

Venkatesh Terikuti

Senior ML Platform Engineer | MLOps Specialist | Data Scientist | Senior Data Analyst
vterikuti@gmail.com | +1 (402) 6302823 | linkedin.com/in/terikuti | github.com/venkateshterikuti

Professional Summary

A growth-focused Data Analyst with 6+ years of experience translating user lifecycle data into actionable strategies for SaaS and subscription-based products. Expert in building scalable Tableau dashboards, developing robust data models with dbt Core and advanced SQL, and collaborating with engineering teams to create a reliable analytics infrastructure. Proven ability to drive user acquisition, engagement, and retention through deep-dive analysis and predictive modeling.

EXPERIENCE

Senior Software engineer – Data Scientist

Nov 2021 – Jan 2023

GlobalLogic India

Hyderabad, IN

- Built deep learning pipelines using TensorFlow and attention-based LSTMs to forecast demand for 100K+ SKUs with 96% accuracy; integrated models via CI/CD using Jenkins and GitLab CI.
- Designed and deployed a centralized MLOps platform on AWS SageMaker, serving batch and real-time features for 10+ models; reduced model deployment latency by 40%.
- Engineered ETL workflows in MySQL, Hive, and Airflow; utilized Databricks for scalable data processing and collaborative analytics.
- Partnered with marketing and supply chain stakeholders to deliver insights and dashboards supporting product inventory and campaign effectiveness.

Senior Data Analyst

Nov 2019 – Oct 2021

Geetham Software (Client: Lytx)

Chennai, IN

- Led development of scalable Tableau dashboards for driver safety, operations, and compliance KPIs across US and LATAM regions; reduced dashboard load time by 40%.
- Built robust SQL-based pipelines on AWS (Lambda, S3) and Databricks Delta Lake to ingest 6M+ daily telematics events.
- Established semantic data models and reusable data marts using dbt Core and Git-based workflows to enable self-service analytics across business units.
- Collaborated with engineers to implement event tracking schema and ensured reliable instrumentation for analytics use cases.
- Enabled experimentation and A/B testing analysis by building cohort-based reporting tools and retention visualizations.

Data Analyst

Mar 2016 – Oct 2019

Geetham Software

Chennai, IN

- Developed automated ETL workflows via AWS Glue and SQL for ingesting CRM, finance, and campaign data; reduced manual reporting workload by 60%.
- Conducted churn analysis and user behavior segmentation using Python (pandas, seaborn); insights helped boost conversion by 12%.
- Designed and maintained daily performance dashboards in Google Data Studio and Tableau, tracking marketing, operations, and customer engagement KPIs. Automated recurring business reports and alerting systems, improving report turnaround time by 40% and ensuring timely insights for quarterly reviews.

ML Research Intern

Feb 2025 – Jun 2025

University of Arizona

Tucson, AZ

- Built a scalable semantic search engine for e-commerce by fine-tuning Sentence Transformers and Google FLAN-T5 on a 1.3M-product catalog, integrating FAISS for fast vector retrieval and enabling real-time, intent-aware search.
- Implemented a full-stack ML pipeline covering data prep, model fine-tuning, and dynamic filter extraction (e.g., price, ratings), enhancing structured query refinement and delivering a 90% increase in Precision@10 over traditional keyword search.
- Led applied research on LLM-powered search, authoring a paper titled "*LLM-based Semantic Search for Conversational Queries in E-commerce*" submitted to CIKM'25 (ACM), showcasing innovative use of Sentence-BERT for user query understanding and product discovery.

EDUCATION

Master of Science in Data Science, The University of Arizona (GPA: 3.89)

Dec 2024

Bachelor of Technology in Mechanical Engineering, Acharya Nagarjuna University (GPA: 3.42)

May 2015

PROJECTS

Predicting Customer Churn in E-Commerce

Dec 2023

- Created churn prediction models using Random Forest, SVM, and XGBoost with >90% accuracy on the 'Online Retail II' dataset (500K+ records).

- Conducted retention cohort analysis and lifetime value estimation; tuned model performance using cross-validation and grid search.
- Delivered Tableau dashboards and actionable insights to simulate real-world e-commerce churn scenarios and KPIs.

Casino Player Intelligence Hub: Driving Engagement & Revenue

May 2025

- Architected comprehensive ML platform for online casino player behavior prediction, implementing end-to-end model lifecycle management including automated retraining, A/B testing, and performance monitoring
- Developed predictive models for player churn (30-day window), LTV forecasting, and fraud detection using advanced machine learning techniques, with integrated model validation and regulatory compliance reporting
- Built real-time inference system with PostgreSQL backend and Python-based ML pipeline, supporting personalized marketing campaigns and risk management with <100ms response times
- Designed interactive dashboards with KPI monitoring, model performance tracking, and business impact visualization, enabling data-driven decision making for marketing and operations teams

Enterprise Knowledge Copilot: Agentic RAG for Internal Q&A & SOP Automation

Jun 2025

- Developed an enterprise-grade RAG platform for knowledge workers to ask natural language questions against internal documentation (SOPs, runbooks, tech specs, policy PDFs).
- Leveraged LangChain with pgVector and Redis to embed, retrieve, and rerank context chunks; OpenAI GPT-4 for synthesis and summarization.
- Designed multi-agent orchestration for use cases like IT support, onboarding, compliance checks, and change management documentation.
- Containerized and deployed the stack on AWS EKS with Terraform and GitOps using ArgoCD; built real-time observability with Prometheus/Grafana.
- Implemented S3-based encrypted storage, IAM-bound access control, and usage audit logs to meet security & compliance standards.

Skills

Programming Languages: Python, C++, R, SQL, Bash

Data Science & Machine Learning: Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch, XGBoost, Sentence-BERT (SBERT), LLMs, Time Series Analysis, Statistical Modeling, Supervised Learning (Regression, Classification), Unsupervised Learning (Clustering), NLP, Reinforcement Learning, PySpark, DataBricks, Computer Vision

ML Platform & Infrastructure: MLOps, Model Lifecycle Management, AWS SageMaker, TensorFlow, PyTorch, Model Validation, MLFlow, CI/CD Pipelines, Docker, Kubernetes

Product & Event Tracking: Amplitude, Mixpanel (Familiar), Segment (Familiar), Event Instrumentation, KPI Standardization, Experimentation Frameworks

Marketing & Customer Analytics: Customer Segmentation, Churn Prediction & Prevention, Lifetime Value (LTV) Analysis, Player Behavior Modeling, Campaign Effectiveness Analysis, A/B Testing Frameworks, Personalization Strategies, Revenue Forecasting, Anomaly Detection, Forecasting Metrics (CAC, Signup, Trial, Revenue)

Cloud & Data Engineering: AWS (Lambda, Glue, S3, EC2, RDS/PostgreSQL), Azure (familiarity), GCP, ETL/ELT Pipelines, Data Warehousing (Snowflake), PostgreSQL, MySQL, MongoDB, dbt Core, Apache Airflow. Data Architecture, Jira, Service Now, Git Workflows, Delta Lake, Redshift,

BI & Visualization: Tableau, Power BI, Google Data Studio

Leadership & Project Management: Team Leadership & Mentoring, Global Team Coordination, Stakeholder Management, Agile Methodologies, Cost Optimization, Project Delivery, Process Improvement, Technical Documentation, Resource Allocation

Certifications: Deep Learning Specialization – Andrew Ng, Coursera