

ACADEMIC QUALIFICATIONS

| Year | Degree /Board | University /Institution | %/CGPA |
|-------|---|---|---------|
| 2025* | Post Graduate Diploma in Business Analytics | IIM Calcutta, IIT Kharagpur, ISI Kolkata | - |
| 2023 | B.S. & M.S. (Dual Degree) Mathematics | Indian Institute of Science Education and Research Bhopal | 8.09/10 |
| 2018 | CLASS XII | Delhi Public School Bokaro | 94.6 % |
| 2016 | CLASS X | Notre Dame Academy Jamalpur | 10/10 |

KEY SKILLS/TOOLS Statistical Inference, ML, Probability Theory, Python, SQL, R, Excel, PySpark, Power BI, MATLAB, LaTeX

AWARDS AND ACHIEVEMENTS

| | |
|-------------------|---|
| Academic Awards | <ul style="list-style-type: none"> Qualified for State Level(III) Exam Vigyaan Vidhyarthi Manthan, organised at CSIR-CIMFR Dhanbad [2016] Selected for Summer Research Fellowship jointly sponsored by IASc-INSANA-SASI under IMSc Chennai ['21] International Rank 57 in National Science Olympiad conducted in 1000 cities & 14 countries across Asia |
| Case Competitions | <ul style="list-style-type: none"> Rank 2/1159 in HRMony, an HR-Organisational Structure in Start-ups case, hosted by IIM Ahmedabad ['23] Rank 4/980 in Scytale, performed analysis on World Bank data to draw insights, hosted by IIM Ranchi ['23] Rank 5/940 in IDB-Analytics 3.0, Surgical materials data analytics competition, hosted by IIM Calcutta ['23] Ranked Top 5/627 in E2A, dashboard creation for Gaadi Dekho using Power BI, hosted by XLRI Delhi ['23] National finalist (out of 283) in Cogentix, predict offer rejection by recruits, hosted by IMT Hyderabad['23] |
| Certifications | <ul style="list-style-type: none"> Completed Python for Everybody specialisation on Coursera offered online by University of Michigan ['20] |

ACADEMIC PROJECTS

| | |
|--|---|
| Credit Risk Modelling (Finance) | <ul style="list-style-type: none"> Predicted the probability of customer loan default for 15 lakh+ applications with 400+ features in 37 files Performed data aggregation, SMOTE for class imbalance, feature engg; Trained LGBM and CatBoost model Created weighted ensemble model using soft voting classifiers; Achieved an AUC score of 0.88 & recall 0.79 |
| Flight Delay Prediction (Classification) | <ul style="list-style-type: none"> Trained a multi-class classification model for flight cancellation and delay prediction on 1 lakh+ flights data Performed EDA, data cleaning and pre-processing; Implemented Random Forest & Decision Tree Classifier Executed Hyper-parameter tuning for XGBoost Classifier using Grid Search; Improved accuracy by 12.2% |
| IP Index Forecasting (Time Series) | <ul style="list-style-type: none"> Forecasted monthly Industrial Production index of electric and gas utilities in the US via time series analysis Used KPSS test to check trend stationary, de-seasonalized data, analysed ACF-PACF plots for SARIMA order Assessed AIC, BIC & ACF of Residuals for optimal SARIMA model selection & achieved MAPE score of 2.4% |
| Employee Salary Prediction (Regression) | <ul style="list-style-type: none"> Predicted Employee Monthly salary for a firm, utilizing MLR models on a dataset with more than 30 features Checked multicollinearity via VIF, feature importance using t-test; Outlier detection by Jack-Knife method Performed dimensionality reduction using PCA; Residual analysis using Quantile-Quantile plot & KS Test Employed Lasso and Ridge Regression, improving R² value to 0.946 from baseline Linear Regression model |
| Master's Thesis | <ul style="list-style-type: none"> Studied Fourier Extensions for smooth non-periodic functions, a technique to address Gibbs phenomenon Applied SVD to manage ill-conditioned matrices & further used Chebyshev points, reducing error by 10⁻⁸ Investigated its applications in computing the convolution of two compactly supported functions using FFT |

ADDITIONAL PROJECTS

| | |
|---|---|
| LLM-based SQL Query Generator (Gen AI) | <ul style="list-style-type: none"> Utilized Google PaLM & LangChain to convert natural language questions into SQL queries for a retail store Implemented few-shot learning with Hugging Face embeddings and ChromaDB for complex SQL queries Developed an easy-to-use web-based UI using Streamlit to facilitate interaction & querying MySQL database |
| Similar Product Recommendation (CNN) | <ul style="list-style-type: none"> Identified similar e-commerce products based on images & their description using non-parametric learning Extracted feature embeddings from images using pre-trained EfficientNet & DensetNet Deep CNN models Achieved F1-score of 0.84; identified top-k visually similar products using KNN & fine tuning the threshold |
| PII Detection (NLP) | <ul style="list-style-type: none"> Automated manual masking of Personal data by building a model on a dataset of 6800+ educational essays Generated POS tags using SpaCy; Tokenized via Bert Tokenizer; RegEx to identify phone nos., emails & urls Employed stacked & ensemble DeBERTa models to classify tokens; Improved F-5 score from 0.69 to 0.953 |
| Bank Customer Segmentation (Clustering) | <ul style="list-style-type: none"> Identified customer segments by examining 10 lakh+ transactions from 800k+ bank customers during 2016 Carried out RFM analysis; Utilized PySpark dataframe to store cleaned dataset & applied VectorAssembler Executed PySpark K-Means and GMM to achieve a silhouette score of 0.60; identified 5 customer profiles |

POSITIONS OF RESPONSIBILITY & EXTRA CURRICULARS

| | |
|------------------|---|
| Coordinator, TT | <ul style="list-style-type: none"> Sports Coordinator, IISERB ['22 -'24] Organized Sangharsh (IISERB Sports Meet) for 100+ participants |
| Sports Accolades | <ul style="list-style-type: none"> Silver Medal in General Championship TT at IIT KGP Runner Up in IISER Bhopal Sports Meet in TT MD Received Emerging Player Medal in TT and secured 1st position in Cricket in Section Wars at IIM Calcutta |
| DS Casebook | <ul style="list-style-type: none"> Devised case solutions & KPIs for Logistics & Healthcare industry in the 2nd version of PGDBA DS Casebook |

ELECTIVES : Baeyesian Methods, FRM, Comp. Fin., Big Data

INTERESTS : TT, Poetry Composition