Jugal Sheth

Senior AI Data Engineer — Cloud Data Platform Architecture — Databricks & Spark Specialist

📞 +1 (857) 869-8235 ✉ jugal.sheth63@gmail.com in LinkedIn GitHub 🌐 Portfolio 📊 Tableau

Professional Summary

Data Engineer with 5+ years in operational and technical roles at fast-growing startups, delivering software products that achieve significant business impact. Proven expertise in product implementation, SQL-based monitoring, and BI tools to ensure optimal product performance and client value. Strong background in cross-functional collaboration with Product, Engineering, and Customer Success teams to enhance product offerings and drive seamless launches. Deep technical proficiency in cloud data platforms (AWS, Azure), API integrations, and building scalable solutions while taking ownership in unstructured environments. Track record of acting as Voice of the Customer, synthesizing feedback across multiple channels to inform product roadmaps and process improvements while maintaining 99.9% system reliability across production environments serving 600+ users.

Professional Experience

VaynerX   **Oct 2025 – Present**
Senior AI Data Engineer   *New York, NY*
- Program managed data platform product launches serving 600+ users, hosting internal knowledge sessions and creating comprehensive enablement materials including technical documentation and implementation training resources that reduced onboarding time by 35% and ensured smooth rollouts across cross-functional teams
- Utilized SQL and BI tools (Tableau, Looker) to monitor health of data products post-implementation, partnering with Support and Operations teams to identify and correct misconfigurations, ensuring clients achieve maximum value while maintaining 99.9% system uptime and delivering proactive insights on product performance
- Owned feedback loop conversations synthesizing data across multiple channels including user surveys, support tickets, and stakeholder meetings, acting as Voice of the Customer to identify pain points and collaborate with Product and Engineering teams to enhance platform capabilities and address market needs effectively
- Led discovery sessions with business stakeholders to understand product requirements and translate needs into technical solutions, taking ownership in ambiguous environments to prioritize effectively and deliver data governance frameworks that reduced data quality issues by 40% while managing SLAs for 50+ critical datasets

VaynerX   **Dec 2023 – Oct 2025**
Data Engineer   *New York, NY*
- Delivered AI-powered conversational assistant product leveraging OpenAI LLMs and RAG architecture that achieved significant business impact, reducing time-to-insight by 50% for 200+ users while taking ownership of the full product lifecycle from requirements gathering through launch and post-implementation support in a fast-paced startup environment
- Utilized SQL to monitor product health and troubleshoot client issues across AWS cloud environment, building fault-tolerant data pipelines on Snowflake and PostgreSQL that delivered 40% reduction in processing time while maintaining 99.9% uptime, demonstrating ability to prioritize effectively and resolve ambiguous technical challenges
- Worked closely with Product and Engineering teams to enhance data platform offerings, conducting thorough discovery sessions to understand client requirements and implementing distributed processing frameworks (PySpark) and CI/CD pipelines that improved forecasting accuracy by 12% and enabled scalable product delivery
- Built comprehensive enablement materials and training documentation while taking on variety of roles including implementation lead, technical support, and product advocate, synthesizing feedback from multiple stakeholders to drive continuous product improvements and ensure maximum client value from data infrastructure investments

VaynerMedia   **Oct 2021 – Dec 2023**
Data Analyst   *New York, NY*
- Operated in generalist role at early-stage division collaborating with cross-functional Product, Growth, and Marketing teams to deliver data products that achieved significant business impact, supporting $200K in budget allocation decisions while demonstrating willingness to take on variety of responsibilities from analysis to implementation to training
- Developed BI dashboards in Tableau with SQL-based monitoring capabilities tracking ROAS and CPA metrics, conducting discovery sessions with stakeholders to understand requirements and synthesizing feedback to enhance product functionality, resulting in improved decision-making speed and self-service analytics adoption
- Built API integration pipelines connecting multiple platforms, taking ownership of product implementations from concept to launch while working in unstructured environment, reducing manual reporting effort by 60% and establishing measurement frameworks that demonstrated clear business value to clients and internal stakeholders

