

Ajay Mekala

AI/ML Engineer | Production ML Platforms, MLOps, Frontier Model Evaluation

Teaneck, New Jersey | (862) 440-8042 | mekalaajayk@gmail.com

[linkedin.com/in/ajaymekala](https://www.linkedin.com/in/ajaymekala) | github.com/mekala27-45

Authorized to work in the US, F-1 OPT, no sponsorship required

SUMMARY

I am an AI/ML Engineer with four years across production machine learning, data platform engineering and frontier model evaluation. I build systems that set prices for one of the largest retailers in the world, and I spend the rest of my time finding out where the newest language models break.

At Walmart I built and operate the ML platform behind promotion and clearance pricing, running on Azure Databricks, Delta Lake and MLflow with infrastructure provisioned in Terraform. It serves over 200,000 predictions per month. PromotionsAI and ClearanceAI sit on top of it and have driven 7.8 million dollars in incremental revenue. Standardizing MLOps across those pipelines cut research-to-production lead time by 40 percent and experiment turnaround by 60 percent.

Alongside that, I evaluate and stress-test frontier models on contract for Handshake AI, Snorkel AI, Mercor and Outlier. I have authored more than 200 golden-solution engineering tasks and over 50 accepted Terminal-Bench environments. I am completing an MS in Data Science at Montclair State University in May 2026.

TECHNICAL SKILLS

Machine Learning

Python, PyTorch, TensorFlow, scikit-learn, XGBoost, LightGBM, Hugging Face Transformers, Pandas, NumPy, SciPy, statsmodels, Feature engineering, Demand elasticity modeling, Time-series forecasting (ARIMA, Prophet), NLP, Embeddings and vector search

MLOps and Infrastructure

MLflow, Weights and Biases, Databricks, Databricks Feature Store, Docker, Kubernetes, Terraform, FastAPI model serving, Blue-green deployment, Model monitoring and drift detection, Azure Pipelines, GitHub Actions, pytest, Automated A/B harnesses, Cookiecutter templating

Data Engineering

Delta Lake, Apache Spark, PySpark, Apache Airflow, Kafka, Azure Event Hubs, dbt, ETL and ELT, Data validation and enrichment, Connector library design, Distributed training on Spark

Cloud and Storage

Azure Databricks, Azure Data Lake, Azure Pipelines, Azure Event Hubs, AWS S3, AWS Redshift, AWS Lambda, GCP BigQuery, GCP Dataflow, GCP Pub/Sub, PostgreSQL, MySQL, Snowflake, MongoDB, Redis, FAISS, pgvector

Evaluation

RLHF, DPO, SFT, Agentic trajectory evaluation, Tool-use grading, Rubric design, Adversarial prompting, Red-teaming, LLM-as-a-judge, Inter-annotator agreement

Analytics

Power BI, Tableau, Matplotlib, Seaborn, Plotly, A/B testing, Cohort and funnel analysis, Hypothesis testing

Languages

Python, SQL, TypeScript and JavaScript, Go, Bash, R

PROFESSIONAL EXPERIENCE

AI/ML Engineer | Walmart | Remote

November 2024 to Present

- Built and operate the ML platform behind promotion and clearance pricing on Azure Databricks, Delta Lake and MLflow with Terraform-provisioned infrastructure, cutting research-to-production lead time 40 percent and serving over 200,000 predictions per month.
- Shipped PromotionsAI, driving 5.5 million dollars in incremental revenue in its first full fiscal year and 2.3 million the following quarter.
- Built and own the MLOps stack for ClearanceAI, lifting sell-through 4.6 percent.
- Standardized MLOps across both systems and the ingestion pipelines, cutting experiment turnaround 60 percent, holding a 99.9 percent SLA and halving incident MTTR.
- Automated ingestion, validation and enrichment with PySpark and Delta Lake and co-authored the reusable connector libraries, cutting new source onboarding 70 percent.

Tools: Azure Databricks, MLflow, Delta Lake, PySpark, Terraform, Kubernetes, Docker, Kafka, XGBoost, Python

AI Training and Evaluation Specialist | Contract | Remote

November 2022 to Present

Clients: Handshake AI, Snorkel AI, Mercor, Outlier AI, Alignerr, Welocalize, AfterQuery, FleetAI, SME Careers

- Authored 200+ golden-solution software engineering tasks for Handshake AI's Project Helix, building multi-file repository problems with hidden test suites and graded difficulty tiers across Python, TypeScript and Go, and passed Helix Screening with a top-decile reviewer rating.
- Evaluated long-horizon coding-agent trajectories on real software workflows, scoring tool-use correctness, sub-goal decomposition, state tracking and recovery from failure, and surfaced systematic failures in tool-call sequencing and shell command grounding that fed directly into agent fine-tuning.
- Authored 50+ accepted Terminal-Bench tasks for Snorkel AI under the Terminus 2 agent scaffold: containerized environments with reference solutions and automated verification tests, calibrated to challenge frontier agents while staying unambiguous and solvable.
- Designed 150+ model-stumping prompts validated against GPT, Claude and Gemini-class baselines, plus graduate-level STEM question-answer pairs for reasoning benchmarks that passed multi-layer expert review.

Tools: RLHF, DPO, SFT, Agentic Evaluation, Terminal-Bench, Rubric Design, Red-Teaming, Python, TypeScript, Go

Software Engineer | Tech Mahindra | Hyderabad, India

May 2023 to July 2024

- Engineered ETL pipelines with Python, PySpark and Apache Airflow processing 50TB+ datasets for Fortune 500 clients, cutting data preparation time 45 percent and roughly 120 thousand dollars per year in cloud compute costs.
- Built ARIMA and Prophet time-series forecasting models at roughly 85 percent accuracy driving demand planning and budget allocation.
- Optimized data models through indexing and partitioning, cutting query execution time 60 percent for 200+ users, and ran A/B, cohort and funnel analyses informing 1.5 million dollars or more in strategic decisions.
- Automated 20+ Power BI and Tableau dashboards with scheduled refresh and row-level security, removing roughly 40 percent of manual reporting.

Tools: Python, PySpark, Airflow, SQL, BigQuery, PostgreSQL, Power BI, Tableau, Prophet, ARIMA

Data Analyst Intern (Part-Time) | Montclair State University | Montclair, NJ

April 2025 to July 2025

- Analyzed 114+ community engagement projects using statistical analysis, NLP and time-series methods.
- Built 20+ interactive dashboards in Python, Tableau and Power BI that cut report preparation time 40 percent.
- Applied text classification and sentiment analysis to 5,000+ qualitative survey responses, producing a strategic report presented at a multi-university conference.

Tools: Python, Pandas, NumPy, SQL, NLP, Tableau, Power BI

EDUCATION

Master of Science, Data Science

Montclair State University, Montclair, NJ | May 2026

Bachelor of Technology, Artificial Intelligence and Data Science

KL University, Hyderabad, India | April 2024