

# Arka Barua

📞 +880 1864-644612

✉ arka.barua.173@gmail.com

📍 Chattogram, Bangladesh

🔗 [arkabarua173.github.io/ark](https://arkabarua173.github.io/ark)

in Arka Barua

🔄 ArkaBarua173

## Summary

---

As a recent Computer Science and Engineering graduate seeking opportunities as a Data Analyst and Machine Learning Intern, I bring a solid foundation in programming languages and a keen interest in data analysis and machine learning. Proficient in Python, Pandas, Matplotlib, Seaborn, Plotly, Scikit-learn, SQL, MS Excel, PowerBI and eager to apply my skills in real-world projects.

## Education

---

**BSC North South University**, Computer Science & Engineering 2017-2023

- **CGPA:** 3.22/4.0
- **Trail:** Artificial Intelligence and Networks
- **Coursework:** Design and Analysis of Algorithms, Database Systems, Machine Learning, Pattern Recognition.

**HSC Chittagong Govt. Model School & College**, Science 2016

- **GPA:** 3.93/5.0

**SSC Chittagong Govt. High School**, Science 2014

- **GPA:** 5.0/5.0

## Certificates & Courses

---

**SQL (Basic) Certificate** July. 2024

(Issued By HackerRank)

[www.hackerrank.com/certificates/e5fc650fe728](https://www.hackerrank.com/certificates/e5fc650fe728) 🔗

**Machine Learning A-Z: AI, Python & R + ChatGPT Prize [2024]** May. 2024

(Issued By Udemy)

[www.udemy.com/certificate/UC-973b796e-f7bf-4729-9ca6-3e96d4301eac](https://www.udemy.com/certificate/UC-973b796e-f7bf-4729-9ca6-3e96d4301eac) 🔗

**Advanced SQL** May. 2024

(Issued By Kaggle)

[www.kaggle.com/learn/certification/arkabarua173/advanced-sql](https://www.kaggle.com/learn/certification/arkabarua173/advanced-sql) 🔗

## Skills

---

**Languages:** C, C++, Python, PHP, HTML, CSS, SQL, JavaScript

**Libraries:** Pandas, Scikit-learn, Matplotlib, Plotly, React, Tailwind CSS

**Software:** Visual Studio Code, Tableau, PowerBI, MS Excel, Microsoft SQL Server, PostgreSQL

## Projects

---

### Interactive Sales Dashboard for an E-commerce Platform

July 2024

- In this project, I visualized key e-commerce insights using Tableau to drive data-driven decisions.
- The dashboard includes KPIs, sales by category, payment distribution, regional sales, and city breakdowns.
- **Tools Used:** Tableau.

### Bike Sales Performance Analysis with SQL and Power BI

June. 2024

- This project leverages SQL and Power BI to analyze and visualize bike sales data.
- It focuses on key metrics such as order status distribution, average processing times, and sales trends over time.
- The analysis provides insights into customer purchasing behavior, product performance, and inventory management, helping to optimize sales strategies and operational efficiency.
- **Tools Used:** Python, PostgreSQL, DAX, PowerBI.

### Abalone Age Regression with Ensemble Method

April 2024

- This project aims to predict the ages of abalone.
- XGBRegressor, CatBoostRegressor and LGBMRegressor are trained.
- A Voting Regressor ( XGBRegressor + CatBoostRegressor + LGBMRegressor ) with soft voting is trained as the final model.
- The primary evaluation metric used is Root Mean Squared Logarithmic Error.
- **Tools Used:** Python, Pandas, Matplotlib, Seaborn, Scikit-learn, XGBoost, CatBoost, LightGBM.

### Binary Classification with an Imbalanced Dataset

Feb. 2024

- This project aims to predict the probabilities of customers exiting the bank.
- XGBoost is used to address the imbalance by increasing the weights of the minority class.
- The model was evaluated using accuracy, precision, F1 score, recall, and AUC-ROC score.
- SHAP (Shapley Additive Explanations) is used to understand the model.
- **Tools Used:** Python, Pandas, Plotly, Scikit-learn, XGBoost, SHAP.

## References

---

### Sumoy Barua

Lead Software Engineer, Cefalo

**Contact:** +880 1610-001914

### Dr. Mahdy Rahman Chowdhury

Associate Professor, North South University

**Contact:** +88 02 55668200