Soham Chakraborty

M.Tech in Computer Science, ISI Kolkata

+91 8240125480 | sohamchakraborty10qw@gmail.com |

53/37, Vidyayatan Sarani, Alambazar, Baranagar, Kolkata - 700035

-Linkedin | Github | Codechef | Leetcode | DOB: 30 December 2002

Introduction

Computer Science Engineer with hands-on experience as a Data Science Intern at Exposys Data Labs. Proficient in C, Python, Java, Machine Learning, SQL etc. Certified in Python programming by IIT Ropar and CMI-NPTEL. Successfully developed projects including a customer segmentation model, a diabetes prediction model, and an autonomous drone navigation system.

Education

Indian Statistical Institute, M. Tech in Computer Science — July 2024 – June 2026

- **Coursework:** Design & Analysis of Algorithms, Discrete Mathematics, Probability and Stochastic Processes

Government College of Engineering and Textile Technology, Serampore, B. Tech in Computer Science & Engineering — Dec 2020 – June 2024

- CGPA: 9.44/10.0 Certificate
- **Coursework:** Machine Learning, Natural Language Processing, Data Structures, Operating System, Database Management System, Design & Analysis of Algorithms

Kendriya Vidyalaya Cossipore, Class 12 - 94.8% — Apr 2019 – Mar 2020

Kendriya Vidyalaya Cossipore, Class 10 - 92.4% — Apr 2017 – Mar 2018

Internship

Data Science Intern, Exposys Data Labs — June 2023 – July 2023

- Analyzed and Preprocessed Customer data.
- Plotted various graphs for feature selection.
- Documentation of the Project

Projects

Autonomous Drone Navigation — Report

B.Tech Final Year Project, Co-authors: Soumajit Roy and Priyanshu Chaurasiya

- Data Preprocessing
- Developed different ML Models and compared them with the LTC, Cfc models.
- Creation of Offline Dataset in a rooftop environment
- Hyperparameters optimization using TPE sampling.
- Extensive real-time testing in various environments

Customer Segmentation Model — Report

- Data preprocessing and Feature Extraction from the Dataset with help of various plots
- Clustering the dataset using k-means clustering algorithm
- Using Elbow plot, Silhouette Score as well as Calinski- Harabasz Score to determine the number of clusters.
- Visualizations, such as scatter plots, were created to visualize the distribution of clusters, highlighting variations in spending behaviours and income levels.

Technical Skills

Programming : Python, C, Java, SQL, HTML, CSS

Operating Systems: Windows, Linux

Interests: Deep Learning, Machine Learning, Data Science, Artificial Intelligence, Competitive Programming