Nikhila Madhunala

nikhilamadhunala01@gmail.com | 304-460-5482 | linkedin.com/in/nikhila-madhunala

Professional Summary

Data Engineering and ML Specialist skilled in big data processing and GCP. Proficient in Python, TensorFlow, and Scikit-learn with experience in streaming pipelines, model training, and large-scale deployments. Committed to optimizing workflows and enabling reliable AI solutions.

Education

Master's in Business Analytics
University of North Texas, Denton, TX – December 2024

Bachelor of Technology in Computer Science
CMR College of Engineering and Technology, India – August 2021

Technical Skills

Programming: Python, Go, PyTorch, TensorFlow, Scikitlearn, SQL, Git

Data Engineering: Model training, big data processing (Apache Flink, Hadoop, Spark)

Databases: Snowflake, BigQuery

Cloud: Google Cloud Platform (BigQuery, Dataflow, Pub/Sub), Amazon Web Services

DevOps: MLOps

Analytics/ML: Machine learning, forecasting

Practices: Optimization, ownership, reliability

Professional Experience

Data & ML Engineer Intern
Next Era Path *Feb 2025 – Present*

- Built high-throughput pipelines with Apache Flink and Google Dataflow. Processed millions of events per minute at low latency, ensuring reliable streaming infrastructure and supporting business-critical real-time analytics needs.
- Implemented feature stores and cataloging systems to improve data discoverability. Enabled consistent reusability of assets across projects, reducing duplicated effort and accelerating model development cycles across teams.
- Partnered with stakeholders to define key data requirements for analytics. Delivered robust and scalable data solutions aligned with business timelines, improving trust and adoption of real-time data-driven decision making.
- Documented pipeline architectures, publishing internal blogs and talks. Shared insights on scalability and reliability, building organizational knowledge while raising engineering maturity across data teams companywide.

Data & ML Engineer
Infosys (Client: Westpac) *Jan 2021 – Dec 2022*

- Trained and evaluated ML models with TensorFlow and Scikit-learn on large datasets. Improved model accuracy and reduced overfitting, ensuring production-ready ML systems optimized for real-world applications.
- Deployed ML models into production with TensorFlow Serving on GKE. Delivered scalable and secure inference systems, enabling reliable predictions and robust serving architecture for end-user applications.
- Automated retraining workflows to evaluate candidate models on new datasets. Ensured production models remained accurate and reduced drift, increasing the trustworthiness of deployed ML systems at scale.
- Created internal evaluation frameworks and collaborated with ML engineers to track performance. Supported rapid iteration and accelerated delivery of high-quality models for critical product features.

Data & ML Engineer Intern
Powersoft Global Pvt Ltd *May 2020 – Aug 2021*

- Optimized SQL queries and storage formats in BigQuery and Snowflake. Reduced query execution costs by 30% and improved performance, enabling efficient analytics on large-scale enterprise data warehouses.
- Partnered with product teams to identify performance bottlenecks in analytics queries. Delivered optimized architectures that reduced reporting delays and supported faster customer-facing insights delivery.
- Designed and maintained cost-optimized data workflows on GCP and AWS platforms. Balanced performance, scalability, and budget goals to deliver sustainable cloud infrastructure for ML applications.
- Mentored engineers on data reliability, optimization strategies, and scalable ML pipeline design. Improved adoption of big data best practices and elevated organizational capability in handling large-scale datasets.

PROJECTS

Capstone: Predicting Vehicle Recalls from Complaint Data Aug 2024 – Dec 2024
Inherited fragmented data pipelines across AWS, Azure, and GCP that caused duplication of effort. Consolidated infrastructure into unified pipelines for training and deployment. Delivered a single platform that improved scalability and reduced maintenance costs.
Implemented CI/CD pipelines integrating Kubeflow and MLflow for continuous training and deployment cycles. Automated testing, validation, and release approvals. Significantly shortened model release cycles and increased business agility.
Drove adoption of observability practices by integrating monitoring with Prometheus and Grafana. Enabled teams to proactively track model drift and infrastructure health. Reduced time-to-detect and resolve issues, enhancing reliability.
Acted as a mentor for engineers transitioning into MLOps roles. Provided hands-on guidance in containerization, distributed computing, and cloud cost optimization. Built a knowledge-sharing culture that elevated overall technical maturity.