# Omdena-Milan

Challenge Report

## Identifying potential areas for urban agriculture in Milan, Italy.

**Leads**

Maria Fisher

Mulugheta T. Solomon

Bernardino Sassoli

July - 2024

# Content

# Summary

The focus of this project was to improve urban agriculture through the utilisation of data from different sources such as satellite imagery, tabular data, and scientific papers to generate synthetic datasets. These datasets were used to analyse and develop strategies for enhancing food security in urban areas. Additionally, our efforts involve integrating AI-powered decision support systems that provide web-based tools to assist users with determining the best locations, crops, and managing pests. This strategy has the potential to assist communities in enhancing urban farming techniques and establishing green areas to combat the adverse effects of climate change.

# Introduction

Sustainable farming practices are being implemented to tackle issues related to food security and to adjust to climate change using sophisticated technological resources such as remote sensing and GIS. This is achieved through the use of advanced technological tools for predictive modelling to optimise crop yields and minimise environmental impact. Remote sensing and GIS technologies play a crucial role in this process, ultimately leading to the development of more resilient agricultural systems capable of adapting to changing climatic conditions and environmental issues. The project's main objective was to identify urban areas suitable for creating or enhancing green gardens to enhance food security in urban settings.

Various methodologies and tools were used in order to meet the project's goals. These included the utilisation of remote sensing technology and data science to evaluate variables such as land use, soil quality, crops, pests, and diseases. The purpose of this was to develop machine learning models for predicting site and crop suitability, as well as determining the most effective management approaches.

The project resulted in the development of data analytics dashboards for visualising and extracting insights on the potential of urban agriculture in Milan. Additionally, web applications were designed to enhance stakeholder engagement and encourage the adoption of sustainable agricultural practices in response to technological advancements and environmental issues.

# Objective

The primary objective of this study was to systematically identify areas in Milan that are optimal for urban agriculture by assessing various factors such as land use, soil health, crop suitability, pest and disease threats, and ecological constraints. Furthermore, the project sought to create web applications to enhance stakeholder engagement and encourage the adoption of sustainable farming practices in response to advancements in technology and growing environmental issues.

# Methodology

Participants were organised into three main teams: **Task1**, **Task2**, and **Task3**, for management purposes. Each team was assigned specific responsibilities such as data collection, data curation, data exploratory analysis, modelling, web application development and deployment, and reporting.

The teams were led by designated leads and co-leads, who were responsible for coordinating the participants within their respective tasks, organising activities, and assisting participants in reaching project objectives. Regular meetings were scheduled and conducted with active participation from the majority of participants, resulting in the successful completion of all proposed activities by the expected deadline.

By the end of the 8-week project, each group had developed a set of tools and resources that were integrated into a comprehensive web application for urban agriculture planning in Milan. These applications will provide users with valuable insights and recommendations for successful farming initiatives in urban spaces.

## Tasks main goals

| | |
|---|---|
| Task1 | • Develop models for the identification of optimal urban locations for community gardens through the analysis of soil and climate data, sunlight exposure, and local conditions using satellite imagery and tabular data. |
| Task2 | • To use machine learning algorithms for the prediction of optimal crop varieties by analysing past climate data and specific local factors.<br>• Create predictive models to forecast crop yields, enhancing strategic planning and efficient resource allocation. |
| Task3 | • Employing AI tools for the prediction of pests and pathogens to enhance agricultural management strategies and mitigate food losses. |

# Overview of weekly activities

**Week 1-2: Raw Data Collection and Preprocessing**

**Satellite Images Database**: High-resolution satellite imagery was sourced from the European Space Agency, NASA, and Google Earth Engine. These images were organised and preprocessed for analysis, enabling the extraction of crucial information such as land use classification, vegetation analysis, surface temperature estimation, topography, land use change detection, environmental monitoring, and infrastructure mapping.

**Tabular Data Compilation**: Comprehensive data on crops suitable for urban agriculture, including climate requirements, soil preferences, and growing seasons, was compiled. Additionally, information on pests and diseases relevant to these crops was gathered to support effective urban agriculture planning and management.

**Week 3-4: Data Preprocessing and Exploratory Data Analysis (EDA)**

**Data Preprocessing:** During the Data Preprocessing phase, techniques like handling missing values, removing duplicates, and standardising data were used to ensure dataset accuracy. This step was crucial for preparing the data for analysis and modelling. In the EDA phase, statistical tools and visualisation techniques were used to explore datasets and uncover patterns. Visualisations like scatter plots, histograms, and heatmaps helped the team understand the data and identify relationships between variables.

**Dashboard:** The EDA dashboard allowed for interactive exploration of datasets, enabling users to generate customised reports. This tool provided stakeholders with insights and recommendations for urban agriculture project locations. Investing in Data Preprocessing and EDA during weeks 3 and 4 laid the foundation for predictive modelling and optimization. Cleaning and analysing the data helped the team extract valuable information for strategic planning and implementation of urban agriculture initiatives.

**Week 5-6: Model Development and Validation**

Tasks 1, 2, and 3 all involve the development and validation of predictive models. The models were created to evaluate the suitability of urban agriculture in Milan, forecast crop yields, and address pest and disease outbreaks.

These models took into account a range of environmental and agricultural factors to deliver precise predictions.

2. The predictive models underwent a validation process to confirm their accuracy and dependability. Following validation and feedback from colleagues in their respective teams, adjustments were made to enhance the models' performance and practicality.

3. A regression model was formulated to predict crop yields, while a classification model was constructed to distinguish between different crop types. The Crop Disease Classification models were developed through a series of steps, including data splitting, data

augmentation, selection of model architectures, training with appropriate techniques, and evaluation of performance using various metrics.

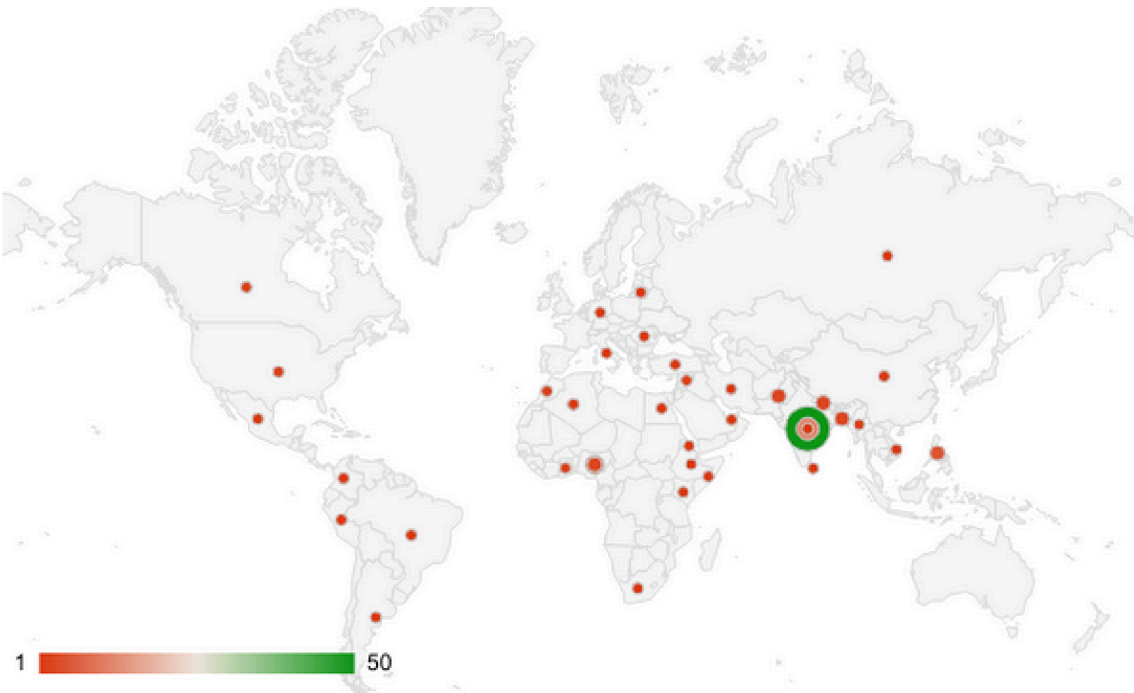
**Week 7-8: Web Application Development and Final Reporting**

**Web Applications**: The web applications were designed with user-friendly interfaces to make it easy for stakeholders to interact with the data and explore different scenarios. The visualisations provide a clear representation of the potential urban agriculture sites, allowing users to easily identify key areas for further investigation.

