JASMIN PATEL

jasminpatel.dev@zohomail.in
in/jasminpatel2110 | github.com/Jasmin-Patel-2110 | peerlist.io/jasmin_patel

PROFESSIONAL SUMMARY

Forward-deployed AI/ML Engineer. Build and deploy ML models and LLM pipelines in Python: vision-language models, prompt design, online learning, RAG, FastAPI. Lead developer of AI bookkeeping platform, built on-site at CA firm. Co-authored Scopus-indexed ML research. Fluent in English, Hindi, Gujarati.

TECHNICAL SKILLS

Languages: Python, JavaScript; Rust (basic)
ML & Deep Learning: scikit-learn, River (online learning), Pandas, NumPy, transformers, data preprocessing, model evaluation
LLMs & GenAI: vision-language models (NuExtract3, Qwen-VL), prompt engineering, structured JSON extraction, RAG, embeddings
Math: linear algebra, calculus, probability & statistics
Model Serving & Cloud: vLLM, MLX, Ollama, quantized models, Docker, AWS EC2, RunPod, DigitalOcean
Backend: FastAPI, REST APIs, JWT auth, Redis job queues, Node.js
Databases: PostgreSQL (with extensions PostGIS and pgvector), MongoDB
Tools: Linux, Git, GitHub Actions, Caddy, Postman
AI-Assisted Dev: Claude Code, Codex, OMP (agentic, spec-driven workflows)

EXPERIENCE

AI/ML Engineer **Apr 2026 - Present**
Kriit Technologies Pvt. Ltd. Surat, Gujarat

- Forward-deployed at CA firm. Lead developer of company's AI bookkeeping platform: model-based extraction, prompt design, invoice processing, Tally auto-posting, client-app APIs.
- Moved production GeoJSON API (~37,600 files, 1.2 GB) for Salesforce mapping app from AWS Elastic Beanstalk to DigitalOcean droplet. Byte-identical responses. Set up continuous deployment (CD) from GitHub. Hosting cost cut from ~$40 to $6/month (85%).

Software Engineer Intern **Dec 2025 - Apr 2026**
Kriit Technologies Pvt. Ltd. Surat, Gujarat

- Built v1 invoice extraction and classification API. Turns invoices into structured data. Cuts manual data entry.
- Built FastAPI services. Helped with product decisions, technical research, deployment.
- Worked with Salesforce CRM and cloud products for company integrations.

PROJECTS

AI Bookkeeping Automation for CA Firms (Kriit Technologies) *Python, FastAPI, NuExtract3, MLX, MongoDB, Redis*

- v1: Real-time invoice extraction and transaction classification API. Qwen-VL served with vLLM. Online ML model (River TF-IDF + Adaptive Random Forest) learns from each user correction. Deployed with Docker on AWS EC2, RunPod.
- v2: Rebuilt as on-premise system. NuExtract3 vision-language model (8-bit, MLX) reads bank statements and invoices (PDF, image, CSV, Excel) and returns structured JSON.
- Designed extraction prompts and JSON templates per document type. Tested and tuned on real client documents. Validation checks catch bad extractions before entries post to TallyPrime accounting software.

Waste Collection Management Platform *FastAPI, PostGIS, React, Expo, OSRM, VROOM, Docker*

- Multi-tenant platform to plan, dispatch and live-track waste pickup trips. FastAPI modular monolith (115 API operations), PostgreSQL/PostGIS, row-level security per tenant, 5-job GitHub Actions CI.
- Route optimization with self-hosted OSRM + VROOM. React admin web: live map, trip replay, CSV reports. Offline-first Expo driver app: background GPS, idempotent sync outbox.

Diabetes ML Pipelines Comparison Tool | GitHub | Live *Python, Pandas, scikit-learn, Streamlit, Matplotlib*

- Interactive tool comparing imputation, scaling, model and tuning choices on Pima and Frankfurt diabetes datasets.
- Evaluated 120 pipelines: 2 datasets, 6 imputation methods, 2 scaling methods, 5 ML models. Best accuracy 0.97 (KNN).

Authentication System | GitHub *Node.js, Express.js, JWT, MongoDB Atlas*

- REST APIs: register, email verification, login/logout, password reset. Postman API docs.

EDUCATION

Bachelor of Engineering - Computer Engineering **2022 - 2026**
Government Engineering College, Bharuch (GTU) **CPI**: 7.65/10 | **CGPA**: 7.73/10

PUBLICATION & LEADERSHIP

- Co-authored and presented "Diabetes Diagnosis Using Machine Learning: High-Accuracy Predictive Modelling with Diverse Data Sources and Methods" at Hinweis 3rd International Conference on AI and Data Science (AIDE). Paper indexed in Scopus.
- Led team research pitch at "IDEATHON 2025 - Innovate for Sustainable Future" (GTU Ventures). Ran chess tournament, Nabham 2024.