Soham Ghosh

Kolkata, West Bengal

 soham7704 | Soham Ghosh | My Portfolio | sohamghosh7704@gmail.com | +91 8777026712

Professional Summary

Results-driven **B.Tech Computer Science Engineering** student passionate about transforming complex data into actionable insights. Skilled in **Python, SQL, Power BI, and Generative AI**, with hands-on experience in **data analytics, data visualization, and AI-driven applications**. Proven ability to integrate analytics and AI for real-world problem-solving. Adept at working with **LangChain and Streamlit** to build intelligent, interactive solutions. Seeking to leverage analytical and programming expertise in a **Data Analytics / GenAI Developer** role.

Core Competencies

Programming & Databases: Python, SQL, Java, MySQL, HTML, CSS

Data Analytics & Visualization: Power BI, Advanced Excel, Data Cleaning, BI Dashboards

Frameworks & Libraries: LangChain, Streamlit, Pandas, NumPy, Matplotlib, Seaborn, HuggingFace Transformers

Tools & Platforms: Jupyter Notebook, Google Colab, Git, VS Code, DBeaver, Netlify

Machine Learning & AI: Generative AI, Retrieval-Augmented Generation (RAG), Time-Series Forecasting, Volatility Estimation, Deep Learning Models

Education

University of Engineering & Management, Kolkata   **2022–2026 (Expected)**
B.Tech, Computer Science Engineering — GPA: 7.86/10 (Till 6th Semester, Equivalent: ~78.6%)

Julien Day School, Howrah, ISC   **2022**
ISC: 74.66%

Projects

RepoSage: AI-Powered GitHub Repository Analysis Tool   GitHub Link
Tech Stack: *Streamlit, Python, LangChain, FAISS, HuggingFace Transformers, Groq API*
- Developed a repository intelligence tool enabling developers to query GitHub repos with **memory-enabled follow-up support**, reducing code exploration time by **40%.**
- Implemented **Retrieval-Augmented Generation (RAG)** with HuggingFace embeddings for accurate, context-aware answers.
- Integrated NLP to interpret developer queries and return real-time, code-aware responses.

AI Career Advisor: Resume & Job Description Matcher   GitHub Link
Tech Stack: *Streamlit, Python, LangChain, PyMuPDF, Groq LLM*
- Built a tool achieving **95% prediction accuracy** in resume-job description match scoring.
- Delivered **85% precise skill-gap detection**, improving candidate job readiness.
- Enhanced resume keyword alignment by **70%** through personalized AI-generated recommendations.

Financial Analysis – Stock Price & Volatility Forecasting   GitHub Link
Tech Stack: *Pandas, NumPy, Statsmodels, YFinance, Plotly, ARIMA, SARIMA, GARCH, XGBoost, LSTM*
- Achieved **85%+ prediction accuracy** for Tata Motors stock trends using time-series forecasting and deep learning.
- Integrated **real-time data pipelines** with 7-day forward forecasting for dynamic investment strategies.
- Performed advanced statistical diagnostics **(ADF, ACF/PACF, Ljung-Box)** to validate model performance.

Cafe Sales Analysis: SQL-Based Retail Analytics   GitHub Link
Tech Stack: *MySQL, DBeaver*
- Reduced trend detection time by **30%** using advanced aggregations, window functions, and time-series analysis.
- Forecasted **monthly revenue growth** and identified **peak sales hours** to optimize branch operations.
- Designed custom **SQL queries and automated reporting workflows**, improving decision-making speed for branch managers.

Certifications & Achievements

- HackerRank SQL (Basic) – Credential ID: [7865/HH/709]
- HackerRank 5★ Gold Badge – SQL
- German Language A1 – Henry Harvin Institut (2024)