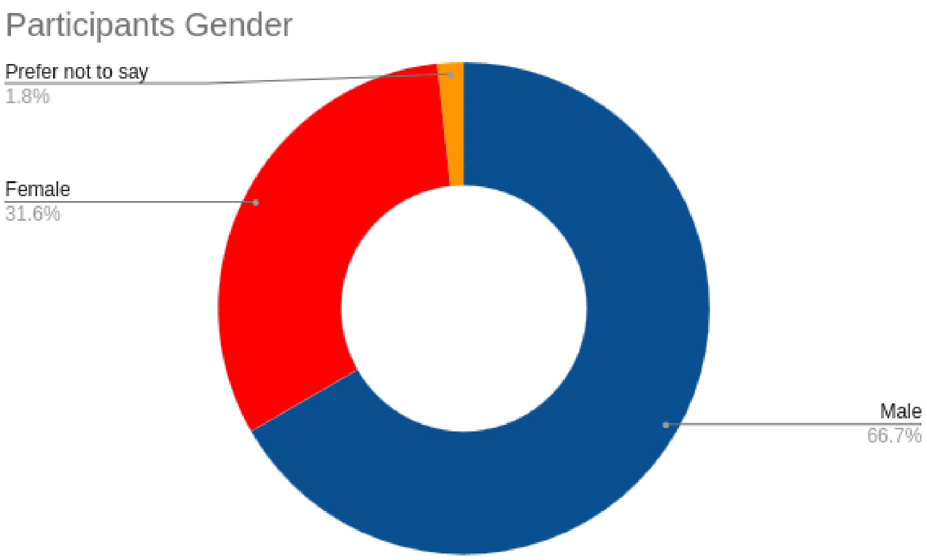
**Final Reports**: The final reports have been created to document every project activity, methodology, and result. A comprehensive overview of all tasks completed, such as data collection, preprocessing, analysis, and model development, can be found in the project repository on Google Drive and GitHub, guaranteeing detailed documentation and transparency.

