Bhanu Teja Kurakula

📞 217-926-2758 | ✉ bhanutk2@illinois.edu | 🔗 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)

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	: AWS, Databricks, Alteryx, Snowflake, Confluent Kafka, SQL Server, Office, Azure
Tools	: Docker, Tableau, Airflow, Terraform, Git, Postgres, Grafana, PowerBI
Libraries	: Numpy, Pandas, Scikit-learn, Matplotlib, Streamlit, Tensorflow, Plotly, Faust
Certifications	: AWS Solution Architect Associate, AWS Machine Learning Specialty

Work Experience - (31+ Months)

AGCO Corporation　　　　　　　　　　　　　　　　　　　　　　　June 2024 – Present
Advanced Analytics Intern (8+ Months)

- **Engineered over 10+ complex ETL pipelines** using various **AWS** services, ensuring seamless data integration and conforming, resulting in standardized tractor sales data across Europe.
- Translated **Alteryx** workflows into **PySpark** jobs, improving efficiency by implementing comprehensive tests to maintain data quality at all stages, also **saving $20K in licensing costs**, making it cost effective.
- Collaborated with stakeholders to build logical data map for migration and documented user manuals to ensure smooth handover. In addition, built **15+ tableau prep** flows and **dashboard** as part of decomission efforts.

Cognizant Technology Solutions　　　　　　　　　　　　　　September 2022 – July 2023
Programmer Analyst (AI & Analytics) (11 Months)

- 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.
- Improved performance by **optimizing SQL queries**, resulting in reduced query execution time. Reduced **tableau** scorecard delay by **20%** 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 improving data quality.

Data Engineer Intern (5 Months)　　　　　　　　　　　　　　　　　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 **2GB**, leveraging **Delta Lake optimizations** to enhance performance and efficiency.

PhysioAI　　　　　　　　　　　　　　　　　　　　　　January 2022 – February 2022
Data Science and Analytics Intern (2 Months)

- Developed a real-time yoga assistant using **Azure, TensorFlow**, and **MediaPipe** to accurately detect and classify over **30+** yoga poses, integrating a voice assistant feature for real-time guidance and pose correction.
- 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.

Ineuron.ai　　　　　　　　　　　　　　　　　　　　　September 2020 – January 2021
Data Science Intern (5 Months)

- Spearheaded the life cycle of **3** categories of a warehouse project by scraping over **10K+** images per category.
- Ensembled, trained and deployed the best **CNN models** to production, enhancing operational efficiency.

Projects

Coupon Accepting Prediction Pipeline | MLFlow, EC2, S3, RDS, Terraform, Docker, Prefect, Github Actions
Built a MLOps pipeline, enabling experiment tracking, orchestration, and automated deployment using CI/CD.

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