

Prithwjit Ghosh

Data Scientist Specialist | Accenture Global Technology

✉ ghoshprithwjit39@gmail.com | 📞 +91-7595986858/9230358950

🐙 GitHub | 📁 Portfolio | 🔗 LinkedIn

Data Scientist at **Accenture** with **3** years of experience and an **IIT Kanpur** Master's in Statistics, delivering predictive analytics solutions across forecasting, risk modeling, customer analytics, and financial planning through production-grade machine learning systems and business-focused insights.

Professional Experience

Data Scientist Specialist || *Accenture Global Technology*
Bengaluru, Karnataka, India

July 2023 – Present

● Sales and Guest-Count Forecasting | *Global Quick Service Restaurant (QSR) Brand*

- » Designed and productionized an ensemble forecasting system predicting sales and guest counts up to **48 months** ahead using *Prophet, Theta, MSTL, LightGBM, and Naïve* models with **macroeconomic** indicators for 6 different countries.
- » Engineered a custom horizon-aware evaluation framework (*trend & seasonal MAPE*) using **450+** forecasting models and ensemble balancing strategy to stabilize long-range forecasts, achieving **97 - 99%** accuracy for 24 months and **95 - 96%** for 48 months.
- » Implemented a production-grade **MLOps** pipeline using **Docker** and **GitHub Actions** for CI/CD, incorporating unit and integration testing along with security scanning (*SonarQube, Snyk*) prior to artifact deployment to **JBFrog** Artifactory
- » Deployed automated monthly forecasting workflows via **Astronomer** through Apache **Airflow DAG** on **AWS Fargate**, enabling reliable large-scale production forecasting.

● Intelligent Collections 3.0 | *Water Treatment Brand*

- » Built customer-level late payment prediction models using **XGBoost** with **50** selected features from **1000+** derived features using **feature importance**; achieved **84%** AUC (due-date) and **90%** AUC (due-month) with **10K+** monthly customer base.
- » Developed calibrated **risk scoring** and **risk categorization** using quantile calibration, reducing overdue amounts by **38%**, lowering open Accounts Receivable by **15%**, and increasing collections by **12%** over next 2 years after go live.
- » Designed short-term AR forecasting models (1 - 6 month horizon) across client segments using *statistical methods* (SMA, EWMA, ARIMA, SARIMA, Theta, MSTL) and *ML/DL models* (XGBoost, LightGBM, CatBoost, TCN).
- » Achieved **95-98%** accuracy across key categories — *Not Yet Due, Current Due, and Over Due*; results delivered through interactive **Power BI** dashboard for leadership and client reporting.

● Proof of Concepts

» Cash Flow Forecasting | *Power Utility Company*

Delivered end-to-end B2C cash flow forecasting using **6M+** customers data, improving Cash-In accuracy to **98%** (from 70%) and Cash-Out to **93%** (from 64%) through advanced ML modeling and enabling business insights via interactive dashboards.

» Marketing Data Analysis | *Internationally Renowned Liquor Brand*

Analyzed customer behavior across **subscription patterns**, marketing campaigns, conversion funnels, up-sell, cross-sell and churn propensity using **RFM** and advanced analytics to drive data-informed growth and retention strategies.

Self-Projects

● Demographic-Aware Recommender System with Facial Embeddings

Jun '25 - Mar '26

Built and deployed a facial-embedding based recommendation engine predicting *age, gender, and race* (up to **97%** accuracy), with a bias-aware **LangChain** agent for personalized suggestions, deployed via a containerized **Streamlit** app on **Hugging Face**.

● Agentic AI Travel Planner*(ongoing)

Apr '26 - Present

Built an end-to-end agentic planning system with **LangGraph**, **FastAPI**, **Streamlit**, hybrid retrieval, and specialist planning agents for routing, budgeting, and itinerary optimization, supported by an automated ingestion and evaluation pipeline.

Skills

- * **Programming Languages:** Python, SQL (MySQL, RedShift, Presto, DuckDB, SQLite), R, Git, GitHub Actions
- * **Tools/ Gen-AI:** MS Excel, MS PowerPoint, Power BI, FastAPI, Redis, VS Code, Cursor, Copilot, Claude Code, OpenAI, Claude
- * **Deployment/ Cloud Tools:** Docker, Kubernetes, Streamlit, Astronomer, Airflow, AWS (Fargate, EMR/ Hadoop, EC2, S3, Cloudwatch)
- * **Libraries:** Pandas, Scikit-learn, XGBoost, LightGBM, Prophet, NLTK, Spacy, Tensorflow, Keras, Langchain, Langgraph, Ollama
- * **Domains:** Data Science, Data Analysis, Predictive Modelling, Classification, NLP, Statistics, Time Series, BI, LLM, AI Agents, RAG

Education

🎓 **M.Sc. in Statistics**, Indian Institute of Technology Kanpur (IITK)

Aug '21 - Jul '23

CGPA: 8.9 | JAM Rank: 7 (AIR) | Department Rank: 6

🎓 **B.Sc. in Statistics**, Bidhannagar College

Jun '18 - Jul '21

CGPA: 9.99 | Department Rank: 1