

Rajveer Jadav

+1 437 608 0339 | r2jadav@uwaterloo.ca | rajveerjadav19@gmail.com | LinkedIn | GitHub

Professional Summary

- AI/ML Engineer and Full Stack Developer with experience in designing, developing, and deploying machine learning and generative AI solutions. Proficient in Large Language Models (LLMs), RAG pipelines, LangChain, LangGraph, and end-to-end MLOps workflows using MLflow, DVC, and Weights & Biases. Skilled in Python, TensorFlow, Keras, and OpenCV for AI/ML model development. Experienced in full-stack web development with React, Node.js, Firebase, and AWS, building scalable, secure, and production-ready applications. Strong background in data analysis, software engineering, and CI/CD pipelines, with published research in image/video denoising.

Technical Skills

AI/ML/GenAI: Python, TensorFlow, Keras, CNN, OpenCV, Generative AI, Mistral/Gemini APIs, Large Language Models (LLMs), RAG, LangChain, LangGraph, MCP, Agentic AI, MLflow, DVC, Weights & Biases, Docker, AWS SageMaker/S3

Web Development: React, HTML, CSS, JavaScript, Bootstrap, Streamlit, Node.js, Express, REST APIs, SEO, Selenium, Jest

Databases: SQL, MongoDB, Firebase

Data Analysis & Visualization: Power BI, Tableau, MATLAB, IBM SPSS

Programming & Algorithms: C#, C++, C, Graph Theory, Multi-threading, CMake, Linux

IT/Tools/Platforms: Jira, Bomgar, TeamViewer, GitHub, Colab, VS Code, Jupyter, Railway, MS 365, Zoom

Education

University of Waterloo

Master of Engineering in Electrical & Computer Engineering, GPA: 3.7/4.0

Waterloo, ON

Aug 2024 – Aug 2025

Chennai, India

Sri Sivasubramaniya Nadar College of Engineering

BTech Information Technology, GPA: 8.8/10

Jul 2019 – May 2023

Experience

ML Graduate Researcher

University of Waterloo

May 2025 – Oct 2025

Waterloo, ON

- Led development of Bimodal Multi-Scale Attention Autoencoder (Bi-MSAAE) for facial image/video denoising under the supervision of Prof. Sagar Naik; integrated multi-scale encoder, noise gate, channel attention, and dual-head decoder for RGB/grayscale modalities.
- Achieved PSNR gains of 6.25 dB (to 31.18 dB) on CELEBA-HQ and 3.85 dB (to 28.57 dB) on FER2013; improved SSIM up to 0.92, outperforming FFDNet and FastDVDnet baselines.
- First-authored paper “Bimodal Multi-Scale Attention Autoencoder (Bi-MSAAE) for Facial Image and Video Denoising” submitted to IEEE Access.

Full Stack Developer

Lead-On-Leads Pvt. Ltd.

May 2023 – Jul 2024

Kathmandu, Nepal

- Designed and developed core features for internal and client-facing web application using React.js, Node.js, and Express, improving system usability and performance.
- Implemented and optimized RESTful APIs, improving data flow efficiency and reducing API response times by up to 25%.
- Collaborated with the QA team to debug issues, write test cases, and stabilize releases before deployment.
- Contributed to containerizing services using Docker and assisted in deploying applications on AWS (EC2, S3, and RDS), streamlining the release pipeline.

Projects

AI-Powered Study Assistant | *LangChain, LLM, RAG, ChromaDB, Firebase, Streamlit, Python*

- Created a personal study assistant that simplifies topics, generates quizzes, and answers syllabus questions to help students learn faster.
- Implemented a full RAG pipeline using LangChain, ChromaDB, and Firebase; built Streamlit UI, vector embeddings, syllabus upload flow, and secure real-time retrieval.

RAG Customer Support Chatbot | *FAISS, FastAPI, Docker*

- Designed a support chatbot that handles multi-turn conversations, retrieves ticket history, and provides cited answers like a real customer support agent.
- Integrated workflows using FAISS vector search and FastAPI backend; containerized with Docker and optimized for multi-step reasoning and tool calling.

MLOps Projects | *DVC, AWS, GitHub, MLflow, W&B, TensorFlow, Scikit-learn*

- Built end-to-end ML pipelines to automate dataset ingestion, model training, evaluation, and artifact tracking for reproducible ML experiments.
- Implemented CI/CD-driven pipelines with DVC + AWS S3 for data versioning, MLflow/W&B for experiment tracking, and automated hyperparameter tuning via GitHub workflows.

Financial AI Agent | *Phidata, Mistral, Open-Source LLM*

- Developed an AI agent that helps users understand financial data, answer queries, and generate insights in natural language.
- Used Phidata agent framework with Mistral models to build tool-enabled reasoning workflows that fetch data, parse results, and deliver structured financial insights.

Generative AI Web Suite | *Python-Streamlit, Mistral API, GitHub*

- Built a suite of AI tools including a cover letter generator, code debugger, chatbot, and story generator enhancing the capabilities and practical applications of LLMs.
- Developed Streamlit interfaces integrated with Mistral APIs; deployed using Github and Streamlit Community Cloud.

Service Provider Web-App (Fixify) | *React, Selenium, Jest, Railway*

- Created a full web platform for booking home maintenance services with real-time updates, service search, and user reviews.
- Engineered React frontend with Firebase auth + Firestore, AWS assets, automated Jest unit tests, and Selenium integration tests; deployed via Railway.

Publications

Bimodal Multi-Scale Attention Autoencoder (Bi-MSAAE) for Facial Image and Video Denoising, Rajveer Jadav, Marzia Zaman, Darshana Upadhyay, Kshirasagar Naik. IEEE Access, under review, 28th Oct. 2025.

Certifications

AI for Everyone — Andrew Ng, DeepLearning.AI 

AWS for Beginners — Great Learning 

Programming Through C++ — NPTEL, IIT Bombay 

Python for Machine Learning — Great Learning 

The Arduino Platform & C Programming — University of California, Irvine 