Kunwar Uday Singh Sikarwar

New York kunwarudaysingh.71@gmail.com 315-886-7577 in/kunwaruday

Summary

Machine Learning Engineer experienced in designing predictive models and scalable ML pipelines across cloud data ecosystems. Skilled in Python, SQL, Airflow, MLflow, Docker, and model lifecycle automation for reproducible, monitored, and production-ready ML systems.

Experience

WPP Media New York, NY

Machine Learning Engineer *November 2024 - Present*

- Built and maintained forecasting + optimization pipelines in Python/BigQuery; orchestrated with Airflow, containerized with Docker, and tracked with MLflow for repeatable training, scoring, and backfills. Supported planning decisions across $8M in tracked volume and contributed to 20% outcome lift (as reported internally).
- Productionized experimentation / holdout evaluation workflows (lift estimation, guardrails, power checks) in SQL + Python; automated run + reporting logic cut manual analysis time 30% and improved consistency across releases.
- Productionized LTV-style models with versioned BigQuery scoring tables and automated QA; added Airflow data-validation gates (schema drift, reconciliation, outliers) to prevent data regressions and stabilize scheduled scoring runs.
- Built GPT-4 (OpenAI API) RAG assistant retrieving governed BigQuery tables/metric definitions; Dockerized + scheduled in Airflow, reducing recurring reporting effort by 30%.

IConsult Syracuse, NY

Data Scientist *November 2023 - October 2024*

- Built reusable training + scoring pipelines (feature tables, evaluation harness, scheduled refresh) in Python/SQL; improved forecast performance by 12% on key targets through backtesting and metric standardization.
- Implemented uplift / causal scoring workflows to estimate treatment impact and rank interventions; published versioned outputs to BigQuery to enable repeatable batch scoring and monitoring.
- Developed automated anomaly detection + data-quality monitoring in Python/SQL; reduced lag in issue identification by 35% and improved alert consistency for stakeholders.

Intradiem Atlanta, GA

AI/ML Engineer *May 2023 - May 2024*

- Built and validated classification + forecasting models for workforce planning and risk signals; improved predictive performance by 22% through feature engineering, cross-validation, and rolling backtests.
- Deployed production LLM + RAG support chatbot integrated into agent workflows; reduced average handle time by 15% by retrieving answers from curated knowledge and operational data.
- Implemented end-to-end MLOps pipelines on AWS SageMaker with Docker + MLflow + Airflow, including automated training, model registry/versioning, scheduled retraining, and monitoring hooks for production stability.

Project

Production Forecasting and Batch Scoring Pipeline (Python, BigQuery SQL, Airflow, MLflow, Docker)

- Built end-to-end training + batch scoring pipeline with Airflow, Docker, and MLflow; implemented rolling backtests, scheduled retraining, and reproducible configs.
- Improved forecast metrics by 10–15% vs baseline on validation windows and reduced manual refresh effort by 30% via automated runs and standardized evaluation.

Real-time Deep Learning Model Serving (PyTorch + FastAPI + MLflow + Docker)

- Served TorchScript PyTorch model through FastAPI inference API with Pydantic request validation and consistent, versioned preprocessing artifacts to ensure training parity.
- Logged experiments and registered model versions in MLflow Model Registry; implemented stage/version-based safe loading and structured request/response logging for reproducible inference.

Experimentation and Uplift Measurement Framework (Python, FastAPI, LangChain, FAISS/Chroma, OpenAI GPT-4 API, Docker, Airflow)

- Built production-style RAG Q&A API using FastAPI + LangChain, retrieving from FAISS/Chroma vector index and generating cited answers via OpenAI GPT-4 API.
- Added caching, rate limiting, and structured JSON outputs with regression tests to catch prompt/model drift; Dockerized service and scheduled index refresh/backfills via Airflow.

Education

Master of Science in Data Science 2024

Syracuse University

Bachelor of Technology in Computer Science & Engineering 2018

Galgotias University

Skills

- **Programming:** Python, SQL, R
- **ML Modeling:** Forecasting/time-series, classification/regression, uplift/causal modeling, anomaly detection, segmentation; scikit-learn, XGBoost
- **LLM / GenAI:** RAG pipelines, LangChain, OpenAI GPT-4 APIs, evaluation/guardrails
- **MLOps / Engineering:** Airflow, MLflow, Docker, Git, Jupyter, CI/CD
- **Cloud / Data:** GCP (BigQuery, Vertex AI), AWS (S3, EC2, SageMaker, RDS, Lambda), Snowflake, dbt
- **Apps / BI:** Streamlit, Tableau, Power BI