

Yogavarshni Ramachandran

San Jose, CA | +1 (408)549-4786 | yogavarshni@gmail.com | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

SUMMARY

Data Scientist with 4+ years of experience in Python, SQL, and Cloud to build scalable machine-learning, NLP, and A/B testing frameworks for GTM, product, marketing, and forecasting analytics, driving retention and data-informed growth across Telecom, AI, SaaS and Fintech.

PROFESSIONAL EXPERIENCE

Data Scientist

June 2025 – Present

Saayam – Machine Learning, Business Intelligence, Fraud Analytics, Automation, Customer Experience.

San Jose, CA

- Modeled a python-based **fraud prevention** ML system to flag **anomalous** accounts with 85% precision, deploying dashboards on **10K+** transactions to improve customer trust, platform security, and ensure transparency.
- Streamlined analytics migration to **Power BI** by optimizing **MySQL** procedures on 18+ tables & formulating corresponding DAX queries.

AI Data Scientist

August 2024 – June 2025

San Jose State University – RAG, GenAI, LLMs, GCP, AI Agents, GitHub Workflows.

San Jose, CA

- Productionized an [AI website generator](#) integrating **GPT-4**, **Gemini**, and **Claude** on GCP (FastAPI), enabling users to create instant B2B online marketing websites through prompts, reducing load time by 0.7s and improving Core Web Vitals.
- Developed an **RAG**-based intelligent [document QA system](#) using LangChain, Llama 3.2, and ChromaDB with **vector embeddings**, enabling analysts to extract instant insights from local PDFs and improving answer precision evaluated via ROUGE metrics.

Data Science Intern

June 2024 – August 2024

Discover Financial Services – Product Pricing & Marketing, Banking, Data Quality, Financial Analytics.

Riverwoods, IL

- Automated competitor rate tracking across 6+ banks using **Airflow and Python**, driving data-backed pricing and marketing initiatives.
- Executed **pricing strategies** and marketing campaigns using market trend analysis, implementing a 0.85bps rate increase on 9M CD products that increased deposit balances by 12% and generated \$800K in income (Finance).
- Enhanced time-series **forecasting models** for deposit by integrating **customer activity** features, improving prediction accuracy by 5%.

Product Data Analyst

March 2022 – August 2023

Freshworks – SaaS, B2B, Go-To-Market Strategy, Product Analytics, Funnel Analysis, Root Cause Analysis.

India

- Joined the **GTM** analytics, built a **data pipeline (Snowflake, dbt)** integrating data points from Customer Success, Sales, Product, enabling data-driven **product insights** and establishing reporting for cross-functional stakeholders aligned with 6+ **OKRs**.
- Delivered **KPI** and self-serve **tableau** dashboard for the sales team, enabling tracking of **product usage**, **sales strategy**, and **growth opportunities**, contributing to a 10% increase in **revenue** through user engagement strategies in Q1 2023.
- Facilitated strategic decision making by developing **predictive models** for churn and revenue forecasting (linear regression, ARIMA), delivering 90% accuracy in identifying high-value accounts and improved customer **retention**.
- Performed analysis of subscriber data, delivering [churn dashboards](#) identifying \$12.57M at-risk revenue with strategies to reclaim 15%.
- Refined dunning strategy using **A/B testing** and user **segmentation** based on payment history and user behavior, reducing customer calls by 20% and improving **operational optimization** and business requirements.

Software Engineer / Data Engineer

August 2020 – March 2022

Infosys – NoSQL, Data Governance, Data Architecture, Data Mining, Agile, Consulting, SDLC.

India

- Designed a streamlined ETL data management pipeline for **AT&T**, involving data collection, data manipulation, data modeling, and loaded **big data** into **distributed databases** (Cassandra and Kafka), reducing data processing time by 50%.
- Optimized **AWS** data warehouse solutions (**S3, EC2, Redshift**) for client-critical applications, elevated **query performance** by 40%.

TECHNICAL SKILLS

Programming	: Python (Pandas, NumPy, Scikit-learn, SciPy, PyTorch, TensorFlow), SQL, Spark, R, SAS
ML/AI	: Supervised, Unsupervised Learning, Time Series, NLP, RAG, LLM, VectorDB, Langchain, GenAI, Deep Learning
Analytics	: A/B Testing, Hypothesis Testing, Machine Learning, Statistics, Experimentation, Data Mining
Databases	: MySQL, Oracle, MongoDB, Snowflake, Hadoop, Hive, Spark, Kafka, Airflow, DBT, Teradata, BigQuery
Visualization	: Tableau, Power BI, Looker, QlikView, Seaborn, Matplotlib, Microsoft Excel, Google Sheets
Tools	: Azure (Certified), AWS (S3, EC2, Redshift, AWS Glue), Docker, CI/CD, Git, MS Office, Six Sigma

EDUCATION

San Jose State University

Master of Science in Data Science

San Jose, CA

August 2023 – May 2025

PROJECTS

Reddit Moral Judgement Dilemmas ([Project Link](#)) – *Extract, Transform, Load (ETL), Data Engineering, Data Visualization.*

- Trained and deployed **ML models** (Random Forest, XGBoost, Regression, Decision Tree) leveraging NLP to classify bias in moral decision making, achieving 72 % accuracy and 74 % precision, providing data-driven recommendations for social insights.

AI Bedtime Story Generator ([Project Link](#)) - Multi-Agent System, Prompt Engineering, LLM Orchestration, REST API

- Engineered and deployed ([app](#)) an end-to-end **multi-agent generative AI system** (Flask, Next.js, GPT-3.5) on Render, integrating storyteller and judge LLMs to generate, evaluate, and refine kid-safe bedtime stories.