Optima Global (Client: The Hanover Insurance Group)   **Jun 2021 – Oct 2021**
Claims Data Analyst   *Remote*
- Optimized PostgreSQL queries processing 100K+ insurance claims records, demonstrating hands-on SQL troubleshooting and self-service capabilities that achieved 60% reduction in fraud detection reporting latency through indexing strategies and query plan optimization, while building Power BI dashboards tracking SLA adherence that improved operational efficiency by 5 days

Data Engineering & AI Solutions Projects

LLM-Powered Conversational Analytics Platform – Production AI Agent   Ongoing
- Building production-grade conversational AI agent for customer self-service analytics, integrating OpenAI GPT-4 with RAG pipeline architecture to enable natural language interactions with BI data, implementing prompt engineering best practices and ensuring high-quality data inputs for LLM performance optimization, targeting 5K+ monthly conversations

- Architecting multi-turn dialogue management system with context retention and intent classification capabilities, deploying on AWS cloud infrastructure with fault-tolerant design patterns to deliver adaptive conversational experiences and 40% reduction in support tickets while maintaining data security and compliance

SKU Master Ware – Enterprise E-Commerce Data Platform   Oct 2025
- Architected and deployed production-scale inventory analytics platform for 3 e-commerce clients managing 10K+ SKUs across multiple channels, designing PostgreSQL database schema with 15+ normalized tables implementing star schema and data vault patterns, Row-Level Security for data governance, and distributed processing for real-time multi-channel data reconciliation
- Built real-time analytics engine computing 40+ KPIs with event-driven ETL orchestration using Airflow, automated reorder workflows, and barcode-driven GRN systems, reducing manual intervention by 80% while engineering serverless DevOps pipeline on AWS achieving 99.9% uptime and sub-second page loads for BI dashboards

GeoPulse AI – LLM-Powered Financial Analytics Application   Feb 2025
- Deployed production OpenAI-powered Streamlit analytics application analyzing global financial news to generate market forecasts, demonstrating LLM integration expertise with RAG architecture, dbt-based modular ETL pipelines, and explainable AI metrics framework ensuring interpretable ML-driven insights with data lineage tracking and auditability for financial services use cases

Education

Northeastern University   Boston, MA
Master of Science in Engineering Management

- **Relevant Coursework:** Data Mining and Machine Learning, Database Management and Design, Computation and Visualization, Algorithmic Digital Marketing, Engineering Probability and Statistics

University of Mumbai   Mumbai, India
Bachelor of Engineering in Mechanical Engineering

Technical Competencies

Product & Solutions: Product Launch Management, Implementation Training, Enablement Material Creation, Product Monitoring & Health Checks, Discovery Sessions, Requirements Gathering, Troubleshooting & Escalation Support, Voice of the Customer, Feedback Loop Management

SQL & BI Tools: SQL (Expert - Monitoring, Troubleshooting, Query Optimization), Tableau, Power BI, Looker, Data Visualization, Self-Service Analytics, Performance Monitoring, Dashboard Development

Operational & Cross-Functional: Early-Stage Startup Experience, Operational/Generalist Roles, Product Management Collaboration, Engineering Partnerships, Customer Success Support, Prioritization in Ambiguous Environments, Ownership & Accountability, Taking on Variety of Roles

Cloud Platforms & Data Engineering: AWS (EC2, S3, Lambda, RDS, Redshift), Azure (ADF, Synapse), Snowflake, PostgreSQL, ETL/ELT Pipelines, Airflow, Data Architecture, Distributed Processing (PySpark), Real-Time Data Processing

Integration & APIs: REST APIs, Webhooks, API Integration Frameworks, Postman, Third-Party Platform Integrations, Data Pipeline Development, Event-Driven Architecture

Programming: Python (Advanced), SQL, PySpark, R, Git/GitHub, Automated Testing, CI/CD Pipelines

AI/LLM & Analytics: RAG Architecture, OpenAI API, LLM Integration, Conversational AI, Data Quality for ML Models, Predictive Analytics

Soft Skills: Strong Communication, Technical Training Delivery, Stakeholder Management, Client-Facing Experience, Documentation, Process Improvement, Problem-Solving in Unstructured Environments