BHANU TEJA KURAKULA

📞 217-926-2758 | ✉ bhanutk2@illinois.edu | in linkedin.com/bhanukurakula | ⌂ github.com/bhanuteja2001 | 📍 United States

EDUCATION

University of Illinois at Urbana-Champaign August 2023 – May 2025
Master of Science in Information Management, (Data Science and Analytics) (GPA: 3.94/4)
Relevant Courses: Data Modeling, Database Design & Prototyping, Data Warehouse and BI

Amrita University July 2018 – June 2022
Bachelor of Technology in Electronics and Communication Engineering (GPA: 3.34/4)

TECHNICAL SKILLS & CERTIFICATIONS

Programming Languages	: Python, SQL, Pyspark, R
Platforms & Frameworks	: Snowflake, AWS, Databricks, Azure, Alteryx, Tensorflow, Kafka, SSMS, Kubernetes
Tools	: DBT, Dagster, Airflow, Terraform, Tableau, Git, Grafana, Postgres, MongoDB, Docker
Certifications	: AWS Solution Architect Associate, AWS Machine Learning Specialty

WORK EXPERIENCE

AGCO Corporation June 2024 – Present
Data Engineering Intern

- **Engineered 10+ complex ETL pipelines** using **AWS Glue, S3,** and **lambda**, ensuring seamless data integration and conforming, resulting in standardized tractor sales data across the European Region.
- Utilized **PySpark** to improve data processing efficiency by translating **Alteryx** workflows into multiple glue jobs, implementing comprehensive **error-handling & unit testing** to uphold data quality across all stages.
- Implemented infrastructure provisioning by adopting **IaC** approach with **CloudFormation** and **YAML**, enhancing scalability, automation, and reducing deployment time while minimizing manual setup errors.
- Developed and orchestrated the workflows using **Glue Workflow** & **Step-functions**, utilizing **SNS** for automated notifications and **CloudWatch** for monitoring, reducing manual intervention.
- Conducted **data discovery** and created **data lineage maps** to facilitate workflow migration. Collaborated with stakeholders to document data flows and business requirements, ensuring a seamless transition through **UAT**.

Cognizant Technology Solutions

Data Engineer September 2022 – July 2023

- Enhanced **production data warehouse** reliability by leveraging **Snowflake**, **DBT**, **Dagster** and **Grafana** to monitor and troubleshoot issues, overseeing 200+ metric dependencies and guaranteeing continuous data flow.
- Improved performance by **optimizing SQL queries**, resulting in 20% reduction in query execution. Reduced tableau scorecard delay by 40% by collaborating with downstream and upstream teams.
- Spearheaded a proof of concept (POC) for detecting fradulent insurance claims using **Machine Learning**, showcasing innovation for future production integration. Demonstrated potential for improving data quality.

Data Engineer Intern March 2022 – July 2022

- Designed and implemented a scalable, fault-tolerant data pipeline for datasets with over 4M+ records, leveraging **Lambda** and **Glue** for ETL transformations, **Redshift** as the Data Warehouse, and S3 as the data source.
- Executed complex data processing and analysis tasks in **Databricks** on datasets exceeding 100GB, leveraging distributed computing and **Delta Lake optimizations** to enhance performance and efficiency.

PROJECTS

Credit Card Rewards Pipeline | Terraform, Docker, ECR, Pyspark, Redshift, Glue, S3, Lambda, Looker Studio

- Developed a **serverless data pipeline** on AWS for analyzing credit card rewards using Open Banking API.
- Implemented ETL processes to transform the data, optimizing the schema for analysis and insights generation.
- Orchestrated the workflow using **Step functions** and used **Looker studio** to visualize.

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

- Developed **MLOps** pipeline using **XGBoost** for coupon prediction, leveraging AWS and Terraform.
- Implemented **MLflow** for experiment tracking, Prefect for workflow orchestration, and Docker for containerization.
- Established CI/CD pipeline with GitHub Actions, automating testing and deployment processes.