

BHANU TEJA KURAKULA

Chicago, IL | 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 with 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.

SKILLS

Programming Languages	: Python, SQL, Spark, R
Platforms & Frameworks	: AWS, Snowflake, Databricks, Alteryx, Azure, Oracle, SQL Server, Confluent Kafka, Office
Tools	: Docker, Tableau, Airflow, Terraform, Postgres, MongoDB, Grafana, PowerBI
Libraries	: Numpy, Pandas, Scikit-learn, Matplotlib, Streamlit, Tensorflow, Plotly, Faust

WORK EXPERIENCE

Data Engineering Intern - AGCO Corporation

June 2024 - Present

- Engineered over 10+ complex ETL pipelines on AWS, standardizing tractor sales data across the EU.
- Translated 20+ Alteryx workflows into PySpark jobs, saving \$20K in licensing costs, making it cost effective.
- Implemented infra provisioning by IaC & other best devops practices reducing deployment time by 30%.
- Developed over 15+ Tableau prep ETL workflows and a dashboard as part of migration effort.
- Collaborated with cross-functional teams to build data maps and documented user manuals to ensure smooth handover.

Technologies Used: Tableau, Prep, Alteryx, SSMS, SQL Server, Oracle, AWS(Glue, S3, Athena, CloudWatch, SNS, RDS)

AWS Data Engineer - Cognizant

September 2022 - July 2023

- Enhanced production warehouse reliability by overseeing 200+ metrics and 100+ jobs ensuring uninterrupted data flow.
- Improved performance by writing optimized SQL queries resulting in 20% reduced query execution time.
- Collaborated with the upstream team to reduce Tableau scorecard delay by 15-20 minutes on a year over year basis.
- Presented a ML POC on fraudulent data claims to over 10+ warehouse managers demonstrating potential integration.

Technologies Used: Snowflake, Grafana, Tableau, SQL, Airflow, Python, MLaaS, Presto

Data Engineer Intern - Cognizant

March 2022 - July 2022

- Engineered an ETL pipeline for a dataset with 4M+ records, leveraging various Azure services as part of the internship.
- Implemented a batch data pipeline using Delta Live Tables (DLT) to process a 2GB dataset, created interactive dashboards using Databricks SQL for data visualization and reporting.

Technologies Used: Databricks, Pyspark, SQL, Excel, Azure (Blob, PowerBI, Synapse, Functions, Data Factory, Cosmos)

Data Science and Analytics Intern - PhysioAI

January 2022 - February 2022

- Enhanced user experience by developing a real-time yoga assistant to accurately detect and classify over 30+ yoga poses.
- Designed and implemented a dimensional data model to track user progress across 40+ real-time performance metrics, optimizing data retrieval for analytics and reporting.

Technologies Used: Postgres, SQL, Python, Tensorflow, Streamlit, Tableau, Dimensional Modeling

EDUCATION

University of Illinois, Urbana-Champaign

August 2023 - May 2025

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

(GPA 3.94/4.0)

Relevant Coursework : Data Modeling & Database Design, Data Warehouse & Business Intelligence, Data Governance, RDBMS

Amrita University

July 2018 - June 2022

Bachelor of Technology - Computer Engineering

(GPA 3.35/4.0)

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 MLOps 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.

Decentralized App for Covid-19 Prognosis using AI | Dash-Plotly, Tensorflow, Flask, MongoDB, Ganache, Solidity

Developed a POC to store medical records, including results generated by CNN models, on to a blockchain & a NoSQL DB.

ACHIEVEMENTS

- HoneyWell Hackathon Winner, 2021 : Secured 1st position out of 180 teams for developing a POC on smart traffic system, winning \$500.
- Rising Star Award - 2022, Cognizant : Awarded for being the Best College Hire based on client feedback and leadership ratings.