Nandini Chinta

New York | nandinic.012@gmail.com | +17169077993 | GitHub | LinkedIn

Summary

- Data Engineer with around **7 years of experience** delivering reliable, scalable, and high-performance data and ML systems across GCP, AWS, and Azure.
- Specialized in real-time data streaming, warehousing, data quality engineering, and cloud cost optimization.
- Proven record improving **latency** (40%), **throughput** (35%), and reducing **cloud costs** by up to **80%** through strong architectural design and reusable platform abstractions.
- Hands-on experience deploying ML systems, LLM-powered applications, and privacy-aware AI features in production.
- *Certified Google Cloud Professional Data Engineer*

Skills

Languages: Python, SQL, Java
Data Engineering: BigQuery, Dataflow, Datastream, dbt, Data Modeling, ETL/ELT, Streaming & Batch Pipelines
ML & AI: LLMs, RAG, Reinforcement Learning, CNN/RNN, XGBoost, PyTorch, TensorFlow
MLOps & ML Infra: Vertex AI, SageMaker, CI/CD, Model Serving, Feature Stores, Experimentation Pipelines
Cloud: GCP (BigQuery, Pub/Sub, Cloud Run, Cloud SQL, Document AI),
AWS (SageMaker), Azure (Data Factory, Synapse)
Systems & Tools: FastAPI, Docker, Terraform, GitLab CI/CD, Prefect, Data Observability, Spring Boot, Git, Jira
Visualization: Looker, Tableau, Power BI, Matplotlib

Experience

Data Engineer - Veolia North America, New Jersey

June 2023 - Present

- Core contributor to Veolia's **enterprise Data & AI Platform**, building governed data warehouses, real-time streaming pipelines, ML infrastructure, and platform observability used across BI and ML teams spanning six business units.
- Served as Interim **Lead, Data Engineering Operations**, owning daily platform health, incident prioritization, and SLA compliance while coordinating with engineering, product, and data stakeholders.
- Designed and standardized **core ingestion architectures** (CDC, API, SFTP, Batch), enabling one-click onboarding of **50+ data sources** and maintaining **hundreds of production tables** in BigQuery.
- Re-architected batch ingestion into **streaming-first pipelines** using Datastream and Dataflow, reducing data availability latency from days to seconds, improving query latency by **40%**, throughput by **35%** across **100+ TB/day** workloads.
- Implemented **enterprise-grade data governance and security** (IAM, row-level security, taxonomies, access expiration policies) and built a **self-service Access Management Platform** (FastAPI, Pydantic, PostgreSQL), reducing BigQuery access provisioning from **2 days to under 2 seconds**.
- **Led Dataplex Data Quality Project** and designed reusable **incremental ingestion**, **deduplication**, and **historical merge frameworks** across **7** enterprise data lakes, improving dataset accuracy to **99.8%** and BI trustworthiness.
- Architected a platform-wide **observability layer** (freshness checks, anomaly detection, SLAs, cost attribution) and optimized BigQuery/dbt models using **partitioning**, **clustering**, and **predicate** pushdown, reducing incident resolution time by **50%** and cloud costs by up to **80%**.
- Designed **distributed backend services** handling **5K+** concurrent requests with **sub-200ms latency**, and deployed distributed ML training pipelines on Vertex AI, reducing training time by **30%** and improving model accuracy by **15%**.
- Architected a **hands-free wake-word voice activation system** using an on-device TFLite model with noise suppression and echo cancellation, buffering audio for Whisper STT, routing queries to a secure internal AI agent, and returning responses via TTS to enable **low-latency**, privacy-aware voice interaction.
- Participated in a cross-functional hackathon delivering voice-controlled navigation for a **Boston Dynamics Spot robot**.
- Mentored junior engineers by defining ingestion templates, data modeling standards, and operational best practices, enabling independent onboarding of new data sources.

Machine Learning Researcher - SUNY Buffalo, New York

Feb 2022 - June 2023

- **Fairness in Multi-Agent Reinforcement Learning:** Investigated fairness in multi-agent systems by developing algorithms that balance individual and collective rewards under constrained environments. Conducted experiments in a custom grid-world, evaluating Q-Learning, DQN, DDQN, and PPO, demonstrating that DDQN and PPO achieve superior coordination, equitable outcomes, and efficient goal completion.

- **Traffic Flow Control using Multi-Agent Reinforcement Learning (MARL):** Researched traffic flow optimization in smart cities by modeling traffic lights as agents and learning policies to maximize overall traffic efficiency. Implemented and evaluated Deep RL algorithms, including Actor-Critic and PPO, using the SUMO simulator, achieving improved coordination and reduced congestion.

Software Engineer - Kaiser Permanente(TCS), India

June 2019 - Aug 2021

Contributed to enterprise healthcare platforms at Kaiser Permanente, building production-grade ML systems, backend services, and data pipelines supporting clinical operations and physician workflows.

- Developed **batch job Microservices**, eliminating **90%** of manual operational effort automating SOP workflows.
- Optimized **SQL queries** and **stored procedures** for high-volume transactional systems, saving **100+** engineering hours per month in downstream processing and reporting.
- Implemented **concurrency and multithreading** strategies to increase service throughput from **500 → 1,100 TPS** (+120%) and reduce batch execution time by **50%**.
- Integrated Python **ML classification models** into production applications for real-time decisions.
- **Led a team** of engineers by setting coding standards, reviewing designs, and driving best practices for backend services and data workflows.

Research & Projects

RAG Applications - Veolia

Python, GCP, LLMs

- Built enterprise RAG applications integrating **vector search** and **LLMs** for document intelligence, improving query accuracy by **40%** and reducing manual search time.
- Deployed optimized pipelines using **LangChain + GCP**, enabling scalable document ingestion and retrieval while lowering latency.
- Developed **Streamlit-based UI applications** for RAG systems, providing user-friendly interfaces for business teams to query documents and access insights seamlessly.

GenAI with LLMs

Python, AWS Sagemaker, LLMs

- Utilized **data parallelism and Microsoft's ZeRO optimizations** to manage large-scale models, reducing computational costs by 40% without compromising performance.
- Applied Reinforcement Learning from Human Feedback (**RLHF**), increasing model alignment with human values by **50%**, as evidenced by user satisfaction.
- Deployed Retrieval-Augmented Generation (**RAG**), ensuring more precise answers and reducing **hallucination**.

Affective State Analysis using VGG-16

Python, Computer Vision, Image Classification

- Built a multi-view user engagement recognition prototype, classifying four engagement levels using an Image Classification System.
- Developed Convolutional Neural Network, fine-tuning **VGG-16** with effective weights, achieving an accuracy of **87%**.

Education

Master of Science in Artificial Intelligence **Aug 2021 – Feb 2023**

University at Buffalo, SUNY, Buffalo, NY, CGPA: 3.6/4

- **Teaching Assistant:** Courses: 574 Machine Learning, 546: Reinforcement Learning

B. Tech in Computer Science and Engineering **Jul 2015 – May 2019**

Vignan University, India, CGPA: 9.05/10

Certifications

- Google Cloud Professional Data Engineer [link]
- Terraform Associate [link]
- Prefect Associate [link]
- GitLab CICD