise deployment / AI-first product design

neovand.github.io | NeoVand | mohsenvand | mmv@mit.edu | +1 857 500 0980

AI/ML engineer and product designer advancing AI adoption at Caterpillar through original products, shared infrastructure and engineering education. I build across model development, evaluation, multimodal UX and enterprise deployment, and led a four-engineer team delivering 12+ applications for 8,000+ technical users. Previously, I led research at BrainCo, developed biosignal models at Apple, and conducted doctoral and postdoctoral research at MIT.

TECHNICAL SKILLS

Programming	Python, TypeScript, MATLAB, Mathematica, GLSL, WGSL
Applications & UI	SvelteKit, React, Three.js / WebGL
Design & UX	Adobe Creative Cloud, UX study design, Design Systems
Cloud & data	Azure, AWS, Docker, progressive delivery, Snowflake, multimodal data pipelines
Data science	Time-series analysis, biosignal processing, feature extraction, scientific visualization
Machine learning	PyTorch, JAX, transformers, self-supervised learning, mechanistic interpretability
Training & alignment	FSDP, DeepSpeed, Ray; SFT, PEFT, LoRA, PPO / DPO, quantization
Inference & serving	Ollama, Transformers.js, WebGPU, ONNX / WASM, vLLM
Realtime AI	OpenAI Realtime, ElevenLabs, Deepgram; low-latency multithreaded harnesses
Agentic engineering	LangGraph, agent harnesses, RAG, tool orchestration, human-in-the-loop workflows
Evaluation	MLflow, DeepEval, benchmark design, automated and human evaluation

PROFESSIONAL EXPERIENCE

Caterpillar Inc. | Apr 2024 - Present

Chief AI Architect, Global Finance | Jan 2026 - Present

- **Tempo:** Built an enterprise investigation platform and designed its meeting-like interview interface, combining stakeholder discovery, scheduling and OpenAI Realtime voice interviews with structured forms. Available to hundreds of internal users for idea intake, requirements gathering and conflict resolution.
- **Aperture:** Built a new agentic image-generation harness for slide decks and visual assets, using terse prompts and hundreds of references to reduce reliance on third-party subscriptions.
- **Touchless Invoice Processing agent:** Co-developed end-to-end orchestration from email ingestion and scanned-invoice OCR through exception handling and Snowflake writes, with a detailed multi-level evaluation framework. Reduced invoices needing human review from **100% to 2%**.
- **Teaching:** Created 10+ workshop apps, including agentic RAG, talking dashboards and a voice-only data-science interface I designed. Teach engineers and senior leaders in the workshop's fifth iteration.
- Developing shared infrastructure for AI application prototyping and deployment across Global Finance.

Neo Mohsenvand

PROFESSIONAL EXPERIENCE / CONTINUED

Caterpillar Inc. | Apr 2024 - Present

Senior Principal Data Scientist, Global Finance | May 2025 - Jan 2026

- **Penman:** Built and deployed an enterprise financial-analysis system using **SvelteKit and Azure** to analyze earnings-call transcripts and regulatory filings for finance leaders.
- Redesigned earnings-call transcript proofreading in Penman to help finance teams identify disclosure risks.
- Built **MLflow / DeepEval** evaluation infrastructure and coordinated organization-wide benchmark development.
- Built an AI-idea prioritization agent with RICE-derived scoring, internal impact metrics and human review.
- Designed adversarial LLM analysis of draft financial disclosures.

Senior Engineering Fellow, Artificial Intelligence | Apr 2024 - May 2025

- Founding technical hire of CAT Tech AI; led four engineers delivering **12+ applications** for **8,000+ technical users** through engineering-analysis automation.
- Created **Magic Table** for million-row spreadsheet/database workflows, combining formulas and iterative local-LLM prompts to reduce cloud compute costs; developed versions with Ollama.
- Built a **ModernBERT-based embedding model** for Caterpillar products and parts; orchestrated data collection, developed a synthetic training-pair generation app, and trained on an **on-premises H200 cluster**.
- Led distributed training, data curation, SFT/PEFT, PPO/DPO alignment, and automated and human evaluation.
- Established an AI curriculum and reading group across 10+ teams, training 300+ engineers.

