Niroj Koirala

6469918577 Nirojkoirala7@gmail.com Queens, NY

Motivated and analytical Computer Science graduate with hands-on experience in data engineering, cloud computing, and analytics. Completed a 7-month Data Engineering Bootcamp and a Data Tech Internship, applying Python, SQL, and Azure Data Factory to build scalable data pipelines. Published first-author IEEE research on IoT-based smart systems. Graduated with High Honors (GPA: 3.96) and recognized as a CIRE Research Scholar, NSF Scholarship recipient with a strong foundation in teamwork, leadership, and academic excellence.

SKILLS

- ETL development
- Data analysis/validation
- Data cleaning
- Amazon web services (AWS) integration
- SQL proficiency
- Real-time analytics
- Python

EXPERIENCE

DATA TECH INTERN **Queens, NY**
TOQC *June 2024 - September 2024*

- Assisted in building and maintaining ETL pipelines using Azure Data Factory and SQL Server.
- Automated data ingestion from APIs and internal systems into Azure Data Lake, ensuring accuracy and reliability.
- Used Python (Pandas) and SQL for data transformation and quality checks prior to analysis.
- Helped develop Power BI dashboards to visualize operational performance.
- Collaborated with senior data engineers to troubleshoot integration and workflow issues.

IT SUPPORT **Queens, NY**
York College, CUNY *October 2024 - May 2025*

- Provided IT support to faculty and students, resolving hardware, software, and network issues.
- Managed user accounts, system setup, and troubleshooting through Microsoft 365 and Active Directory.
- Authored internal documentation to improve help desk efficiency and knowledge sharing.

EDUCATION

BACHELOR OF SCIENCE IN COMPUTER SCIENCE, BUSINESS ADMINISTRATION **GPA 3.89**
York College, City University of New York (CUNY)

Awards & Honors

- Dean's List
- Summa cum laude
- Excellent Project Award
- NSF Scholar
- Phi theta Kappa Honor society

Extracurricular Activities

- NJCAA Student Athlete
- Technology Student Association (TSA)
- Peer Mentor

CERTIFICATIONS

- AWS Certified Data Engineer – Associate (DEA-C01)
- Microsoft Certified: Azure Data Fundamentals (DP-900)
- AWS Certified Cloud Practitioner (CLF-C02)
- IBM Data Engineering Professional Certificate

- Takeo Data Engineering Bootcamp Certificate

RESEARCH AND PUBLICATIONS

- Smart Traffic Light Control with Emergency Vehicle Detection, **IEEE** Conference Publication.
- Conducted a year-long research project addressing emergency vehicle delay caused by inefficient traffic light systems.
- Designed an IoT sensor–based and computer vision system to detect emergency vehicles and dynamically optimize traffic lights.
- Served as First Author on a peer-reviewed IEEE paper demonstrating measurable improvements in emergency response simulation.
- Conducted under the CIRE (CUNY Interdisciplinary Research Experience) program; received a $6,000 research award.
- Presented findings at SUSTech (California), EVEN (Atlanta), and multiple CUNY symposiums.
- Utilized Python, Arduino, and real-time data processing frameworks for modeling and analysis.

TRAINING

Takeo Data Engineering Bootcamp, 03/01/25, 09/01/25

- Completed a 7-month, project-based data engineering program focused on data pipeline design, orchestration, and data warehousing across AWS and Azure.
- Built ETL/ELT pipelines using Python (PySpark), SQL, and Apache Spark to ingest, clean, and transform structured and semi-structured data.
- Designed and implemented the Bronze–Silver–Gold data lake architecture using Azure Data Factory (ADF), Azure Data Lake, and Azure SQL Database.
- Integrated real-time data streams using Kafka and Spark Structured Streaming, enabling near real-time analytics.
- Automated ingestion and transformation in AWS using Glue, Snowpipe, and Lambda, improving pipeline efficiency.,
- Applied star/snowflake schema modeling and implemented data warehousing in Snowflake for analytics.,
- Improved performance on simulated datasets by ~60% through partitioning, caching, and Parquet optimizations.
- Used Airflow to orchestrate batch and streaming workflows, ensuring pipeline reliability.
- Implemented data validation with Great Expectations to enforce schema consistency and data quality.
- Created Power BI dashboards from curated (Gold layer) datasets to visualize KPIs and insights., Practiced Git-based version control and CI/CD concepts for collaborative development.
- Developed strong understanding of data governance, metadata management, and cloud security best practices.