Lakshya Khandelwal

lk3068@columbia.edu | (646) 372-9473 | LinkedIn | GitHub | lakshya-khandelwal.com

Education

Columbia University **New York, NY**
Master of Science in Data Science *Sep 2025 – Dec 2026*

Professional Work Experience

Calix **New York, USA**
AI/ML Engineer Intern *Jun 2026 – Sep 2026*

- Built a conversational **virtual infrastructure ticketing agent** using **Google ADK**, **Gemini 2.5 Pro**, and **MCP**, automating VM, storage, and network issues/requests through Jira Service Management with a draft-review-confirm framework.
- Developed a **multi-agent LLM cost analytics platform** serving **1,200+ developer clients**, integrating LiteLLM spend data with organizational hierarchies and routing natural-language queries to specialized agents for analysis and track AI Adoption.
- Engineered scalable backend API services for an internal **AI agent platform supporting 7+ agents**, integrating secure authentication, agent routing, session management, and automated Jira/Confluence token refresh for enterprise AI workflows.
- **Skills: React, Typescript, LiteLLM, Agentic AI, Safety & Security, CI/CD Pipeline, Software Development, Kubernetes**

Anheuser-Busch InBev GCC **Bengaluru, IN**
Data Scientist *Jan 2024 – May 2025*

- Built Python-based **web scraping pipelines** to ingest distributor sales data into SAP databases, cutting **manual collection time by 70%** and enabling **daily ingestion of 50,000+** records for downstream analytics.
- Implemented an XGBoost model with lag and rolling features for SKU-level **distributor demand forecasting**, replacing **SARIMA** and reducing **MAPE from 28% to 11%**, improving accuracy and inventory planning.
- Collaborated with **supply chain, distributor, and operations** teams and led **exploratory and root cause analysis** on inventory and shipment data, **reducing stockouts by 18%** and **improving on-time fulfillment by 12%**.

Data Science Intern *Jul 2023 – Dec 2023*

- Created three interactive **Power BI dashboards** displaying automation KPIs, bot performance, and operational metrics, empowering 15 stakeholders to make data-driven decisions and improving process visibility.
- Streamlined data pipelines using SQL, Python, and cloud-based platforms, **automating data ingestion**, transformation, and loading processes, resulting in **60% reduction in manual hours** for internal customers.
- **Skills: Python, MySQL, Machine Learning, Time-series, Business Intelligence, Data Modeling, A/B Testing, Mathematics.**

Project Experience

BrandBuddy | Multi-Agent AI Content Creation Platform **New York, NY**

- Engineered a **12-node LangGraph** product pipeline with a supervisor orchestrating **trend discovery, content generation, and automated publishing** agents, enabling end-to-end Instagram and email campaign creation from a **single prompt**.
- Integrated **iterative content refinement and rubric-based evaluation**, enabling automated retries, real-time **refinements to align with brand personalization**, and structured checkpoints for user intervention without restarting the pipeline.
- **Skills: FastAPI, LangGraph, Typescript, NLP, Foundation Models, Cloud Infrastructure, Experimentation**

LENS | Multi-Agent Data Analysis Platform **New York, NY**

- Orchestrated an **agentic AI Application for retail product analytics** over a 3-table dataset with **10M+ records, generating structured, citation-backed insights, visualizations**, and actionable recommendations through data retrieval and exploration technology.
- Designed a **4-agent iterative reasoning architecture (SQL → EDA → Hypothesis → Supervisor loop)** leveraging LangGraph with up to 8 reasoning iterations for exploratory data analysis, statistical analysis, and data quality checks.
- **Skills: LangGraph, OpenAI SDK, , Text-to-SQL, LLM Orchestration & Evaluation, Data Validation.**

FinancialRAG | End-to-End RAG System **New York, NY**

- Architected and deployed a Financial RAG chatbot for **15 public companies** by building a retrieval pipeline with **token-based chunking, metadata filtering, and vector indexing in ChromaDB**, supporting semantic search with citation support.
- Implemented **multi-model evaluation** across GPT, Claude, Gemini, and Llama, identifying Gemini 3 Flash achieving 96.88% **accuracy at 1/10th the cost**, and **benchmarked 98.44%** peak accuracy to optimize model selection and reduce costs.
- Launched **FastAPI backend** on AWS Lambda and frontend on Streamlit for serverless, cost-efficient query handling.
- **Skills: RAG, FastAPI, Streamlit, ChromaDB, OpenAI, LLM Fine Tuning, Docker, Information Retrieval.**

Technical Skills

Languages & Libraries:Python, Java, Tensorflow, PyTorch, Scikit-learn, HuggingFace, Pandas, NumPy, Matplotlib, Keras, R
Data Science & ML: Data Mining, Deep Learning, Generative AI, LLM, LangChain, Computer Vision, Research, CUDA
Data Engineering & Databases: SQL, PostgreSQL, MongoDB, Scala, GitHub, Hadoop, Hive Spark, Cloud Computing
Cloud Platforms & DevOps: Amazon AWS, GCP, Snowflake, Tableau, Looker, Azure, Power BI, SaaS, MLOps, Matlab, Git
Coursework: Computer Science, Robotic Process Automation (RPA), Natural Language Understanding, Artificial Intelligence