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 (UIUC)   August 2023 – May 2025
Master of Science in Information Management, (Data Science and Analytics)   (GPA: 3.96/4)
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	: Azure, AWS, Databricks, Snowflake, Alteryx, Confluent Kafka, SQL Server
Tools	: Docker, Tableau, Git, Grafana, Postgres, MongoDB, Airflow, DBT, Terraform
Certifications	: Data Engineering with Databricks, AWS Solution Architect Associate

Work Experience

AGCO Corporation   June 2024 – Present
Data Engineering Intern

- **Engineered 10+ complex ETL pipelines** using Glue, S3 and RDS, ensuring seamless data integration and conforming, resulting in standardized tractor sales data across multiple regions in Europe.
- Utilized **PySpark** to improve data processing efficiency by translating **Alteryx** workflows into multiple glue jobs, implementing comprehensive **error-handling & unit testing** to maintain data quality in all stages.
- Improved infrastructure provisioning by adopting a **IaC** with **CloudFormation** and **YAML**, enhancing scalability, automation, and reducing deployment time by 30% 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 **Logical data maps** to facilitate workflow migration. Collaborated with stakeholders to document data flows and business requirements, ensuring a seamless transition through **UAT**.

Cognizant Technology Solutions   September 2022 – July 2023
Data Engineer

- Enhanced **production data warehouse** reliability by leveraging **Snowflake**, **DBT**, **Airflow** and **Grafana** to monitor and troubleshoot issues, overseeing 200+ metric dependencies and guaranteeing continuous data flow.
- Spearheaded a proof of concept (POC) for detecting fradulent insurance claims, showcasing innovation for future production integration. Demonstrated potential for improving data quality.

Data Engineer Intern   March 2022 – July 2022

- Engineered a fault-tolerant data pipeline on **Azure** for datasets with over 4M+ records, leveraging **Functions** and **Data Factory** for ETL transformations, **Synapse** as the Warehouse, and **Blob Storage** as the source.
- Executed complex data processing and analysis tasks in **Databricks** on datasets exceeding 100GB, leveraging distributed computing and **Delta Lake optimizations** to improve performance and efficiency.

PhysioAI   January 2022 – February 2022
Data Science and Analytics Intern

- Developed a yoga assistant using Azure, TensorFlow, and MediaPipe to detect and classify over **30+** yoga poses.
- Captured user progress through **40+** metrics, managed data on PostgreSQL & created a tracker using Tableau.

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.

Credit Card Rewards Pipeline | Terraform, Docker, ECR, Pyspark, Redshift, Glue, S3, Lambda, Looker Studio
Developed a serverless pipeline, transformed data, optimized schema, orchestrated workflow & visualized insights.

Coupon Accepting Prediction Pipeline | MLFlow, EC2, S3, RDS, Terraform, Docker, Prefect, Github Actions
Built an MLOps pipeline, enabling experiment tracking, workflow orchestration, and automated deployment.