

DHRUBOJYOTI HAZRA

Email: dhrubojyotihazra@gmail.com | Phone: +91-8100105934 | Kolkata(West Bengal)

LinkedIn: <https://www.linkedin.com/in/dhrubojyoti-hazra-a71a9a325>

GitHub: <https://github.com/dhrubojyotihazra>



OBJECTIVE

Third-year B. Tech Data Science student with a 9.0 CGPA and hands-on experience in ML model development, RAG architectures, and data visualization. Seeking a Data Analytics or Data Science internship to build scalable analytical pipelines, structured data workflows, and drive data-backed business decisions.

EDUCATION

- **Brainware University** – B.Tech in Computer Science and Engineering (Data Science) *CGPA: 9.0/10 | Expected June 2027*
- **Bagati Ram Gopal Ghosh Higher Secondary School** –WBCHSE Board 12th (Pure Science) *Percentage: 70% | 2023*
- **Tribeni Tissues Vidhyapith** –ICSE Board 10th (With Computer Applications) *Percentage: 86% | 2021*

SKILLS

- **Programming Languages:** Python, Java, C, SQL
- **Databases:** MySQL, Oracle SQL, Supabase, SQLite
- **Machine Learning & AI:** Random Forest, KMeans Clustering, Scikit-learn, Hugging Face Transformers, RAG architectures, Gemini API
- **Tools & Frameworks:** Power BI, Tableau, Excel, Git/GitHub, React, FastAPI, Streamlit, Plotly, Pandas, NumPy
- **AI-Assisted Development & Workflows:** Proficient in rapid prototyping and development using Cursor, Windsurf, Replit Agent, Lovable, Claude Code, and n8n automation.
- **Core Fundamentals:** Data Structures & Algorithms, Operating Systems, DBMS, Computer Networks

EXPERIENCE AND TRAINING:

- **Data Analytics Trainee | Euphoria GenX, Kolkata | Jan 2026 - Present**
 - Participating in an industry training program focused on structured data workflows and business intelligence.
 - Building data pipelines and developing interactive analytical dashboards using Power BI to derive actionable insights from complex datasets.
- **Data Analytics & Visualization Simulations | Forage | Aug 2025**
 - Completed the Deloitte Australia Data Analytics Job Simulation, focusing on real-world data problem-solving.
 - Completed the TATA Data Visualization program, creating dashboards to empower business stakeholders with effective insights.
- **Software Development Trainee | Samsung Innovation Campus | Oct 2025**
 - Successfully completed intensive training in coding, programming, and algorithmic problem-solving.

PROJECTS

Automatic Attendance System & Smart Curricular Platform (SIH 2025) | Sep 2025

- Led a **team of 6 members** as Project Leader at **Smart India Hackathon 2025**.
- Designed an **AI-powered automatic attendance system** with face recognition for real-time validation.
- Oversaw **task delegation, architecture design, and sprint workflow management**.

Health Tracker – Vital Analytics Dashboard

(Samsung Innovation Campus Hackathon 2025) | Oct 2025 | [GitHub](#)

- **Tech Stack:** React, TypeScript, Flask, Pandas, HuggingFace Transformers, Google Gemini, Plotly.
 - Developed a **Streamlit-based** health tracking web application during a **3-day hackathon**.
 - Implemented **RandomForest** ML models using **Scikit-learn** to predict diabetes and heart disease from Kaggle datasets.
 - Built data preprocessing and **analytics** workflows using **Pandas and NumPy** with **interactive visualizations**.
 - Enabled **CSV upload** and generated **downloadable PDF health reports**.
- Tech Stack:** Python, scikit-learn, Pandas, NumPy.

NovaTrade AI

(Personal Project) | Feb 2026 | [GitHub](#)

- Built a **crypto trading analytics dashboard** merging historical trade execution data with the Fear & Greed Index to uncover sentiment-driven trading patterns.
- Applied **KMeans clustering** to segment traders into behavioral archetypes (High-Volume, High-Frequency, Cautious) and trained a **RandomForest classifier** to predict next-day profitability using lagged sentiment and PnL features.
- Delivered an interactive **4-tab Plotly dashboard** with real-time sidebar filters and actionable strategy recommendations derived from ML insights.
- **Tech Stack:** *Python, Streamlit, Scikit-learn, Pandas, NumPy, Plotly.*

CodeGalaxy

(XiBit Hackathon – Texibition Event) | Mar 2026 | [GitHub](#)

- Developed an **interactive 3D WebGL visualization tool** that renders GitHub repositories as explorable universes, with files positioned by semantic similarity using UMAP and Sentence-Transformers.
- Integrated **Tree-sitter deep code parsing** to compute Fan-In, Fan-Out, and Cyclomatic Complexity, with AI-generated architectural summaries via OpenRouter.
- **Tech Stack:** *React, React Three Fiber, Vite, FastAPI, Python, Tree-sitter, UMAP, Sentence-Transformers, SQLite.*

OpinionIQ

(College Project – VAC: Basic AI Tools & Applications | Under Development) | [GitHub](#)

- Building an **enterprise-grade AI sentiment analysis web app** that ingests customer feedback CSVs and auto-detects the feedback column using Gemini 2.5 Flash.
- Classifies sentiment using **DistilBERT (HuggingFace)** and generates interactive Plotly dashboards (pie charts, keyword bar charts) with an AI-written executive summary.
- Features a **dual-agent chat interface** supporting both data queries (quantitative analytics) and report generation (qualitative content) using RAG-based context.

CERTIFICATES

- **Voices of Tomorrow'-A Speech Competition on India of the future – Brainware University (May 2025)**
- **Deloitte Australia Data Analytics Job Simulation – Forage (Aug 2025)**
- **Data Visualization: Empowering Business with Effective Insights – TATA on Forage (Aug 2025)**
- **Internal Hackathon - Smart India Hackathon 2025 – Brainware University (Sep 2025)**
- **Samsung Innovation Campus-Coding and Programming – Brainware University (Oct 2025)**