Rijo S Lal

Machine Learning Engineer

+91 8848808198 slalrijo2005@gmail.com LinkedIn GitHub

Career Overview

Self-taught Machine Learning Engineer with practical experience across the ML lifecycle, from data preprocessing to model deployment. Proficient in modern machine learning frameworks and tools, with a solid understanding of core algorithms, neural networks, and evaluation techniques. Demonstrated success delivering real-world solutions through personal and open-source projects. Committed to continuous learning and applying ML to impactful, scalable problems

Experience

Freelance Machine Learning Developer *Remote*
Self-Employed *Jun 2024 – Present*

- Delivered ML-based solutions for B2C businesses, including model development and deployment.
- Handled end-to-end workflows: data preprocessing, model training, evaluation, and iteration.
- Communicated with non-technical stakeholders to translate business problems into ML tasks.

ML Consultant *Remote*
Karakonam Medical College (Research Collaboration) *Jan 2025 – Feb 2025*

- Collaborated with medical research students to explore ML applications in diabetic neuropathy studies.
- Provided guidance on data handling, model design, and experimental setup.
- Assisted in interpreting results and aligning technical outcomes with research goals.

Projects

Quizzy *Demo* GitHub

- Developed an AI-driven virtual interviewer that evaluates technical communication skills and emotional responses (confidence, tension, nervousness) through facial expressions. Integrated **RAG** with **Llama3-Groq-70B** for AI conversations, **Llama3.2-2B** for data extraction, and **MaxBai embeddings** for **ATS** estimation. Used a **fine-tuned MobileNet** model for emotion detection and **Mediapipe** for posture analysis. **Transformer**-based models summarize candidate profiles.Built with Django, hosted through **Cloudflared** tunnel, with **MLflow (hosted on DagsHub)** for ML training and tracking. Employed **BeautifulSoup** for LinkedIn job scraping, **EdgeTTS** for text-to-speech, **Whisper** for speech-to-text, and **DVC** for data version control.
- Technologies: Python, Django, MLflow, DagsHub, Llama3-Groq-70B, Llama3.2-2B, MaxBai, MobileNet, Hugging Face Transformers, LangChain, ChromaDB, Mediapipe, EdgeTTS, Whisper, BeautifulSoup, DVC.

Mono-Kit *Demo* GitHub

- Developed a **similarity retrieval library** capable of retrieving similar audio, images, and documents, with user-configurable fine-tuning. Employed **VGGish** as the default audio model, **ResNets** for image embeddings, and custom Siamese network architectures as user-tunable models. Utilized semantic-text-splitter for contextual chunking and **all-MiniLM-L6-v2** for document embeddings. The library is a developer-focused tool that enables developers easy implementation of custom hum-to-search, Google Lens-like functionality, and RAG applications.
- Technologies: TensorFlow, TensorFlow Hub, ChromaDB, Setuptools.

Mini Projects

Corix *Demo* GitHub

- Developed an ML-powered cardiovascular risk prediction service using wearable data (Random Forest) and lab test data (LSTM). Designed as an API service with user authentication and a daily token limit to serve

models. Implemented automated training, evaluation, and model versioning using MLflow (DagsHub), with data tracked via DVC. Built a modular FastAPI service, containerized and deployed on AWS.

- Technologies: Python, FastAPI, AWS (EC2), MLflow, DVC, DagsHub, MongoDB (Atlas), TensorFlow, scikit-learn, Docker.

Sentio *Demo* GitHub

- Developed a real-time sentiment and toxicity classification web application using LSTM-based models built with TensorFlow. Implemented a dual NLP pipeline to classify text as Positive, Negative, or Neutral and detect toxic content using multi-label classification. Integrated preprocessing, tokenization, and sequence handling into an end-to-end inference workflow. Deployed with Streamlit for interactive, real-time input and display
- Technologies: Python, TensorFlow, Streamlit, Pandas, NumPy, Scikit-learn, Pickle-mixin

BrainTumorDetection (MLOPS) *Demo* GitHub

- Built an X-ray image–based brain tumor classification system using a Vision Transformer (ViT) model implemented in TensorFlow. Designed a complete MLOps pipeline with automated training, evaluation, and model tracking using MLflow and DVC, integrated with DagsHub for remote experiment and data management. Exposed the model through a FastAPI application containerized with Docker.
- Technologies: Python, TensorFlow, FastAPI, MLflow, DVC, DagsHub, Docker, OpenCV.

Publications

Beyond Top-K: Contextual Retrieval through Cluster Traversal *Jun 2025 - Jul 2025*
Zenodo

- Investigated the limitations of traditional Top-K vector retrieval and proposed a cluster-based hierarchical retrieval framework leveraging HDBSCAN and representation vectors derived via PCA and Ridge regression.
- Designed and implemented a context-aware retrieval pipeline that models inter-document semantic structure, enabling query-to-cluster matching for targeted search within large-scale corpora.
- Validated the proposed method through empirical evaluation, demonstrating improved retrieval precision (+11% MRR) and efficiency (46% lower latency) over baseline vector similarity search techniques.

Technical Skills

Languages: Python, SQL, C/C++, Javascript

Technologies: TensorFlow, Keras, Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, Hugging Face Transformers, YOLO, OpenCV, NLTK, MLflow, DVC, FastAPI, Streamlit, Django, BeautifulSoup, ChromaDB, MongoDB Atlas, PostgreSQL, Docker, AWS (EC2), Git/GitHub

Education

MVHSS Arumanoor *June 2022 – March 2024*
12th Grade – Science with Computer Science

- Percentage: 91.08%
- **Coursework:** Programming (C++), Algorithms, OS Fundamentals, Database Management (SQL)

Brototype *June 2024 – Jul 2025*
Machine Learning Bootcamp *Certificate*

Certifications

- Machine Learning Specialization — Stanford University (Coursera)
- The Complete Python Developer in 2023 — Udemy
- Python for Machine Learning and Data Science — Udemy