

ADITYA JINDAL

Bangalore, IN · +91 8529754162 · jaditya8889@gmail.com ·  ·  · 

PROFILE

Full Stack AI Engineer with 4 years of experience building production-grade distributed systems, agentic AI and ML systems. Experienced in fast-paced, high ownership environments; have owned end-to-end product across backend, infra, and AI platform.

TECHNICAL SKILLS

Languages:	Python, TypeScript, Rust, SQL, C++
Backend & Data:	FastAPI, Django, PostgreSQL, Apache Doris, ClickHouse, Redis, Kafka, MongoDB
AI Agents & LLM:	LangGraph, Strands, Langfuse, PyTorch, TensorRT, Triton
Infrastructure & Observability:	AWS, Docker, Kubernetes, OpenTelemetry, Prometheus, Grafana, ElasticSearch
Frontend:	React, Next.js, TanStack Query, Zustand

EXPERIENCE

Simbian.ai | **Senior Software Engineer - AI Agents, Backend & Platform** · Mountain View, CA · Remote · Sep 2025 – Present

- Built a secure, multi-tenant self-service analytics platform on Apache Doris, adopted across 100+ customers, where they use SQL or Ask AI to create custom widgets, dashboards, and reports for complete visibility into the agents and their platform operations.
- Built our own agent observability and monitoring stack on Doris, replacing the Langfuse/ClickHouse deployment.
- Built a vendor-agnostic data service layer for agents across 50+ integrations, processing ~50K alerts/day, 10-50 tools/alert, with integration-health monitoring, a notification system for alerting, and customer-extensible MCP tools.
- Worked on the agent harness on tools, permissioning, provider specific prompt caching, and skills, enabling customer-specific workflows; reduced cost 10x and investigation time 3x with a ~95% prompt-cache hit rate.
- Built multi-tenant AuthN/AuthZ for 10+ enterprise customers: OTP/password login, SAML/Okta SSO, custom hierarchical RBAC, audit logs, database schemas and query optimizations, external API, and security-vulnerability fixes.
- Built a PostgreSQL->Doris CDC pipeline, with monitoring and alerting for data freshness and pipeline failures. Separately cut production PostgreSQL CPU by 30% and reduced critical-query p95/p99 latency by 20x.

Wells Fargo | **Software Engineer — Full Stack AI & Automation** · Bangalore, IN · Jun 2023 – Aug 2025

- Built a GenAI document-processing platform for classification, extraction, human-review, processing 20K documents/day across 100+ businesses and 45K+ document types; replaced Kofax/UiPath and saved \$2M+ in licensing and operating costs.
- Reduced new-document onboarding from ~3 months of bespoke Kofax UI, model training, and creating locators to providing representative documents; improved extraction accuracy from 20% to 90%, with near-100% classification accuracy.
- Built an AI secondary-review workflow for stop-payment disputes, replacing manual QA reducing resolution time and review cost.
- Built 5+ AI/RPA automations, saving \$1M+, including LDAP-based entitlement validation across 100+ groups for 10K users/month.
- Led development of a Records Retention Bot processing 3M+ files across 100+ sites, automating retrieval, archiving, and deletion.

Maya Labs (YC S22) | **Software Engineer — ML & Platform** · Remote · Jun 2022 – Dec 2022

- Built connectors for applications like Gsheet, Gmail, Notion, etc and workflow logic for converting natural language using LLMs into customisable automation workflows (like n8n).
- Built augmentation generators for Flatland to create more complex 2D drawing programs for self-learning model experiments.

Solvio.AI | **Software Engineer Intern — AI/ML** · Bangalore, IN · Sep 2020 – Sep 2021

- Developed MathOCR for converting printed and handwritten math images to LaTeX; built a 10M+ augmented image dataset using NumPy, Matplotlib, and OpenCV and transformer training/evaluation/inference pipeline using GCP/AWS.
- Used TensorRT for model optimisation, Triton Server for high-throughput deployment (5000 req/min on 1x NVIDIA T4), and Prometheus + Grafana for metrics and alerting. Deployed for 5 EdTechs companies.

UVx Technologies | **Software Developer Intern — Embedded AI** · Bangalore, IN · Jun 2020 – Aug 2020

- Built and deployed a real-time mask, temperature, and face recognition system on Raspberry Pi & Jetson Nano; generated 100K+ image dataset from augmented scraped images and trained models.
- Converted model to TensorRT achieving 50 ms latency; used PyQt for UI, MySQL for data storage, and Grafana for logging.

PROJECTS

GameTools · Rust, Python, FastAPI, MongoDB, Redis, AWS, gRPC, TCP/UDP, WebSockets · Dec 2021 – Present

- Reverse-engineered EA's undocumented wire protocol and gRPC schemas from packet captures and binary frames; built a high-performance Rust client for low-latency communication with game servers.
- Built a FastAPI service over the resulting APIs using MongoDB, Redis, and AWS, handling 1M req/day and 50K+ players.
- Automated deployment and management of 100+ VMs with Hyper-V and PowerShell.

Facial Expression Recognition · Brain & Cognitive Society, IIT Kanpur · Python, ML · Apr 2020 – Jul 2020

- Implemented ResNet & U-Net masking blocks inspired model in PyTorch with highest accuracy — 77% on FER2013 dataset.
- Optimised model using TensorRT for low latency, real-time video emotion detection; used LFFD & OpenCV for face detection.

Comparing DNN with Human Representations · Brain & Cognitive Society, IIT Kanpur · ML · Apr 2020 – Jul 2020

- Extracted image features using pre-trained CNN models (VGG16, ResNet) to compare similarity predicted by deep learning models with human representation. Won Best Research Award from SnT Council, IIT Kanpur.
- Designed behavioural experiment in HTML, CSS, and JS to study human similarity judgment between images of different animals.

EDUCATION

Indian Institute of Technology, Kanpur - B.Tech · Jul 2019 – Jul 2023

AWARDS & RECOGNITION

- Spotlight Award — Wells Fargo (2024 & 2025)
- Best Research Award — SnT Council, IIT Kanpur
- JEE Advanced 2019: AIR 3364 among 160K candidates
- JEE Mains 2019: AIR 99.63 percentile among 1.2M candidates