Rajagopal Somaskandan

Chennai, India | +91 7200288242 | rajagopal-somaskandan.netlify.app | somaskandan931@gmail.com |
https://linkedin.com/in/somaskandan-rajagopal | github.com/Somaskandan931

PROFESSIONAL SUMMARY

Applied Machine Learning Engineer specializing in explainable ML and NLP systems, with experience building production-ready fraud detection, semantic search, and ranking pipelines using XGBoost, LightGBM, and transformer-based models. Strong focus on model evaluation, SHAP-based explainability, and FastAPI deployment for real-world business use cases.

TECHNICAL SKILLS

Programming & Databases : Python, Java, SQL (PostgreSQL), MongoDB
Machine Learning : Supervised and Unsupervised Learning, Feature Engineering, Model Training, Hyperparameter Tuning, Evaluation Metrics
ML Algorithms : Linear Regression, Logistic Regression, Decision Trees, Random Forest, XGBoost, LightGBM
NLP & Generative AI : Transformers, spaCy, NLTK, Sentence Embeddings, Semantic Search, Retrieval-Augmented Generation (RAG), Prompt Engineering, LangChain, ChromaDB
Data Analysis & Visualization : Pandas, NumPy, Matplotlib, Plotly, Streamlit, Power BI
APIs, Deployment & Tools : FastAPI, Flask, REST APIs, Docker, Git/GitHub, MLflow, AWS (Basics)

EDUCATION

Sathyabama Institute of Science and Technology	**B.Sc. Computer Science (AI Specialization)**	First Class	*2025*
Sathyabama Institute of Science and Technology	**M.Sc. Computer Science**		*2027*

PROJECTS

FinGuard Pro: Explainable AI for Financial Fraud Detection | *Python, SQL, XGBoost, SHAP, FastAPI*   *2025*
- Architected an end-to-end fraud detection pipeline on highly imbalanced UPI-style transaction data, achieving ***94% recall and ROC-AUC 0.995*** .
- Integrated SHAP-based, transaction-level explainability to support auditability, regulatory compliance, and analyst trust.
- Reduced false positives by combining ML predictions with rule-based AML heuristics (velocity abuse and structuring patterns).

CareerGenie: RAG-based Resume–Job Matching & Skill Gap Analysis | *Sentence-Transformers, ChromaDB, LLMs, LaTeX* 2026
- Designed a Retrieval-Augmented Generation (RAG) pipeline that semantically matches resumes to job descriptions beyond keyword-based ATS systems.
- Implemented embedding-based indexing with Sentence-Transformers and ChromaDB, enabling query-, role-, and location-aware filtering.
- Generated AI-driven skill gap analysis and automated ATS-ready resumes by converting structured profiles into LaTeX via a modular resume generator (Re-Boost).

PaperMind: AI-Powered Academic Paper Recommender | *OpenAI Embeddings, FAISS, FastAPI, React*   *2026*
- Built a full-stack semantic paper recommendation system using dense embeddings and FAISS for high-precision vector similarity search.
- Developed a FastAPI backend supporting scalable indexing and LLM-based relevance explanations for retrieved papers.
- Implemented a React frontend enabling topic-based search and PDF uploads for semantic matching against academic abstracts.

S.A.F.E: Real-Time Crowd Anomaly & Risk Detection System | *OpenCV, Optical Flow, Isolation Forest, Streamlit*   *2026*
- Developed a privacy-preserving, real-time crowd risk monitoring system using optical flow–based motion signals instead of individual tracking.
- Implemented zone-based (2×3) spatio-temporal anomaly detection to identify panic signatures, congestion, and abnormal flow patterns.
- Incorporated camera shake detection via global motion analysis to suppress false positives from unstable video feeds and improve alert reliability.

SourceUp: Explainable Procurement Intelligence Platform | *Sentence-BERT, FAISS, LightGBM, FastAPI, Redis*   *2026*
- Built a semantic supplier discovery engine using Sentence-BERT embeddings and FAISS for intelligent procurement search.

- Improved ranking relevance with a LightGBM learning-to-rank model achieving **NDCG@10 0.7439 (+9.2%)** .
- Enabled low-latency retrieval for 10k+ suppliers via FastAPI microservices and Redis-based caching.

EXPERIENCE

Data Analyst Intern | *India Advocacy* *July 2024 – October 2024*
- Cleaned, filtered, and structured **over 1M web records** and curated 15–20k high-quality datasets to support targeted outreach campaigns.
- Built automated data extraction pipelines using Java and Selenium to improve data collection efficiency.
- Worked with real-world data quality issues, implementing validation checks aligned with business usability requirements.

Security Trooper (National Service) | *Singapore Armed Forces* *January 2020 – January 2022*
- Supported security operations, operated authorized transport vehicles under regulated operational environments.
- Assisted Ministry of Health COVID-19 task force initiatives through structured data collection and reporting.

Sales Associate (Part-Time) | *Bata* *February 2022 – August 2022*
- Worked in a high-volume retail environment, managing customer interactions and daily operational responsibilities with strong communication, time management, and teamwork skills.

PUBLICATIONS & CONFERENCES

FinGuard Pro: Explainable AI for Financial Fraud Detection | *ICAAIC 2025* *ISBN: 979-8-3315-6587-9*
Real-Time Explainable Fake News Detection using BERT-LSTM | *ICASET 2025*

CERTIFICATIONS

AI and Machine Learning | *SpaceZee Technologies* *June 2024*
Python Core | *SpaceZee Technologies* *February 2024*