

Chandru Ganeshan

chandruganeshan24@gmail.com | +91 9087535918 | India | github.com/itzmechandruganeshan | linkedin.com/in/chandru-g24
| itzmechandruganeshan.github.io

Summary

AI/ML Engineer with 2+ years building production GenAI systems: LangGraph multi-agent orchestration, hybrid RAG and semantic search, and NL-to-SQL analytics across media, fintech, automotive, and clinical domains. Builds deterministic, guardrailed agents — typed Pydantic tool contracts, safety-filter recovery, hallucination containment — validated by regression and LLM-eval suites.

Skills

Languages & Backend: Python, FastAPI, Uvicorn, asyncio, RESTful APIs, Server-Sent Events (SSE), NL-to-SQL, SQL, Apigee
GenAI & Agentic Systems: LangGraph, LangChain, Google ADK, Vertex AI Agent Engine (Reasoning Engines), Gemini 2.5, Amazon Bedrock, Mistral-7B, Veo 3.0, Structured Outputs (Pydantic), Tool Calling, Supervisor-Worker Orchestration, Checkpointing, DeepEval

RAG, Search & ML: Hybrid Search (Dense Vector + BM25), FAISS, Vertex AI Search, Sentence Transformers, RapidFuzz, spaCy, Scikit-Learn, SciPy (Signal Processing), Pandas, NumPy

Cloud & DevOps: Google Cloud (Cloud Run, BigQuery, Cloud SQL, Cloud Trace, GCS, Agentspace), Amazon Web Services (Bedrock, ECR), Docker, Kubernetes, Jenkins CI/CD, Git

Experience

Searce Inc

May 2026 – Present

AI / ML Engineer (Client: Mediacorp Singapore — mewatch OTT platform)

- Re-architected the Linear EPG sub-agent as a LangGraph **StateGraph** with structured Pydantic tool contracts (~7.6K LOC), replacing regex pre-checks with a typed clarification enum and driving one bounded repair from field-level **ValidationError** output.
- Hardened against Gemini 2.5 safety dropouts by routing RECITATION and empty-candidate finish reasons to a recovery node, and blocked invented channels by grounding every search call in exact/alias/fuzzy resolution over the 121-channel catalog.
- Migrated EPG retrieval from legacy NL-to-SQL to Vertex AI Search over a rolling 15-day BigQuery window with a constraint-relaxation ladder, and fixed day-part filters mis-mapped on 6 of 24 hours spanning 10,945 of 44,135 rows (24.8%) plus interval-overlap matching for broadcasts crossing day-part boundaries.
- Built the date-aware response contract: per-card **CURRENTLY_PLAYING** / **UPCOMING** / **PAST** status, carousels restricted to live programmes with the remainder demoted to context cards, BigQuery hydration of program IDs for deep links and artwork, and sports classification isolating live fixtures from highlights and replays.
- Built the multilingual routing node (English, Singlish, Mandarin, Malay, Tamil) with script gating, Indonesian/Malay disambiguation and English fallback; fixed a Gemini list-content crash killing every request and an SSE leak exposing classifier output in the UI.
- Implemented conversational pagination across content search and EPG: prior tool calls with matching arguments are mined for served titles, and the exclusion set applies to both Vertex AI Search filters and BigQuery predicates, eliminating repeats.
- Shipped a Cloud SQL to BigQuery pipeline decoding msgpack LangGraph checkpoints into a nested conversation schema with initial-dump and incremental sync modes; covered the agent with a DeepEval **GEval** suite and 1,379 lines of regression tests.

Cloudside Technologies Pvt Ltd

Jan 2025 – Apr 2026

AI / ML Engineer

- Built an autonomous sales-prospecting agent on Google ADK, deployed to Vertex AI Agent Engine and Agentspace, chaining lead discovery, four qualification filters, batched Apollo enrichment, and Pipedrive v2 CRM sync with lookup-before-create dedup.
- Built Tpay, a bilingual (English/Arabic) finance assistant on Gemini function calling: a two-pass NL-to-SQL chain that generates then verifies and corrects each BigQuery query, plus per-language **FAISS** retrievers and 16-category cash-flow forecast extraction.
- Engineered a scene-continuous Veo 3.0 ad-video pipeline that seeds each scene with the prior clip's OpenCV-extracted final frame and stitches the sequence with MoviePy, orchestrated by a five sub-agent marketing hierarchy.

ISPG Technologies Pvt Ltd

May 2024 – Dec 2024

GenAI / ML Engineer Intern

- Built a 3-stage hybrid search engine for automotive parts over ~1.39M fitment records: Mistral-7B on Amazon Bedrock for keyword and SKU extraction, thread-parallel dense retrieval across separate title and category indexes (**all-mpnet-base-v2**) gated at a 0.55 cosine floor, and a BM25Okapi lexical fallback for zero-recall queries.
- Designed an async FastAPI API for PNG Jewellery parsing colloquial Indian pricing ("2 lakhs", "under 50k") via a typo-tolerant multiplier grammar and spaCy POS keyword extraction, with concurrent **asyncio** retrieval and a RapidFuzz relevancy gate.

Microbiological Laboratory Pvt Ltd

Jan 2024 – Apr 2024

Data Scientist Intern

- Automated Rotor-Gene RT-PCR interpretation for an 8-target meningitis/encephalitis panel using Savitzky-Golay smoothing, linear-regression baseline fitting to locate amplification take-off (C_T), and SciPy peak/prominence analysis of melt curves classified against per-pathogen T_m threshold bands with explicit borderline handling.
- Generated clinical PDF reports with amplification and melt-curve plots (FPDF, Plotly, Matplotlib), replacing manual review.

Projects

PyMLRS – Clinical RT-PCR Pathogen Detection Framework | *Open-Source PyPI Package*

- Extracts raw RT-PCR fluorescence signals, derives C_T take-off points via baseline regression, classifies 8 pathogen targets by melt-temperature band, and emits verified PDF test summaries.

Rextractor – Proprietary RT-PCR Binary Data Parser | *GitHub / Open Source*

- Reverse-engineered XML/binary parser for the proprietary Rotor-Gene **.rex** container, emitting per-target amplification, melt (HRM), and cycle data as clean Pandas DataFrames.

Education

Master of Science (M.Sc.) in Data Analytics | Bharathiar University

Oct 2024

Bachelor of Science (B.Sc.) in Statistics | Government Arts College

Jul 2022