

# BHANU TEJA KURAKULA

Chicago, IL (Open to Relocation) | k.bhanuteja2001@gmail.com | [Linkedin](#) | [Github](#) | +1-(217)926-2758



## SUMMARY

**Data Engineer** with **2.5+ YoE** working across Banking, Finance, Healthcare, AgTech and EduTech sectors. Expertise in designing **robust data models**, developing scalable **ETL/ ELT** data pipelines delivering high-quality data through rigorous quality checks. Experience collaborating with technical and non-technical stakeholders, driving impact, clarity and actionable insights. Proficient in **SQL, Python, Tableau, Data Modeling, Snowflake, Databricks** and **Airflow**.

## WORK EXPERIENCE

### Data Engineer - AGCO Corporation

June 2024 - August 2025

- Developed data models, designed **Tableau** dashboard & built a SQL based **ELT** pipeline supporting **40+ stakeholders** to track **80+ product metrics** & other operational KPIs.
- Collaborated with various Product Managers in defining product metrics, data quality standards, SLAs and other recommendations. Looped the stakeholder in every step, resulting in better business insights.
- Owned 10 ETL data pipelines on AWS**, standardized large volumes of data from various data sources across the EU with schema enforcement & data quality checks, used CloudFormation to automate deployment.
- Led migration of **10 legacy Hadoop** batch data pipelines to **AWS**, collaborated with stakeholders to set data validation standards, documented and supported the end users from business in adopting the migrated workflows.
- Analyzed the market trend for tractor sales, automated report generation producing monthly summary using **NLP** and **statistical analysis** from incoming **qualitative & quantitative** data. Communicated analysis to stakeholders.

### Data Analyst - Polygon Health Analytics

May 2025 - August 2025

- Integrated LLM-based **classification** and deduplication feature into an existing Java vaccine-newsletter service using OpenAI, Gemini APIs and rolled it out after rigorous **statistical validation** to improve tag quality.
- Implemented near-duplicate detection with OpenAI embeddings, performed **cosine similarity** and utilized **HDBSCAN** to reduce redundancy of articles, validated with backfill runs before a staged rollout to production.
- Gathered business requirements and developed a **FinOps dashboard** for leadership to track token utilization, cost per API call, and usage trends, with drill-down capabilities by model, user, and service.

### Data Engineer - Cognizant

September 2022 - July 2023

- Ensured Snowflake data warehouse reliability by monitoring **200 key metrics** and **10 Airflow** DAGs, actively resolving job failures on-call and ensured timely report generation ahead of business use, meeting SLAs.
- Developed a machine learning model for fraud claims, providing **predictive analytics** for clients, performed EDA on a complex internal dataset, performed **A/B testing** to deploy the best model.
- Identified recurring bottlenecks, collaborated with stakeholders to **optimize SQL queries**, resulting in a **20 percent** reduction in query execution time, reducing delay by **20 minutes** on a year over year basis.

### Data Engineer Intern - Cognizant

March 2022 - August 2022

- Built an ETL pipeline on **Databricks** to retrieve and process airline safety datasets, utilizing **Delta Lake** for scalable data storage and Spark SQL/PySpark for efficient data transformation and processing.
- Automated pipeline orchestration through **Databricks Jobs**, while delivering actionable insights through interactive data visualizations, comprehensive error handling, and data validation mechanisms.
- Engineered an **end-to-end data pipeline on AWS**, used **Glue** for transformation, **S3** and **Redshift** for storage, **CloudWatch** for logs & monitoring, **Lambda**, **SNS** and **EventBridge** for notifications.

## EDUCATION

### University of Illinois, Urbana-Champaign (UIUC)

August 2023 - May 2025

Master of Science - Information Management

**Relevant Coursework** : Data Modeling & Database Design, Data Warehouse & Business Intelligence, Data Statistics, Visualization

### Amrita University

July 2018 - June 2022

Bachelor of Technology - Electronics and Communication Engineering

## SKILLS

|                                   |                                                                                                           |
|-----------------------------------|-----------------------------------------------------------------------------------------------------------|
| <b>Programming Languages</b>      | : Python, Advanced SQL, Spark, Java                                                                       |
| <b>Platforms &amp; Frameworks</b> | : AWS, Snowflake, Databricks, Kafka, Airflow, Tensorflow                                                  |
| <b>Tools</b>                      | : Tableau, PostgreSQL, Docker, Github Actions(CI/CD),                                                     |
| <b>Miscellaneous</b>              | : Data Modeling, OOPs, Data Structures, Algorithms, APIs                                                  |
| <b>Certifications</b>             | : <u><b>AWS Solution Architect Associate, AWS Machine Learning Specialty, Databricks Certified DE</b></u> |

## PROJECTS

### Kafka-Based Real-Time Data Processing Pipeline | Confluent Kafka, docker

Developed a Kafka-based pipeline, processed data, ensured privacy compliance, and provided real-time insights.

### Coupon Accepting MLOps Pipeline | MLFlow, EC2, S3, RDS, Terraform, Docker, Prefect, Github Actions

Built a MLOps pipeline, enabling experiment tracking, orchestration and automated deployment using CI/CD.