Samuel Joshua

Bangalore, India | samuel.21ai25@gmail.com | +91 9880144285 | ksj047.github.io

linkedin.com/in/samueljoshua47

Education

Visvesvaraya Technological University, Bachelor of Engineering (AI & ML) – Karnataka, India — Expected May 2025

- Major in Artificial Intelligence and Machine Learning with 8.0/10 GPA, ranking in top 15% of cohort
- Completed advanced coursework in Python, Deep Learning, Django, and ML Algorithms with 90%+ scores
- Earned 8+ industry certifications including Google Python specialization and IIT Bombay robotics program

Work Experience

Artificial Intelligence Engineer, Es Magico AI Studio – Hybrid, Bangalore — May 2025 – Currently working

- Architected end-to-end Gen AI workflows using N8N, reducing video generation pipeline time by 40% and enabling automated multimedia content creation for 500+ daily user requests
- Developed Multi-Agent System using Google ADK that improved video generation accuracy by 35% as measured by user satisfaction scores, deployed in production serving active users
- Built Churn Prediction ML model for a Pharma Products which increased profits by 20% for the pharma company.

Artificial Intelligence Intern, Es Magico AI Studio – Hybrid, Bangalore — Feb 2025 – May 2025

- Built 5+ POCs for online interview cheating detection using pre-trained vision models and multimodal LLMs, achieving 92% accuracy in fraud detection and reducing false positives by 60%
- Implemented secure OAuth2.0 flows for Google and LinkedIn integration, supporting 1000+ concurrent users with zero security incidents and 15% improvement in user onboarding completion rates
- Deployed asynchronous Celery task queues, eliminating API call blocking and improving response time by 70% for real-time model inference serving 500+ concurrent requests

Software Developer Intern, Es Magico AI Studio – Hybrid, Bangalore — Sep 2024 – Dec 2024

- Engineered RAG chatbots processing 10,000+ vehicle manual and banking document queries, achieving 89% answer accuracy and reducing customer support tickets by 45%
- Implemented Neo4j graph database with semantic chunking for 50GB+ document corpus, improving retrieval precision by 32% and reducing query response time from 3.2s to 1.1s
- Developed scalable multipart video upload system for Azure Blob Storage handling 500MB+ files, supporting AI interviewer platform with 95% upload success rate
- Contributed to Django-based ML application serving 2000+ daily API calls with 99.5% availability and sub-200ms response times

Machine Learning Intern, Unified Mentor – Remote, Bangalore — Jun 2024 – Sep 2024

- Developed ML models achieving 94% accuracy in thyroid cancer recurrence prediction, 87% precision in fraud transaction detection, and 12% MAPE in vehicle price prediction using XGBoost and ensemble methods
- Processed 100K+ medical records and 500K+ financial transactions, implementing advanced feature engineering that improved model performance by 18% over baseline approaches
- Deployed production-ready models with comprehensive evaluation metrics, reducing prediction inference time by 55% through model optimization techniques

Python & ML Development Intern, Karunadu Technologies Private Limited – Bangalore — Oct 2023 and Nov 2023

- Completed intensive 2-month training program covering Python, Data Science, Django, DBMS, Power BI, and ML algorithms with 95% assessment scores
- Built end-to-end ML application for galactic object classification using Decision Trees on Django framework,

achieving 91% classification accuracy on 5000+ astronomical data points

- Mastered production-level Python technologies and database management, contributing to team productivity improvements of 20%

Projects

Web Research Agent (HuggingFace Deployed) May 2025

- Developed autonomous research agent using Tavily API integration, processing 1000+ web queries daily with 85% relevant result accuracy and generating comprehensive reports 60% faster than manual research
- Implemented iterative LLM evaluation pipeline that improved information synthesis quality by 42% through refined query generation and multi-source validation
- Deployed production system handling 200+ concurrent users with sub-3-second response times and 99% uptime

Privify AI - Privacy-Focused Facial Recognition System Feb 2025

- Final year capstone project implementing FaceNet architecture with Neo4j graph database, achieving 96.5% face recognition accuracy while maintaining complete data privacy for 10K+ user profiles
- Optimized graph-based relationship modeling reducing query complexity by 40% and enabling real-time recognition processing under 500ms per frame
- Implemented privacy-preserving techniques ensuring zero data leakage with encrypted feature storage and secure authentication protocols

Enterprise RAG Chatbot for Vehicle & Banking Documents Oct 2024

- Processed 2000+ pages of Bajaj motorcycle manuals and banking documents, creating optimized vector embeddings with 91% semantic similarity accuracy and reducing manual document search time by 80%
- Architected graph-based chunk storage system enabling contextual retrieval with 35% better relevance scores compared to traditional vector-only approaches
- Deployed scalable solution handling 500+ queries per hour with average response accuracy of 87% validated against subject matter experts

Production ML Model Suite Jul-Sep 2024

- Vehicle Price Prediction: Achieved 89% accuracy ($R^2 = 0.89$) using ensemble methods on 50K+ vehicle records, reducing price estimation error by 23% compared to traditional appraisal methods
- Fraud Detection System: Built classification pipeline processing 100K+ transactions daily with 95% precision, 87% recall, reducing financial fraud losses.
- Medical Prediction Model: Developed thyroid cancer recurrence classifier with 94% sensitivity, potentially enabling early intervention for 200+ high-risk patients monthly

Hand Gesture-Controlled Stock Analysis Platform Jul 2024

- Built computer vision system using PyAutoGUI and OpenCV achieving 92% gesture recognition accuracy, enabling hands-free navigation of financial dashboards with 15+ gesture commands
- Integrated real-time stock data from yfinance API processing 500+ securities, creating automated Tableau visualizations updated every 30 seconds
- Reduced analyst workflow time by 40% through gesture-based interface, supporting seamless data exploration for portfolio management decisions

Technical Skills & Achievements

Programming & AI/ML: Python (1+ years, expert-level), scikit-learn, XGBoost, TensorFlow, Keras for end-to-end ML pipelines; LLM integration (GPT, Gemini), multimodal AI, computer vision with 90%+ model accuracies across projects

Data & Infrastructure: SQL, Neo4j CypherQL, Azure Blob Storage, vector databases; Processed 1M+ records across projects; Django backend serving 10K+ users; Git/GitHub with 563+ contributions

Production Systems: Deployed 8+ ML models in production; N8N automation workflows; Celery task queues; 99.5% uptime achievement; OAuth2.0 security implementation

Certifications: Google Python Specialist, IIT Bombay E-Yantra, 8+ industry certifications in AI/ML technologies