

# Sunni Kumar

AI/ML Engineer | Full-Stack Expertise | IIT Madras

📍 Bengaluru, KA, IN | ✉ [sunnikumar.arya@gmail.com](mailto:sunnikumar.arya@gmail.com) | ☎ +91 6394540805

🌐 [sunni-kumar](#) | 🔄 [sunny-arya-codes](#) | 🔗 [sunni-kumar.vercel.app](#)

## Summary

AI/ML Engineer with a proven track record of deploying production-grade computer vision and generative AI solutions. Expertise in building scalable backend systems (Spring Boot, Python) and currently pursuing a B.S. in Data Science from IIT Madras.

## Experience

### Freshmenu

Sep 2023 – Present

Software Engineer

- **Reduced P95 latency by 40%** for core transaction APIs by optimizing Spring Boot data access patterns and query logic, directly supporting a **50% increase in daily order volume** without requiring additional infrastructure..
- **Owned the end-to-end integration of third-party delivery platforms** (Zomato, Swiggy, etc.) via REST APIs, architecting a unified data transformation layer that increased daily order volume by **50%** and expanded market reach..
- **Architected and delivered a full-stack HR automation platform** (Next.js, Spring Boot, MySQL) that eliminated **60%** of manual processing tasks, saving an estimated 120 HR man-hours per month and improving data accuracy..
- **Enhanced scalability and reliability** by deploying applications on **AWS (EC2, S3)**, ensuring **99.9%** uptime and production-readiness.

## Education

### IIT Madras

May 2022 – Apr 2026 (Expected)

B.S. in Data Science

- **Coursework:** DBMS, DSA, Software Engineering, Machine Learning, Deep Learning, Computer Vision
- **Projects:** Grocery APP(Flask, Vue.js, SQLite), AI powered LMS(FastAPI, Vue.js, Gemini, PostgreSQL), CrimeCast-Forecasting(ML)

## Projects

### Smart Product Listing Tool (Computer Vision + ML Deployment)

[GitHub Repo](#)

- **Tech Stack:** Python, TensorFlow, OpenCV, FastAPI, Docker, AWS.
- **Engineered and deployed a scalable inference API using FastAPI and Docker**, enabling real-time metadata generation in a production environment and reducing manual data entry time by over **90%**.

### AI-Powered Resume & Job Match Assistant (Generative AI + NLP)

[GitHub Repo](#)

- **Tech Stack:** Next.js, FastAPI, Python, LangChain, FAISS, Gemini, PostgreSQL, Docker, AWS EC2..
- **Developed an LLM-driven recruitment tool using LangChain and FAISS vector embeddings** to parse and match resumes against job descriptions, achieving **85% accuracy** in skill-to-requirement mapping during validation tests. Deployed the full-stack application on AWS EC2, demonstrating end-to-end system ownership.

## Skills

**AI/ML:** Python, TensorFlow, PyTorch, Scikit-learn, Pandas, NumPy, OpenCV, LangChain, Gen AI, Machine Learning, DL.

**Backend & Deployment:** Java, Spring Boot, FastAPI, Flask, Node.js, Microservices, REST APIs, Docker, AWS.

**Databases:** PostgreSQL, MySQL, MongoDB, FAISS.

**Frontend:** JavaScript/TypeScript, React.js, Next.js, Vue.js.