JAIVAL PATEL

jaival.patel602@gmail.com | linkedin.com/in/jaival-patel602 | github.com/jaival0602

EDUCATION

University of Southern California August 2022 - May 2024
Master of Science in Electrical and Computer Engineering (Machine Learning and Data Science) GPA: 3.4/4.0
Coursework - Machine Learning, Cloud Computing, Object-Oriented Programming, Deep Learning, Database Systems

University of Mumbai August 2018 - May 2022
Bachelor of Engineering in Electronics and Telecommunication Engineering GPA: 3.7/4.0
Minors - Artificial Intelligence and Machine Learning (IBM) GPA: 3.9/4.0

WORK EXPERIENCE

Software Engineer | Kopacetic, Inc. | Los Angeles, CA May 2025 - Present
- Built a **multimodal LLM** service for a meeting-intelligence platform that ingests vocal and facial data and auto-annotates, scores behaviors from 0 to 100, streams results to **BigQuery** and is deployed on **GCP Cloud Run**, improving annotation accuracy by 35%
- Led a team of 3 to build an **Agenda-Checker** Natural Language Processing (NLP) pipeline using LLMs plus rule-based checks to extract agenda items and completion signals, and delivered a **Node.js REST API** and a **React** dashboard for agenda tracking
- Designed **GKE Autopilot** to cut cold-start tails and enable fine-grained autoscaling with **Terraform** + **Helm**, Workload Identity, NetworkPolicies, and HPA, and a staging cluster via **Argo CD**

Data Scientist | Gino Corporation | Los Angeles, CA October 2023 - August 2024
- Automated weekly sales and inventory reporting for ops leadership by replacing manual spreadsheets with a **Python** plus **MySQL** program and a scheduled validation, cutting report generation time by 96%
- Built an **Extract, Transform, Load (ETL)** pipeline to migrate 2,000+ products into Odoo with field mapping, deduplication, and pre and post-load checks for reliability, plus batched, incremental loads and indexed queries for scalability
- Developed an inventory-forecasting **Machine Learning (ML)** module on **AWS SageMaker** using **XGBoost** with a custom **ARIMA** baseline, achieving **12% lower WAPE** and **3% bias** than baseline on 2 to 4 week horizons across top SKUs

Software Developer Intern | Pheme Software Pvt. Ltd. | Mumbai, India June 2021 - July 2021
- Engineered a **FastAPI REST** backend in Python for online workspace platform using **SQLAlchemy** and **PostgreSQL** while implementing **JWT** auth, **RBAC**, and input validation, achieving 350 RPS (requests per second) and zero downtime
- Shipped a **React** and **TypeScript** dashboard with **WebSockets** for real-time workflow updates for 300 concurrent users

PROJECTS

Smart Supply Chain Planner – Agentic Logistics System July 2025 - September 2025
- Built a multi-agent AI system using the **Model Context Protocol (MCP)** to automate demand forecasting, routing, and procurement decisions in a simulated logistics network spanning 80+ orders and 2 distribution centers
- Implemented autonomous agents for Demand, Inventory, Procurement with **FastAPI, LangGraph**, **and Python OR-Tools**, deployed via **AWS ECS, Lambda, and RDS** with S3 for data storage and a **Next.js** dashboard for human-in-the-loop approvals
- Delivered a scalable orchestration platform that reduced total route distance by 40%, cut stockouts by 30%, and lowered simulated logistics costs by 20%, while enabling real-time replanning through agentic coordination

Retrieval-Augmented Generation Based Mental Health Chatbot December 2024 - February 2025
- Built a supportive **mental-health chatbot** with a **FastAPI** and **Uvicorn** service integrating **LangChain** and OpenAI to deliver fast, grounded answers, reducing median response time to 1.1s
- Optimized a **RAG pipeline** on **Azure PostgreSQL** with OpenAI embeddings, deployed via **Docker** on Azure App Service, achieving 98.33% unique document retrieval

E-commerce Management System July 2024 - September 2024
- Implemented **RESTful APIs** with **Spring Boot** for product, order, and user management, handling up to 10,000 RPS
- Devised a shopping cart and order system with real-time inventory validation, reducing inventory update errors by 25%
- Designed database models using **Hibernate ORM** and **MySQL** to ensure data consistency and optimized query performance

SKILLS

Languages: Python, C++/C#, HTML/CSS, JavaScript, TypeScript, SQL, Java
Framework and DB: FastAPI, PostgreSQL, MongoDB, Spark, React, Angular, Airflow, SQLAlchemy, Redis, BigQuery
Python Libraries: PyTorch, scikit-learn, TensorFlow, NLTK, JAX, LangChain, NumPy, Pandas, Matplotlib
Tools and Technologies: Tableau, Power BI, Git, GitHub, DVC, AWS, GCP, Jira, MLFlow, Docker, Jenkins, Kubernetes