RESEARCH & ENGINEERING EXPERIENCE

BrainCo

Head of Research | Nov 2021 - Mar 2024

- Led **11 researchers** developing wearable AI and brain-computer interfaces; named inventor on **five provisional patent applications**.
- Led conditional-GAN reconstruction of breathing patterns from **25 Hz wearable PPG**: one-second breathing-curve feedback and **1.47 breaths/min MAE** across ten test subjects.
- Led soft-speech recognition research achieving **35% word error rate** with contact microphones versus **99-100%** with MEMS microphones in tested music/noise conditions across eleven participants.

MIT Media Lab

Postdoctoral Research Scientist (part-time) | Nov 2021 - Jun 2022

- Researched multimodal memory systems and physiological data, building on doctoral work in continuous-capture memory interfaces and self-supervised EEG representations.

Apple

Machine Learning Research Intern | Jun - Sep 2021

- Developed **parameter-efficient transformers** for biosignal event detection and continuous time-series processing.
- Built methods to inspect attention heads and activation patterns in continuous time-series models.

Cortico | Information Design & Data Visualization Intern | Summer 2017
Built a browser WebGL engine for dynamic social-network graphs.

University of British Columbia | Data Scientist & Researcher | Sep 2013 - Aug 2014
Developed a Python library to retrieve, compose and visualize Allen Brain Atlas connectivity data in 3D.

Neo Mohsenvand

EDUCATION

Massachusetts Institute of Technology | PhD, Media Arts and Sciences
Completed 2021; MIT degree record: Feb 2022
Advisor: Pattie Maes. Co-advisors: Tomaso Poggio and Ed Boyden.
Thesis: Classifying and Displaying Brain-Waves through Self-Supervised Learning.
Doctoral research: first author of SeqCLR (2020), contrastive EEG learning evaluated on emotion recognition, sleep staging and abnormal-EEG detection.

University of Oxford | MSc, Mathematical Modelling and Scientific Computing | 2013
Advisor: Jared Tanner. Co-advisor: Mason Porter.
Thesis: Sparse-Signal Recovery and Compressive Sensing for Optical Functional Brain Imaging.

Tehran Polytechnic / Amirkabir University of Technology
BSc, Biomedical Engineering; BSc, Electrical Engineering | 2011
President, IEEE Student Branch.

HONOURS & AWARDS

NTT Data Fellowship | MIT, for Rhizome, a smart memory book (2020).
First place and Koch AI Prize | MIT Grand Hack, for Memoroom, VR for Alzheimer's (2017).
Bill Mitchell Design Award | MIT Media Lab (2014).
Khwarizmi Prize | First in mathematics and physics; third the previous year (2006).
Semifinalist | National mathematics and physics olympiads (2004).

PUBLIC ENGINEERING & INTERACTIVE EDUCATION

- **harnessXray** | Agent observability: streamed request capture, record/replay and persistent checkpoints.
- **Voicebook** | Voice-driven document reader with AI-controlled scrolling and highlighting; browser ONNX inference and optional cloud models.
- **Tissue** | Browser transformer training and interpretability: activation capture and neuron ablations.
- **LangX** | Executable agent workflows, checkpointing and human-approval demonstrations.
- **jaxverse** | Interactive deep-learning book with models that train live in the browser.
- **Embedding Playground** | Hands-on retrieval, classification and clustering experiments with browser inference.

SELECTED PUBLICATIONS

Soft Speech, Loud World: Bone Conduction Microphones Enhance Voice Assistant Interaction
Principal investigator. IEEE International Conference on Consumer Electronics (2024).

Generating Breathing Patterns in Real-Time: Low-Latency Respiratory Phase Tracking From 25 Hz PPG
Principal investigator. HealthyIoT / HealthWear, Springer LNICST 530 (2024).

Contrastive Representation Learning for Electroencephalogram Classification
First author. ML4H NeurIPS Workshop, PMLR 136 (2020).

Optical-flow analysis toolbox for characterization of spatiotemporal dynamics in mesoscale optical imaging of brain activity
Coauthor. NeuroImage 153, pp. 58-74 (2017).

Rumor Gauge: Predicting the Veracity of Rumors on Twitter
Coauthor. ACM Transactions on Knowledge Discovery from Data 11(4), Article 50 (2017).

Spontaneous cortical activity alternates between motifs defined by regional axonal projections
Coauthor. Nature Neuroscience 16, pp. 1426-1435 (2013).

SELECTED MEDIA COVERAGE

MIT Technology Review | Profile of wearable memory research (2018)
Quartz | Mini-documentary: Wearable technology that records memories (2018)
Fast Company | DataVRse: immersive data exploration (2016)