

ABHITAY SHINDE

(646) 225-0802 | abhitayshinde@gmail.com | New York, NY | [LinkedIn](#) | [Portfolio](#)

EDUCATION

Columbia University M.S. Data Science, GPA: 3.7/4.0 • Relevant Courses: Agentic AI, Fintech and Data Economy (PhD elective), Big Data Analytics, Statistics, and Applied Machine Learning.	Sep 2024 - Dec 2025 New York, NY
NMIMS University B.Tech Honors in Computer Engineering and Data Science/Analytics, GPA: 3.9/4.0	Jun 2020 - May 2024 Mumbai, IN

PROFESSIONAL EXPERIENCE

Julius Baer Data Science Researcher (Generative AI and Personalization) • Built an LLM personalization chatbot using LangGraph, Vertex AI, and prompt chaining, automating content selection, cutting curation by 35%. • Ran causal and exploratory analysis on customer engagement with Python and SQL, driving a 25% lift in customer content engagement metrics. • Established an experimentation metrics framework, enhancing A/B test accuracy, guardrails, and reproducibility across cross-functional teams.	Jun 2025 - Sep 2025 New York, NY
TD Bank Data Scientist Consultant (Customer Analytics and Segmentation) • Executed causal inference and A/B testing on 100M+ transactions in PySpark and Databricks, validating fraud-detection strategies at scale. • Designed XGBoost fraud propensity scoring, raising base PR-AUC from 0.20 to 0.67 (235% gain) with SHAP explainability for stakeholders. • Partnered with senior management and risk teams to inform customer segmentation, money laundering reduction, and strategy via data insights.	Jan 2025 - May 2025 New York, NY
Ask2AI Quantitative Research Apprentice (Acquisition Analytics and Growth Optimization) • Deployed uplift modeling and experimentation to isolate incremental acquisition impact, clarifying ROI attribution across channels by 22%. • Built regression and MILP optimization for acquisition allocation using PyGurobi, raising customer LTV by 12% and reducing churn by 25%. • Synthesized forecasting-driven causal insights to guide growth strategy and executive decisions, boosting acquisition rates across key segments.	Jan 2025 - May 2025 New York, NY
Columbia University Teaching Assistant (Algorithms to Data Science) • Advised 250+ MS students on ML algorithms, product analytical reasoning, experiment design and inference in structured weekly sessions. • Designed case studies and assignments on A/B test design, modeling, and data-driven decision-making for applied data science coursework.	Sep 2024 - Jan 2025 New York, NY
Navin Fluorine International Limited Data Scientist Consultant (Predictive Maintenance and Operational Analytics) • Architected predictive maintenance with XGBoost, cutting equipment downtime 35% and reducing KPI ingestion latency 40% via Airflow. • Deployed anomaly detection using XGBoost and Isolation Forest to surface high-risk events, lifting recall 15% and shortening repair cycles.	May 2023 - Nov 2023 Mumbai, IN
E-Revbay Pvt. Ltd. Founding IT Intern (Acquisition and Retention Analytics) • Led customer acquisition analytics pipelines using Equifax data, SQL and Python, improving qualified customer leads by 50% per quarter. • Built real-time dashboards and root-cause analysis pipelines in Tableau, eliminating 25% of manager intervention time and speeding response.	Dec 2021 - May 2022 Mumbai, IN

SELECTED PROJECTS

IBM Agentic Library - Graph RAG for Document Q&A <i>LangGraph, RAG, Neo4j, FAISS, ChromaDB, LLMs, Python, OpenAI APIs</i> • Built a Graph RAG pipeline for IBM's open-source agentic library using Neo4j and FAISS for document ingestion and structured retrieval. • Modeled documents as knowledge graphs capturing entity relationships, improving retrieval precision 30%+ over naive vector search. • Designed agentic orchestration using LangGraph for multi-step reasoning, decomposing complex queries into sub-tasks via CoT prompting.
Social Media Analytics - AI Detection & Source Credibility <i>Python, NLP, Fine-Tuning, PyTorch, Hugging Face, Web Scraping, CLIP, BLIP</i> • Built an end-to-end agentic pipeline ingesting 2TB+ of Reddit data daily, flagging misinformation using fine-tuned LLM and BLIP analysis. • Developed an AI-generated content detector for human vs. AI images using BERT and CLIP classifiers, achieving 89% accuracy on GCP. • Implemented art style classification via ResNet transfer learning, with Airflow-orchestrated batch inference cutting attribution latency by 40%.
Causal Product Analytics Suite <i>Python, Causal Inference, A/B Testing, Propensity Score Matching, DiD, Regression, MILP</i> • Built a Media Mix Model estimating incremental LTV across channels; causal optimization revealed different channel rankings vs. last-touch. • Applied PSM-based DiD to a feed change with 110% lift; post-matching, 7-day retention was ~0%, preventing an incorrect ship decision. • Ran causal analysis on a 15.4% post-launch ESAU decline; isolated cohort mix and seasonality as primary drivers, preventing incorrect rollback.

LEADERSHIP EXPERIENCE

- Founding team member for analytics products (MyLoanBhai, TestMyPolicy, FinQy 2.0), scaling operations across 35 cities, 3 countries.
- Lead author on 4+ peer-reviewed IEEE papers with 15+ citations, spanning predictive maintenance, NLP-SQL automation, and ML modeling.
- Winner of global AI hackathon (1,311 participants, 19 countries), placing 1st out of 750 teams with international recognition for applied AI.

TECHNICAL SKILLS

- **Generative AI & LLMs:** RAG, Graph RAG, LangChain, LangGraph, Agentic AI, Prompt Engineering, Hugging Face, FAISS, ChromaDB.
- **Machine Learning:** Regression, XGBoost, Random Forest, LSTM, GNN, Clustering, PCA, Uplift Modeling, Forecasting, A/B Testing, SHAP.
- **Data Engineering:** Python, SQL, PySpark, Hive, Hadoop, Pandas, NumPy, Airflow, Tableau, Qlik Sense, Docker, Git, Jupyter, Databricks.
- **Analytics & Experimentation:** Causal Analysis, Root Cause Analysis, Metric Design, Hypothesis Testing, Statistical Modeling, PSM, DiD.
- **Cloud & Infrastructure:** AWS, GCP, Vertex AI, BigQuery, MLOps, Data Pipelines, Databricks, Docker, Notion, Excel.