

GHAZI RAHMAN SHAIK

Los Angeles, CA | (424)-471-8817 | ghazishaik5@gmail.com | <https://www.linkedin.com/in/ghazi-rahman-shaik/>

Education

- Master of Science | California State University, Northridge** *Aug 2022 - Dec 2024*
Computer Science
- Bachelor of Technology | K L University, India** *Jun 2017 - May 2021*
Electronics and Communication Engineering

Work Experience

Associate Data Engineer | Cognizant Technology Solutions *Dec 2020 – Aug 2022*

- Developed ETL workflows using AWS Glue and Apache NiFi increasing data accuracy by 15% and saving the company approximately **\$50,000** annually in data correction costs.
- Implemented Informatica MDM solutions, standardizing and integrating data from multiple sources, enhancing data consistency and reliability by 25%, leading to an estimated cost saving of **\$75,000** per year.
- Developed complex SQL queries and managed database performance, optimization, and troubleshooting, reducing query response time by 30%.
- Designed and implemented data pipelines to handle large-scale data ingestion, processing, and storage, supporting data-driven applications and analytics.
- Collaborated with data scientists and analysts to understand data needs and deliver high-quality datasets for machine learning models and analytics.
- Optimized and maintained data lakes and data warehouses, improving data retrieval speed by 30% and reducing storage costs by 20%, resulting in annual savings of **\$50,000**.
- Employed data governance best practices, including data quality checks, lineage tracking, and security protocols, to safeguard data integrity and compliance.
- Automated data workflows using Apache Airflow, reducing manual intervention and improving workflow efficiency by 40%.

Data Analyst Intern | Airports Authority of India *March 2019 – May 2019*

- Collaborated with Air Traffic Control to collect and analyze communication data, improving safety protocols and reducing communication errors by 15%.
- Assisted in developing and optimizing data logging systems for advanced navigational equipment, including Instrument Landing System (ILS) and DVOR, ensuring accurate data recording and retrieval.
- Conducted data quality assessments and implemented data cleaning procedures, enhancing the reliability of navigational data by 20%.

Technical Skills

Languages	Frameworks and Libraries	Database Management	DevOps/Cloud
Python, Java SQL, JavaScript	Pandas, Scikit-learn, Matplotlib, TensorFlow, PySpark	MySQL, PostgreSQL, ETL, Data Warehousing	Docker, Kubernetes, Git, CI/CD AWS - S3, EC2, Redshift

Projects

- Integrated Data Warehouse for Hospital Management Systems:** Developed an integrated data warehouse model using Snowflake, Informatica PowerCenter, SQL, and Spark to streamline patient data management, reducing data retrieval time by 30% and improving reporting accuracy by 20%.
- Speech Detection Using Smartphone Gyroscope:** Conducted research using Java (Android Studio) and Python (Jupyter Notebook) to detect speech via smartphone gyroscope, achieving 85% accuracy in identifying speech patterns with Machine Learning (DTW).

Certifications

- Java J2EE by Wipro, Informatica PowerCenter, Advanced DevOps by Visual Path.