# Participants

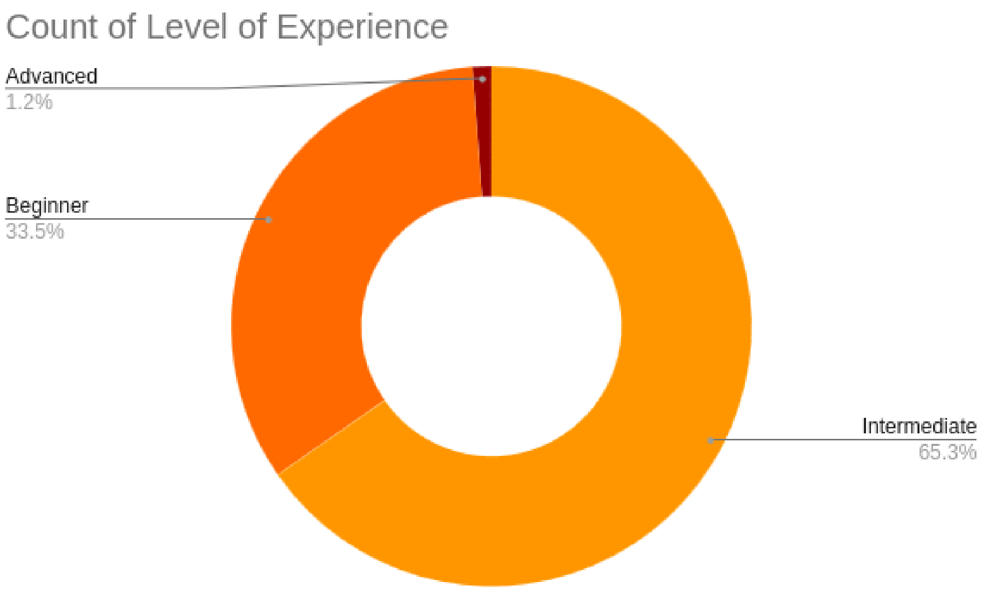
Challenge participants distribution



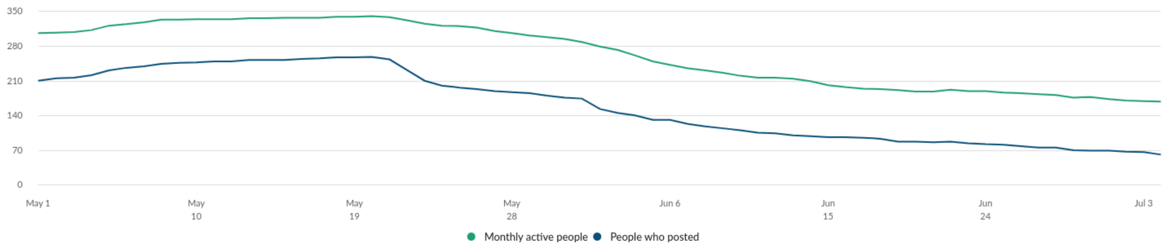
# Gender




# Level of expertise



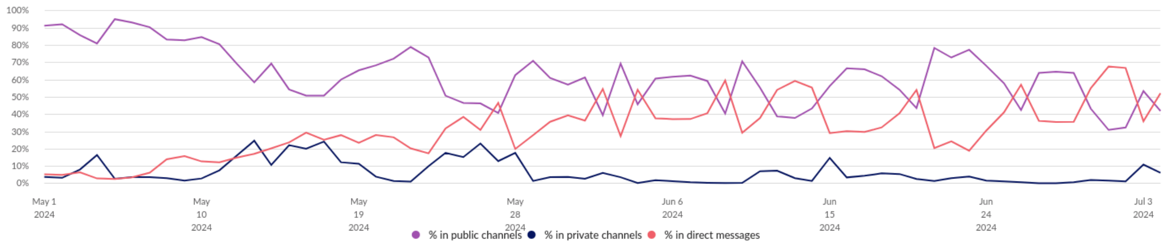
# Participation and engagement

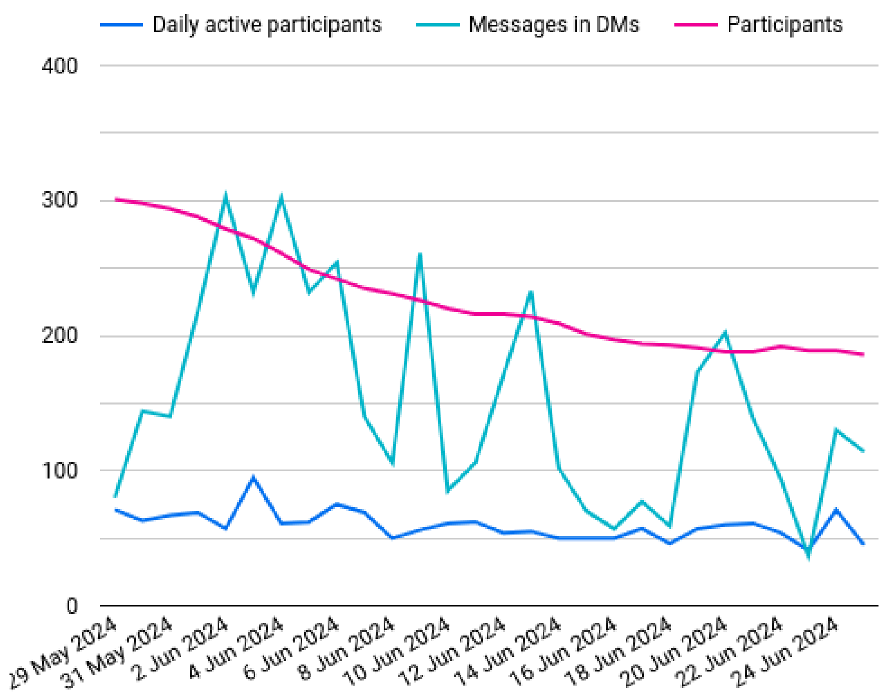


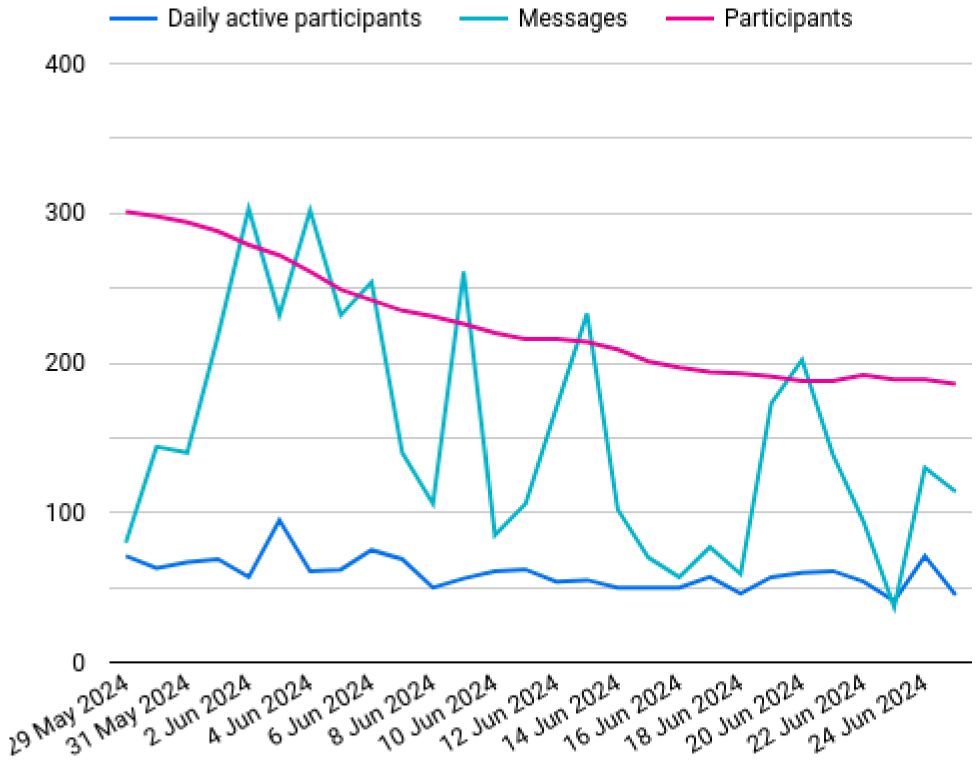

**69%**
In public channels

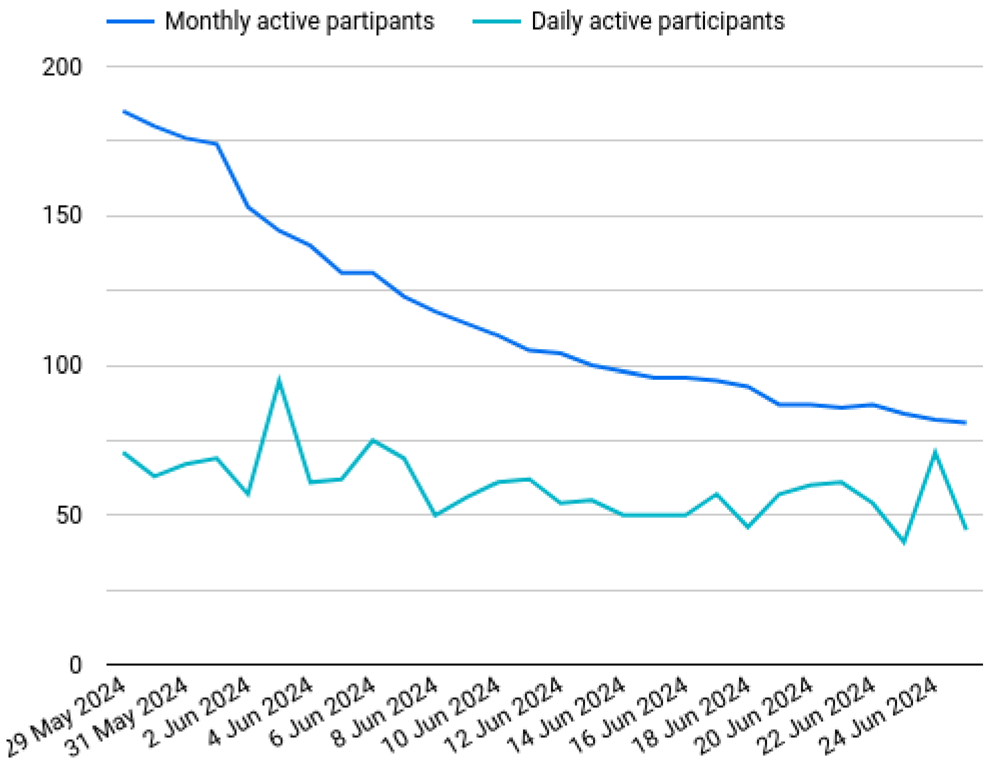
**7%**
In private channels

**24%**
In direct messages




Daily active participants
Messages
Participants
400
300
200
100
0
29 May 2024
31 May 2024
2 Jun 2024
4 Jun 2024
6 Jun 2024
8 Jun 2024
10 Jun 2024
12 Jun 2024
14 Jun 2024
16 Jun 2024
18 Jun 2024
20 Jun 2024
22 Jun 2024
24 Jun 2024


Monthly active partipants
Daily active participants
200
150
100
50
0
29 May 2024
31 May 2024
2 Jun 2024
4 Jun 2024
6 Jun 2024
8 Jun 2024
10 Jun 2024
12 Jun 2024
14 Jun 2024
16 Jun 2024
18 Jun 2024
20 Jun 2024
22 Jun 2024
24 Jun 2024

# Results

The project resulted in the creation of datasets, such as satellite imagery, detailed soil type information, climate data, and data on pests and pathogens. These datasets were used to create and improve predictive modelling, allowing for more accurate identification of suitable areas for urban farming through GIS mapping tools and applied to web applications. These findings play a crucial role in directing future strategic actions towards sustainable urban agriculture initiatives, promoting resilient agricultural practices, and supporting global climate adaptation and environmental sustainability efforts.

# Task1 - Selecting Areas for Urban Farming

# 1. Data Collection

**Task Leaders:** Hilda Posada and Ashwitha Kassetty
**Task Team Leads:** Pujan Thapa, Michael Darko Ahwireng, Madhav Agrawal

---

### Summary

Task 1 focused on the collection of satellite, climate, infrastructure, and social data for selecting areas for urban farming in Milan, Italy. The task was structured by first identifying the geographical zones of Milan, with a focus on Zones 4 (Porta Garibaldi) and 9 (Porta Vittoria). Teams were formed based on the number of active members and their expertise, with subgroups assigned specific data types. Members with intermediate experience in satellite data provided guidance and leadership to these subgroups.

The data collection was successfully completed over two weeks for both Zone 4 and Zone 9.

---

### Key Accomplishments

1. **Data Collection:** Successfully collected satellite, climate, infrastructure, and social data for Zones 4 and 9.
2. **Educational Resources:** Developed video tutorials on Python search/download from Google Earth Engine (GEE), climate, and geospatial data collection.
3. **Meetings:** Regular meetings were held, ensuring active participation and engagement from all team members.
4. **Motivated Teams:** Maintained motivated and engaged teams throughout the project.

## Challenges Faced

1. **Data Storage:** Managing vast amounts of data led to the decision to use Google Drive for storage.
2. **Data Collection:** Difficulty in identifying sources for CO2, social, and climate data was resolved with guidance from Maria Fisher.
3. **Member Participation:** Efforts were made to engage non-active members, which included sending multiple update messages.
4. **Time Zones:** Different time zones posed a challenge; resolved by adhering to UTC and using Slack and Huddle for communication.
5. **Beginners' Involvement:** Difficulty in involving beginners in geospatial tasks was mitigated by creating detailed tutorials.
6. **Participant Identification:** Inconsistencies in participant names across lists and Slack profiles were addressed by thorough searches in multiple channels.

