VARDHAN RAJ MODI

vrym2@student.le.ac.uk | 7824380067
Leicester LE39HH Contact - LinkedIn | GitHub

PERSONAL STATEMENT

Satellite data science graduate and an experienced python code developer passionate about utilising data analytics to understand environmental challenges, offering expertise from work and various projects which involved data analysis, preparation, transformation, and visualisation. Using exceptional interpersonal and relationship-building abilities to communicate complex analytical concepts.

WORK HISTORY

Machine Learning Intern | Open Climate Fix – London 11/2022 - 03/2023
OCF is a non-profit ML lab, utilising vast amounts of satellite data (NWP), UK power network data, etc. to forecast (Nowcasting) the solar PV generation on a huge scale.

- My duties include pre-process various client data (ex: UKPN) and perform operation in a production environment using CI/CD tools.
- Assisting in the streamlining various components of data configurations, to ingest the data for the ML model training.
- Working on the computations that are performed in the company's physical server's aptly named (Leonardo and Donatello)

Earth Observation Intern | OilX Ltd – London 06/2022 - 09/2022

Secured a SPIN (Space Placement with INdustry), with OilX Ltd., funded by the UK Space Agency for a project to monitor and effectively quantify the greenhouse gas emissions from oil infrastructure across Europe.

- Undertake research with literature review and for tangible data and potential methodologies to adapt.
- Develop pythonic modules (through version control) from the research of automatic download of satellite data (Sentinel-5P and meteorological wind data (ERA-5) data) and researching potential weather numerical models.
- Test and debug python modules provided by the supervisor.

EDUCATION

University of Leicester, Leicester Expected in 09/2023
Master of Science: Satellite Data Science

Project 1 – NLP Location Text Frequency analysis of Wikipedia articles (R-Studio)

- The project deals with the Geo-Text analysis of the Wikipedia articles. Analysing word frequency, and building a Nave Bayes and Logistic regression model to predict if a given Text would be in a Wikipedia article based on its location.

Project 2 – Urban Air Monitoring System and Health (UAMS-H) (R-Studio)

- Lead the development of an air quality monitoring prototype web application as the Project Manager, effectively assimilating Satellite data acquisition and visualisation through 'R' programming. Link to the prototype and the technical report.

Project 3 – Satellite Data Analysis on California Dixie Fire (2021) (Google-Colab)

- Automatic data acquisition (with Google Earth Engine and Copernicus Data Hub), pre-processing, analysis, and visualisation of pre- and post-Dixie fire have been conducted in Colab (python programming). For a brief look into the project, please open this link to the report and for code.

Project 4 – U.S. Diabetic patient data analysis and readmit prediction (Google Colab)

- Worked on U.S patient data to predict whether a patient going to be readmitted, utilizing K-Means clustering, Random Forest, Logistic Regression, and other ML algorithms. Link to the Colab notebook and the report.

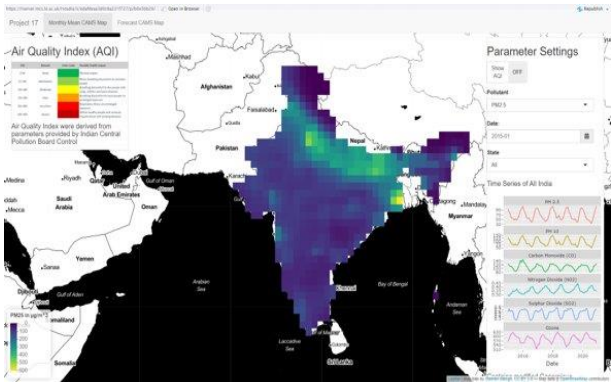
TECHNICAL SKILLS

- Modular Python Programming – Proficient in Packages (numpy, pandas, xarray, netcdf, scipy, matplotlib, torch)
- Structured Query Language (MySQL, PostgreSQL, PostGIS)
- Bash scripting (Linux, PowerShell)
- IDE's (VS Code, Google Colab)
- R programming (R Studio)
- Linux User (Ubuntu, Centos)
- GitHub version control

REFERENCES

- Jacob Bieker – Machine Learning Engineer, Open Climate Fix- jacob@openclimatefix.org
- Tansey, Kevin J. (Prof.) - Professor of Remote Sensing, University of Leicester-kjt7@leicester.ac.uk

Regarding a group project (UAMS-H), idea is to effectively utilise automatic extraction of satellite data with CDS (Climate Data Service) API of Copernicus which are available in NetCDF format and target specific gases (in this case) pollutants and plot the time series according to the user's required time range. It also has the feature of Air quality index which estimates the state of pollution. To check out web-application prototype- https://bit.ly/eo_p17



California Dixie fire was one of the largest wildfires in the U.S. history. As part of my Satellite Data Analysis in Python module, I was tasked with visualising the severity of the burned regions as well as extent of the wildfire throughout its course. My work with the project has been shared through this link and a visual map of burned severity can be seen here.