SREEKANTH GOPI
\lanta Metropolitan Area, GA 30062 ▪ 646-875-3366 ▪ sree0912555@gmail.com
www.linkedin.com/in/s-gopi

OBJECTIVE

Advanced Data Scientist / Applied AI / Forward Deployed Engineer / AI-ML Engineer: Agentic Systems & Production AI Platforms specializing in LLM systems, retrieval-augmented generation (RAG), agentic AI workflows, and applied machine learning. I focus on designing evaluation-driven, production-ready AI systems that integrate scalable retrieval pipelines, semantic memory, re-ranking, and context-engineering. I bring peer-reviewed research, enterprise deployment experience, and hands-on ML delivery to build AI products that are reliable, explainable, and performance-measured in real-world environments.

CAREER HIGHLIGHT

Extensive expertise in AI, ML, and data science with 5+ years of global experience across analytics, solutions architecture, and fintech project delivery. Recognized as Top Prompt Engineer & AI Innovator of the Year 2025 (IAOTP) and Global Top 30 Trailblazer, with wins at the Intel Transatlantic Hackathon (1st, defect detection) and Llama-2 Hackathon (2nd, financial AI). Published and presented multiple peer-reviewed AI research works in IEEE, ASEE, and AIMC across LLMs, EEG/HRV, agentic AI, and generative music. Designed agent-oriented workflows and context-engineering strategies to ensure reliability, evaluation rigor, and knowledge continuity in real-world production environments. Designed and deployed production-grade AI platforms integrating agentic workflows, retrieval-augmented reasoning, and evaluation harnesses, delivering measurable operational improvements across enterprise systems. Experienced in sensor-driven, time-series–heavy data (EEG/HRV, real-time signals) directly transferable to industrial IoT, asset monitoring, and autonomous operations.

EDUCATION

Kennesaw State University, Marietta, GA — Jan 2022 - Dec 2023
Master of Science in Computer Science | Thesis Model | Artificial Intelligence — GPA: 3.80

EDHEC Business School, Nice, France — Aug 2001 - May 2002
Master of Business Administration: Information Technology and Entrepreneurship — GPA: 3.12

Bharathiar University, Coimbatore, India — June 1995 - March 1999
Bachelor of Engineering in Mechanical Engineering — GPA: 3.22

SKILLS

Programming: Python, Java, SQL
Mathematics: Statistical methods, distributions, regression, maximum likelihood estimators, calculus, linear algebra
ML Algorithms: Neural Networks, Decision Trees, Random Forests, XGBoost, Nearest Neighbors, Naive Bayes, SVM, Decision Forests
ML Libraries: Scikit-learn, Keras, TensorFlow, PyTorch, PySpark
AI Platforms: LangChain, LlamaIndex, Semantic Kernel, ClarifAI, AssemblyAI, MCP, A2A, OpenAI libraries
Agentic AI: Agent workflows, n8n, Langflow, Flowise, Copilot, Railway automations
Vector Search: Pinecone, ChromaDB, Vectara, Weaviate, Qdrant, Astra
Generative AI: GPT-4, Gemini, Claude, DeepSeek, zero-shot prompting, RAG workflows, fine-tuned LLM usage
LLM Frameworks: Hugging Face Transformers, DeepSpeed, LoRA, PEFT, HF Accelerate, Weights & Biases, OpenLLM, Ray, RLHF frameworks
Deep Learning: LSTM, image recognition, video processing, face recognition, health signal data
Data Science & ML Systems: Time-series modeling, anomaly detection, signal processing, sensor data analytics, supervised/unsupervised learning, reinforcement learning foundations
Cloud & Dev: AWS, Google Cloud, Azure, Railway, Vercel, n8n, Linux environments
Personalization: User modeling, prompt optimization, AI-driven personalization, GenAI recommendation systems
Prototyping: ReactJS, Roo Code, Trae, Bolt.new, Lovable, HTML/CSS/JS
Search & Ranking: Scalable LLM search, recommendation, ranking optimization
Research & Writing: Paper publishing, philosophy, psychology, cognitive science, conversational AI design

PROFESSIONAL EXPERIENCE

Morgan Stanley July 1st 2024 - present

Senior Prompt & AI/ML Engineer | RAG & Agentic AI | LLM Evaluation & Retrieval Optimization | Atlanta, GA