---

## Tasks Completed

**Meetings, Deadlines, Reminders, Structure, Planning, Strategy, Timeline:**

- **Leads and Team Leads:** Ashwitha Kassetty, Hilda Posada, Michael Darko Ahwireng, Pujan Thapa, Madhav Agrawal, Vivek Vardhan, Rajesh KR

**Satellite Data Collection:**

- **Team Leads:** Madhav Agrawal, Michael Darko Ahwireng, Pujan Thapa
- **Team Members:** Denis Mwangii, Venura Pussella, Pratiksha Naik, Birtukan Birawo, Jasmin, Nitesh Kumar Kesharwani, Nishthaa, Deepa R, Shovan Raj Shrestha

**Infrastructure Data Collection:**

- **Team Lead:** Pujan Thapa
- **Team Members:** Aditya Trivedi, Bipan Neupane, Harsh Khurana, Yasmine Medjadba, Venura Pussella, Jasmin, Kamia Salango, Jesi Martin Maglana

**Climate Data Collection:**

- **Team Members:** Nhat Minh Nguyen, Havish B, Opeyemi Ebunoluwa, Saisree Gownipalli Ramanaiah, Nitesh Kesharwani, Venura Pussella

**Social Data Collection:**

- **Team Members:** Leila Javanmardi, Rupak R, Michael Darko Ahwireng

**Report Compilation:**

- **Leads and Team Leads:** Hilda Posada, Ashwitha Kassetty, Pujan Thapa, Michael Darko Ahwireng, Saisree Gownipalli Ramanaiah, Madhav Agrawal

---

## Technical Updates

- **Satellite Data:** Used OSMnx library to extract and analyse data from OpenStreetMap for Zone 9. Data types included 'amenity' and 'land use' tags. Data processed and visualised, and saved as .tiff files.
- **Social Data:** Collected from OSM and saved as CSV files.
- **Infrastructure Data:** Collected from OSM and saved in .tiff and shapefile formats.
- **Climate Data:** Downloaded from dataverse.harvard.edu, covering January 1, 2023, to May 14, 2024, saved as CSV files.

---

## Communication and Collaboration

The team used Slack and Huddles for communication and collaboration. Regular Zoom meetings were held for updates and addressing issues. Decisions were made collaboratively, with highly skilled members sharing knowledge and beginners reaching out for help. Team engagement was high, and new members who joined later were directed to appropriate resources and tutorials.

---

## Team Feedback

- **Positive Feedback:** Members appreciated the team structure, meetings, and clarity of the project.
- **Tutorials:** Helped increase participation from beginners.
- **Collaboration:** Inter-team collaboration eased task completion and reduced difficulties.

---

## Next Steps

On Monday (05/20/24), the collected data will be uploaded to Google Drive and passed to the EDA team for pre-processing and analysis.

---

## Additional Notes

Organising the data collection task and setting up logistics took time, causing a one-week delay. Future efforts will focus on reducing setup time and making quicker decisions to repurpose time for actual data collection.

---

### Key Metrics

- **Satellite Imagery Resolution:** 1-10 metres
- **Temporal Resolution for Satellite Data:** Monthly
- **Climate Data Collection Range:** January 2023 - December 2023

---


### Acknowledgments

This project's success is attributed to the dedication of all team members and leads. Special thanks to:

- **Zone 9 Members:** Maria Fisher, Ashwitha Kassetty, Hilda Posada, Maz, Michael Darko Ahwireng, Pujan Thapa, Nitesh Kersharwani
- **Zone 4 Members:** Harsh Khurrana, Birtukan Birawo, Venura Pussella, Kamia Salango, Nishthaa, Pratiksha Naik, Nitesh Kesharwani


Attachments with collected data and tutorials can be found in the respective folders in Google Drive.

# 2. EDA

**Leads/Co-Leads:** Hilda Posada, Nitesh Kumar Kesharwani, Ashwitha Kassetty
**Team Leads:** Deepa Ravindran, Daikukai Bindah, Madhav Agrawal, Birkutan Birawo, Muddassir Mujawar, Pujan Thapa, Vidhi Gupta, Opeyemi E. Abodunrin, Harsh Khurana, Vivek Vardhan, Prachi Kumari

### Summary

Exploratory data analysis (EDA) focused on data extracted from satellite images, which included 12 variables such as the Soil Adjusted Vegetation Index (SAVI), Land Surface Roughness, and Land Slope. This data, formatted in CSV, contained null values that necessitated cleaning. The resolution of images, collected at varying resolutions, was standardised during the data cleaning process.

### Key Accomplishments

1. **Data Collection:**
    - Gathered additional features not included in the initial phase.
    - Collected data suitable for urban agriculture from 2014 for model training.
2. **Data Cleaning:**

    - Cleaned CSV data for zones 4 and 9, addressing null values and ensuring correct data formats.
    - Standardised image resolution to 30m and preserved outliers.
3. **Exploratory Data Analysis (EDA):**
    - Conducted EDA on features such as Land Slope, NDWI, NDVI, and Sunlight Exposure for zones 4 and 9.
    - Determined threshold values for each feature and developed plots for areas exceeding these thresholds.
    - Created descriptive statistics and graphical representations for better data understanding.
4. **Tutorials Creation:**
    - Developed tutorials on data loading, preprocessing, visualisation, and basic EDA to assist team members.

## Challenges Faced

- **Team Structure and Planning:** Ensuring effective communication and task planning across different time zones.
- **Data Management:** Integrating and managing image data collected at varying resolutions.
- **Team Participation:** Re-engaging inactive team members through regular updates and progress communication.
- **Tracking Contributions:** Difficulty in tracking contributions, especially from members not active on Slack.

## Tasks Completed

- **Additional Data Collection:**
    - Maz, Vidhi Gupta, Birtukan Birawo, Pratiksha Naik, Opeyemi Ebunoluwa Abodunrin, Harsh Khurana, Nitesh Kumar Kesharwani.
- **Data Cleaning and EDA:**
    - Varun Garg, Birtukan Birawo, Deepa Ravindran, Soham Das, Maz, Vidhi Gupta, Harsh Khurana, Pujan Thapa, Bipan Neupane, Opeyemi Ebunoluwa Abodunrin, Pratiksha Naik, Christine Elizabeth Escano, Jumacaq, Yasmine Medjabda.

## Upcoming Tasks

- **Dashboard Creation:** Using Google Looker Studio.
- **Model Development:** Using the cleaned and pre-processed data.

## Technical Updates

- **Missing Data Collection:**
    - Collected previously unrecorded features or those at different resolutions.
- **Data Cleaning:**
    - Filled null values using data from nearby regions and ensured correct data formats.

- **Exploratory Data Analysis (EDA):**
    - Developed descriptive statistics and graphical representations to enhance data comprehension.

## Communication and Collaboration

The team used various communication channels to discuss tasks and engagement. Meetings and strategies were primarily decided by the co/leads with input from team leads and members. Tasks were allocated through polls and participation sign-ups on Google Sheets. Slack was used as the primary communication tool. Inactive members were re-engaged by reaching out directly and providing necessary resources.

## Team Feedback

Meetings between task leads and team leads helped clarify expectations and align plans with project goals. Regular messages in the channel kept the teams updated and engaged. Tutorials and recordings were beneficial for participants at all levels.

## Next Steps

- **Data Upload:** Cleaned and pre-processed data will be uploaded to Google Drive starting Monday (06/03/2024).
- **Dashboard Creation:** Will commence using Looker Studio on Monday (06/03).

## Additional Notes

Ambiguities in task objectives delayed data collection and organisational challenges further contributed to the week-long delay. Future efforts will focus on streamlining task setups and decision-making processes to prevent similar delays.

## Key Metrics

- **Spatial Resolution:** 30m
- **Data Format:** Tabular
- **Data Structure:** Longitude, Latitude, Zone, Feature value

## Attachments

- **Pre-processed data**
- **Cleaned data**
- **Visualisation**
- **Link to the drive structure:**
    - Sunlight Exposure (zone 9)
    - Sunlight Exposure (zone 4)
    - Built-up area (zone 4)
    - Built-up area (zone 9)
    - Zone 4 & 9 Built-up Area
    - Sunlight Exposure for Zone 4 and 9

- **EDA Links:**
    - NDVI-EDA-zone4&9.ipynb (Zone 4 and 9 Built up area)
    - Sunlight_exposure_zone4_eda.ipynb (Zone 4)
    - EDA_built_up_area_Zone4.ipynb (Zone 4 Build-up Area)
    - EDA_built_up_area_Zone9.ipynb (Zone 9 Build-up Area)
    - Task1_EDA.ipynb (built-up area and climate, zone 4 and 9)
    - EDA python (Visualisation: sentinel 2 data)
    - Milan_Urban_Farming_eda.ipynb
    - CSV_Zone4_BUILD_UP&BARE.zip
    - EDA_Zone_4_Slope_&_Roughness.zip (slope, roughness, zone 4)
- **Tutorial Link:**
    - Data Analysis tutorial for beginners.mov
- **Individual Reporting:**
    - Contributions (Vidhi Gupta)
    - EDA Report.docx (Harsh Khurana)
    - Summarised Report on Sunlight Exposure in Zone 9 and Milan (Bipan Neupane)

### Acknowledgments

The project's success is attributed to the dedication and hard work of both team members and leads. Their combined efforts were crucial in overcoming challenges and achieving milestones. Special thanks to Maria Fisher, Ashwitha Kassetty, Nitesh Kesharwani, and Hilda Posada for their exemplary leadership and guidance throughout the project.

