

ANKITA BOSE

Senior Data Scientist

Senior Data Scientist with 5+ years of end-to-end ML/AI experience in Pharmaceutical and Marketing Analytics. Proven track record of designing predictive Machine Learning models, NLP-powered pipelines, and automated BI solutions that helps data-driven real-time decision making by turning data into insights, optimising Engagements, deliver measurable ROI. Expertise in Python, SQL, DataBricks, Machine Learning, Deep Learning (LSTM/RNN), LLM-integrated analytics, and ThoughtSpot Model & Live-board development. Additionally 2+ years as a Subject Matter Expert & Guest Lecturer in Applied Statistics. Adept at translating complex data into strategic business decisions across Fortune 500 clients.



EXPERIENCE



Aug 2024 - Present

GENPACT INDIA PVT. LTD.

Assistant Manager Data Science (Full-Time)

- Project : AI-Powered ThoughtSpot Data Model for Insights Advantage data, PBM metrics Automation :** Currently developing *LLM-integrated ThoughtSpot Data Model & Live-boards* with drill-down visualisations for transactional, financial, and operational PBM metrics, while integrating *Generative AI-powered Natural Language Querying (NLQ)* to enable non-technical Value & Access stakeholders to interrogate complex pharmacy claims data conversationally — surfacing real-time insights on prescription drug utilisation, claims reimbursement patterns, rebates and cost optimisation trends across a unified, self-serve analytics platform.
- Project : AI-Powered ThoughtSpot Data Model, HS Analytics & 340B Chargeback Automation:** Architected and deployed *LLM-integrated ThoughtSpot data models and Liveboards*, enabling *Natural Language Querying (NLQ)* and *AI-driven, self-serve insights* for stakeholders — reducing ad-hoc data requests by automating HS Standards, Regional performance, channel performance, Account Performance and 340B chargeback analysis. *Leveraged ThoughtSpot's Spotter 3 LLM-powered search layer and conducted Spotter Coaching* to transform raw commercial data into interactive, real-time true insights, accelerate decision-making timelines by eliminating manual reporting bottlenecks.
- Project : Automated Health System Segmentation and Targeting Pipeline Development:** Led the transition of manual Excel and SAS-based processes to an automated, scalable system using *Python and SQL* for a Health System *Segmentation, and Targeting solution*. Integrated 340B density, CMP win, Medicaid density, clinical partnerships, and patient/provider metrics into a unified segmentation framework. **Delivered ~\$500K annual savings** by eliminating vendor dependencies and **enabling robust performance tracking with automated tiering logic**.
- Project : MMx and ROI Automation:** Built an automated Python-based pipeline for *MMx modeling and ROI analysis*, replacing manual Excel/SAS workflows to improve reproducibility and efficiency. Designed *linear mixed-effects models* with advanced *feature engineering (outlier treatment, lagged variables, adstock transformations)* for accurate demand forecasting. Performed segment level and national level evaluations using *MAPE, contribution Analysis to quantify effectiveness*. Assessed *KAM & SAE impact* at account levels, measuring *Retrospective ROI* and *optimized NRx (excluding Medicaid)*.
- Project : Developed an end-to-end predictive analytics automated pipeline to optimize messaging strategies for HCPs, enhancing TRx performance:** Integrated multiple data sources (sales, activity logs, suggestions), engineered features (rolling windows, message frequencies, static HCP attributes), and built a *CatBoost quantile regression model* (tuned with Optuna) for TRx forecasting. Leveraged *SHAP* for feature insights, *Monte Carlo simulations* to optimize messaging sequences with business constraints, and *cosine similarity* for imputing missing data. This approach drove personalized, compliant communication strategies, improving HCP engagement and prescription outcomes.



June 2022 - Aug 2024

AXTRIA INDIA PVT. LTD.

Senior Associate Decision Scientist (Full -Time)

- Part of Data Science Team, involved in Data Processing, dealt with missing values, potential outliers and highly skewed features, performed normalization & scaling, conducted Feature Engineering, applied Machine Learning & Deep Learning Models, executed different Ad-hoc Statistical Analysis, worked in different tools – MS SQL, Redshift, Python, Databricks, MS Excel, MS PowerPoint, R etc. helped in real-time Data Driven Decision Making and participated in client interactions across projects for big pharmaceutical companies.
- Project: Developed & automated an AI orchestration project focused on sales prediction and impressions optimization** using *Databricks (Python and SQL)*. Integrated multiple datasets at the customer x year-week level over 2 yrs. Developed *LSTM models* for different segments with unique target variables, employing extensive feature engineering (training spans, lagged variables, mean, std, etc.). Optimized model architecture using various layers, activation functions, and optimizers, and evaluated performance with *R2, RMSE, MSE, and classification metrics*. Used *Genetic Algorithm (GA)* to optimize channel impressions, recommending an optimized HCP Pool based on model outcomes, GA recommendations, and business rules.
- Project: Customer Segmentation and Targeting Strategy for a vaccine portfolio**, to grow market share across IDN and HCP customers types: Applied different business rules to create a **master database** at account level, performed *Promotion Response Model* for accounts with good Sales and Call data to estimate *customer valuation and promotional sensitivity* and hence *Segment the HCPs* based on 5 relevant dimensions and statistical methods, *identified potential target list* based on desired impactable sales and available workload or budget. Finally, provided optimized call frequency recommendations.
- Project: RBM Team Sizing for a new Eye Injection (drug) launch:** Worked on developing analytical solution for Call Planning and RBM Team Sizing, actively involved in different Ad-hoc analysis to **maximize the RBM's productivity and efficiency**. Interviewed different stakeholders across teams to identify the key pain areas of RBM as well as methodology of **account prioritization, workload parameters**, performed *Machine Learning Models* to measure their impact on historical RBM workload, selected statistically significant parameters, hence *segment the accounts into 3 workload Segments (H/M/L) based on consolidated workload index* considering model weights. Multiple Levers were adjusted to finally arrive at 3 good scenarios of RBM sizing recommendation.
- Project: Test-Control Analysis using PSMPy :** Conducted test-control analysis with *PSMPy*, matching treated and control groups via propensity scores to ensure balanced covariates and accurate treatment effect estimation. Analyzed pre- and post 3-month NBRx data to calculate lift (DoD) and assess the impact of alert campaigns.

