

SATHISH KUMAR V

AI Solutions Architect & Team Lead | AI/ML Engineer

Chennai, India | Open to Remote (UK/US Timezone) | Willing to Relocate
+91 8056770991 | sathishkumar05@gmail.com | LinkedIn | GitHub | Portfolio

PROFESSIONAL SUMMARY

AI Solutions Architect and Team Lead with 4+ years delivering production AI systems for enterprise clients. Lead a team of 7 engineers building document intelligence, RAG pipelines, and agentic AI solutions across PE firms, manufacturing, and SaaS. Own 15 internal AI products including enterprise chat platform and MCP integrations. Key achievement: contract extraction system reduced processing from 260 hours to 26 minutes (99% reduction), cutting costs 15x vs GPT-4 approach. Won 5 client deals through technical pre-sales and solution architecture.

TECHNICAL SKILLS

AI & ML:	LangChain, LangGraph, RAG, Multi-Agent Systems, Agentic Workflows, Azure OpenAI, Claude/Anthropic, AWS Bedrock, PaddleOCR, pgvector, Azure AI Search, Prompt Engineering
Cloud & Infra:	Azure AI Foundry, Microsoft Fabric, AWS (SageMaker, Bedrock), Docker, Kubernetes, Terraform, GitHub Actions
Programming:	Python (Expert), TypeScript (Advanced), FastAPI, React, Node.js, SQL (Expert), C#
Data & LLMops:	dbt, PostgreSQL, Redshift, Power BI, Azure Data Factory, Langfuse, MLflow, LLM Evaluation
Enterprise:	SharePoint, Salesforce, Azure DevOps, Microsoft Graph, Copilot Studio, MCP Protocol, RBAC, GDPR, AI Guardrails

CERTIFICATIONS

Azure AI Engineer Associate (AI-102) – In Progress | AWS Solutions Architect – Planned Q2 2026

PROFESSIONAL EXPERIENCE

AI Solutions Architect & Team Lead <i>JMAN Group · Chennai (Remote UK/US Clients)</i>	Oct 2025 – Present
- Lead Cluster 5 (7 engineers) in AI Centre of Excellence; own 15 internal AI products, architect client solutions, drive pre-sales - Architected enterprise RAG system using Azure OpenAI + pgvector enabling 200+ employees to query documents with RBAC - Designed MCP server extending Microsoft Fabric with Excel export/dashboards; deployed on Azure Container Apps - Led PE firm AI context layer using Claude Enterprise + MCP with citation tracking and audit trail - Won 5 deals through solution architecture; manage 4 concurrent client projects; conducted 16 interviews Products: Enterprise AI Assistant (200+ users), AI Backend Platform, Consulting Agent, Power BI Semantic Assistant	
AI Engineer → Product Lead <i>JMAN Group</i>	Apr 2024 – Oct 2025
Contract Intelligence Platform: Architected extraction pipeline using PaddleOCR + Qwen 7B for 200-page contracts, reducing 260 hours → 26 minutes (99% reduction) - Achieved 85% straight-through automation with confidence-based human routing at 85% threshold - Reduced per-page cost from \$0.03 (GPT-4V) to \$0.002 via hybrid-model approach; grew PoC to 6-month production - Designed AI architecture for PE firm investment workflows; built contract insights chatbot; launched enterprise chat platform	
Data Engineer <i>JMAN Group</i>	Apr 2023 – Apr 2024
- Built medallion architecture platform processing 5M+ records daily using dbt, Redshift, Azure Data Factory - Designed SCD Type 2 across 20+ entities for compliance; reduced data incidents by 40% with quality monitoring	
Software Engineer <i>JMAN Group</i>	Feb 2022 – Apr 2023
Built RESTful APIs using Node.js, NestJS, FastAPI serving 1M+ daily requests. Docker-based CI/CD reducing release cycle by 60%.	
Full Stack Developer <i>Various Companies · Chennai</i>	2017 – 2022
Enterprise systems using C#, .NET, SQL Server for retail and financial services.	

PROJECTS & HACKATHONS

SapienSync Talent Recommendation Engine	AI PoC
- Built AI engine ranking employees for roles using 6-step pipeline: semantic matching, weighted scoring (Skill 40%, Competency 25%, Availability 25%), GPT-4o enrichment, pgvector ANN search - Developed 5 forecasting models using Holt-Winters achieving 4.4% MAPE with P10/P50/P90 confidence bands	
Lead Generation Network UK PE Sales Intelligence	Hackathon 2024

- Built pipeline targeting UK PE firms from Companies House, FCA Register; Medallion Architecture with entity resolution
- GPT-4o-mini enrichment; weighted lead scoring; Power BI dashboards for sales intelligence

EDUCATION

MCA · University of Madras · 2017–2020 | **BCA** · Jairam Arts and Science College · 2014–2017

AWARDS & PUBLICATIONS

- Equity stakeholder (shares awarded) | 2× Hackathon 3rd place | Published: "Contract Intelligence at Scale" — Medium
-

Notice Period: 60–90 days | **Work Authorisation:** Open to UK/US sponsored visa