

Vikas Reddy

vikasreddyk@gmail.com | +1(630) 696-7013 | <https://github.com/Vikster123> | US Citizen

EDUCATION

The Ohio State University, B.Eng. Computer Science | Columbus, OH August 2021 - May 2024

WORK EXPERIENCE

Marriott International(HR & Financial Services), Data Engineer January 2023 - Present

- Migrated 8 HR analytics pipelines to Oracle Data Flow (PySpark) and standardized 25 curated ADW datasets, improving scalability and ensuring consistent metrics across HR analytics reporting.
- Collaborated with cross-functional data engineers and analysts to translate business requirements into production-grade data solutions supporting compliance and reporting needs.
- Stakeholder teams and persisting 12+ months of history in ADW to improve auditability and trend analysis.
- Owned and productionized weekly labor-forecast ingestion by building PySpark validation + transforms that published 14 trusted forecasting tables/attributes into ADW, sustaining 98%+ successful runs to improve reliability for planning/reporting.
- Designed and optimized Spark-based ETL pipelines aligned with Hadoop ecosystem principles, improving scalability, reliability, and cost efficiency across enterprise analytics workloads.
- Rebuilt 10 Talent Analytics workflows as reusable PySpark ETL jobs, improving maintainability and enabling repeatable cloud execution for daily/weekly reporting.
- Validated and supported daily/weekly Spark production runs, ensuring 20 ADW target tables stayed accurate and up to date for downstream HR reporting.
- Built 6 ODI pipelines to schedule and trigger Data Flow jobs with standardized dependencies/retries, improving job reliability and reducing manual intervention.

Marriott International, Global Tech Intern, Global Operations Team- Bethesda, MD June 2022 - August 2022

- Built 4 API endpoints to automate server health re-checks, reducing manual revalidation and improving operational response time for support teams.
- Integrated Forescout API ingestion and normalization, standardizing 15 health signals per device/server to improve consistency and usability for downstream monitoring.
- Designed and delivered a request-streamlining internal app connecting 20+ vendors and on-site associates, presented to senior leadership, and supported operational workflows across 9,500+ properties.

PROJECTS / LEADERSHIP

East Liberty Honda Plant Capstone Project- August 2023 - May 2024

- Led a 4-person team to automate brake-light inspection using a Cognex vision camera + PLC integration, enabling real-time pass/fail decisions on the production line.
- Deployed a high-throughput inspection workflow capable of processing up to 950 vehicles while improving inspection efficiency by ~15% and reducing manual checks.
- Presented performance results and ROI to the Plant General Manager, contributing to a 23% increase in budget allocated to AI and automation initiatives.

BreedVision (AWS DeepRacer League by Amazon Web Services (AWS) - May 2022 - June 2022

- Developed an end-to-end dog-breed recognition tool using 1,500 labeled images, benchmarking custom CNN, AlexNet, and ResNet architectures to optimize classification performance.
- Implemented preprocessing and augmentation (resize/normalize/flip/rotate) and targeted 3 key facial regions (ears/nose/eyes) to reduce overfitting and improve generalization.
- Produced an experiment report benchmarking CNN/AlexNet/ResNet and achieved 96.54% test accuracy, with ResNet outperforming AlexNet/CNN by 14% to select the best production candidate.

Shanti Mental Health App by MARS Inc - Feb 2021 - July 2021

- Built and deployed a web app for 30 students supporting 3+ workflows and 400 submissions/week, improving submission tracking and reducing manual coordination.
- Integrated a third-party API and stored 300 records in Excel to enable reporting and simplify admin review.
- Deployed on AWS EKS, shipped release via a repeatable containerized deployment pipeline.

TECHNICAL SKILLS

Big Data & Analytics: Hadoop, HDFS, YARN, Hive, Python, PySpark, Structured Streaming with Kafka

Data Engineering: ETL/ELT Pipelines, Data Modeling, Incremental Loading, Data Warehousing, Batch & Near Real-Time Processing

Programming & Querying: Python, PySpark, SQL, Scala

Cloud Platforms: AWS, Oracle Cloud Infrastructure (OCI Data Flow, ADW, ODI, OAC), Google Cloud Platform

Orchestration & DevOps: Airflow, Docker, Git, Linux/Unix

Databases: Oracle SQL Server, MS SQL Server, MongoDB

ML / Advanced (Exposure) MLOps concepts, AI/ML lifecycle, Computer Vision (CNN, AlexNet, ResNet)