- **Enterprise AI Delivery**: Led prompt engineering, retrieval design, and agent-oriented AI programs inside a highly regulated global investment bank, embedding governance, auditability, and reliability into all deployments.
- **Confidential Research Execution**: Operated within the bank's restricted Research division, collaborating end-to-end across Research, Development, QA, Architecture, Risk, and Production to ship compliant, production-grade AI systems.
- **Cross-Functional Teamwork**: Worked closely with global data science, ML engineering, architecture, and product teams in structured Agile environments, aligning technical design with formal change-control standards.
- **Technical Communication Leadership**: Delivered clear technical briefings and design walkthroughs 3–4 times weekly to engineering and leadership forums, covering retrieval strategies, evaluation results, and production readiness.
- **Entity Prediction & Classification**: Entity Prediction & Signal Extraction: Developed a near-real-time ML entity prediction system mapping user queries to companies, tickers, and industry entities using synonym expansion and learned query patterns, enabling precise signal extraction for downstream AI workflows.
- **Agentic AI Architecture:** Designed and deployed multi-agent AI systems coordinating retrieval, reasoning, ranking, and decision logic, improving system-level accuracy and reliability across large-scale, production environments.
- **Production AI Engineering & MLOps (Applied):** Implemented model and prompt versioning, evaluation pipelines, and CI-based deployment workflows for LLM and agentic systems, using MLflow-based observability, controlled experiments, and QA→prod promotion via Linux environments and pull-request–driven releases.
- **Operational Reliability (AI Systems):** Ensured regression stability, deterministic behavior, and traceability across agentic and retrieval-augmented systems through structured testing, prompt evaluation harnesses, and monitored production rollouts in regulated enterprise and startup environments.
- **Production-Ready Architecture**: Built indexing, vector storage, re-ranking, and entity-extraction pipelines aligned with enterprise performance, monitoring, and resiliency requirements used across financial research platforms.
- **Structured Evaluation Rigor**: Developed controlled evaluation harnesses for prompt-conditioning and retrieval benchmarking, ensuring deterministic behavior, explainability, and regression stability at scale.
- **Division-Level Recognition**: Selected to showcase applied GenAI work at the firm's internal Technology Expo, representing the Research division and demonstrating measurable business impact and secure design.
- **Quantified Outcomes**: Delivered **95%+ structured-prompt accuracy**, improved chatbot response relevance **from ~55%→80%**, and strengthened retrieval reliability across millions of user interactions.

NeuroHeart.AI - Published Research Papers App Dec 1st 2023 - Dec 2025

AI Architect - Full Stack

- **End-to-End Delivery:** Designed and implemented an end-to-end conversational AI system using Copilot/Codex inside VS Code, driving rapid iteration from architecture design to production-ready code.
- **Backend Architecture:** Built and owned the backend platform using FastAPI, including RAG pipelines, agent orchestration, vector search, session memory, and persistent data storage.
- **Frontend Integration:** Developed and integrated a React-based frontend, enabling real-time conversational interaction, session continuity, and personalized AI-driven user flows.
- **Deployment & Agents:** Deployed the complete system on the Lovable platform, fully coded via vibe-coding and refined with Codex, and integrated external AI agents for live meditation audio and generative music synthesis.

TalentGenius, FL Feb 2024 - June 2024

GenAI DataScientist

- **GenAI Platform Delivery**: Led end-to-end development of generative-AI capabilities for a career-management platform, prioritizing grounded LLM workflows over generic ML pipelines.
- **Prompt & Context Engineering**: Designed structured prompt frameworks and retrieval-aware conditioning across multiple foundation models to ensure stable, explainable responses.
- **RAG System Design**: Built ingestion, chunking, embedding, and Postgres/PGVector–based retrieval with re-ranking to improve profile-to-job semantic matching.

- **Evaluation & Model Analysis**: Benchmarked multiple LLMs for cost, latency, grounding fidelity, and task accuracy, creating repeatable evaluation harnesses.
- **Cross-Functional Delivery**: Collaborated with Engineering and Product in Agile sprints to scope, deploy, and monitor AI features in production environments.
- **Data & SQL Engineering**: Authored complex Postgres queries and weighted-similarity scoring logic supporting real-time recommendations.

