Daivya Shah

New York, NY | (917) 847-4174 | daivya.shah@nyu.edu | linkedin.com/in/daivya-shah/ | daivyashah.com

Education

New York University **Expected May 2026**
B.A. in Computer Science and Data Science, Minor in Business *New York, NY*

- **Relevant Coursework:** Fundamentals of Machine Learning, Data Management and Analysis, Principles of Data Science, Causal Inference, Regression and Forecasting Models, Probability and Statistics, Linear Algebra, Data Structures, Calculus III

Experience

Newmark Jun 2025 – Aug 2025
Software Engineer Intern *New York, NY*

- Developed a responsive Commercial Real Estate (CRE) stacking plan using **React** and **TypeScript** to visualize floor-by-floor tenant and vacancy data through interactive layouts, filters, and color-coded lease expirations for faster leasing decisions.
- Built an interactive time-based activity feed using **React Query** to merge notes, tours, proposals, and deal updates into one chronological view, improving collaboration and reducing information retrieval time.
- Created a Kanban-style deal pipeline with **@dnd-kit** for drag-and-drop deal tracking, supporting inline editing, search, and customizable columns to help brokers manage leads and track deal progress efficiently.

Harvest Feb 2025 – Mar 2025
Founding Software Engineer *New York, NY*

- Developed an AI-powered manufacturer discovery platform using Next.js, FastAPI, **Firebase**, and **Pinecone** for vector search.
- Scraped and parsed **500+** manufacturer websites and catalogs using **Selenium** and **Unstructured.io**, normalizing **100K+** records in Firestore and embedding them with GPT-4 in Pinecone to power hybrid **RAG** manufacturer retrieval.
- Built a learning-to-rank recommendation engine using **scikit-learn** and **XGBoost Ranker** on feedback (clicks, shortlists), computing cosine similarity and feature scores (MOQ, region) to personalize manufacturer results and improve **AUC** + **NDCG@5** relevance scores.

Everise Jul 2024 – Dec 2024
Data Engineer Intern *New York, NY*

- Automated ETL pipelines using Python, SQL, and **Apache Airflow** to integrate **500K+** data points from multiple sources into a centralized warehouse (MISDW) and built interactive **Power BI** dashboards with **DAX** to track B2P%, QA, and SLA performance.
- Conducted root cause analysis on customer service KPIs using SQL queries, Python transformations, and Power BI (segmentation, drill-downs, and DAX measures) across **100K+** user interactions to identify inefficiencies and improve customer satisfaction scores.
- Built a forecasting pipeline using **ARIMA** and **LightGBM** models to predict client-level Bill-to-Pay (B2P%) from time-series billing data with **93%** accuracy, automated via Airflow DAGs and visualized in Power BI to improve revenue forecasting.

eMeasurematics - Industrial Autonomous Solutions Jun 2024 – Jul 2024
Data Analyst / Machine Learning Intern *Chicago, IL*

- Collaborated on a **Flask** predictive maintenance system for industrial vehicles using **SQLAlchemy**, **MySQL**, and ensemble models (**Random Forest**, **XGBoost**) on sensor data, achieving **91%** accuracy in predicting maintenance needs and reducing downtime.
- Developed a **Tableau** dashboard with calculated fields, parameters, and real-time alerts to support predictive maintenance planning.
- Configured, calibrated, and updated firmware on 16 radar sensors for precise slab positioning across 4 steel facility cranes.

Extracurricular Activities & Projects

NYU Machine Learning Club Jun 2024 – Present
Vice President *New York, NY*

- Led workshops on regression, feature engineering, and model evaluation, training **60+** students for DS/ML interviews.
- Organized industry talks with leaders from Meta, Datadog, and other top tech firms, expanding membership to **1,100+** students.

NYU Stern Business Analytics Club Sep 2024 – May 2025
ML Team Analyst *New York, NY*

- Studied deep learning architectures and models such as **RNNs** for N-way K-shot image classification and **transformers**, and how to build them using common frameworks like **PyTorch**; applied these skills in the "Optiver – Trading at the Close" competition.
- Engineered a stock closing reference price tracker for **400+** Nasdaq stocks using **LSTM** and ARIMA models in Python to predict auction imbalances and price directions from order book and closing auction data, achieving **27%** lower MAE than baseline predictions.

Trustworthy AI Lab x GES UCLA Hackathon Jun 2024

- Won **1st place**/46 teams by developing a Data Clean Room using **Azure Confidential VMs**, **TPM2-tools**, and **Key Vault** to enable secure data sharing between advertisers and publishers for predictive analytics; awarded an internship at the Trustworthy AI Lab, UCLA.
- Improved CTR prediction accuracy by **20%** using Random Forest models trained on **GAN-generated synthetic data** (**99.87%** fidelity).

Technical Skills

Languages: Python, SQL, JavaScript/TypeScript, Java, C/C++, HTML/CSS
Libraries: pandas, NumPy, scikit-learn, PyTorch, TensorFlow, LightGBM, XGBoost, statsmodels, LangChain, PySpark, SQLAlchemy
Frameworks & Tools: React, Next.js, Node.js, Flask, FastAPI, Redis, Apache Airflow, Apache Spark/Kafka, Databricks, Firebase, Docker, Azure Cloud, Pinecone, Selenium, MySQL, PostgreSQL, MongoDB, Power BI, Tableau, Git, JIRA, Agile