# Vraj Shah

+19029941436 | shahvraj2016@gmail.com | linkedin.com/in/vraj-shah | github.com/shah-vraj | portfolio | Blog

## Professional Summary

Highly motivated Data Engineer & Software Developer with expertise in Python, SQL, ETL pipelines, and cloud technologies. Skilled in designing scalable data processing solutions, optimizing database performance, and automating data ingestion and transformation workflows. Passionate about leveraging Azure Databricks, Synapse, and Power BI to extract meaningful insights and drive data-driven decision-making. Currently deepening expertise in big data analytics and Azure data services.

## Technical Skills

**Programming & Data Processing**: Python (Pandas, NumPy), SQL, FastAPI, Kotlin, Java, JavaScript

**Data Engineering & ETL**: Data Cleaning, Transformation, ETL Pipelines, Data Warehousing

**Databases**: PostgreSQL, MySQL, SQL Server, MongoDB, SQLite, Oracle

**Cloud & DevOps**: Azure (Learning Databricks & Synapse), AWS, GCP, Docker, Kubernetes, CI/CD (GitHub Actions, GitLab CI)

**Big Data & Analytics**: Apache Spark (Databricks), Data Governance, Power BI (Learning)

**Testing & Debugging**: JUnit, Mockito, API Performance Tuning, Debugging, Git

**Soft skills**: Strong communication, Problem solving, Teamwork and collaboration, Adaptability, Attention to detail, Critical thinking, Time management, Ability to learn continuously

## Work Experience

**Simform Solutions** **Dec 2021 - Nov 2023**

*Software Engineer* *Ahmedabad, India*

- **Designed and developed scalable ETL pipelines** using Python (Pandas) and SQL, efficiently processing and transforming large datasets, **improving data ingestion and transformation speed by 40%**.
- **Developed high-performance REST APIs** using FastAPI and Spring Boot, facilitating seamless real-time data processing and data integration between multiple services.
- **Optimized complex SQL queries** and indexing strategies in PostgreSQL & MySQL, **reducing execution times by 30%** for analytics workloads and enhancing database efficiency.
- **Implemented data validation and transformation automation**, ensuring data consistency and integrity across multiple data sources and reducing manual intervention.
- **Streamlined CI/CD pipelines** using GitHub Actions & GitLab CI, automating deployments, minimizing downtime, and **enhancing deployment efficiency by 50%**.
- **Collaborated with cross-functional teams**, translating business requirements into efficient data engineering solutions, **leading to better insights and decision-making processes**.

## Education

**Dalhousie University** **Jan 2024 - Sep 2025**

*Master of Applied Computer Science (Co-op Candidate)* | ***GPA: 4.07/4.3*** *Halifax, Canada*

**Relevant coursework**: Database Management & Warehousing, Cloud Computing, Cloud Architecting

**Charusat University** **Jun 2018 - Apr 2022**

*Bachelor of Computer Engineering* | ***GPA: 8.7/10*** *Gujarat, India*

## Projects

**ETL on Tweets Data** | *Source Code* **JAVA | OpenCV | MySQL | GCP**

- **Built a distributed transaction system** using Java, enabling dynamic redirection of queries across multiple MySQL VMs in GCP, supporting **seamless ETL processes for efficient data processing and real-time interactions**.
- **Implemented horizontal and vertical data fragmentation techniques** for a distributed MySQL database ("SocialMedia"), optimizing query performance and ensuring scalability in a cloud-based environment.
- **Integrated and processed a large dataset** of tweets from Kaggle, **automating the ETL pipeline and enhancing system reliability**, while ensuring robust and scalable backend architecture that aligns with distributed database management best practices.

**K8s Microservices** | *Source Code* **Spring Boot (Kotlin) | Kubernetes | Docker | CloudBuild**

- **Implemented Kubernetes-based microservices architecture**, creating a highly available system with two Spring Boot containers, **ensuring secure communication and fault tolerance between services**.
- **Deployed microservices using Kubernetes** and automated deployment processes with **CloudBuild**, **improving deployment speed and scalability by 40%.**
- **Worked with Docker containers** to facilitate easy application deployment, **increasing the speed of testing and production cycles** while ensuring a consistent environment across development and production.

**ServiceHub** | *Source Code* **ReactJS | Spring Boot (JAVA) | MySQL**

- Developed a scalable web application using **ReactJS**, **implementing modular UI components for a seamless user experience**. Integrated lazy loading and performance optimizations, **reducing page load time by 40%**.
- **Applied Object-Oriented Programming principles** (e.g., classes, objects, inheritance) to design modular and reusable code for the backend, ensuring maintainability and scalability.
- **Architected backend using Java Spring Boot and integrated it with a ReactJS frontend, achieving 99.9% uptime** and enabling seamless real-time interactions among over 1,000 active users.