

Venkata Sai Bharadwaj Velamakanni

Jersey City, NJ | venkata242000@gmail.com | +1 714 860 2987 | Portfolio | [linkedin.com/in/venkatavelamakanni](https://www.linkedin.com/in/venkatavelamakanni) | Github

Professional Summary

Results-driven **Data Engineer** with 2+ years of experience in **SQL, Python, Spark, and Airflow**. Expertise in **building scalable data pipelines, real-time streaming, and cloud infrastructure**. Passionate about **optimizing ETL workflows, data modeling, and automation for data-driven decision making**..

Experience

Senior Business Intelligence Engineer, BulkMagic – Jersey City, NJ Jan 2025 – Present

- Engineered a real-time data pipeline using Apache Kafka, AWS Lambda, and Spark reducing data latency by 50%.
- Designed and optimized ETL workflows in Airflow & AWS Glue, improving data processing efficiency by 40%.
- Built scalable data lake architectures using Delta Lake & S3, supporting high-volume analytics and storage optimization.
- Developed real-time reporting dashboards in Tableau & Power BI, improving executive decision-making speed by 30%.

Business Analyst II (Graduate Teaching Assistant) , CSUF – Fullerton, CA Aug 2024 – Jan 2025

- Taught regression, Monte Carlo simulation, and forecasting, improving students' data interpretation skills, leading to a 15% rise in test scores.
- Developed interactive Excel-based decision models using PivotTables, Solver & scenario analysis, enabling students to optimize business forecasting & risk modeling.
- Designed a data visualization guide for selecting optimal chart types, enhancing project clarity by 20%.
- Led SQL-based data manipulation workshops, covering joins, subqueries, and window functions, strengthening students' analytical proficiency.

Data Engineer, Vistex & Suven Consultants – India Jan 2021 – Dec 2022

- Built an audience segmentation model using Python, SQL, and K-Means clustering, improving ad personalization and increasing CTR by 25%.
- Automated data pipelines in SQL & SAP, optimizing rebate and compensation models for enterprise clients, reducing processing errors by 30%.
- Designed A/B testing frameworks & Power BI dashboards to monitor marketing campaigns, decreasing budget by 15%.
- Developed standard analytics for GTMS cloud systems, enhancing data accuracy and dashboard reporting success by 90%.

Projects

End-to-End Data Engineering & AI/ML Pipeline for Fraud Detection github.com/fraud-detection

- Developed a real-time fraud detection pipeline using Apache Kafka & PySpark, processing millions of transactions daily.
- Deployed a Random Forest model on AWS SageMaker, achieving 94% precision in fraud detection
- Automated workflows with Airflow & Docker, reducing manual processing time by 70%.

NYC Taxi Data Pipeline

- Built an ETL pipeline using Azure Data Factory & Databricks, reducing data processing time by 30%.
- Transformed raw taxi trip data into Parquet format, improving query performance by 20%.
- Designed Power BI dashboards to provide insights on fare trends, trip patterns, and operational efficiencies.

Technologies

Programming: Python (Pandas, PySpark), SQL (PostgreSQL, Redshift, Snowflake), R

Data Engineering Tools: Apache Kafka, Spark, Airflow, Databricks

Cloud & Infrastructure: AWS (S3, Glue, Lambda, EC2), Azure Data Factory

Big Data & ETL: Data Warehousing, Delta Lake, Streaming Pipelines

Visualization: Tableau, Power BI

DevOps & Automation: Docker, Kubernetes, Git, CI/CD

Education

California State Univeristy, Fullerton, M.S. in Information Systems - Business Analytics Jan 2023 – Dec 2024

- Coursework:** Computer Architecture, Comparison of Learning Algorithms, Computational Theory