

Mayukh Sen
(609)-429-0555

Piscataway, NJ

[Portfolio](#)

mayukh.sen13@gmail.com

SUMMARY

Data Science graduate student with advanced modelling and data analysis skills in R and Python. Keen interest in statistical modelling, data science & AI. Seeking entry-level jobs as data scientist to solve complex business problems.

EDUCATION

Rutgers University, New Brunswick, NJ

M.S in Statistics – Data Science

5/2025

Kalyani Government Engineering College

B. Tech, Electronics and Communication Engineering, 9.13 on a scale of 10

2023

PROFESSIONAL EXPERIENCE

Merck & Co.

Data Science Co-op (Neural Network Interpretability)

2024-present

- Collaborating with Merck's research group and Rutgers Statistics Department for evaluation of Neural Networks.
- Review of pretrained models on MNSIT dataset. Thorough analysis and evaluation of the performance on the data.
- Evaluating and summarizing interpretable Machine Learning algorithms like SHAP, Integrated gradients.
- Finetuning of hyperparameters to improve accuracy and performance of models on real life datasets.

M N Dastur & Co (P) Ltd

Data Engineering Intern

2023

- Implemented robust data pipelines using BigQuery, Snowflake and Cassandra in Cloud to support MLOps, enabling seamless integration of machine learning models into production.
- Troubleshooting issues in existing data pipelines and models. Doing ETL and building Data Warehouses and Data Lakes.
- Collaborated with clients in the energy sector to design and optimize CI/CD workflows, ensuring efficient data ingestion, and data consistency and reliability for machine learning applications in Google Cloud.

MinD Webs Ventures

Data Science Intern

2022

- Collaborated with the data science team to develop and refine machine learning models, applying algorithms such as Random Forest, XGBoost and AutoEncoders for recommendation systems, to perform Remote Sensing
 - Assess business strategies using A/B Testing Procedure and Causal Inference. Used Survival Analysis to finetune Lifetime Value Models and used Random Survival Forest Learning to classify into risk groups.
 - Worked on ML Ops, model deployment, handling distribution shifts and post-production monitoring.
-

RESEARCH PROJECTS

[Automatically Auditing LLMs by Discrete Optimization](#)

2024

- LLMs is automatically searched for input-output pairs. Audit models by specifying and solving a discrete optimization problem. We look for non-toxic prompt and toxic output pairs, French prompt- English output pairs by reversing LLMs.
- Used Coordinate Ascent (ARCA) and GBDA for reversing LLM and log probability loss to assess LLM performance.
- Using gpt-2, Mistral-7b ,BERT , and Civil Comments Dataset to perform experiments and detect toxic outputs.
- Observed similar performance over all LLMs denoting the effectiveness and robustness of ARCA for auditing LLMs.

[Auto EDA Library](#)

2024

- Library to automate ML workflow. Generates Time series plots, seasonal components, and trend automatically on call.
 - Generates Violin Plots, does Kernel Density Estimation, seasonal catplots and generates Fourier Terms automatically.
 - Performs Harmonic Regression using Fourier terms. Performs PCA and produces Scree plots to show variance explained
 - Packed as a class. Containerizing using Docker and to be published into python PyPi for open sourcing and public usage.
-

TECHNICAL AND OTHER SKILLS

Core Domain Expertise: Data Visualizations, Exploratory Data Analysis, Machine Learning, Data Preprocessing, Logistic Regression, Hypothesis Testing, Data Mining, Natural Language Processing, Deep Learning, LLM, Transformers, Clustering, Predictive Modelling, XG Boost, Random Forest, CatBoost, PCA, adaBoost, Pandas, NumPy, Scikit-learn
Computing and Programming: Python, R, SQL, Excel, TensorFlow, PyTorch, CI/CD Pipeline, DevOps, Huggingface
Communication: Business Proficiency in English, Fluent in Bengali and Hindi, Public Speaking, Conflict Resolution

HONORS AND ACTIVITIES

[Bloomberg 2023 Global Trading Challenge](#) (Rutgers University, New Brunswick)

2023

Led a team of 3 peers in the Bloomberg 2023 Global Trading Challenge, a university-based trading competition. Algorithm was based on following business news and investing in large variety of options. Finished the competition with a profit of \$98,112, ranked globally in the top 5% in a participation pool of 2000+. Finished as highest ranked team from Rutgers.

[Academic Success Mentor, Statistics – \(Rutgers University, New Brunswick, Thrive SSS\)](#)

2023-present

Provide statistics tutoring and academic support to first generation college students and to those in need of support. Assist students in developing improved study habits and a deeper understanding of subject matter through one-on-one sessions. Hands on tutoring on hypothesis testing, ANOVA, t-tests, probability distributions, statistical inference, Linear Regression.