

# KAMATCHI SOMESH S

Tamil Nadu, India • +91 96558 41515 • [kamatchisomesh@gmail.com](mailto:kamatchisomesh@gmail.com) • GitHub: [Somesh-S-Dev](#) • LinkedIn: [kamatchi-somesh](#)

**Machine Learning & Generative AI Engineer** specializing in RAG systems, NLP pipelines, and production-grade LLM applications with hands-on experience in speech AI, time-series forecasting, and scalable AI-powered web integrations. Strong expertise in model optimization, multilingual conversational AI, and end-to-end ML pipeline development.

## Experience

DEC 2025 – PRESENT

**Artificial Intelligence Intern | WinsSoft Innovations | Tamil Nadu**

- Shipped a production-ready **RAG-based NLP-to-SQL system** transforming a free ERP analytics dashboard into a monetizable AI-powered analytics platform.
- Enhanced system performance with a **metadata-driven orchestration pipeline** (without LangGraph), reducing **response latency from 120–180 seconds to 20–30 seconds** (~83% improvement).
- Engineered **schema-aware prompting using Gemini API** for accurate structured SQL generation across multi-table ERP datasets.
- Reduced hallucinations and irrelevant data retrieval through modular schema routing and controlled table-level retrieval strategies.
- Expanded **validation layers and guardrails** to ensure secure, reliable, and business-consistent query execution in production workflows.

JUN 2024 – NOV 2024

**Machine Learning Intern | Sayvai | Coimbatore**

- Constructed an **end-to-end speech intelligence pipeline** using Whisper **with speaker diarization and automated meeting summarization**.
- Enhanced and evaluated time-series **forecasting models (LSTM, ARIMA, SARIMA)** for Amazon product sales prediction and seasonal trend analysis.
- Created a multilingual conversational voice bot integrating LLM APIs to support Tamil language interactions.
- Conducted **research on low-resource language** adaptation using SpeechT5 for STT/TTS fine-tuning experiments.
- Automated AI project workflows** via Notion API for structured experiment tracking and data orchestration.

## Education

AUG 2021 – MAY 2026

**M.Sc. Artificial Intelligence and Machine Learning | Coimbatore Institute of Technology | Coimbatore**

GPA: 7.9 / 10

## Certifications

- Agents & Agentic AI Fundamentals – Intel
- Generative AI Essentials – IBM SkillsBuild
- Cyber Threat Management – Cisco
- Cybersecurity Fundamentals – IBM SkillsBuild

# Projects

## Stealth Thief Escape RL

- Formulated a custom Gymnasium-compatible reinforcement learning environment simulating a stochastic 10×10 city grid with dynamic police agents, surveillance zones, and obstacles.
- Accelerated a Dueling Double DQN agent to learn optimal escape policies under constrained and hazardous navigation conditions.
- Built persistent training infrastructure with replay buffer serialization and automatic checkpointing for continuous learning across sessions.
- Created real-time visualization and evaluation loops to monitor policy improvement and agent behavior.

## CodeDoc – AI Documentation Generator

- Shipped a privacy-first AI documentation generator using local LLMs (Ollama) to convert source code into structured technical documentation without external APIs.
- Optimized prompt design and chunking strategies for efficient inference on 8B parameter models running on consumer hardware.
- Integrated a modular backend–frontend architecture enabling scalability from single-file analysis to full codebase documentation.
- Implemented real-time streaming preview with syntax highlighting and export functionality to PDF and Word formats.

## Financial Anomaly Detection System

- Developed an end-to-end fraud detection platform with Flask dashboard for real-time anomaly analysis and model comparison.
- Trained multiple models (Isolation Forest, One-Class SVM, Random Forest, XGBoost, Autoencoders) with PyTorch GPU support.
- Upgraded the preprocessing pipeline with feature engineering, scaling, encoding, and imbalance handling using SMOTE/ADASYN.
- Designed automated evaluation and SHAP-based explainability using Precision, Recall, F1, ROC-AUC, and PR-AUC metrics.

# Skills

**Languages:** Python, SQL, C#, ASP.NET

**AI/ML:** Machine Learning, Deep Learning, NLP, Time Series Forecasting, Transformer Models

**Generative AI & LLMs:** RAG, Prompt Engineering, LangGraph, Gemini API, Hugging Face, Vector Databases

**Speech & Multilingual AI:** Whisper, SpeechT5, ElevenLabs, SarvamAI, Conversational AI

**Frameworks & Tools:** FastAPI, Docker, Git, REST APIs, GCP Basics

**MLOps & Pipelines:** Model Evaluation, Data Preprocessing, Experiment Tracking, Automation Workflows