

Detail-oriented Data Engineer with a master's in data science and 3 years of experience designing and optimizing ETL pipelines, real-time analytics systems, and cloud-native deployments. Proven ability to deliver scalable, low-latency data workflows using Python, PySpark, SQL, and distributed frameworks on AWS/GCP. Skilled in streamlining ingestion-to-insight lifecycles and automating performance monitoring for production-grade systems.

PROFESSIONAL EXPERIENCE

Data Scientist, Megh Computing

August 2020 – March 2023

- Reduced AI inference latency by 40% by optimizing object detection pipelines using TensorRT and OpenVINO, enhancing real-time surveillance capabilities.
- Improved anomaly detection accuracy by 25% through fine tuning YOLOv5 and Faster R-CNN models, facilitating quicker threat identification.
- Enhanced deployment efficiency by 30% by engineering cloud-edge AI solutions on AWS and GCP, reducing operational overhead in bandwidth sensitive environments.
- Increased system throughput by 60% via FPGA/GPU accelerated model pipelines, ensuring consistent performance for high volume video feeds.
- Decreased integration time by 35% by leading the development of Megh's Video Analytics SDK, enabling clients to embed custom AI capabilities seamlessly.
- Streamlined post-deployment monitoring by automating CI/CD driven performance benchmarking pipelines, cutting debugging cycles by 50%.
- Delivered tailored AI solutions across retail, finance, and smart city sectors, aligning deployments with sector specific KPIs.
- Led technical onboarding processes, establishing knowledge transfer frameworks that reduced new hire ramp up time.

TECHNICAL SKILLS

- Languages:** Python, R, SQL, PySpark, Bash, Java, C++
- Statistical Methods:** Regression, ANOVA, Chi-Square, Hypothesis Testing, A/B Testing, Time Series Analysis
- ML & Analytics:** scikit-learn, TensorFlow, XGBoost, LightGBM, PCA, Clustering, Predictive Modeling
- Data Analytics Tools:** Pandas, NumPy, SciPy, Statsmodels, JMP Pro, Tableau, Power BI
- Cloud Platforms:** AWS (EC2, S3, SageMaker), GCP, Databricks, Docker, CI/CD
- Visualization & Reporting:** Tableau, Power BI, Matplotlib, Seaborn, ArcGIS, StoryMap
- Databases & Engineering:** MongoDB, PostgreSQL, MySQL, Redis, Apache Spark, ETL Pipelines

EDUCATION

- Master of Science in Data Science**, Rochester Institute of Technology May 2025
- Bachelor of Engineering in Information Science and Engineering**, BNM Institute of Technology August 2020

PROJECTS

Autonomous Vehicle Safety (Explainable AI) | Python, OpenCV, TensorFlow, CRAFT, TCAV

Collaborators: Toyota Research Institute, University of Florida, University of California Irvine

- Developed XAI pipelines using CRAFT and TCAV to interpret object detection models in autonomous vehicles.
- Reduced false positive braking events by providing model transparency, improving safety critical decisions.
- Enhanced pedestrian detection accuracy, contributing to safer AV navigation systems.

Statistical Analysis of Online Sales Data | Python, Pandas, SciPy, ANOVA, Chi-Square

- Applied regression, ANOVA, and chi-square tests to identify significant sales trends and customer behavior patterns.
- Delivered actionable insights supporting data driven marketing and pricing decisions, leading to improved revenue forecasting.

Marketing Strategy Optimization (GIS + Data Viz) | ArcGIS, StoryMap, Python, Data Visualization

- Transformed raw marketing data into spatial insights using StoryMap, identifying high engagement zones and campaign gaps.
- Drove improved customer targeting and regional strategy optimization, directly increasing marketing ROI and engagement rates.

ACCOMPLISHMENTS

Winner – SCB Business Analytics Competition 2025

- Recommended Claude 3 as RIT's AI platform by evaluating LLMs (Claude 3, GPT-4, Mistral) across 7 strategic benchmarks, including cost, performance, privacy, and fairness.
- Led technical analysis and stakeholder alignment, resulting in a first-place win for presenting a scalable, compliant AI roadmap tailored to higher education.

Published Research: Application to Detect Skin Cancer using CNN | IJLTET, 2020— [Link to Paper](#)

- Achieved 82% classification accuracy by developing and training a MobileNet based deep learning model for melanoma detection, leveraging dermoscopic image datasets to assist early clinical diagnosis.
- Demonstrated real-world applicability by validating the model's performance against dermatological benchmarks.