**JOURNAL - Karandeep Singh Mann**

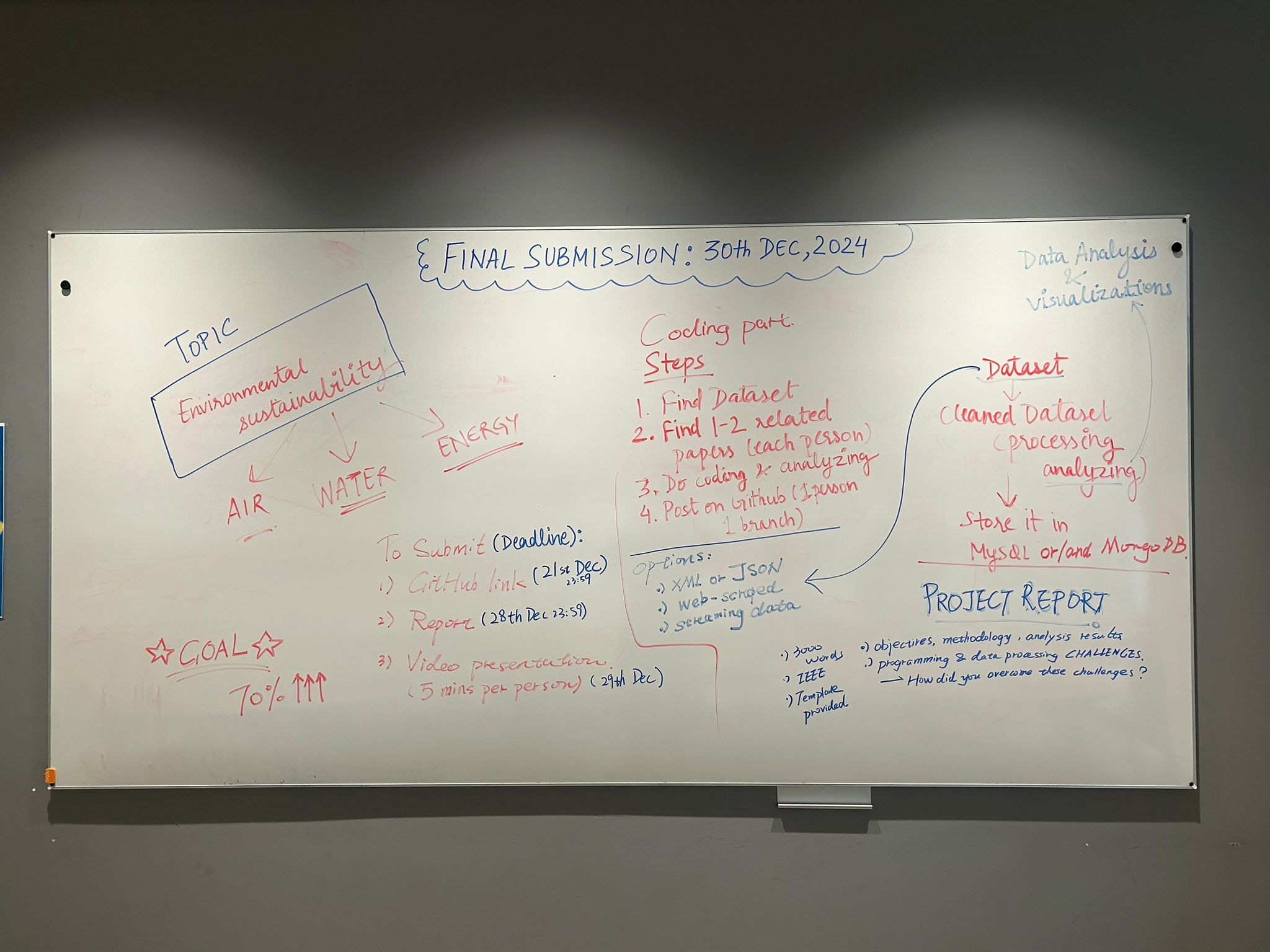
The following journal discusses my contribution for the work that we have done for successful completion of our Programming for AI project.

