

Yogavarshni Ramachandran

San Jose, CA | +1 (408)549-4786 | yogavarshni@gmail.com | [linkedin.com/in/yogavarshniramachandran](https://www.linkedin.com/in/yogavarshniramachandran) | [Portfolio](#)

SUMMARY

Data Scientist with 4+ years of experience in Python, SQL, and Cloud to build scalable machine-learning and A/B testing frameworks for GTM, product, pricing, marketing, and forecasting analytics, driving growth and actionable insights across Research, SaaS and Fintech.

PROFESSIONAL EXPERIENCE

Data Scientist

June 2025 – Present

Saayam – Machine Learning, Business Intelligence, Fraud Analytics, Feature Engineering, Root Cause Analysis. **San Jose, CA**

- Modeled a **fraud detection** system to flag **anomalous** accounts with 85% precision, deploying dashboards on 10K+ transactions to improve customer trust, platform security, and ensure transparency.
- Migrated data analytics to **Power BI** by analyzing existing **MySQL** procedures on **18+** tables & formulating corresponding **DAX** queries.

Data Scientist, Research

September 2024 – June 2025

San Jose State University – Data Analysis, B2B SaaS, GenAI, GTM, Product Management, Product Launch. **San Jose, CA**

- Built an **AI-driven interface** to auto-generate websites, reducing build time 80%, reducing **startup costs**, and enhancing **UX insights**.
- Implemented a **RAG** pipeline to fine-tune **GPT-4**, **Gemini**, and **Claude**, delivering real-time context and reducing website load time by **0.7s**, improving lighthouse by 85% and **SEO** by 90%, enhancing **user experience** and **engagement**.

Data Scientist Intern

June 2024 – August 2024

Discover Financial Services – Product Pricing & Marketing, Banking, Data Quality, Risk Analysis. **Riverwoods, IL**

- Monitored competitor deposit rates by **web scraping** data from 6+ banks (BoFA, Ally Bank, Amex) using **MySQL**, **Python**, and **Airflow**, supporting rate adjustment strategies and marketing initiatives to improve competitiveness.
- Drove **pricing strategy** and deposit business growth by executing a **0.85 bps** rate increase on the **9M CD** through market trend analysis, resulting in a **12%** uptick in CD balances and **\$800K** in NII.
- Enhanced time-series **forecasting models** for deposit changes by adding features based on customer activity, boosting accuracy by **5%**.

Data Analyst, CS & Sales

March 2022 – August 2023

Freshworks – SaaS, B2B, Go-To-Market Strategy, Product Analytics, Funnel Analysis, Customer Experience. **India**

- Joined the **GTM** analytics, built a **data pipeline (Snowflake, Sheets)** integrating data points from Customer Success, Sales, RevOps, Product, enabling data-driven **product insights** and establishing reporting for cross-functional stakeholders aligned with 6+ **OKRs**.
- Led the creation of automated **KPI** tracking systems for a team of **50+**, providing insights to **VPs of Sales**, influencing **roadmap** priorities.
- Delivered **self-serve tableau** dashboard for the sales team, enabling independent tracking of **product usage**, **sales strategy**, and **business growth** opportunities, contributing to a **10%** increase in revenue through improved user engagement strategies.
- 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** churn data, delivering dashboards identifying \$12.57M at-risk revenue with strategies to reclaim 15%.
- Refined dunning strategy using **A/B testing** and **segmenting users** by payment history and behavior, reducing customer calls by **20%** and improving **operational optimization** and business requirements.

Data Engineer

August 2020 – March 2022

Infosys – Oracle SQL, NoSQL, Data Governance, Data Architecture, Data Mining, Agile, Consulting, SDLC. **India**

- Designed a streamlined ETL data management pipeline for client **AT&T**, involving data collection, refinement, data modeling, and seamlessly 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), SQL, Spark, R
Analytics	: A/B Testing, Hypothesis Testing, Machine Learning Algorithms, NLP, Experimentation, Deep Learning
Databases	: MySQL, Oracle, MongoDB, Snowflake, Hadoop, Hive, Spark, Kafka, Airflow, DBT, Star Schema, Big Data
Visualization	: Tableau, Power BI, Looker, QlikView, Seaborn, Microsoft Excel, VBA, Google Sheets, Slides
Tools	: Azure (Certified), AWS (S3, EC2, Redshift, AWS Glue), Salesforce, CRM, Git, Jira, Confluence, PowerPoint, Six Sigma

EDUCATION

San Jose State University

Master of Science in Data Science

San Jose, CA

August 2023 – May 2025

PROJECTS

E-commerce Data Warehouse System ([Project Link](#)) – Extract, Transform, Load (ETL), Data Engineering, Data Visualization.

- Architected a dual-database system (**normalized** operational and **star schema** analytical) with **Python-SQL ETL pipelines** for real-time tracking and historical sales-shipment analysis enabled warehouse **demand forecasting** to enhance **operational efficiency**.

Reddit Moral Judgement Dilemmas ([Project Link](#)) – Random Forest, SVM, XGBoost, Regression, Decision Tree, Recommendations.

- Trained **ML** models using **NLP** features to detect user bias in moral reasoning on Reddit, achieving **72%** accuracy and **74%** precision.