Soumyadip Ghorai

CONTACT NO : +91 7872817970
EMAIL : ghorai.soumyadip33@gmail.com
Portfolio : https://soumyadipghorai.github.io

Education :

Christ University, Bangalore, 2023	MSc Data Science	3.52/4 (till date)
Indian Institute of Technology, Madras, 2024	BSc Data Science (distance)	7.6/10 (till date)
University of Calcutta, 2021	BSc Statistic	7.74/10

Experience:

Data Analyst Intern @Tweek Labs, Bangalore   MAR 2022 – AUG 2022
- Implemented methods like **moving avg, selective scaling** to remove fluctuations in sensor data to get accurate results.
- Setup a separate notebook of **interactive charts** to check for anomalies in various parameters of athletes using **plotly.**
- Implemented new features like **max hip rotation velocity**, **max shoulder speed** to satisfy customer demand.
- 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.
- **Setup pipeline** to store data in google sheet using **google API,** made interactive dashboard using **Meta Base** to track **KPIs.**
Skills : python, Meta Base, Google Sheets, Google API, PyCharm, VS Code, GitHub

Summer Intern @Ericsson, 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 that might reduce the overall processing time upto **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 :

Personal Finance : [github] [applink]   Feb 2022 – Feb 2022
- Collected data on a daily basis of my expenses in a **spreadsheet** and created an interactive **Streamlit** web app.
- 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, treemap and Q&A** to help make my financial decisions.
Tech : python, streamlit, plotly express, heroku, HTML, CSS, markdown, VS Code, git

Networth Predictor : [github]   NOV 2021 – DEC 2021
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** in 100+ categories from around the world using **Beautiful Soup**.
- Preprocessed the raw data, performed **EDA** to understand the data and done **feature engineering and feature scaling** .
- Applied **Random Forest Regressor** to predict net worth based on available features with **MAE** of **1.4.**
- Created frontend using **HTML, CSS and jinja code** and in the backend used **flask.**
Tech : Machine learning, python, Flask, jinja code, HTML, CSS, VS Code, git.

News Scraper web app : [github]   AUG 2021 – AUG 2021
- A **responsive** news app that shows news on various categories by scraping popular news sites in **real time**.
- Users can see the **thumbnail**, **title and a short description** and can be **redirected** to the actual article.
Tech : Flask, Beautiful Soup, HTML, CSS, VS Code

Skills :

Python   SQL   Tableau   Machine Learning   web scraping   HTML   CSS   VScode   PyCharm

Achievements :
- **3X Expert at Kaggle** with 5+ EDAs having 100+ upvotes
- **3star at codechef**, 200+ problems solved on **leetcode**.

Other Activities :
- Google Developers student club **Machine Learning Lead.**
- Represented college in multiple **state level athletic meets.**