SOMPARTHA SINHA

+91-8293186877 | sinhasompartha@gmail.com | Sompartha Sinha | sompartha

EDUCATION

JADAVPUR UNIVERSITY Expected-2027
Bachelor of Engineering in Construction Engineering **CGPA-7.5/10**

IIT MADRAS Expected-2027
Bachelor of science in Data Science and Applications **CGPA-8.5/10**

EXPERIENCE

AI/ML Intern
GrowGlobal • Supervised by Mr. Anirban Roy
Feb 2025 – June 2025

- Built a high-performance CAPTCHA-solving system using PyTorch, achieving **>97% character-wise accuracy** on a real + synthetic dataset.
- Developed a CRNN model with CTC loss and a custom CaptchaDataset class to handle variable-length, multi-character labels.
- Applied data augmentation (rotation, distortion, noise) to improve generalization across diverse CAPTCHA styles.
- Deployed the model using a **Flask API** and integrated it into a **Chrome Extension** for sub-100ms real-time CAPTCHA solving.
- Conducted extensive testing and optimization, including hyperparameter tuning, label encoding strategies, and synthetic data generation.
- 🏆 **Awarded Best Intern** for outstanding performance and impactful contributions to the automation pipeline.

Research Intern — Indian Statistical Institute (ISI), Kolkata
Under Prof. Agnimitra Biswas, mentored by Ankit Lodh | Ongoing Project
Duration: June 2025 – Present

- Working on a machine learning–based crop yield prediction system using environmental features such as reservoir levels, rainfall, and temperature.
- Conducting a detailed analysis to predict crop yields for a target year using historical data on rainfall, reservoir levels, and other environmental factors..
- Merged crop-wise yield datasets into a unified structure and engineered features to analyze rainfall patterns and state-wise variability.
- Trained multiple machine learning models (including Random Forest Regressors) to evaluate different modeling strategies — crop-wise vs state-wise, with and without log transformation.
- Performed year-wise prediction experiments to assess model accuracy by comparing predicted yield with actual recorded values for that year.
- Investigated the effect of skewed yield distributions across crops and implemented log scaling and state-level encoding to improve generalization.

SKILLS

Languages	Java, C++, python, C
Expertise	Data Structure and Algorithm(DSA), Problem Solving, Competitive Programming(CP)
Coursework	SQL, Bash Shell, Object-Oriented Programming, Tableau, Power BI
Databases	PostgreSQL, MySQL, MongoDB
Web Development	HTML, CSS, JavaScript, TypeScript, Flask, Django, Node.js, Vue.js, React.js,Tailwind CSS
Operating Systems	Windows, Ubuntu, Kali Linux
Machine Learning	ML Algorithms, scikit-learn, Fine-Tuning Large Language Models (LLMs)

CERTIFICATION

- **Cloud Computing** by IIT Kharagpur on NPTEL(certificate)
- **Investment Banking** by 365 Financial Analyst(certificate)
- **Financial Market** by Yale university(ongoing)

PROFILES

LEETCODE PROFILE GEEKSFORGEEKLS PROFILE CODEFORCES PROFILE GITHUB

PROJECTS

Stock Analyzer – JavaFX Desktop App for Real-Time Stock Tracking **(GITHUB LINK)**

- Developed a desktop application in JavaFX to fetch and display real-time stock prices using the Alpha Vantage API.
- Designed a clean and interactive GUI with search input, live price updates, and 7-day historical price charts using JavaFX LineChart.
- Implemented robust API integration and JSON parsing with Gson to handle live financial data for NSE/BSE-listed companies.
- Successfully demonstrated real-time stock monitoring and chart visualization through a responsive Java desktop application.

Capstone Project: Business Data Management (Hashtago Digital Marketing Company) **(GITHUB LINK)**

- Collect the raw sales and invoice data from the company and analyzed **sales and invoice data** using Google Collab to identify revenue trends, late payment patterns, and customer retention rates, improving actionable insights by **25%**.
- Designed interactive dashboards in **Tableau** and **Power BI** to visualize sales performance, retention metrics, and payment timelines, reducing overdue invoices by **20%**.
- Developed predictive models to forecast sales trends and customer churn, enabling stakeholders to improve retention strategies and optimize cash flow management, increasing operational efficiency by **15%**.

Built My Own Linear Regression Model from Scratch **(GITHUB LINK)**
Python, NumPy, Gradient Descent, California Housing Dataset

- Developed a complete Linear Regression model **from scratch using NumPy**, without relying on machine learning libraries for model training.
- Implemented **manual gradient descent** for parameter optimization, gaining deep understanding of how linear models learn.
- Learned the importance of **feature scaling**, **cost function minimization**, and how to evaluate models using **RMSE, MAE, and R^2 score**.
- Visualized training loss over epochs and compared **actual vs. predicted values** to assess model performance.
- Strengthened my foundations in linear algebra, optimization, and machine learning workflows.