Saikat Bera

Kolkata, West Bengal, India | (+91) 9062794564
berasaikat | bsaikat | saikatb@cmi.ac.in | cmi.ac.in/ saikatb

Profile

I did my bachelor's in statistics and am currently pursuing my master's degree from CMI. I enjoy solving interesting problems and developing innovative solutions. I have a strong willingness to learn, analyze and execute strategies, and construct quantitative models that leverage advanced data-driven AI/ML approaches to develop inventive solutions.

Education

- **Chennai Mathematical Institute** — **Data Science**
  Master of Science, 2022 - 2024(Ongoing) — *CGPA: 8.81/10.00*
- **Presidency University, Kolkata** — **Statistics**
  Bachelor of Science, 2019 - 2022 — *CGPA: 7.31/10.00*

Work Experience

Data Science and Artificial Intelligence Intern - Shree Cement ltd.
Summer Internship, Kolkata — May 2023 - July 2023

- **Project:** Analytic Driven Exception Management for Influencer Transactions
- **Tools used:** Python, Scikit-learn, Oracle, OpenAPI, PyCaret
- **Description:** Utilizing the Material Transactions Ledger and Mitra Master, I formulated an unsupervised machine learning strategy designed to detect anomalous transactions. Subsequently, deployed this using OpenAPI giving access to the entire sales hierarchy.
- **Assisted:** I had the privilege to contribute to the projects "Advanced Alerts for Dealer Drop Out", "Dealer Segregation" and "Time Series Forecast of Sales" as part of my responsibilities.

Academic Projects

Predicting Unplanned ICU Admissions: An NLP-Based Approach Link
CMI, Chennai / Python, scikit-learn, matplotlib, nltk, keras, transformers — August 2023 - September 2023

- This project is dependent on the MIMIC-III database, which encompasses clinical data from patients admitted to critical care units at the Beth Israel Deaconess Medical Center between 2001 and 2012.
- I conducted a comparative analysis of Machine Learning, Deep Learning, and Natural Language Processing models applied to discharge text notes to predict unplanned hospital readmissions.

Artistic Image Fusion: A Code Implementation Study of Painterly Harmonization Algo Link
CMI, Chennai / Python, matplotlib, pandas, tensorflow — May 2023 - June 2023

- Code implementation of Deep Painterly Harmonization (Luan et al., 2018), an innovative algorithm for seamless integration of photos into paintings that enhances local statistical consistency, enabling superior results compared to traditional photo compositing and global stylization techniques.

Exploring Supervised Learning Techniques: A Comparative Study of Decision Trees, Naive Bayes, and Ensemble Models Link
CMI, Chennai / Python, scikit-learn, matplotlib, pandas — January 2023 - February 2023

- Conducted a performance analysis of two classifiers and two ensemble classifiers: one utilizing bagging with Decision Trees and the other employing boosting with Decision Stumps.
- This analysis was carried out on the "Bank Marketing Data Set". The primary objective of this classification task was to predict whether clients would subscribe to a term deposit.

Interactive World Happiness and Corruption Dashboard: An Academic Endeavor in Data Visualization Link

CMI, Chennai / R, ggplot, tidyverse, R-Shiny — October 2022 - December 2022

– Conducted graphical analysis and developed visualizations using a Shiny Dashboard to gain deeper insights into the World Happiness and Corruption dataset spanning from 2015 to 2020.

Visualisation and Exploratory Data Analysis on different data sets in R Link

Presidency University, Kolkata / Statistics, R — April 2022 - June 2022

– Outlier detection, searched for heteroscedasticity, autocorrelation, multicollinearity, the assumptions of the simple (/Multiple) linear regression model and tried fitting different models on the data sets *(Chicago, econ, autompg, Abalone, SENIC and birthwt)* based on different diagnostic plots.

Technical Skills

Languages	Python, R, MySQL, LaTeX (Overleaf / R Markdown).
AI/ML	Pandas, NumPy, Matplotlib, Scikit-Learn, PyTorch, TensorFlow, ggplot2, tidyverse, Shiny.
Misc.	Tableau, Power BI, Lyx, Microsoft Office (Excel and PowerPoint), Git.

Key Courses

Postgraduate	Programming and Data Structures (Python), Mathematical Methods – Analysis, R (Visualisation), RDBMS and SQL, Data Mining and Machine Learning, Design and Analysis of Algorithms, Distributed Computing and Big Data, Predictive Analytics*, Advanced Machine Learning*, Natural Language Processing*, Financial Modelling with Python*.
Undergraduate	Probability and Probability Distributions, Linear Models, Statistical Inference, Econometrics, Regression Analysis, Survey Sampling, Statistical Quality Control, Time Series Analysis, Stochastic Processes, Operational Research (OR), Real Analysis, Metric Space, Linear Algebra, Design of Experiments.

* *indicates ongoing courses*

Training and Certifications

DeepLearning.AI, Stanford University	Machine Learning Specialization Link
DeepLearning.AI	Deep Learning Specialization Link
DeepLearning.AI	Natural Language Processing Specialization Link
DeepLearning.AI	Generative Adversarial Networks Specialization Link

Awards and Achievements

2019 - Ongoing	**Innovation in Science Pursuit for Inspired Research (INSPIRE)**, Issued by Department of Science and Technology, Government of India.
2019 - 2022	**SVMCM: Swami Vivekananda Merit-cum-Means Scholarship**, Issued by Govt. of West Bengal

Positions of Responsibility

PREDICT Presidency University Statistics Society

Social Media Manager, PR Team

– Tried to promote a statistical community to help each other grow, develop and enrich ourselves collectively.

– Participated in organizing more than 10 different lectures on diverse topics by some eminent professors, researchers.