

# BHANU TEJA KURAKULA

United States | bhanutk2@illinois.edu | [linkedin](#) | [github](#) | [Portfolio](#) | +1-(217)926-2758  
AWS Solutions Architect-Associate | AWS Machine Learning - Specialty | Databricks Certified Data Engineer

## SUMMARY

Data professional who turns chaos into clarity. Experience across the full data spectrum—from engineering pipelines to delivering predictive insights that matter. Not just educated, but battle-tested through real-world challenges across multiple industries.

## WORK EXPERIENCE

### Data Engineering Intern - AGCO Corporation

June 2024 - Present

- Engineered over **10+** complex ETL pipelines using various AWS services like Glue, S3, RDS and lambda, ensuring seamless data integration and conforming, resulting in standardized tractor sales data across the European Region.
- Translated **Alteryx** workflows into **PySpark** jobs, improving efficiency by implementing comprehensive error handling & unit tests to maintain data quality across all stages, also **saving \$20K in licensing costs**, making it cost effective.
- Implemented infrastructure provisioning by adopting **Infrastructure as Code(IaC)** approach using Cloud Formation, enhancing scalability, automation, and reducing deployment time by **30%** while minimizing manual setup errors.
- Collaborated with stakeholders to build logical data maps for migration and documented user manuals to ensure smooth handover. In addition, built over **15+ tableau prep flows** and a **dashboard** as part of decommission efforts.

### Data Engineer - CTS

September 2022 - July 2023

- Enhanced production data warehouse reliability by leveraging **Snowflake**, **DBT**, **Airflow**, and **Grafana** to monitor and troubleshoot issues, overseeing **200+** metric dependencies and ensuring uninterrupted continuous data flow.
- Improved performance by writing **optimized SQL queries** resulting in 20% reduction in query execution time. Reduced tableau scorecard delay by **15-30 minutes** on a yearly basis by collaborating with downstream and upstream teams.
- Spearheaded a proof of concept (POC) for detecting **fraudulent insurance claims** using **machine learning**, showcasing innovation for future production integration. Demonstrated potential for AI Integration into the production warehouse.

### Data Engineer Intern - CTS

March 2022 - July 2022

- Engineered a fault-tolerant data pipeline on Azure for a dataset with over **4M+** records, leveraging **Data Factory** and **Functions** for ETL transformations, **Synapse** as the Warehouse, **Blob Storage** as the source & **Power BI** for dashboard.
- Implemented a batch data pipeline on databricks using **Delta Live Tables (DLT)** to process and transform a **2GB** dataset, and created interactive dashboards using Databricks SQL for data visualization and reporting.

### Data Science and Analytics Intern - PhysioAI

January 2022 - February 2022

- Developed a real-time yoga assistant using **Azure**, **TensorFlow**, and **MediaPipe** to accurately detect and classify over **30+** yoga poses, integrated a voice assistant for real-time guidance, enhancing user experience and improving posture accuracy.
- Captured user progress through **40+** metrics and effectively managed user data on **Postgres**, used **Tableau** to create a progress tracker. Presented the POC during the monthly incubator meet to secure investors interest

## 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, Prefect

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

## EDUCATION

### University of Illinois at Urbana-Champaign

August 2023 - May 2025

Master of Science - Information Management (Data Science & Analytics)

(GPA 3.94/4.0)

**Relevant Coursework** : Data Modeling, Database Design, Data Warehouse and Business Intelligence, Data Governance, Statistics

### Amrita Vishwa Vidyapeetham (Amrita University)

July 2018 - June 2022

Bachelor of Technology - Electronics and Communication Engineering

(GPA 3.35/4.0)

## SKILLS

### Programming Languages

: Python, SQL, Pyspark, R

### Platforms & Frameworks

: AWS, Databricks, Alteryx, Snowflake, SQL Server, Confluent Kafka, Office, Azure

### Tools

: Docker, Tableau, Airflow, Terraform, Git, Postgres, MongoDB, DBT, Grafana, PowerBI

### Libraries

: Numpy, Pandas, Scikit-learn, Matplotlib, Streamlit, Tensorflow, Plotly, Faust

## ACHIEVEMENTS

**Honeywell Hackathon Winner** : Secured **1st position** out of 180 teams for developing a POC on smart traffic system, **winning \$500**.