Goutham Kishore Krishnamoorthy

Arlington, TX | +1-682-376-8423 | gouthamkishore.k@gmail.com | LinkedIn | GitHub | Portfolio

Professional Summary

Data Engineer with 4+ years of experience building scalable data pipelines, ML-driven analytics solutions, and cloud-based architectures across banking and enterprise domains. Expertise in Python, SQL, Apache NiFi, and AWS services (S3, Lambda, Glue, Redshift) with hands-on experience in machine learning, NLP, and emerging AI technologies including prompt engineering and LLM integration. Proven ability to translate complex business requirements into production-grade data solutions, supporting advanced analytics, predictive modeling, and AI-powered applications. Strong collaborator with experience working in agile environments and delivering measurable business outcomes.

Technical Skills

- **Programming & Query**: Python (Pandas, NumPy, Scikit-learn), SQL, R, Scala, Java
- **ML & AI**: Machine Learning, Deep Learning (TensorFlow, PyTorch, Keras), NLP (TF-IDF, Transformers), XGBoost, Statistical Analysis, LLM Integration, Prompt Engineering, RAG (Retrieval-Augmented Generation)
- **Data Engineering**: Apache NiFi, ETL/ELT Pipelines, Data Warehousing, Apache Kafka, PostgreSQL, Oracle, MongoDB
- **Cloud & DevOps**: AWS (S3, Lambda, Glue, Redshift, RDS, EC2, SageMaker), Docker, Kubernetes, CI/CD (Jenkins), Helm, Git
- **BI & Visualization**: Tableau, Power BI, Apache Superset, Plotly, MicroStrategy
- **Frameworks & Tools**: React, Node.js, TypeScript, Express, LangChain, Spring Boot, Auth0, JIRA

Additional Skills & Interests

- **Version Control:**Git, GitHub, GitLab with experience in collaborative development workflows
- **Agile Methodologies:**Scrum, Kanban, Sprint Planning, Daily Standups
- **Communication:**Technical documentation, stakeholder presentations, cross-functional collaboration
- **Continuous Learning:**Active contributor to data engineering communities, regular attendee of AI/ML meetups

Professional Experience

Compass Group USA **Aug 2024 - Dec 2025**

Business Intelligence Engineer

- Led end-to-end analytical initiatives by translating ambiguous business questions into measurable objectives, integrating Oracle POS and OrderMaestro data using Python and SQL to produce enterprise-grade datasets covering sales, inventory, labor, and financial performance
- Delivered measurable business impact by building demand forecasting and baseline regression models in Python (scikit-learn, XGBoost), reducing stockouts and excess inventory by 15% across market locations through predictive analytics
- Engineered 30+ features from time-series data, performing deep exploratory data analysis to uncover demand seasonality, cost drivers, and workflow bottlenecks that directly informed inventory and labor planning decisions
- Influenced executive decision-making by presenting clear, data-driven insights to senior leadership through interactive dashboards (Power BI, Plotly), directly supporting budgeting, profit optimization, and weekly inventory governance
- Automated reporting workflows using Python scripts and AWS Lambda functions, reducing report generation time by 60% and enabling near-real-time decision support

CGI **Aug 2021 - Jan 2024**

Data Engineer

- Built and automated 20+ production-grade data pipelines using Apache NiFi and Python, reducing manual data processing effort by 40% and enabling near-real-time data ingestion across SQL databases (Oracle, PostgreSQL), file systems, and Kafka streams
- Leveraged AWS cloud services extensively for data storage (S3), transformation (Glue, Lambda), and analytics (Redshift), designing scalable ETL workflows that processed 10M+ records daily for banking clients
- Modernized enterprise data workflows by migrating and transforming data across Oracle and PostgreSQL using Python-driven ETL processes, supporting analytics and regulatory reporting at scale
- Improved reporting accuracy and speed by refactoring transformation logic and optimizing SQL queries, increasing dashboard accuracy by 25% and cutting report generation time from hours to minutes
- Developed data quality frameworks using Python and SQL to validate data integrity, implementing automated testing and monitoring that reduced data quality incidents by 30%
- Enabled business-ready analytics by designing executive-facing dashboards in Apache Superset and Power BI, delivering financial and operational insights consumed by C-level stakeholders
- Collaborated on AI/ML initiatives by preparing clean, analytics-ready datasets for data science teams, supporting risk modeling and predictive analytics projects for banking clients
- Supported revenue and growth initiatives through technical pre-sales collaboration, conducting client demos of data solutions, while maintaining strict data governance and compliance standards (GDPR, SOX)
- Utilized Git for version control and participated in agile ceremonies, contributing to CI/CD pipeline improvements using Jenkins