# 3. Modelling

**Leads/Co-Leads:** Nitesh Kumar, Hilda Posada, Muhammad Mazhar, Harsh Khurana, Opeyemi Ebunoluwa Abodunrin

**Team Leads:** Vidhi Gupta, Soham Das, Micheal Darko Ahwireng, Sneha Mahata, Birtukan Birawo, Varun Garg.

## Summary:

In the modelling phase, supervised and unsupervised learning models were developed and evaluated using datasets from Zone-4 and Zone-9 (year 2014). The supervised learning models included Random Forest Classifier, Decision Trees, BiLSTM, XGB Classifier, TabNet Classifier, and Logistic Regression, aimed at identifying the best model for the task. In parallel, several unsupervised learning models were explored, including KMeans clustering, Principal Component Analysis (PCA), Self-Organizing Maps (SOM), Graph Neural Networks (GNN), and Density-Based Spatial Clustering of Applications with Noise (DBSCAN), to analyse and extract meaningful patterns from the data. Two models were selected, one from supervised and one from unsupervised learning, for deployment.

## Key Accomplishments:

- Developed and trained a total of 15 supervised and 10 unsupervised learning models.
- Used Zone-4 and Zone-9 data (years 2014 and 2023) for modelling.
- Kept track of various performance metrics on supervised (MLflow) and unsupervised models.
- Provided tutorials on techniques of supervised and unsupervised modelling.
- Model validation.
- Finalised two models, one from each supervised and unsupervised, for deployment.

## Challenges Faced:

- Different time zones and other commitments made it challenging to gather all team members together.
- Initially, it was difficult to get beginners to take up tasks until tutorials and guides were made available.
- Inactive team members were encouraged to participate through regular updates in the channel, which helped re-engage them and boost their interest in the project.
- The mismatch between names on the participation list and Slack profiles made it difficult to identify team members.
- Tracking the work and progress of all team members was challenging.
- The contributions list was based solely on Slack messages and discussions; if someone was not discussing on Slack and not sharing their work, their names might be missing.
- The project faced delays of 3 days due to the challenges mentioned above in the first week.

## Tasks Completed:

- Unsupervised Modelling: Opeyemi Ebunoluwa Abodunrin, Sneha Mahata, Birtukan Birawo, Varun Garg
- Model Evaluation: Opeyemi Ebunoluwa Abodunrin, Sneha Mahata, Birtukan Birawo
- Visualisation: Opeyemi Ebunoluwa Abodunrin, Sneha Mahata, Birtukan Birawo, Varun Garg
- Tutorials: Varun Garg
- Reporting: Sneha Mahata

## Upcoming Tasks:

- Web Application Development & Deployment for Task 1

## Technical Updates:

### Supervised Learning:

Various models such as Random Forest Classifier, Decision Trees, BiLSTM, XGB Classifier, TabNet Classifier, MLP Classifier, Sequential ANN, and Logistic Regression have been developed and evaluated using accuracy, precision, recall, ROC, and AUC criteria. Performance metrics are tracked for these models to facilitate the selection of the most suitable one.

### Unsupervised Learning:

Various models including KMeans clustering, Principal Component Analysis (PCA), Self-Organizing Maps (SOM), Graph Neural Networks (GNN), and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) have been developed and evaluated using Inertia, CHI (Calinski-Harabasz Index) Score, and DBI (Davies-Bouldin Index) Score. Performance metrics are also tracked for these models to aid in choosing the most appropriate one.

## Communication and Collaboration:

The team collaborated primarily using Slack messages and channels to discuss tasks, project engagement, and strategies. Co/leads, with input from the leads and members, decided on strategies. Experience-based tasks were determined through polls and participation sign-ups on Google Sheets, ensuring a fair distribution. Tasks were allocated accordingly. Communication was fostered through consistent Slack messages and updates, ensuring ample exchange of information for discussing issues despite the minimized calls and huddles for this phase.

## Team Feedback:

Inter-team collaboration among team leads helped reduce the difficulty in completing tasks, making the tasks for this project a shared responsibility among the teams. Team leads' timely messages in channels for tasks also kept the teams updated and engaged with each other's progress.

## Next Steps:

- Development phase

## Key Metrics:

### Supervised Learning:

- Model/Algorithm
- Hyperparameters & Their Values
- Accuracy
- Precision
- Recall
- AUC

- Model Performance with CV

### Unsupervised Learning:

- Model/Algorithm
- Hyperparameters & Their Values
- Inertia
- CHI Score
- DBI Score

## Acknowledgments:

I extend my sincere gratitude to the entire team for their hard work and dedication during the ML modelling phase. Special thanks to Deepa R for collecting 2014 data quickly; Yasmine Medjadba, Soham Das, Harsh Khurana, Shivang Singh, and Vidhi Gupta for their expertise and guidance in supervised models; and Muhammad Mazhar, Sneha Mahata, Birtukan Birawo, and Opeyemi Abodunrin for their crucial contributions to unsupervised models. Your collaborative efforts and commitment have been instrumental in our success. Thank you all for your outstanding contributions.

# 4 . Web App

**Leads/Co-leads:** Nitesh Kumar, Hilda Posada, Harsh Khurana, Shivang Singh

**Team Leads:** Sneha Mahata, Vidhi Gupta, Varun Garg, Pujan Thapa

## Summary:

Developing the web application involved three main components: backend development, frontend development, and deployment. For the frontend, contributors focused on creating a user-friendly interface for input collection, with model outputs displayed on a Folium map. The layout was designed with aesthetically pleasing elements and dashboards were integrated for clear information presentation. Frontend developers collaborated closely with frontend and backend leads to refine designs and integrate data display using Python scripts.

On the backend, contributors established robust server-side logic using Python, ensuring strong API development and session management. They successfully integrated the XGBClassifier and K-means Clustering models to provide predictive functionalities. Efficient data handling processes were implemented, including data ingestion, cleaning, transformation, and database management, to ensure optimised backend operations.

The application was deployed on Streamlit Cloud, making it publicly accessible via a web URL.

## Key Accomplishments:

- Developed a user-friendly interface for input collection and displayed model outputs.
- Integrated dashboards perfectly on the webpage for clear and accessible information.
- Applied suitable colour combinations to enhance user experience and readability.
- Integrated XGBClassifier and K-means Clustering models into the backend.
- Developed Python scripts for backend logic and prediction display.
- Merged 2014 and 2023 datasets across zones, enhancing data access and EDA.
- Integrated Looker Studio charts into the EDA page and refactored code in the prediction file.
- Held frequent meetings with the web app development team.
- Deployed the application on Streamlit Cloud.

# Challenges Faced:

- Coordinating across different time zones and managing other commitments made synchronous collaboration challenging.
- Initial difficulty in task assignment to beginners was resolved through tutorials and guides.
- Encouraging participation from inactive team members required consistent updates and project progress sharing.
- The contributions list relied on Slack messages and discussions, potentially omitting contributions from less active members.

# Tasks Completed:

## Front-End Development:

- UI/UX design: Amritangshu Daskar
- Web Page design: Sneha Mahata, Deepa Ravindran
- Dashboard Integration: Pujan Thapa
- Looker Dashboards Development: Anidipta Pal, Ashwitha Kassetty, Harsh Khurana, Deepa Ravindran

## Back-End Development:

- ML Models Integration: Vidhi Gupta, Sneha Mahata
- Backend Logic: Vidhi Gupta, Sneha Mahata
- EDA Improvisation: Pujan Thapa, Vidhi Gupta

## Web App Deployment:

- Deployment: Nitesh Kesharwani

## Reporting:

- Back-End Development: Shivang Singh
- Front-End Development: Harsh Khurana

- Final Report: Ashwitha Kassetty

# Technical Updates:

## Back-End Development:

For the backend, supervised (XGBClassifier) and unsupervised (K-means Clustering) machine learning models were successfully integrated. Python scripts were developed to manage backend logic, enabling efficient data input and prediction display. Datasets from 2014 and 2023 were merged across different zones, improving data accessibility and EDA. Looker Studio charts were incorporated into the EDA page, and code in the prediction file was refactored for enhanced performance and maintainability.

