

Soumyadip Ghorai



Professional Goals

MSc Data Science graduate with work experience in **MNCs, startups** and mentored **150+ students in Analytics**, looking forward to learn, assist, collaborate and grow with the best minds in data science and analytics to create more industry-ready experts in Data Science.

Contact Information

Mobile: +91 7872817970

Email: ghorai.soumyadip33@gmail.com

Other Links :

Portfolio : [/soumyadipghorai.github.io](https://soumyadipghorai.github.io)

kaggle : [/soumyadipghorai](https://soumyadipghorai)

Code Chef : [/sghorai_2000](https://sghorai_2000)

LinkedIn : [/soumyadip-ghorai](https://soumyadip-ghorai)

GitHub : [/soumyadipghorai](https://soumyadipghorai)

Tableau : [/soumyadip.ghorai6618](https://soumyadip.ghorai6618)

LeetCode : [/sghorai](https://sghorai)

Skills

Languages

Python | SQL | HTML | CSS

Technical

Machine Learning | Web Scraping | Data Visualization | Web Development | Data Analysis | Natural Language Processing

Tools

Tableau | VS Code | GitHub | MS Office | Jupyter Notebook | Excel | figma

Academic History

Christ University

Bangalore | 2021 - 2023

MSc Data Science - 8.87/10

Indian Institute of Technology Madras

Online | 2021 - 2023

BSc Data Science - 7.2/10

University of Calcutta

Kolkata | 2018 - 2021

BSc Statistics - 7.74/10

Achievements

- 3X Expert @**Kaggle** with 4 data sets, 10+ notebooks, & 1000+ views.
- 3star @**codechef** and 475+ problems solved on **leetcode** using python.

Other Activities

- Google Developers student club **Machine Learning Lead**.
- Represented my college in multiple **state-level athletic meets**.
- **Cultural head** of the department and hosted 2 departmental events.
- Mentored **200+** students in **Python, SQL, ML, and Analytics**.

Work Experience

KPMG India

Consultant Data Scientist | Bangalore [[certificate](#)]

FEB 2023 - May 2023

- Successfully completed training in **Python, SQL, machine learning, Linear Programming**, and other consulting skills and got selected for full-time.
- **Won 2 consecutive coding competitions** hosted by KPMG.

Skills: Python, SQL, Machine learning, linear programming VS Code, MS Office

Tweek Labs

Data Analyst Intern | Bangalore [[certificate](#)]

MAR 2022 - AUG 2022

- Implemented methods like **moving avg**, and **selective scaling** in python to remove fluctuations in sensor data to get accurate results.
- **Implemented new features** like max hip rotation velocity, max shoulder speed and **developed aggregated scoring** methods to rank players according to their stats to help organizations **create leaderboards**.
- Applied **KNN** to predict ground contact with an average accuracy of **11ms** to get leg parameters more accurately.
- Set up the pipeline to store data in google Sheets using **Google API**, and made an interactive dashboard using **Meta Base** to track **KPIs**.

Skills: Python, Meta Base, Google Sheets, Google API, PyCharm, VS Code, GitHub

Ericsson

Summer Intern | Kolkata [[certificate](#)]

SEP 2021 - OCT 2021

- **Root Cause Analysis project:** Predict the root cause and recommend possible resolutions from the error messages from ENM upgradation logs to reduce the overall processing time by **up to 50%**.
- Wrote a **generalized Python parser** to parse all XML log files and create a clean JSON file to upload on elastic search.

Skills: Python, VS Code, MS Office

Projects

Net worth Predictor : [[GitHub](#)] [[App Link](#)]

AUG 2022 - SEP 2022

Flask-based alternative of FB games to predict your potential future net worth.

- Scraped the data of HNIs using **Beautiful Soup** consisting of people from **100+ categories** and **130+ countries**.
- Preprocessed the raw data, performed **EDA** to understand the data, and did **feature engineering** and **feature scaling** using python.
- Used **Regularization, KNN & Random Forest** models to predict net worth based on available features with an **MAE of 1.06**.
- Created frontend using **HTML, CSS, Bootstrap, jinja code**, backend using **flask**, and deployed using **Heroku**.

Tech: Machine learning, Python, Flask, jinja code, HTML, CSS, Bootstrap, VS Code, git, Heroku, Google Analytics

Personal Finance : [[GitHub](#)] [[App Link](#)]

FEB 2022 - MAR 2022

- Collected data on a daily basis of my expenses in a **spreadsheet** and created an interactive **Streamlit** web app to keep track of my expenses.
- Uploaded on GitHub and deployed on Heroku and the project got **featured** on the Streamlit **community forum**.
- Basic features include **Pie charts, bar graphs, spending trends, treemaps**, and **Q&A** to help keep my avg monthly expenses within **5000/-**.

Tech: Python, streamlit, plotly, Heroku, HTML, CSS, markdown, VS Code, git

Customer churn : [[GitHub](#)] [[App Link](#)]

OCT 2022 - NOV 2022

- Used **feature engineering & feature selection** to get the **6 best** features.
- Used a **Stacked classifier** model combining **SVC, Random Forest**, and **Gradient Boosting** to predict churn using 6 features with **76%** accuracy.
- Created Web app using **HTML, CSS, Python** and deployed using **Heroku**.

Tech: ML, python, Flask, HTML, CSS, Bootstrap, VS Code, GitHub, Heroku