

KAREEB SADAB

Data Scientist • Machine Learning & AI Engineer • Security Practitioner

Chattogram, Bangladesh | +880 1636 466 777 | kareebsadab@gmail.com

[linkedin.com/in/karib-sadab](https://www.linkedin.com/in/karib-sadab) | github.com/Solidx74 | [Portfolio](#) | [Data Science Site](#) | [Data Analysis Site](#) | [Medium](#)

SUMMARY

CS undergraduate at CUET specializing in data science, machine learning, and AI engineering, with a parallel track in application security as a former Chief Information Security Officer. Has shipped 15+ end-to-end projects spanning predictive modeling, NLP/LLM-powered tools, and Dockerized ML APIs — including production security work that took a Web3 platform through audit to token launch. Comfortable across the full data science lifecycle: data collection and cleaning, EDA, feature engineering, model training and evaluation, and deployment behind REST APIs.

TECHNICAL SKILLS

Programming: Python, SQL, C, C++, Java

Machine Learning & AI: Scikit-learn, TensorFlow, PyTorch, Pandas, NumPy, NLP, LLM Integration, Prompt Engineering, RAG

Data Science & Analytics: Exploratory Data Analysis, Data Cleaning & Wrangling, Feature Engineering, Predictive Modeling, Cross-Validation

Visualization & BI: Power BI, Tableau, Matplotlib, Seaborn

Backend & APIs: Flask, FastAPI, REST APIs, Authentication

Databases: PostgreSQL, MySQL, MongoDB, Supabase

DevOps & Tools: Docker, Git/GitHub, Linux, Jupyter Notebook

Cybersecurity: Application Security, Threat Modeling, Vulnerability Assessment, Secure SDLC, SIEM, Burp Suite, OWASP

EXPERIENCE

Chief Information Security Officer (CISO) — W3D (Web3 Discord Community Agency)

Jun 2025 – Jul 2026

- Owned application security across multiple community-driven Web3 projects — including Harold Health, Undead Blocks, and LightningWorks — running threat models and vulnerability assessments before anything shipped.
- Worked directly with development teams to bake security into the SDLC rather than treat it as a final checkbox; Harold Health passed review and reached token launch, now listed on Dexscreener.
- Led a security and SEO audit of the agency's own web presence, producing a prioritized action plan and following through on the fixes.
- Translated technical risk findings into clear, non-technical recommendations for stakeholders and developers.

FEATURED PROJECTS

AI Resume Analyzer / BetteResume-ai | Python, FastAPI, React, MongoDB, LLM APIs, NLP

- Built a REST API that parses resumes against a job description using NLP/LLM calls, scores ATS fit, and returns targeted feedback.
- Extended into a full-stack AI-agent platform with a React frontend and MongoDB-backed persistence that generates ATS-optimized resumes end to end — shipped and live.

→ [Live Demo](#) | [GitHub](#)

Customer Churn Prediction | Python, Scikit-learn, Flask, Docker

- Ran the project end-to-end — data cleaning, feature engineering, and Gradient Boosting model training with SMOTE-based class-imbalance handling — then deployed behind a Dockerized Flask API.
- Kept the ML layer separate from the serving layer so the model can be retrained or swapped without touching the API contract.

→ [GitHub](#)

SIEM Log Pipeline + AI Anomaly Detection | Python, Pandas, Isolation Forest, One-Class SVM

- Built a real-time log parsing and feature extraction pipeline feeding Isolation Forest and One-Class SVM models to flag anomalous activity — bridging the ML and security skill sets.

→ [GitHub](#)

Academic Performance Prediction | Python, Scikit-learn, Random Forest

- Built a full ML pipeline predicting student outcomes from behavioral and academic data — data cleaning, feature engineering, and evaluation on a Random Forest classifier.

→ [GitHub](#)

Loan Approval Prediction | *Python, Scikit-learn, SVM, Decision Tree*

- Built a classification pipeline for credit-risk assessment, comparing SVM and Decision Tree models to automate a manual approval check.
→ [GitHub](#)

PitchIt AI — Sponsorship Outreach Generator | *LLMs, Next.js, Supabase, Firecrawl, Tavily*

- Built a tool that drafts sponsorship outreach by pulling live research through Firecrawl and Tavily first, so the output isn't generic LLM text — storage handled through Supabase.

Additional data work: SQL-based analysis (COVID-19 dataset exploration, Nashville Housing data cleaning in SQL Server), web-scraping pipelines for public/private company data (BeautifulSoup, Pandas), an automated crypto-market tracker (CoinGecko API), and published Tableau Public / Power BI dashboards on data-professional salary benchmarking. Full write-ups on the [Data Science](#) and [Data Analysis](#) portfolio sites.

EDUCATION

B.Sc. in Computer Science & Engineering — Chittagong University of Engineering & Technology (CUET) *In Progress*

CERTIFICATIONS

- Google Cybersecurity Certificate
- ISC² Certified in Cybersecurity (CC)
- LetsDefend — SOC Fundamentals

PUBLICATIONS & WRITING

- **W3D / Harold Health Security Report** — identified and remediated a critical broken-access-control vulnerability in a patient-provider credential registry. [Read on Medium](#)
- AI/ML and security articles published on [Medium \(@kareebsadab\)](#)

HACKATHONS & COMPETITIONS

- Competed in the IEEE Computer Society Hackathon, collaborating in a team to design and build a working software solution under time constraints.
- Selected for an upcoming DevOps Hackathon.
- Participated in multiple Kaggle Datathons, applying data cleaning, EDA, feature engineering, and model evaluation to real-world datasets.
- Competed in multiple CTF competitions spanning web security, cryptography, reverse engineering, forensics, networking, and OSINT — ranked #938 worldwide in picoCTF 2026.