Kennesaw State University, Marietta, GA — Jan 2022 - Dec 2023
Data Scientist GRA – AI in Education Technology — Computer Science department

- **Conversational AI Systems**: Designed LLM-driven conversational tutoring and advising assistants focused on prompt-first, retrieval-grounded interactions rather than generic ML pipelines.
- **Prompt-Pattern Research**: Developed pedagogical prompt and dialog-flow patterns to support reasoning, guidance, and explanation consistency in educational use-cases.
- **RAG-Based Recommendation Systems**: Built course-recommendation and advising assistants using retrieval augmentation, semantic search, and structured prompt conditioning.
- **Quiz & Content Generation**: Created GPT-4 powered interactive quiz and tutorial generators enabling personalized assessment experiences for STEM learners.
- **End-to-End System Engineering**: Implemented full-stack platforms including Python back-end services, MongoDB chat-memory persistence, Qdrant vector retrieval, and a TypeScript/React UI.
- **Research Contribution**: Supported publication-aligned experimentation, benchmarking, and documentation across LLM evaluation and educational AI design.

Tensonite Pvt Ltd, India — Nov 2016 – July 2019
Data Analyst

- Research market trends, competitors, and business opportunities.
- Analyze data for estimation, quoting, and contract preparation.
- Predict sales figures and raw material needs using data analytics.

Verseidag-Indutex GmbH, Krefeld Germany. — March 2014 - Oct 2016
Country Regional Manager - Data Analyst

- Spearheaded a data analytics initiative for a tensile textile manufacturing company, focusing on analyzing market trends and customer behavior to enhance marketing and sales strategies.
- Conducted detailed statistical analysis of sales data to identify market opportunities, leading to optimized marketing campaigns and a notable increase in sales revenue.

TAIYO Middle East LLC, / Fiobco Factory LLC, Dubai, U.A.E. — March 2003 – December 2011
Product / Project Manager

- Led design engineering team by using strong problem-solving, analytical reasoning, and critical thinking skills to execute structural analysis using TechNet software.
- Data analytics for sales and market analysis and prediction.
- Closed deals, signed construction contracts, and exceeded company sales and marketing targets with a total sales volume of over $50 million for large stadium roofs spanning over 1.5 million sqft.

AI / ML Peer-reviewed Published papers / Journals:

Time-Series & Signal Intelligence: Designed and deployed neural and ensemble models on high-frequency sensor data (EEG/HRV), including feature extraction, anomaly detection, and predictive modeling—methods directly applicable to industrial sensors, asset health monitoring, and predictive maintenance.
EEG Modeling: Implemented KNN, Neural Networks, Random Forests, and XGBoost on EEG datasets to enhance ADHD prediction accuracy to 95% while analyzing EEG/HRV signals from 15 in-person study participants
Raga LSTM: Generative AI LSTM Music Research: Built multi-track Indian classical music generation models using LSTM/Transformer architectures; presented at INFORMS 2023 (Denver) and published in AIMC 2023
RMMM Model: Designed the **RMMM** framework to generate layered, raga-structured compositions with enhanced authenticity and emotional depth
Agentic Learning: Adaptive Learning & Agentic AI: Designed long-context RAG pipelines and agent workflows to improve adaptive MCQ generation and learning relevance (FIE 2024 & 2025)

Meditation Science: Stress Reduction via Meditation: Analyzed EEG/HRV signals to quantify physiological effects of meditation-based interventions (ASEE 2024, MSCS Thesis)
Health AI: Real-Time Health Monitoring: Built rPPG + LLM feedback system for live physiological signal interpretation and guidance (ASEE 2024)

HACKATHON WINS:
Intel Winner: Intel Transatlantic Hackathon: Led a 6-member team to build a deployable AI defect-detection solution using computer vision and edge-deployable models, reducing scrap by 30% (1st Prize)
Llama-2 Runner-Up: Clarifai Llama-2 Hackathon: Built an AI-driven Telegram bot offering financial guidance through structured reasoning workflows and multimodal insights (2nd Prize out of 1700+ participants)

Github Repository: https://github.com/datasci888