

ARNAB ROY

Data Science

☎ +91-8617740451 ✉ arnabroy7640@gmail.com 🔗 <https://www.linkedin.com/in/arnab-roy-7b5bba22b/>

Examination	Majors	Institute	Year	CPI/%
Post-Graduation	Applied Statistics and Informatics	IIT Bombay	2022-2024	6.92
Graduation	Statistics	St. Xavier's College	2019-2022	8.3
Intermediate	Science	Sarada Vidyapith	2017-2019	93.8%

SCHOLASTIC ACHIEVEMENTS

- Secured a pre-placement offer from AXTRIA for the Data Analyst position (2023)
- Secured **AIR-77** in IIT JAM Mathematical Statistics exam conducted by IIT Roorkee (2022)
- Recipient of the **DST-INSPIRE-SHE** Scholarship, awarded to top 1 percentile students pursuing science

PROFESSIONAL EXPERIENCE

Data Analyst Intern | Axtria – Ingenious Insights (May'23-Jul'23)

Optimizing Promotional Efforts Across Bricks Using Clustering

- Employed **advanced clustering techniques** to **optimize promotional strategies** across market segments
- Developed a user-friendly interface using **Python's Dash framework** to apply various clustering methods
- Simplified the process for clients to input data and execute cluster analyses through **button-based interactions**
- Enhanced efficiency by enabling clients to directly **finalize and download** clustering models from the interface

ACADEMIC PROJECTS

Comparison of Portfolios | Course Project (2024)

Guide- Prof. Ashish Das, Mr. Prashant Sarkar | IIT Bombay

- Enhanced portfolio management by ranking stocks on **CAGR, ROCE, Beta & PE** using a multi-criteria approach
- Demonstrated improved profitability by evaluating 5-year portfolio performance using **XIRR and RAR**
- Achieved significant portfolio performance by benchmarking across various **NIFTY and BSE indices**

Classification of Brain Stroke Dataset | Course Project (2023)

Guide- Prof. Sanjeev V. Sabnis, Prof. Radhendushka Srivastava | IIT Bombay

- Conducted **Exploratory Data Analysis (EDA)** on **4981 entries** using Python libraries for valuable insights
- Applied **Standard Scalar normalization** to normalize the dataset and incorporating **SMOTE** for data balancing
- Implemented **Logistic Regression, LDA, Naive Bayes, kNN, Decision Tree, Random Forest and XGBoost** to compare their performances on the basis of measures such as **Sensitivity, Specificity and ROC**

Exploring Optimal Regression Models through Variable Model Selection | Course Project (2023)

Guide- Prof. Siuli Mukhopadhyay | IIT Bombay

- Strived to identify the **optimal model** for predicting the response variable using a **12-variable dataset**
- Performed **univariate analysis**, utilized **added variable plots** to gauge the marginal importance of predictors and selected the best subset of predictors on the basis of **Adjusted R^2 , Mallows's C_p , AIC and BIC criterion**
- Implemented **Stepwise Regression method**, performed **residual and outlier analysis** for model assessment and Developed a **multiple linear regression** model comprising 6 predictors, achieving an **accuracy of 80%**

Diabetes Data Analysis: Dimension Reduction and Clustering | Course Project 2023

Guide – Prof. Siuli Mukhopadhyay | Department of Mathematics, IIT Bombay

- Conducted a detailed Exploratory Data Analysis on **Pima Indian diabetes** dataset having 768 diabetic patients.
- Implemented techniques like **Principal Component Analysis** and **Factor Analysis** for reduction of features
- Explored various Clustering methods like Hierarchical and Non-Hierarchical i.e., **Agglomerative** and **K-Means**

KEY COURSES UNDERTAKEN

Academic and Online Courses

- Regression and Time Series Analysis, Derivative Pricing, Statistical Data Mining, Statistical Epidemiology
- Explore ML Workshop Beginner & Intermediate track by *Google AI*, Introduction to R, SQL by *Datacamp*

TECHNICAL SKILLS

Programming Language and Tools Python, R, SQL, Microsoft Excel, Powerpoint, Word, Tora, Latex

POSITIONS OF RESPONSIBILITY

Logistics Coordinator | *Mathematics Olympiad'24* | *Mathematics Association, IIT Bombay* (2024)

- Part of a **8+** core member team responsible for driving the **Logistics** of the Mathematics Olympiad '24.
- Managed logistics for the Math Olympiad, ensuring smooth operations for over 400 participants by coordinating venues, transportation, and accommodation

HOBBIES & INTERESTS

- Serving as a SME in Statistics and Probability on Chegg. India, assisting global students with their queries.
- Cricket, Football, Carrom, Watching Movies and TV series, Cycling, Paper Quilling.