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 Docker and MLflow containers in AWS to support MLOps, enabling seamless integration of machine learning models into production.
- Maintenance and improvement of existing data pipelines. Troubleshooting issues in pipelines and models.
- Collaborated with clients in the energy sector to design and optimize data workflows, ensuring efficient data ingestion, processing, and data consistency and reliability for machine learning applications in AWS Cloud Data Platform.

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, XG Boost, KNN for recommendation systems, to solve real-world business challenges.
- Conducted Exploratory Data Analysis (EDA) on diverse datasets. Applied advanced feature engineering methods, like dimensionality reduction, and categorical encoding, to optimize model input data and improve accuracy.
- Overseeing ML Ops, model deployment and model development ,troubleshooting and post production monitoring .

RESEARCH PROJECTS

Chicago Crime Dataset Time-Series Analysis   *2023*
- Performed a Time Series Forecasting on Chicago Crime Dataset from 2008 to 2023 of 22 columns and 4million rows.
- Performed EDA, dropped NA values, feature elimination, and feature engineering. Obtained informative plots, heatmaps.
- Analyzed trends and seasonal components of the data. Performed Spatial analysis and identified the crime hotspots.
- Forecasted crimes per day in the hotspots. Employed ARIMA, SARIMA, TBATS, Holt-Winters, LTSM models for the time series analysis and forecasting. Evaluated models based on RMSE and MAPE.
- LTSM consisted of two layers, and it produced the best forecasts on test data with minimal RMSE and MAPE.

Identification and Classification of Diabetic Retinopathy (Final Year Project)   *2023*
- Built a classification model utilizing Convolutional Neural Networks (CNN) like MobileNet2, EfficientNet to effectively classify fundus images on diabetic retinopathy severity.
- Obtained a classification accuracy of 96.1% and AUC of 0.936 despite the challenge of unbalanced class labels.
- Containerized the model using docker. Deployed the model on a user-friendly Web Interface.

TECHNICAL AND OTHER SKILLS

Core Domain Expertise: Data Visualizations, Exploratory Data Analysis, Machine Learning, Data Preprocessing, Data Management, Code Review, Data Mining, Natural Language Processing, Deep Learning, Databricks, pySQL, Git
Computing and Programming: Python, R, SQL, MS Excel, Tensorflow, PyTorch, keras, scikit-learn, pySpark, Docker,
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. Design workshops on probability distributions and other topics to enhance students' approach to statistical problems.