

HIMANSHU

 waitasecant |  waitasecant |  waitasecant.github.io |  waitasecant@gmail.com

WORK EXPERIENCE

Roni Analytics LLP - Data Analyst Intern

May'24 - Aug'24

- **Implemented** prod-grade cron jobs to quantify granular & token level **TVL** for Pendle & Kelp protocols.
- **Migrated** the **on-chain data pipelines** to **Dune**, **validated** using **DelfiLlama** ensuring data integrity.
- **Implemented** multi-threaded **websocket** solution to track and process **Bybit** order and tradebooks data to compute granular volume metrics for bid-ask and buy-sell analysis for spot and futures trading.

PROJECTS

Real-Time Air Pollution Dashboard

[github dashboard](#)

- **Automated pipeline** to scrape, prepare, validate data for 24 stations in Delhi using **GitHub Actions**.
- **Feature engineered** AQI using **regression tree** with RMSE 3.79 trained on two year data corpus.
- **Developed** an **R-Shiny** dashboard hosting monthly **choropleth maps**, **real-time** time series plots.

Multi-Document Based RAG Chatbot

[github](#)

- **Implemented hierarchical tree-based** indexing to multiple documents for **efficient retrieval**.
- **Implemented** query expansion and retrieval methods combining **BM25** and **Sentence Transformer**.
- **Built** a question answering pipeline to generate accurate context-aware answers.

Development of BestModelSubset.jl

[github](#)

- **Developed** a Julia Package which implements **model selection algorithms** like best subset, forward step-wise selection using **AIC criterion** to select the subset of predictors most significant to response.
- **Optimized** the package for performance and usability, ensuring efficient computation.

SKILLS

Languages & Frameworks: Python, R, Julia, SQL; Flask, Shiny, Plotly, PySpark

Tools & Libraries: Git, Docker, Databricks; Pandas, Scikit-Learn, PyTorch, NLTK

HACKATHONS & OPEN SOURCE

- **AmEx Campus Challenge 2024 - Decision Track:** Ranked 4 in the Econ/Stats cohort in Round 1.
- **It's Data Time - IIIT NR Machine Learning Hackathon 2022:** Ranked 8 out of 236 participants.
- **F1alyze - IEEE MUJ Formula 1 Datathon 2024:** Ranked top 40% (bronze medal) on Kaggle.
- **modis-tools:** A Python library to download MODIS data from the NASA Earthdata. [Link to PR](#)

EDUCATION

Chennai Mathematical Institute

M.Sc. Data Science

2023-25

CGPA: 7.75

Institute of Chemical Technology

M.Sc. Engineering Mathematics

2021-23

CGPA: 8.46

Sri Venkateswara College, University of Delhi

B.Sc. (Hons.) Mathematics; Minor: Physics

2018-21

CGPA: 8.30