## Front-End Development:

The frontend team focused on creating a user-friendly interface for input parameter collection and directly displaying prediction results on the webpage. Modern fonts and visually appealing backgrounds were used, with appropriate colour schemes to improve readability and user experience. Model outputs were visualised on a Folium map, and dashboards were integrated seamlessly into the webpage layout. The Contributor's page was also developed to display relevant data using Python scripts. Regular meetings ensured frontend designs aligned well with backend functionalities.

# Communication and Collaboration:

Communication primarily took place via Slack messages and channels. Although minimal calls and huddles were held, effective communication channels facilitated collaboration and issue resolution. Regular updates and discussions kept the team informed and engaged with project progress.

# Team Feedback:

Collaboration among team leads significantly eased task completion challenges by fostering shared responsibility. Timely updates from team leads in Slack channels ensured teams remained informed and involved in each other's progress.

# Attachments:

- Github Link
- Personal Repo


# Acknowledgments:

- Harsh Khurana (Co-Lead)

- Shivang Singh (Co-Lead)
- Vidhi Gupta (Team-Lead)
- Varun Garg (Team-Lead)
- Sneha Mahata
- Pujan Thapa
- Rupak R
- Arpit Sengar
- Yasmine Medjadba (Jasmin)
- Ashwitha Kassetty
- Anidipta Pal
- Amritangshu Daskar
- Sulagna Parida
- Deepa Ravindran
- Birtukan Birawo
- Nitesh Kesharwani

# Task2-Selecting suitable crops for urban farming

# 1. Data collection

**Task Leaders:** Swathy Ramakrishnan, Sai Deepak Namburu, Chukwudi Emmanuel Asibe

## Summary:

Task 2 involved the data collection of vegetables, fruits, and salads that grow in Milan, Italy. The team's initial focus was on identifying the relevant crops, dividing into subgroups to efficiently gather data. Subgroups were assigned to collect information on 10 types each of vegetables, fruits, and salads, leading to the creation of a synthetic dataset based on essential crop features. The data collection phase was successfully completed within two weeks.

## Key Accomplishments:

- Cohesively collected data based on identified features and created a synthetic dataset.
- Scheduled and attended meetings effectively, ensuring alignment between leads and team members.

## Challenges Faced:

- Assembling the entire team simultaneously due to varying time zones necessitated adherence to UTC time for general project meetings. Communication primarily relied on Slack and Huddle platforms.
- Initial ambiguity regarding task objectives led to delays and confusion in project execution.
- Effective communication across diverse team members posed challenges in ensuring universal understanding of project updates and instructions.
- Limited access to essential resources like specialised software hindered efficient task execution.

## Tasks Completed:

- Collection of vegetable dataset and creation of synthetic dataset.
- Collection of fruits dataset and creation of synthetic dataset.
- Collection of salads dataset and creation of synthetic dataset.

## Upcoming Tasks:

- All data collection tasks have been completed.
- The collected data will be passed to the EDA team for further analysis.

## Technical Updates:

- **Vegetable Data:** Collected primarily from FAO EcoCrop, EPPO Global Database.
- **Fruits Data:** Collected from FAO EcoCrop, ISTAT.
- **Salads Data:** Collected from FAO EcoCrop, Lettuce Farm SLAM Dataset.

## Communication and Collaboration:

The team used various communication channels like Slack, Huddles, and Google Meet to collaborate effectively. Discussions and decisions regarding meetings, strategies, and deadlines were inclusive, involving leads and team members through polls or consensus. Tasks were allocated based on skill levels, ensuring a balanced distribution of responsibilities. Slack Huddles facilitated direct communication among team leaders, who then disseminated information to all members.

The proactive engagement of initially inactive members, guided by team leads, significantly contributed to the successful completion of data collection tasks. Their enthusiasm bolstered data gathering efforts from multiple sources.

## Team Feedback:

- The team demonstrated strong collaborative skills, effectively communicating and working together towards achieving project objectives.

- Knowledge sharing among highly skilled members enhanced overall team development and learning.
- Despite initial challenges, the team displayed resilience in problem-solving, maintaining project momentum.
- New and returning members exhibited high engagement levels, actively participating in tasks and contributing to the data collection process.
- Streamlining decision-making processes is recognized as essential to minimise delays and enhance project efficiency in future tasks.

## Next Steps:

- Data collection files were uploaded to Google Drive on Monday (05/20/24) and subsequently transferred to the EDA team for preprocessing and analysis.

## Additional Notes:

- Ambiguity in task objectives initially delayed organisation and logistics setup, causing a one-week delay. Future tasks aim to streamline setup processes and decision-making for improved efficiency.

## Key Metrics:

- Metrics considered for data collection:
    - Vegetable
    - Salads
    - Fruits

## Acknowledgments:

This project's success is attributed to the dedication and effort of all team members and leads. The first phase of the project has been a valuable learning experience for everyone involved.

- **Vegetable Dataset:** Sai Deepak Namburu, Swathy Ramakrishnan, Nirupamaa Gurusubramanian
- **Salads Dataset:** Ashik C Babu, Sai Deepak Namburu, Swathy Ramakrishnan, Nirupamaa Gurusubramanian
- **Fruits Dataset:** Sai Deepak Namburu, Swathy Ramakrishnan, Nirupamaa Gurusubramanian

# 2. EDA

**Task Co-leaders:** Swathy Ramakrishnan, Chukwudi Emmanuel Asibe, Sai Deepak Namburu

## Summary:

An initial step in the project involved conducting an Exploratory Data Analysis (EDA) for crop selection and yield prediction. To begin, a synthetic vegetable dataset was formed by using both normal and uniform data distributions, with the normal distribution being deemed more appropriate. Following this, synthetic data for fruits and salads were also produced. The datasets, each containing 100, 300, 500, and 700 data points for every crop, were merged into a single dataset for a thorough EDA.

## Key Accomplishments:

1. Created synthetic vegetable dataset and conducted initial EDA to ensure data quality.
2. Expanded synthetic dataset to include all crop types with varying data point counts (100, 300, 500, and 700).
3. Combined datasets from all crops and performed extensive EDA to derive insights.

## Challenges Faced:

Throughout the project, the team encountered and successfully managed several challenges:

- Initial confusion over choosing between normal and uniform data distributions.
- Coordination across varying time zones necessitated adopting UTC as the standard for meetings and relied heavily on asynchronous communication via Slack and Huddle.
- Addressing team member engagement through consistent updates and project progress sharing in channels.

## Tasks Completed:

- Additional Data Collection: Sai Deepak Namburu, Swathy Ramakrishnan
- Data Cleaning and EDA: Kamia Salangho, Nirupama Gurusubramanian, Nishta, Ashwitha Kassetty, Swathy Ramakrishnan, Chukwudi Emmanuel Asibe

## Upcoming Tasks:

The preprocessed and cleaned data will be handed over to the Modelling team for further analysis and model development.

## Technical Updates:

- **Feature Selection:** Identified and collected min-max values of selected features, using these to generate synthetic data.
- **Data Cleaning:** Ensured numerical and categorical data formats were correct.
- **Exploratory Data Analysis (EDA):** Conducted univariate and bivariate analysis, developed descriptive statistics, and created graphical representations to enhance data comprehension.

## Communication and Collaboration:

Effective collaboration was facilitated through various channels such as Slack, Google Meet, and Huddle. Decisions were made collaboratively, leveraging polls and consensus among team members. Tasks were assigned based on expertise levels, encouraging knowledge sharing among team members.

## Team Feedback:

The team demonstrated strong collaborative skills, effectively communicating and coordinating efforts to achieve project goals. Experienced members played a crucial role in mentoring and guiding less experienced members, fostering a supportive environment.

## Next Steps:

Starting Monday (03/06/2024), the cleaned and pre-processed data will be uploaded to Google Drive and transferred to the Modelling team for subsequent analysis and model development.

