

VARUN VADDI

varunvaddi30600@gmail.com | Houston, TX | (713) 539-6996 | [LinkedIn](#) | [Portfolio](#)

SUMMARY

Data Scientist (MS, Data Science) with 3+ years building ML models, GenAI-powered solutions, and BI dashboards that turn large-scale data into business decisions. Experienced applying LLMs, prompt engineering, and statistical modeling to accelerate insight generation and support cross-functional product and analytics teams.

EDUCATION

University of Houston

Master of Science (MS), Engineering Data Science | GPA: 3.7/4

TX, USA

Dec 2025

WORK EXPERIENCE

Data Science Intern, Wild Genomics, CA

May 2025 - Aug 2025

- Designed and ran experiments across 4+ pipeline architectures evaluating accuracy and scalability tradeoffs, increasing analysis-ready data from 30% to 80% and enabling research team to make data-driven decisions 40% faster.
- Improved classification accuracy from 85% to 95% through iterative hyperparameter experiments and threshold optimization - directly informing 2 conservation research decisions including early detection of bacterial presence across ranch ecosystems.
- Presented weekly findings to CEO and CSO, translating complex genomic model outputs into plain language - directly informing 2 conservation research decisions and supporting grant-funded research priorities.

Data Analyst, University of Houston, TX

Jul 2024 - Dec 2025

- Built automated data pipelines integrating 3 admissions systems to support predictive graduation model analyzing 40K+ application/semester, enabled 70% reduction in manual processing and 5-day faster decisions through reliable data ingestion, quality validation, and data instrumentation across 3 integrated systems.
- Designed Tableau dashboards with application funnel metrics and risk segmentation, enabling admissions officers to monitor real-time application flow, prioritize high-risk cases, and improve decision consistency across team.

Data Analyst, Accenture, India

Sep 2021 - Dec 2023

- Optimized SaaS analytics workflows for 80M+ records by streamlining deployment pipelines and improving latency by 30%, enabling Sales & Product teams to identify high-value leads faster and inform territory strategies.
- Drove 10% conversion lift (+\$2.5M annual impact) by defining product adoption metrics, building real-time Tableau dashboards, and identifying 3 funnel drop-off points through user behavior analysis that informed CRM roadmap.
- Led CRM data migrations and developed SQL analytics pipelines identifying conversion bottlenecks across 5-stage funnel, optimized lead scoring, contributing to \$25M revenue growth through improved prioritization.

PROJECTS

AI-Powered Policy Intelligence System

- Built AI-powered self-service compliance assistant using hybrid RAG (BM25, FAISS, embeddings) to classify customer-facing ad policy intents in under 1 second - achieving 80% accuracy and 78% Recall@5, reducing manual review dependency by 60% and scaling to 100+ decisions/min.
- Implemented structured LLM outputs with confidence scoring and bias-aware evaluation framework to ensure responsible AI policy decisions - achieved 95% policy coverage through systematic evaluation, reducing decision violations by 40%.

Customer Retention & Revenue Intelligence Platform

- Built end-to-end ML system on 10.6M orders addressing poor customer retention - achieved 80% AUC (reorder), 90% R² (LTV), 77% AUC (churn) through feature engineering & hyperparameter tuning, identified \$4.65M annual opportunity via segmentation.
- Engineered 45+ behavioral features via PySpark/Databricks and performed RFM cohort analysis (7 segments), deployed dashboards with A/B test planner and ROI calculator enabling Product teams to evaluate experiments and prioritize roadmap initiatives.

Ride-Share Demand & Pricing Intelligence

- Built end-to-end analytics pipeline (Snowflake, dbt, Airflow, Kafka) for 8.6M+ trips, reduced engineering time by 95% with automated validation (37 checks), enabled demand forecasting that optimized driver allocation and reduced wait times.
- Analyzed \$180M+ quarterly revenue patterns using SQL and Tableau, uncovered 30% peak-hour demand lift and 18% payment method premium through user behavior segmentation - directly informing dynamic pricing experiments analogous to surge pricing.

SKILLS

Visualization & Reporting: Power BI, Tableau, Excel, PowerPoint.

Product Analytics: A/B Testing & Experimentation, Experimental Design, Forecasting, KPI Definition, Cohort Analysis, Causal Inference.

Programming: Python, SQL, R, NumPy, Pandas, PySpark.

ML & AI: Scikit learn, PyTorch, TensorFlow, Machine Learning, LLM, RAG, Prompt Engineering, Generative AI, Predictive Modeling.

Platforms & Tools: AWS, Azure, Snowflake, Databricks, Spark, Git/GitHub, CI/CD, Agile.