

KARAN KHATAVKAR

AI/ML Engineer · LLM, NLP & Deep Learning

work.karankh@gmail.com | [linkedin/karan-khatavkar](https://www.linkedin.com/in/karan-khatavkar) | [github/karankhatavkar](https://github.com/karankhatavkar) | Pune, India

TECHNICAL SKILLS

Languages	Python, JavaScript (React/Node), SQL, Java, Bash
GenAI & LLM	RAG, LangChain, LLMs (Llama, Mistral, GPT-4), Hugging Face, Ollama, prompt & context engineering, Claude Agent Skills
NLP & Deep Learning	PyTorch, TensorFlow/Keras, Transformers, BERT/BioBERT, CNNs (ResNet), VAEs, attention (CBAM, self-attention), Mixture Density Networks
ML & Data	Scikit-learn, XGBoost, Random Forest, Pandas, NumPy, SHAP, signal processing (CWT scalograms, PSD)
Vector & Databases	FAISS, Chroma, PostgreSQL, MongoDB, MySQL, InfluxDB (time-series)
Cloud & MLOps	Azure (AI-900, DP-900, AZ-900), Docker, Kubernetes, FastAPI, Streamlit, GitHub Actions, Git

EDUCATION

M. S. Ramaiah University of Applied Sciences <i>M.Tech, Artificial Intelligence and Machine Learning; GPA: 9.4/10.0</i>	January 2025 – Present Bengaluru, India
Dr. Vishwanath Karad MIT World Peace University <i>B.Tech, Electronics and Communication; GPA: 8.97/10.0</i>	August 2018 – July 2022 Pune, India

WORK EXPERIENCE

Amdocs Development Centre <i>Software Developer</i>	August 2022 – January 2025 Pune, India
<ul style="list-style-type: none">Built a RAG pipeline for technical-support queries in LangChain; dynamic chunking cut token usage and improved answer accuracy on internal documentation.Fixed context loss in multi-turn chats with a question-rephrasing step that rewrites follow-ups using prior history before FAISS retrieval; held up through UAT.Developed a dashboard to trace retrieval failures and tune BM25 hybrid-search weights, which cut hallucinated answers.	
Sisai Technologies Private Limited <i>Project Intern</i>	March 2021 – June 2021 & October 2021 – May 2022 Pune, India
<ul style="list-style-type: none">Engineered an InfluxDB pipeline ingesting real-time telemetry from 1,000+ industrial sensors.Automated sensor-fault detection with Z-score anomaly detection.Developed a React dashboard with role-based access to track operational data and sensor health.	

PUBLICATIONS

Khatavkar, K., Chindanur, N.

“MixVal: Continuous Embeddings with Probabilistic Decoding for Multimodal Inference.” *Under Review*, 2026.

- Overcame regression limitations in multi-modal systems by engineering **MixVal (Continuous Encoding + MDN)**, which leveraged **Homoscedastic Loss Balancing** to deliver a **≈90% improvement** in latent precision.

PROJECTS

Early Detection of Alzheimer’s Disease (Explainable AI & Deep Learning)

- Converted EEG recordings into time-frequency **scalograms** and benchmarked CNNs with **CBAM, squeeze-and-excitation, and self-attention** against PSD-feature tree models (AD/CN/FTD), using **saliency maps and SHAP** to tie predictions to specific frequency bands.

Medical Named Entity Recognition (BioBERT)

- Fine-tuned **BioBERT** for token-level clinical NER on the **MACCROBAT2018** corpus (~40 entity types, 78 BIO tags), evaluated at entity level (precision/recall/F1) in PyTorch + Hugging Face.

Remaining Useful Life (RUL) Prediction from Acoustic Signals

- Built a **digital-twin simulator** for run-to-failure gunshot audio, then predicted recoil-spring RUL with **XGBoost** (acoustic features), a **2D CNN** (CWT scalograms), and a **1D CNN** (raw audio); dataset published on Hugging Face.

Generative Image Restoration (VAE)

- Trained convolutional **VAEs** to inpaint masked regions of **COCO and CelebA** images, scoring reconstructions with **PSNR/SSIM** and probing the latent neighborhood for learned structure.