**10th Dec, 2024: Finalizing a topic for the project**

* First Group meeting held to discuss and finalize the topic of the project.
* Taking from “Data Governance and Ethics” module, we chose sustainability as our main topic
* Under Sustainability, we finalized the topic: **“Data Driven Sustainability: Tracking Air, Water and Energy for a Greener Future”** where each person will handle one aspect of sustainability.
* I chose to do the “Air” part
* Time Spent: 1-1.5 hrs

**11th Dec, 2024: Finalizing Work Distribution and Deadlines**

* Second Group meeting to discuss in detail about the tasks to be done for our project.



* Deadlines are also finalized for final submission of our project (30th December, 2024)

***UPDATE: Due to some unforeseen circumstances, the work got delayed and the deadline got shifted to 4th January, 2025.***

* Division of the report part is also discussed. It was decided that each person will be writing each component of report for their topic, i.e. Abstract, Introduction etc and at last everything will be merged. For final version, I chose to do merging of some portion of Related work, Methodology, Bibliography and final editing of the report
* Time Spent: 2 hrs
* Challenges: Conflict of interest to choose which part of the report will be done and deciding on deadlines because of other module’s work involved.

**13th Dec, 2024: Paper Research for finalizing approach to the Air Part**

* Explored various papers on google scholar to formulate the approach that I can use within my project.
* Found one paper by M.Rahman et. Al. that focussed on key pollutants and understanding the impact of them which forms the base of my study for this project.
* Time Spent: 1-2 hrs

**14th Dec, 2024: Relevant Dataset search**

* Research to find a dataset started. Various Sites were explored such as UCI Machine Learning Repository, Kaggle, OpenWeatherMAP, OpenAQ, EPA etc to find the right dataset.
* Suitable dataset is found on Open Meteo API which contains information about key pollutants in the air like PM2.5, Carbon Monoxide, Ozone etc.
* Time Spent: 3-4 hrs
* Challenges: Access restriction because of API keys, data not available in JSON/XML format, incomplete information and non relevant data and some APIs not returning correct results

**16th Dec, 2024 - 21st Dec 2024: Coding for the dataset**

* This period involved the coding part. It involved data processing and transformation on the dataset followed by analyzing the insights through visualizations and saving the dataset in MySQL. To make it more impactful, machine learning models are developed to predict PM2.5 levels in the air alongside a Flask & Streamlit based web application to show real-time monitoring of pollutants.
* Uploading code for my air section on github.
* Time Spent: 5-6 days (a few hours each day)
* Challenges:

i) Choosing the correct plots to find relevant insights within the data

ii) Carefully analyze if we could remove the outliers from the dataset.

iii) Handling various SQL syntax errors such as: “1054 (42S22): Unknown column 'NaN' in 'field list'” etc. while saving data within MySQL database.

iv) The graphs in the Streamlit application were changing rapidly hence a slider has to be implemented to adjust the timings.

v) “SQL\_PASSWORD” environment variable wasn’t working properly because my SQL password has a special character which needed to be encoded.

**23rd Dec, 2024 - 25th Dec 2024: Doing my part for final version of report**

* Combined each teammate’s contribution to their component for Methodology to create a final full fledged Methodology section containing all three components: Air, Water and Energy alongside writing some portion of the related work.
* Flow Charts created describing the work done within each component
* Time Spent: 3 days (a few hours each day)
* Challenges: To fully understand the work done by teammates so as to create a correct flow chart diagram and explaining the steps performed by each teammate.

**27th Dec, 2024: Preparing Presentation Slides for video demonstration**

* PPT Slides are created for the “air” section of the project for video demonstration of 5 minutes to be done.
* A 10 pages PPT slideshow is created briefly explaining all the work that I have done for the air section.
* Time Spent: 2 hrs

**30th Dec, 2024: Video Shoot and Report Compilation**

* Video was shot for the Air part and sent to the other team member for compilation
* Report Section was compiled into one final report document by me
* Time Spent: 30 mins for video shoot, 1-1.5 hours for report compilation on the given template in project description