## Additional Notes:

The setup of synthetic data creation and logistical arrangements took longer than anticipated due to initial task objective ambiguities, resulting in a one-week delay. Moving forward, efforts will focus on streamlining task setups and decision-making processes to optimise efficiency and time management.

## Key Metrics:

- **Name:** Name of the Crop
- **Soil Type:** Type of Soil (e.g., Sandy Loam)
- **Fertility:** Soil fertility level for optimal growth
- **Photoperiod:** Length of Day Period
- **N-P-K Ratio:** Optimal ratio of Nitrogen, Phosphorus, and Potassium
- **Temperature (°C):** Optimal temperature range for growth
- **Rainfall:** Amount of rainfall in millimetres

- **pH:** Optimal pH range
- **Light Hours:** Number of Light Hours needed per day
- **Relative Humidity (Rh):** Desired humidity levels
- **Nitrogen Requirement:** Recommended nitrogen fertiliser application rate
- **Phosphorus Requirement:** Recommended phosphate fertiliser application rate
- **Potassium Requirement:** Recommended potassium fertiliser application rate
- **Yield:** Expected yield in kg per plant or per square metre
- **Category pH:** Optimal pH range for growth
- **Season:** Recommended growing season

## Attachments:

- Data
- Processed Data
- EDA Dashboard

## Acknowledgments:

The successful completion of this project is attributed to the dedication, hard work, and collaboration of all team members and leads. Their commitment and effort were instrumental in overcoming challenges, achieving milestones, and realising the project's objectives in the second phase.

# 3. Modelling

**Leaders:** Swathy Ramakrishnan, Nirupama Gurusubramaniam, Chukwudi Emmanuel Asibe

**Summary:** The focus of Task 2 in our project centred on developing models for crop classification and yield prediction. Through extensive exploratory data analysis (EDA), we identified and processed a dataset comprising 700 data points encompassing various vegetables, fruits, and salads. This dataset was selected for its comprehensive representation of crop types and rich feature set, facilitating robust model development. Our EDA provided crucial insights and patterns within the data, laying a strong foundation for accurate model training and predictions. This phase aimed to enhance agricultural planning and decision-making by delivering reliable insights into crop types and their potential yields.

**Key Accomplishments:**

- Developed a regression model for predicting crop yields.
- Built a classification model for identifying different crop types.

**Challenges Faced:** Throughout this phase, the team encountered and successfully addressed several challenges:

- Coordination across different time zones necessitated adherence to a unified meeting schedule based on UTC time.

- To combat team member inactivity, frequent updates and progress reports were shared via Slack channels, re-engaging members in project activities.

**Tasks Completed:**

- **Regression Model Development for Yield Prediction:**
  - Swathy Ramakrishnan
  - Chukwudi Emmanuel Asibe
  - Shaheer Abdullah
  - Ashik C Sabu
- **Classification Model Development for Crop Selection:**
  - Swathy Ramakrishnan
  - Chukwudi Emmanuel Asibe
  - Nirupama Gurusubramaniam
  - Ashwitha Kassetty
  - Utkarsh Sen

**Upcoming Tasks:**

- Finalise and hand over the pickled model from the classifier to the web app development team.

**Technical Updates: Model Development:**

- **Crop Classification:** Used Random Forest model for its high accuracy and robust feature handling capabilities.
- **Yield Prediction:** Employed a Linear Regression model due to its simplicity and effective performance during testing.

**Data Preparation:**

- Conducted thorough EDA to select a robust dataset of 700 data points, ensuring data cleanliness and relevance.

**Evaluation Metrics:**

- Classification model evaluated using accuracy, precision, recall, and F1-score.
- Regression model assessed using Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared score.

**Model Training and Validation:**

- Implemented stratified train-test splits and cross-validation techniques to ensure model generalizability and prevent overfitting.

**Hyperparameter Tuning:**

- Optimised model parameters through grid search and random search methods for improved performance.

**Model Deployment:**

- Pickled models to be uploaded to Google Drive for integration into the web application.

**Communication and Collaboration:**

- Leveraged Slack and Google Meet for seamless communication and collaboration.
- Active participation and task assignment ensured efficient workflow and progress tracking.

**Team Feedback:**

- Demonstrated strong collaboration and knowledge-sharing among team members, fostering a supportive and productive environment.

**Next Steps:**

- Proceed with finalising models and ensuring their successful integration into the web application.

**Additional Notes:**

- Focus on optimising task setup processes to streamline decision-making and improve efficiency.

**Key Metrics (if applicable): Classification Metrics:**

- Accuracy, Precision, Recall, F1-Score, Confusion Matrix

**Regression Metrics:**

- Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R-squared ($R^2$) Score, Mean Absolute Percentage Error (MAPE)

**Attachments:**

- Preprocessed data
- Notebooks
- Pickled Model: Link to GitHub repository

**Acknowledgments:** The success of this project is attributed to the dedication and hard work of all team members and leads. Their commitment and collaborative efforts were instrumental in overcoming challenges and achieving the project's objectives.

# 4. Web app

**Leaders:** Chukwudi Emmanuels Asibe, Swathy Ramakrishnan

## Summary:

The developed web application includes both frontend and backend components using Streamlit, enhancing user accessibility to the AI application. Users can input parameters in defined fields to obtain predicted results for Yield Prediction or Crop Selection and Yield Prediction options. The application also features an Insight page that visually illustrates the relationship of dimensions to target outcomes.

## Key Accomplishments:

- Developed the frontend of the web app using Streamlit.
- Implemented the backend application using Python.
- Successfully integrated the frontend and backend components.
- Designed and implemented the web app framework, including pages for Yield Prediction, Crop Selection and Yield Prediction, and Insight.
- Addressed challenges related to passive participant engagement by organising effective meetings to encourage active contribution.
- Identified and resolved errors in Python scripts that initially hindered progress, collaborating with team members to debug and ensure script functionality.

## Challenges Faced:

- Passive engagement of participants on the Task.
- Organised meetings to encourage participants to contribute actively.
- Observed errors in Python scripts that hindered task completion.
- Collaborated with team members to debug scripts effectively.

## Tasks Completed:

- Designed the web application interface.
- Developed the web app framework (Yield Prediction, Crop Selection and Yield Prediction, and Insight pages).
- Developed Python scripts for the backend.
- Integrated the backend with the frontend.
- Organised a meeting for the team to review and test the app.

**Participants:**

- Nirupama Gurusubramanian
- Swathy Ramakrishnan

- Chukwudi Emmanuel Asibe

## Technical Updates:

- **User-Friendly Interface:** Designed an intuitive interface for collecting input parameters and displaying prediction results directly on the webpage.
- **Visual Design:** Used modern fonts, visually appealing backgrounds, and suitable colour combinations to enhance readability and user experience.
- **Model Outputs:** Displayed model outputs seamlessly integrated into the webpage.
- **Collaboration:** Held frequent meetings with frontend and backend leads to refine designs and ensure seamless integration with backend functionality.

## Communication and Collaboration:

The team used Slack and Google Meet as primary communication channels for collaboration. Regular meetings facilitated discussions on project involvement, strategies, and deadlines. Decision-making involved all team members through polls or general consensus. Tasks were assigned based on skill levels, with experienced members mentoring beginners and encouraging active participation.

## Team Feedback:

The team demonstrated strong collaboration skills, effectively communicating and working together to achieve project objectives. Experienced members shared knowledge generously, contributing significantly to the team's collective learning and progress.

## Additional Notes:

Team members are committed to sustained involvement in the project, continuously adding value to the Web App Development process, recognizing it as an ongoing endeavour.

## Key Metrics (if applicable):

- **Performance of the Web App (Responsiveness):** Good
- **User Experience:** N/A
- **Technical glitches:** N/A

## Attachments:

- Crop Selection Web App
- GitHub Repository

## Task Acknowledgments:

The success of this project is attributed to the dedication, hard work, and determination of both the team members and the leads. Their collective efforts and unwavering commitment played a pivotal role in overcoming challenges, reaching milestones, and ultimately realizing the objectives of the project's second phase.

# Task3- Pests and diseases relevant to urban farming

## 1. Data collection

**Task Leaders:** The project was led by Swathy Ramakrishnan, Vino Gupta, Akshara Joshi, Nitesh Kesharwani, and Dixit Aryak.

