PRIYANSHU SRIVASTAVA

Gurugram - INDIA | +91-6307405607| priyanshuai.dev | priyanshusrivastava@icloud.com

AI/ML Developer | Prompt Engineer

Professional Summary

Results-driven professional with 4+ years in B2B relationship management and channel sales, now focused on AI/ML. Skilled in data-driven decision-making, stakeholder collaboration, and agile project delivery. Completed hands-on projects in generative AI, LLM fine-tuning, and scalable ML pipelines. Seeking a role where I can combine business acumen with technical expertise to deliver impactful AI solutions.

Skilled in LLM fine-tuning **(LoRA/QLoRA)**, Retrieval-Augmented Generation (**RAG**), and multi-agent orchestration using frameworks like **LangGraph, Agno, AutoGen, CrewAI, Hugging Face** and **Ollama**. Experienced in full-stack app development, with a passion for building tools and autonomous systems that address real-world challenges.

CORE SKILLS

PROGRAMMING & MACHINE LEARNING

- **Languages & Libraries**: Python, NumPy, Pandas, scikit-learn, PyTorch, TensorFlow
- **Modeling & Fine-tuning:** Hugging Face Transformers, LoRA/QLoRA, Retrieval-Augmented Generation (RAG), prompt engineering
- **Agentic Systems & Orchestration:** LangGraph, Letta OS, AutoGen, CrewAI, Agno, Model Context Protocol (MCP), LangChain (multi-agent orchestration, tool use)
- **Inference APIs & Models:** OpenAI API, Groq API, Gemini API, Anthropic Claude API, Whisper-large-v3 Turbo, Silero VAD, LiteLLM, Ollama

DATA ENGINEERING & MLOps

- **Pipelines & Workflow Tools:** Apache Airflow, SQL, MLflow
- **Infrastructure & Deployment:** Docker, Docker Compose, Kubernetes, FastAPI, Node.js, Nginx
- **MLOps:** CI/CD for ML, GitHub Actions, model versioning, monitoring/logging, basic observability practices
- **Cloud & Platforms:** GCP, AWS, Azure, Runpod, Railway, Vercel, Render, Netlify, Ubuntu

TOOLS & FRAMEWORKS

- **Dev & Collaboration Tools:** Git, GitHub, Jupyter, Postman
- **APIs & Integration:** REST APIs, Websockets, external API integrations, FastAPI backends

- **Databases & Vector Stores:** PostgreSQL, Supabase (Auth, Realtime), Pinecone, Vector DBs, Prisma ORM
- **Frontend Development:** JavaScript, TypeScript, React, Next.js (App Router), Vite, Tailwind CSS, Chakra UI, shadcn/ui, Zustand (with persistence)

SOFT SKILLS

- Cross-functional collaboration with product, design, and engineering teams
- Project ownership and agile development practices
- Analytical thinking and data-driven problem solving

AI/ML PROJECTS

AI/ML Developer & Prompt Engineer - Freelance | 2023 – Present

Worked on real-world, open-source AI systems and platforms from the ground up from model orchestration to frontend deployment.

Stateful AI Voice Assistant

Self-Hosted Voice AI System with Persistent Memory and Real-Time Agentic Inference

- Built a self-hosted voice agent with memory, **STT, TTS**, and Agentic Orchestration.
- Powered by **Letta OS**, built on top of **LLaMA, Mistral,** and **GPT** models, with **Whisper-large-v3 Turbo** for **STT**, **Groq** and **Gemini** for fast inference, and **Silero VAD** for accurate voice activity detection.
- **Real-time** response cycles via backend trigger on **utterance-complete** detection.

AI-Powered Agentic Startup Team

Simulating an autonomous startup team with specialized AI agents handling marketing, tech, legal, and support tasks.

- Multi-agent orchestration with shared memory and collaborative workflows.
- Built on **MCP**, **Agno**, **Letta OS**, and **E2B** for secure, isolated execution.
- Integrates fast inference models (**OpenAI, Groq, Gemini models**) with tool-calling and real-time coordination.

   Goal: A self-operating AI co-founder system for entrepreneurs.

Restro-fi – AI Restaurant SaaS Platform

AI-powered platform streamlining restaurant operations from table to kitchen with real-time insights and automation.

- - QR-based table mapping for seamless customer order routing.
- - Persistent cart and real-time dashboard for smooth dining experience.
- - Admin panel for dynamic menu and category management.
- - Tech Stack: **React + TypeScript, Tailwind, Zustand, Supabase (PostgreSQL, Auth, Realtime), FastAPI**

WORK EXPERIENCE

Manager Dealer Relationship

Creative Webmedia Pvt. Ltd (Autoportal) - Gurugram | March 2018 – August 2019

As a Manager Dealer Relationship at Autoportal, I am responsible for building and maintaining strong partnerships with automobile dealerships. My role focuses on enhancing dealer engagement, streamlining communication, ensuring seamless inventory updates, and driving mutual growth through strategic collaboration.

- Coordinated onboarding workflows for 140+ dealers, using data analysis to prioritize regions and streamline integration processes.

- Analyzed engagement metrics to identify retention drivers implemented targeted communication strategies that boosted dealer retention by 37%.

- Collaborated with internal teams (product, marketing) to design and iterate upsell programs based on dealer feedback and performance data, driving measurable revenue growth.

Channel Sales Officer

Wheelseye Technology India Pvt. Ltd - Gurugram | November 2019 - March 2022

Handled B2B channel sales for GPS-based fleet management solutions, focusing on partner acquisition, relationship management, and revenue growth.

- Orchestrated partner onboarding at scale documented processes, identified friction points via surveys/data, and automated routine tasks for faster integration.

- Developed training materials and dashboards (e.g., tracking partner performance metrics) to optimize partner enablement, boosting product adoption by 35%.

- Surpassed targets by applying data-driven territory planning and performance monitoring—skills applicable to evaluating ML model metrics and iterating strategies.

EDUCATION

Bachelor of Commerce

  CSJM University , DAV College, Kanpur, UP - INDIA

Graduation Year : March 2010

HOBBIES

Building AI projects and jamming on the guitar. Enjoy combining creativity and logic in every form.

Dated : 15-June-2025   PRIYANSHU SRIVASTAVA