KARTHVIK SARVADE
AI Engineer

Karthvik.s@gmail.com
https://github.com/Karthvik-s
+1(929)-308-6751
New York, NY

AI/ML Engineer with expertise in designing scalable machine learning systems, model training frameworks, feature stores, and ML prediction services. Proficient in Python, microservices, cloud platforms (AWS, GCP, Azure), MLOps, and LLM serving infrastructure with real-time data pipelines.

Work Experience

Machine Learning Engineer
Aug 2023 - Present
DataQuad Inc | Remote

- Designed and deployed LLM (GPT-3.5, LLaMA, T5)-based applications to automate customer service and domain-specific Q&A. Implemented a real-time NLP pipeline to reduce manual support efforts by 40%.
- Built scalable ML systems, including model deployment and real-time feature stores using AWS Lambda, GCP Cloud Build, and Azure ML Pipelines.
- Developed and optimized cloud-native data pipelines (AWS Glue, GCP Dataflow, Azure Data Factory), increasing data ingestion efficiency by 60%.

AI/ML Intern
Jun 2023 - Aug 2023
Delta Air Lines | Atlanta

- Integrated GPT-3.5 and LLaMA models into Delta's IT support chatbot, reducing ticket resolution time by 30%. Improved chatbot accuracy by 25% using retrieval-augmented generation (RAG) for better response quality.
- Developed predictive maintenance models using Python, TensorFlow, and Scikit-learn to anticipate aircraft component failures, reducing unplanned downtime by 20%.
- Applied data cleaning and transformation techniques to prepare large-scale datasets for model training

Data Analyst
May 2021 - Jul 2022
DBS Tech | Hyderabad

- Designed and optimized SQL-based ETL pipelines using PostgreSQL and MySQL to process transactional data, reducing data ingestion time by 40%.
- Developed ML models Scikit-learn, XGBoost for fraud detection, deploying real-time prediction services.
- Created interactive Power BI dashboards for finance teams to monitor loan default risks, customer churn probabilities, and revenue forecasting, improving decision-making speed.

Core Skills

Programming: Python, Java, SQL; Machine Learning Systems: Feature Store, Model Training Frameworks, ML Prediction Services, LLM/LMM Serving Infrastructure; Cloud Platforms: AWS (SageMaker, Lambda, EC2, Glue), Azure (Databricks, MLflow, Azure Functions), GCP (Vertex AI, BigQuery, Dataflow); Microservices & Backend: Docker, Kubernetes, CI/CD Pipelines; Data Engineering: Real-Time Data Pipelines, ETL, Data Ingestion, Big Data (Hadoop, Spark)

Education

New York University
Aug 2022 - May 2024
Master of Computer Science AI & ML

Jain University
Jun 2018 - May 2022
Bachelor of Electronics and Communications