

Soumili Saha

+91 63718 57773 | logitechsoumili@gmail.com | linkedin.com/in/logitechsoumili | github.com/logitechsoumili

EDUCATION

Institute of Engineering and Management

B.Tech in Computer Science Engineering (AI & ML)

- CGPA: 8.79/10 — Minor in Cybersecurity

Kolkata, India

Jul 2024 – Present

The Newtown School

CBSE: Higher Secondary (Class XII): 87%, Secondary (Class X): 93.83%

Kolkata, India

2024, 2022

TECHNICAL SKILLS

Languages: Python, SQL (Relational Databases), Java, C, JavaScript

Machine Learning: TensorFlow, Keras, Scikit-learn, XGBoost, CNN

Data Science: Pandas, NumPy, Matplotlib, Seaborn

Frameworks & Web: Flask, Streamlit, HTML, CSS, REST APIs

Tools: Git, GitHub, Jupyter Notebook, Google Colab, VS Code

PROJECTS

Environmental Health Risk and Advisory System | *Python, TensorFlow, Flask, Llama 3.2* | GitHub Mar 2026

- Built an AI-powered system that analyzes environmental images (e.g., stagnant water, garbage accumulation) to detect disease risk factors and generate public health insights, implementing modular, tested backend components.
- Benchmarked CNN architectures (MobileNetV2, EfficientNetB0) using transfer learning and selected EfficientNetB0 as the final model based on superior performance.
- Integrated a Retrieval-Augmented Generation (RAG) pipeline with a local Llama 3.2 model and managed the project codebase using GitHub for version control and collaborative development.

Multi-Hazard Disaster Risk Prediction System | *Python, Scikit-learn, XGBoost, Streamlit* | GitHub Feb 2026

- Built a machine learning system to predict risks of natural disasters such as floods and heatwaves using meteorological and environmental datasets.
- Trained and evaluated XGBoost-based models for flood and heatwave prediction, achieving high accuracy (95–98%) based on cross-validation and test data.
- Integrated real-time environmental data from OpenWeather and OpenMeteo APIs and developed a Streamlit web app to visualize environmental indicators and generate location-based predictions.

EXPERIENCE

Machine Learning & AI Intern

May 2026 – Present

BeeSkilled | AICTE

Remote

- Completed hands-on projects in data preprocessing, supervised and unsupervised learning using Python and Scikit-learn.
- Built House Price Prediction, Iris Clustering, and Heart Disease Prediction models using Linear Regression, Logistic Regression, K-Means, and PCA.

AI & ML Intern – IBM SkillsBuild

Jan 2026 – Feb 2026

Edunet Foundation | AICTE

Remote

- Completed structured training in Artificial Intelligence and Machine Learning, covering ML algorithms, data preprocessing, and model evaluation.
- Developed an end-to-end disaster risk prediction application using Python, Scikit-learn, and XGBoost, integrating data preprocessing, model training, and deployment through a user-facing interface.

ACHIEVEMENTS

- Problem Solving:** Solved 75+ algorithmic problems across LeetCode, GeeksforGeeks, and HackerRank, focusing on arrays, hashing, sliding window, and string; achieved **5-star proficiency in Python** on HackerRank.
- Smart India Hackathon (SIH) 2025:** Led a team to design and submit a technical solution; advanced to the pre-final stage and ranking among the **top 50 teams** from the college.
- NPTEL:** Earned **Elite + Gold + Topper (Top 1%)** certification in *Programming in Java* with a **100% score**, and **Elite + Gold (90%)** certification in *The Joy of Computing Using Python*.
- Certifications:** IBM Python for Data Science & AI, DeepLearning.AI RAG, AWS Cloud Practitioner Essentials