Skills

Regression & Classification

Random Forest

K-means Clustering

Decision Tree

LSTM, RNN models

Hypothesis Testing

CatBoost, XGBoost

LSTM, RNN models

DataBricks, ThoughtSpot, Snowflake, Redshift

Pharma Analytics, HCP engagement, Segmentation, Omni-channel Orchestration

Python, SQL, R, MS Excel, PPT

Education

M.SC. IN STATISTICS
University of Kalyani

2018 – 2020

B.SC. (HONS.) IN STATISTICS
Vivekananda College,
Affiliated to University of Calcutta

2015 - 2018

Personal Information

☎ : +91 9007341211 || +91 6291734316

✉ : abose5894@gmail.com

📍 : Bangalore, India

🌐 : [linkedin.com/in/abose1998](https://www.linkedin.com/in/abose1998)

ACHIVEMENTS & AWARDS

• Certificate of Recognition (Dec, 2025) - GENPACT

For consistent error-free outputs, exemplary performance and significant contribution in driving end-to-end HS Modelling, automation in Python/SQL within short span of time.

• Bravo Award (Mar – April, 2024) - Axtria

For consistent delivery of quality work in stringent timelines, received positive client feedback and for helping new team members get on board with KT.

• Game Changer Award (H2, 2023) - Axtria

For significant contribution towards overall objective of the team/project & display of passion and innovation which leads to measurable impact.

• Applause Award (Sept. 2021) - IPSOS

Entrepreneurial Spirit, Innovative Ideas, Partnership with US team and helping in standing up a new product at IPSOS

• Star of the Month Award (July 2021) - IPSOS

Younger one within the team, although taking a faster transition of a big project at IPSOS



Jan 2021 – June 2022

- **IPSOS RESEARCH PVT. LTD.**
Business Analyst (Full -Time)
 - Led **Multi-Touch Attribution (MTA) model** development for Fortune 500 advertisers across e-commerce channels, focused on **Optimizing Digital Marketing techniques** and investments, developed a **Data Analysis Framework** to perform **Data Pre-processing, Feature Engineering, Feature Selection**. Executing **Statistical Machine Learning Models** to Predict channel **Revenue and Conversion Effectiveness in real-time** across all marketing channels.
 - Analysis on Digital Media Investments for the clients in the e-commerce areas implementing **Statistical Techniques** by using **MS SQL, SQL Workbench (Amazon Redshift), R, Python, SAS**.
 - Leveraging cloud-based environments like **AWS, Google Big Query, Facebook Business Manager** in transforming and collecting client’s data and integrating them into modelling framework.



Jan 2020 – Dec 2020

- **CHEGG INDIA PVT. LTD.**
Subject Matter Expert in Statistics (Part -Time)
 - Provide expert solutions to different problems using Statistical Approach in R, Python, MS Excel



Jul 2019 – Dec, 2022

- **HOWRAH VIVEKANANDA INSTITUTION**
Guest Lecturer in Applied Statistics (Part-Time)
 - Delivered lectures on *Descriptive and Applied Statistics*, guiding students in solving statistical problems using *R and Python*.
 - Conducted classes on *advanced statistical techniques*, enhancing students' practical skills in data analysis.
 - Mentored students on *data science projects*, promoting critical thinking and problem-solving abilities.

Languages

English

Bengali

Hindi

Hobbies

Singing

Drawing

Travelling

Teaching



CERTIFICATIONS



- Several courses on **Data Science & Machine Learning & Python** from Udemy



- Training on - **DataBricks & ThoughtSpot**



- National Workshop on “**AI & ML**” from CIET approved by AICTE.



- Course on “**Data Science using R**” from StepUp Analytics.



- 21 days Summer Training Course on "**Data Science**" from **University of Kalyani**.



- Workshop on “**Data Science : Decoding the Black Box**” from **University of Calcutta**.



- National Workshop on "**Statistics and its Application**" from **Presidency University**.