

# Abhishek Jain

Email: [abhishekdm4@gmail.com](mailto:abhishekdm4@gmail.com)

Portfolio: <https://abhishekdm4.github.io/>

Mobile: +91 9560579902

LinkedIn: <https://www.linkedin.com/in/abhishekdm4/>

## SUMMARY

---

Data Scientist with 7+ years of experience, I excel at converting complex datasets into strategic insights that drive business value. My technical proficiency encompasses the complete data lifecycle, including data mining, visualization, and machine learning model development. Having worked with Senior Product Managers and Digital Marketers within Retail, E-Commerce, CPG and Logistics domains has helped me garner skills like strategic thinking and problem solving. A strong communicator, able to bridge the gap between technical teams and non-technical stakeholders, ensuring alignment and effective decision making.

## EXPERIENCE

---

### • Amazon

Hyderabad, India

Senior Data Scientist

Jan 2024 – Present

- Developed and deployed Topology tool using **XGBoost-based Demand Forecast and Optimization models**, that drives all new pickup location launch decision across US and EU5. Achieved accuracy of ~90% and **drove \$60M in CapEx investments** in 2025 while maintaining throughput guardrails.
- Productionized a **text classification framework using BERT** for ~150K/month SDS contacts and 10K+ feedback responses, achieving **82% accuracy (ROC)** on a 12-level taxonomy to surface systematic PARP delivery defects. Scaled the methodology to 5+ delivery streams, including Hub, DSP and Flex drivers.
- Designed a self-serve Text-to-SQL platform leveraging **GenAI** and a RAG-enabled knowledge base to translate natural language into complex customer targeting queries, achieving **~90% execution accuracy** and reducing ad-hoc analytics requests by **~65%**.
- Analyzed 12M customer interaction and transaction records to test hypotheses around engagement decay, behavioral volatility, and negative experience-driven churn. Engineered 150+ recency, frequency, concessions and stability features and built explainable **Logistic Regression and LightGBM models**, achieving early churn risk detection with a **recall of 85%**. Applied **SHAP-based** explainability to ensure transparent and ethical model behavior.
- Partnered with country heads to analyze the impact of shipping cost savings on overall contribution profits (CP), recommending strategic changes in eligibility criteria that drove **~8% incremental orders** resulting in increase of PARP CP by **23% YoY**. Unblocked 5 GL categories that further drove CP by 4% achieving Director Goal.
- Collaborated with marketing director to re-engage customers lost due to decommissions, built a self-serve tool to find nearest alternative pickup points, driving **~67% of customer retention** across 20 decommissions in EU and FE regions.

### • The Math Company

Bangalore, India

Senior Data Scientist

June 2021 – Dec 2023

- Designed and implemented a modular time-series forecasting framework in Databricks with an integrated feature store, forecasting hierarchical demand across region, marketplace, product group, brand, and SKU levels. Reduced experimentation runtime by ~80% and enabled rapid comparison across statistical and ML models (ETS, ARIMA, AutoARIMA).
- Developed a tire quality prediction and process-optimization tool using XGBoost and a Power BI-based front end, simulating the impact of machine-setting changes on defect rates. Estimated annual client savings of €4M.
- Built a self-service analytics solution for sealant compound application using high-frequency laser sensor data, applying FFT and convolutional matrix reduction techniques to detect anomalies. Improved operational efficiency by 97%.
- Created repeat-customer proxy logic using statistical tests and behavioral signals (purchase frequency, geography, seasonality), enabling leadership to quantify repeat vs new customer demand and improve planning accuracy.

- Developed a commodity price forecasting and decision-support tool to predict natural rubber prices using macroeconomic indicators, supporting hedging strategies. Augmented forecasts with NLP-based sentiment scoring of relevant market news.

- **Wipro**

Analyst

Noida, India

Nov 2019 – Jun 2021

- Built an end-to-end rolling 13-month workflow health dashboard in Tableau, from data ingestion and metric definition to deployment, tracking 350+ metrics across 10+ operational domains.
- Designed and optimized ETL workflows that reduced manual reporting effort by 40% and improved data reliability.
- Developed a VBA-based quality management tool with version control, increasing automated quality checks by 74% and improving reporting consistency.

- **Nurtr; Thought Roots Pvt. LTD.**

Data Analyst

Gurugram, India

Sep 2018 – Oct 2019

- Developed a lead classification machine learning model to predict hot vs cold leads, increasing conversion rates from ~40% to ~70%.
- Designed and implemented a lightweight CRM solution to support sales and business development teams, contributing to a 20% increase in upselling.
- Delivered executive-level dashboards and insights (GMV, revenue, collections) to support sales target planning and tracking.

## EDUCATION

---

- **Liverpool John Moores University**

Master of Science in Data Science

UK

Mar 2025 – Mar 2026

- **International Institute of Information Technology**

PG Diploma in Data Science

Bangalore, India

Mar 2020 – Mar 2021

- **Inderprastha Engineering College**

B. Tech CE

NCR, India

Aug 2013 – Jun 2017

## SKILLS

---

- **Programming:** Python, SQL, R
- **Machine Learning & Stats:** Supervised & Unsupervised Learning, XGBoost, CatBoost, Random Forests, Decision Trees, Logistic/Linear Regression (Ridge, Lasso, Elastic Net), K-Means & Hierarchical Clustering, Dimensionality Reduction (PCA), A/B Testing, Time Series Forecasting (ARIMA, SARIMA, SARIMAX, VAR)
- **Deep Learning & NLP:** PyTorch, Neural Networks (CNN, RNN, LSTM, GRU), Transfer Learning, Natural Language Processing (NLP), Anomaly Detection
- **Generative AI & LLMs:** Transformer Models (LLaMA, Mistral, BERT, ViT), RAG (Retrieval-Augmented Generation), PEFT/LoRA Fine-tuning, Model Quantization (FP32 to 8/4-bit), Hugging Face, LangChain, vLLM, Prompt Engineering
- **MLOps & Data Engineering:** Databricks, Flask, Vector Databases (Pinecone, Milvus), Git, Docker, CI/CD
- **Cloud & Visualization:** AWS, Azure, Tableau, Power BI, QuickSight, Data Studio, Excel

## AWARDS & ACHIEVEMENTS

---

- **Mechanism Maven Award 2024** at Amazon for building DSI Causal Inference framework.
- **Appreciation Award 2022** at The Math Company for building tire quality prediction model.
- **Appreciation Award 2021** at The Math Company for turning around client experience.