Education

University of Texas at Arlington **Jan 2024 - Dec 2025**
Master of Science, Data Science
- **GPA:** 3.8/4.0
- **Coursework:** Machine Learning, Deep Learning, Big Data Management, Cloud Computing, Analytics & Visualization, Natural Language Processing, Distributed Systems

Vel Tech High Tech Dr. RR Dr. SR University **Jun 2017 - Jul 2021**
B.Tech, Computer Science Engineering
- **GPA:** 3.08/4.0
- **Coursework:** Foundations of Computing, DBMS, Design Analysis and Algorithm, Data Structures, Artificial Intelligence, Information Security and Retrieval Techniques

Key Projects

AIO AI Operator Platform
- Architected production-grade AI platform delivering real-time insights, automated summaries, and performance metrics through event-driven architecture with TypeScript/Express backend (8+ REST endpoints)
- Integrated LLM capabilities using prompt engineering techniques to generate contextual summaries and automated insights from operational data
- Developed React dashboard with interactive visualizations (Plotly, Recharts) and secure Auth0 OAuth2/JWT authentication
- Deployed on cloud infrastructure (Render/Vercel) with production-ready error handling, monitoring, and CI/CD pipelines

Sentiflux - AI-Powered Sentiment Analysis Platform
- Built full-stack sentiment analysis application using React, Node.js, and Python, leveraging NLP models to analyze text sentiment in real-time
- Implemented RAG (Retrieval-Augmented Generation) architecture to enhance sentiment context understanding by integrating vector database (FAISS) for semantic search and retrieval
- Developed deep learning models using TensorFlow and PyTorch for multi-class sentiment classification, achieving 78% accuracy on customer review datasets
- Created RESTful API backend with Express.js and integrated multiple ML models for comparative sentiment scoring and trend analysis

ServiceNow Ticket Analytics & Predictive Assignment Modeling
- Led comprehensive analysis of 3 years of ServiceNow data (50K+ tickets) using Python and NLP techniques (TF-IDF, BERT transformers) to predict optimal assignment groups and reduce ticket reassignment rates by 22%
- Built machine learning pipeline using scikit-learn and TensorFlow for text classification and feature engineering from unstructured ticket descriptions
- Developed interactive dashboards (Plotly, Power BI) revealing workflow inefficiencies, average resolution times, and SLA compliance metrics that informed process optimization initiatives

Retail Sales Forecasting & Demand Planning
- Built ensemble predictive models (Random Forest, XGBoost, ARIMA) achieving 92% accuracy for monthly retail sales forecasting using 5 years of historical transaction data
- Experimented with deep learning approaches using Keras and LSTM networks for time-series forecasting, comparing performance against traditional statistical methods
- Designed comprehensive Power BI dashboards with drill-down capabilities for demand planning insights, enabling inventory optimization and promotional planning decisions
- Performed feature engineering from time-series data including holiday effects, seasonality decomposition, and external economic indicators

Certifications & Training

- **Deep Learning Specialization - TensorFlow, Keras, Neural Networks: DeepLearning.AI, Coursera**
- **BI Tools Mastery: Tableau, Power BI, MicroStrategy, ArcGIS: Udemy, Coursera, YouTube**
- **Scala & Functional Programming: Coursera**
- **Machine Learning Specialization: Coursera**
- **Prompt Engineering for Developers: DeepLearning.AI**