

SURYA A

Data Scientist | Machine Learning Engineer | AI Specialist

■ arulsurya05@gmail.com | ■ +91-7904578344 | ■ [LinkedIn](#) | ■ [GitHub](#) | [Medium](#) | ■ Chennai, India

PROFESSIONAL SUMMARY

Data Scientist & AI Engineer with 5+ years of experience transforming complex data into actionable ML solutions across automotive and aerospace industries. Expert in statistical modeling, predictive analytics, and production ML deployment using Python, GCP Vertex AI, and modern MLOps practices. Delivered 80+ comprehensive analyses and built scalable AI systems reducing manual effort by 80%+ while improving accuracy by 35%. Finalist in AIXcelerate Hackathon 2025 for innovative automotive AI solution. Published AI thought leader on Medium. Seeking Data Scientist positions at 24+ LPA to drive data-driven innovation.

TECHNICAL SKILLS

AI & Machine Learning: Google Gemini API, GPT-4, Claude, BERT, LLMs, RAG Systems, TensorFlow, PyTorch, Scikit-learn, Hugging Face, LangChain, NLP, Computer Vision, Deep Learning, Prompt Engineering, Fine-tuning

Cloud & MLOps: GCP Vertex AI, BigQuery, Cloud Storage, Docker, Git/GitHub, MLflow, FastAPI, REST APIs, Model Deployment, A/B Testing

Programming & Data: Python (Expert), SQL, JavaScript, Pandas, NumPy, Statistical Analysis, Predictive Modeling, Time Series, Data Mining

Databases & Vector Stores: MySQL, SQLite, FAISS, ChromaDB, Pinecone, Large-scale databases

Domain Expertise: Automotive AI, Election Tech, Healthcare AI, Travel Tech, Master Data Management, PLM/ERP Systems

Methodologies: Agile, Scrum, Responsible AI, Explainable AI, Vibe Coding, Rapid Prototyping

PROFESSIONAL EXPERIENCE

Deputy Manager - AI & Master Data Management

Mahindra & Mahindra Ltd (Automotive Financial Services) | September 2023 - Present

- Led development of **production-grade RAG pipeline** processing **11,000+ automotive part descriptions** with **99.7% cache efficiency** using Google Gemini API, FAISS embeddings, and semantic search—reducing manual effort by **80%** and API costs by **65%**
- Architected **AI-powered cognitive knowledge base "IntelliPart"** achieving **4-167 relevant parts retrieval** from complex multi-criteria searches using advanced semantic search and fuzzy matching algorithms
- Built and deployed **LLM-based document extraction system** processing **10,000+ technical PDFs monthly** with **95%+ accuracy** using Azure Document Intelligence and Gemini API
- Developed **automated data quality framework** with near-duplicate detection and validation, flagging **90%+ inconsistent records** before downstream use
- Implemented **monitoring and feedback loops** for model drift detection, reducing low-confidence outputs from **15% to 3%** through continuous quality scoring and retraining
- Deployed **scalable ML pipelines on GCP Vertex AI** with BigQuery integration for real-time data processing and model serving
- Designed **FastAPI-based REST APIs** for ML model serving, handling **1M+ requests/month** with **<100ms latency**
- Reduced **supplier identification time from 12 months to <3 months** through AI-powered intelligent search supporting BEV transition
- Collaborated with cross-functional teams to translate business requirements into **4+ ML-powered features**, improving decision velocity by **2x**

Senior Data Analyst - Aerospace Engineering

Infosys Limited, Chennai | May 2021 - September 2023

- Completed **80+ comprehensive ML-driven analyses** on team productivity and engineer performance using regression/classification models
- Developed **predictive models** for aircraft engine demand forecasting with **MAPE of 8.5%**, enabling **12% cost reduction**
- Built **Python-based ETL pipelines** (Pandas, NumPy, Scikit-learn) reducing data processing time by **90%** and automating monthly reporting
- Created **Power BI dashboards** for real-time SLA tracking across 200+ engineers; developed **NLP tool** reducing review time by **70%**
- Mentored 2 junior analysts on Python and ML fundamentals

NOTABLE AI PROJECTS & OPEN SOURCE

1. **AI-Powered Mental Health Support Chatbot** | Oct 2025 | Tech: Python, Gemini API, NLP

- Developed conversational AI chatbot with sentiment analysis and context-aware crisis detection using Google Gemini API and privacy-first architecture

2. Tamil Nadu 2026 Election Fact-Checker Platform | Oct 2025 | Tech: JavaScript, BERT, Computer Vision

- Architected AI fact-checking platform with BERT-based NLP (92% accuracy) and deepfake detection for combating election misinformation

3. SmartStay AI - Intelligent Travel Booking Platform | Sep 2025 | Tech: JavaScript, LLM Integration

- Built AI-powered travel optimization engine with NLP query processing, saving 30%+ time and 20%+ cost for business travelers

4. IntelliPart 1.5 - Automotive Parts Intelligence | Jul 2025 | Tech: Python, Flask, NLP, Vector DB

- Enhanced AIXcelerate Hackathon finalist project with conversational search across 100K+ components, production-scale caching and real-time API

5. 50-Day AI Challenge & 5 Types of Quantized LLMs | Sep 2025 | Tech: Python, Quantization (INT8, INT4, GPTQ)

- Completed intensive 50-day AI program building 10+ GenAI/RAG projects; implemented 5 quantization techniques achieving 4x memory reduction at 95%+ accuracy

THOUGHT LEADERSHIP & PUBLICATIONS

Medium Publications | @arulsurya05 | 5 Followers

- "From Acceptable to Deterministic: Why Architectural Clarity Matters in AI-Powered Systems" (Aug 2025)
- "From Analyst to Applied AI Engineer: How Sri Nithya's 50-Day AI Challenge Accelerated My Career" (Sep 2025)
- "Is OpenAI in Decline? Examining the GPT-5 Backlash, Talent Wars, and Rise of Claude" (Aug 2025)
- "AI's Pivotal Moment: Comparing Today's Revolution to Nuclear and Space Eras" (Aug 2025)

CERTIFICATIONS & ACHIEVEMENTS

■ **Finalist - AIXcelerate Hackathon 2025** | Mahindra.AI & Data Community - Recognized for 'IntelliPart' AI innovation

■ **Dale Carnegie Communication & Leadership Program** | 2025

■ **Mahindra Internal AI Training** | 2025 - Responsible AI, XAI, GCP Vertex AI

■ **Google Advanced Data Analytics Professional Certificate** | **GCP Vertex AI Training (Internal, 2025)**

■ Infosys Certified Data Professional | Python Associate | Power BI Developer | Scrum Foundation (SFPC™) | Kaggle Data Viz & SQL

EDUCATION & CONTINUOUS LEARNING

B.E. Aeronautical Engineering | KCG College of Technology, Chennai | 2016 - 2020

Intern at HAL Bangalore (May-Aug 2019), NCC Participation, CII TNGIM 2019

Current AI Focus: GPT-4 Turbo, Claude 3.5, Gemini 2.0 • MoE, State Space Models • GCP Vertex AI deployment • Vibe coding with GitHub Copilot • Quantization (GGUF, AWQ, GPTQ)

Expected Salary: 24+ LPA | **Notice Period:** Less than 30 days | **Relocation:** Open for right opportunity

Key Differentiators: Production AI Experience (11K+ items processed) • End-to-End ML Ownership • Domain Expertise (Automotive + AI) • Rapid Learning (Following AI evolution real-time) • Thought Leadership (Medium writer) • Vibe Coding (3x productivity) • Business Impact Focus