Nikhila Madhunala

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

Professional Summary

Machine Learning Data Engineer skilled in Python, Java, and Scala with expertise in building data pipelines, feature engineering, and deploying ML services on AWS, Azure, and GCP. Experienced in optimizing cloud infrastructure to deliver scalable, reliable machine learning 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, Java, Scala, Docker, Kubernetes, Git, Pytorch, Scikit learn

Data Engineering: Data pipelines, feature engineering, model evaluation

Databases: Redshift, BigQuery, Snowflake

Cloud: Amazon Web Services (S3, Glue, Redshift), Google Cloud Platform (BigQuery, Dataflow), Microsoft Azure

DevOps: CI/CD, MLflow, MLOps

Analytics/ML: Machine learning

Practices: Communication, optimization, problem solving, prototyping, reliability, research, scalability

Professional Experience

Machine Learning Data Engineer Intern
Next Era *Feb 2025 – Present*

- Designed and maintained scalable ETL pipelines with Apache Spark and Kafka. Reduced data latency, improved quality, and ensured seamless delivery of clean datasets for ML workflows supporting predictive and operational solutions.
- Developed reusable feature engineering frameworks in Python, enabling standardized transformations across teams. Improved overall model performance and reproducibility while streamlining handoff between engineering and science functions.
- Migrated fragile legacy pipelines into modern, cloud-native platforms across AWS and GCP. Reduced operational overhead and improved pipeline resilience, scalability, and fault tolerance for critical data ingestion processes.
- Mentored junior engineers on engineering workflows and cloud-native services. Conducted training sessions that accelerated team maturity, improved delivery consistency, and built a stronger culture of shared best practices.

Machine Learning Data Engineer
Infosys (Client: Westpac) *Jan 2021 – Dec 2022*

- Deployed and monitored ML models on AWS and GCP leveraging SageMaker and managed services. Configured autoscaling for high availability and resilience, ensuring production-grade inference environments at scale and reliability.
- Built automated pipelines for evaluating and benchmarking candidate models. Compared metrics systematically to identify the best-performing model. Accelerated experimentation cycles while improving operational efficiency and confidence in results.
- Established monitoring and alerting solutions to track both model and data metrics. Provided early detection of anomalies, supported rapid remediation, and reduced downtime, strengthening the reliability of production ML services.

- Collaborated with data scientists to design hybrid online and offline serving architectures. Ensured consistency between training and inference workflows, reducing divergence and supporting seamless ML model lifecycle management.

Machine Learning Data Engineer Intern
Powersoft Global Pvt Ltd *May 2020 – Aug 2021*

- Migrated high-maintenance workloads to modern AWS, Azure, and GCP infrastructure services. Improved reliability, reduced operational overhead, and enabled more efficient scaling strategies across multiple ML and data projects.
- Partnered with architects to optimize infrastructure resource allocation and design cloud-native pipelines. Delivered cost-effective and scalable workflows that aligned technical design with business objectives across global teams.
- Implemented CI/CD with MLflow and MLOps best practices for end-to-end ML lifecycle management. Automated deployments, improved reproducibility, and accelerated overall release cycles while reducing manual engineering effort.
- Defined and communicated best practices for cloud architecture through documentation and workshops. Enhanced cross-team knowledge sharing, increased engineering efficiency, and ensured sustainable long-term ML operations at scale.

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.