Soumyadip Ghorai

Professional Goals

MSc Data Science graduate with work experience in **MNCs, startups** and mentored **100+ students in Python**, 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
kaggle : /soumyadipghorai
Code Chef : /sghorai_2000
LinkedIn : /soumyadip-ghorai
GitHub : /soumyadipghorai
Tableau : /soumyadip.ghorai6618
LeetCode : /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 | PyCharm | GitHub | MS Office | Jupyter Notebook | Excel

Academic History

Christ University
Bangalore | 2021 - 2023
MSc Data Science - 3.54/4

Indian Institute of Technology Madras Online | 2021 - 2023
BSc Data Science - 7.5/10

University of Calcutta
Kolkata | 2018 - 2021
BSc Statistics - 7.74/10

Achievements

- 3X Expert @**Kaggle** with 10+ notebooks, 4 data sets & 100+ upvotes
- 3star @**codechef** and 300+ 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 **100+** students on **Python, ML, and Analytics.**

Work Experience

KPMG
FEB 2023 - PRESENT
Consultant Data Scientist | Bangalore

- Undergoing training in **Python, SQL** and other consulting-related skills.

Skills: Python, VS Code, MS Office

Tweek Labs
MAR 2022 - AUG 2022
Data Analyst Intern | Bangalore [certificate]

- 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
SEP 2021 - OCT 2021
Summer Intern | Kolkata [certificate]

- **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 data-driven alternative of FB games to predict your potential future net worth in millions using **machine learning**.

- 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