# SARVESH **GANESAN**

India | +91 8668063705 | sarveshganesan2002@gmail.com | **WWW:** linkedin.com/in/sarvesh-ganesan09 | **WWW:** github.com/Sarvesh-GanesanW | **WWW:** sarvesh-ganesanw.github.io/sarvesh-portfolio/

## Professional Summary

AI Engineer with over 2.5 years of experience building production-grade generative AI systems, including autonomous agents and RAG pipelines. Strong expertise in Python, AWS, LangChain, vector databases, and scalable API development. Proven ability to design and optimize cloud-native data and AI platforms, with a focus on low-latency analytics and performance optimization. Currently, we are building a unified, cloud-native data analytics platform with embedded generative AI capabilities.

## Technical Skills

- Python
- Generative AI
- Data engineering (Apache Spark, Apache Iceberg)
- Vector databases (PGVector, Pinecone, FAISS)
- Cloud infrastructure (AWS: Bedrock, EC2, ECR, EKS, S3, Lambda)
- API development (FastAPI, REST APIs, gRPC)
- ML experimentation and deployment (MLFlow, Docker)

## Experience

**AI Research Engineer** **09/2023 to Current**
**Groundzero Software Private Limited** **Chennai**

- Architected a TB-scale Data Lakehouse: Designed and deployed a unified, cloud-native analytics platform using Apache Spark and Apache Iceberg, capable of processing and querying multi-terabyte datasets.
- Reduced RAG Latency by 90%: Optimized Retrieval-Augmented Generation (RAG) pipelines, reducing query response times from approximately 10 seconds to 500-1000 milliseconds through hybrid retrieval (dense/sparse) and vector database tuning (PGVector, Pinecone).
- Engineered high-scale infrastructure with Karpenter: Orchestrated production-grade AWS EKS clusters utilizing Karpenter for just-in-time node provisioning, ensuring 99.9% reliability for heavy, distributed ETL workloads.
- Led enterprise LLM orchestration: integrated AWS Bedrock and LangChain to build autonomous agents for SQL automation, BI visualizations, and analytics, significantly reducing operational costs.
- Optimized Performance on 100M+ Records: Benchmarked and delivered sub-10-second query execution times for complex analytical joins on 100M+ row tables using containerized runtimes.
- Standardized API Infrastructure: Developed an extensible internal API catalog using FastAPI and gRPC, accelerating cross-team feature development by 30%, and ensuring secure, scalable service interfaces.
- Automated Dataset Preparation: Built AI-driven agents that automate end-to-end data preparation and dashboard creation, enabling low-latency, interactive data workflows.

**Software Trainee** **06/2023 to 08/2023**
**SCI-BI Software Solutions Private Limited** **Chennai**

- Developed and maintained data visualizations and dashboards using Power BI and Tableau
- Participated in client meetings to gather requirements and translate them into technical specifications
- Created reports and presentations, enhancing clients' data-driven decision-making

## Education

**Bachelor of Technology**: **Computer Science and Engineering, Specialization in Artificial Intelligence and Machine Learning** **05/2023**
SRM Institute of Science and Technology Chennai
GPA: 9.08

## Projects

AgentSynapse: Enterprise AI agent platform enabling intelligent data analytics automation through specialized agents (SQL, BI, ETL, Analytics) with contextual memory system and comprehensive tool integration for unified data operations
CloudeasyML: An open-source software to deploy and run ML predictions and fine-tune LLMs using AWS Cloud Infrastructure (In-Progress).

## Languages

- English (fluent) | Tamil (Native) | Telugu (Fluent) | Hindi (intermediate)