**Team:** Beenaa Motiram Salian, Neringa Pannem, Sahil Hadke, Denis Surnin, Chinedum Onyema, Varun, William Green, Shivang Singh, Sai Deepak, Ashwitha Kassetty, Sulagna Parida, Antara Tewary, Kamia Salango, Harsh Khurana, Havish B, Nhat Minh Nguyen, Dinesh Kattunga, John Paul Curada, Muhammad Nadhif Taher Ahmad, Héctor Maya, Arpit Sengar, Denis Mwangi, Amandeep Singh.

**Data Collection Summary**

Data collection focused on pest and disease information for the following crops: Apples, Oranges, Strawberry, Kale, Basil, Parsley, Potato, Carrots, Zucchini. The collected data included detailed insights into:

- Pest life stages and reproduction rates.
- Seasonal appearances and feeding habits of pests.
- Damage patterns and visual markers for early identification.
- Environmental factors influencing pest management.
- Crop-specific insights on pest challenges and management strategies.

**Crop Pest Management Report**

The Crop Pest Management Report provided a comprehensive overview of pest management strategies tailored to each crop. Key aspects covered in the report included:

- Detailed crop and pest information.
- Specific damage patterns caused by pests.
- Integrated pest management strategies incorporating cultural, biological, and chemical controls.
- Recommendations for sustainable pest management practices.

The report emphasised the importance of integrated pest management strategies in mitigating pest populations effectively while reducing environmental impact and pesticide resistance. By combining various control measures, farmers can adopt sustainable practices that promote crop health and productivity.

# 2. EDA

Image Data exploration encompassed several key activities:

1. **Visual Inspection**: Samples were visually inspected to assess their quality.
2. **Image Size and Pixel Distribution Analysis**: Detailed analysis was conducted on image sizes and pixel distributions.
3. **Metrics Calculation**: Metrics were calculated to assess image blurriness and noise.
4. **Identification of Outlier Images**: Outlier images were identified and analysed to understand their impact on the dataset.

Datasets explorations encompassed several key activities:

In our primary datasets, namely "Pest Data" and "Disease Data," there exists a column labelled "temperature_range," which specifies the optimal temperature range for the life cycle progression of each pest or disease. This dataset offers valuable insights into the average temperature conditions in Milan and their implications for pest and disease prevalence.

The average summer temperature in Milan exceeds 25 degrees Celsius. By referencing the 'temperature_range' column from "Pest Data" and "Disease Data," which typically ranges from 20 to 30 degrees Celsius, it indicates that Milan's climate is conducive to the development of various pests and diseases.

Historical Climate Data provides a critical insight: the average summer temperatures align closely with the ideal temperature ranges documented in our "Pest Data" and "Disease Data" tables. This alignment substantiates the potential utility of these datasets for generating recommendations tailored to local climate conditions.

To further validate the utility of our datasets for recommendation purposes, we turn our attention to the "Seasonal_Appearance" column in both "Pest Data" and "Disease Data" datasets. This column indicates the seasons during which each life stage of pests or diseases is most active. Visualising these seasonal patterns allows us to correlate historical data on pest and disease occurrences with the documented seasonal appearances.

Upon analysis, it becomes evident that most of the pests and diseases selected in this study are predominantly active during summer, with some activity also noted in spring and fall. This observation corresponds closely with the seasonal appearance data captured in our datasets, reinforcing their reliability for predictive and management purposes.

Furthermore, comparing the overlap of crops between the "History Pest and Disease Data" and our "Pest Data" and "Disease Data" reinforces the consistency and applicability of our findings. The datasets predominantly share the same set of crops, further validating the relevance and reliability of our analyses and recommendations.

In conclusion, the alignment of historical climate data with the documented temperature ranges and seasonal appearances in our datasets supports their robustness for providing informed recommendations on pest and disease management strategies in Milan. These insights underscore the practical application of our datasets in agricultural decision-making processes.

# 3. Modelling

The Crop Disease Classification models were developed using the following approach:

- Initial data splitting into training, validation, and test sets.
- Application of data augmentation techniques to training sets to enhance model generalisation.
- Selection and customization of model architectures such as EfficientNet B1, ResNet, and others.
- Model training with appropriate loss functions and optimizers, utilising techniques like early stopping and hyperparameter tuning.
- Evaluation of model performance on test sets using metrics such as confusion matrix, precision, recall, F1 score, and accuracy.

# 4. Web App Development

The Crop Disease Classification System was a web-based application designed to classify crop diseases using image classification techniques. Key components included:

- **Model Training Page**: Facilitated model training and evaluation with user inputs for various parameters like crop name, augmentation techniques, epochs, and batch size. Outputs included EDA summaries and model performance metrics.
- **Prediction Page**: Allowed users to upload images for disease prediction, displaying predicted classes and actionable insights.

**Directory Structure**

The project directory structure facilitated efficient organisation and navigation of project files, ensuring clarity and ease of access to data, models, and other resources.

The Crop Disease Classification System demonstrated an effective integration of data science and agriculture, providing robust tools for pest management and disease identification. It underscored the importance of data-driven approaches in modern agricultural practices, aiming for sustainable and efficient crop production.

This report concludes Data Collection Task 3, highlighting achievements in pest management and crop disease classification for improved agricultural outcomes.

# Challenge Outcome

The project led to the development of a comprehensive database on potential cultivation areas across the final phase, the project outcomes were implemented for enhancing predictive models that took into account various factors such as climatic conditions, soil types, in urban contexts. These results are instrumental in refining our predictions and improving decision-making processes.

## Tutorials and Resources

**Introduction** This report provides an overview of tutorials and resources related to geospatial data analysis, machine learning, and remote sensing techniques. The resources include tutorials on various tools and platforms such as Google Earth Engine (GEE), QGIS, and machine learning frameworks.

**Tutorials and Resources**

1. **Tutorial Recording on KMeans Clustering**
   - **Author:** sensorfusion.research
   - **Description:** A tutorial recording on KMeans clustering, likely using Google Earth Engine for geospatial data analysis.
2. **Tiff image creation QGIS video**
   - **Author:** Deepa R
   - **Description:** Video tutorial on creating TIFF images using QGIS, a popular open-source geospatial software.
3. **Satellite Imagery - Google Earth Engine (GEE)**
   - **Author:** Maria Fisher
   - **Description:** Resource focused on using satellite imagery via Google Earth Engine for environmental monitoring and analysis.
4. **Merging Datasets with Different Longitude and Latitude**
   - **Author:** Maria Fisher
   - **Description:** Guide on merging geospatial datasets with varying coordinates, likely within the context of GEE.
5. **KMeans Clustering Google Earth Engine**
   - **Author:** sensorfusion.research
   - **Description:** Another resource by sensorfusion.research, specifically on KMeans clustering but tailored for Google Earth Engine.
6. **How to download Data from Copernicus**
   - **Author:** Maria Fisher
   - **Description:** Tutorial on downloading data from the Copernicus program, which provides Earth observation data.
7. **How to create account Google Earth Engine**

- **Author:** Maria Fisher
- **Description:** Step-by-step guide on creating a Google Earth Engine account, essential for accessing and using its tools.

8. **How to Add channel task-Slack**
    - **Author:** Maria Fisher
    - **Description:** Tutorial on adding a channel task in Slack, possibly related to collaborative work or project management.
9. **GEE API Setup Steps**
    - **Authors:** pratikshaofficial3, sensorfusion.research
    - **Description:** Steps for setting up the Google Earth Engine API, crucial for programmatic access and automation.
10. **Tutorial mlflow model**
    - **Author:** Vidhi Gupta
    - **Description:** Tutorial or resource related to MLflow, a machine learning lifecycle management tool, authored by Vidhi Gupta.
11. **Weather Data Collection of Milan**
    - **Author:** Ashwitha
    - **Description:** Resource focusing on weather data collection in Milan, likely involving data sources and methodologies used.
12. **Using QGIS and Quick OSM plugin to get infrastructure data**
    - **Author:** Ashwitha
    - **Description:** Guide on using QGIS and Quick OSM plugin to obtain infrastructure data, showcasing practical applications of GIS in data collection.
13. **Using Python to get GEE images**
    - **Author:** Ashwitha
    - **Description:** Tutorial on using Python programming language to retrieve images from Google Earth Engine, demonstrating automation possibilities.
14. **Downloading Sentinel dynamic world**
    - **Author:** Kaaymyke
    - **Description:** Resource focused on downloading Sentinel satellite imagery for dynamic world monitoring and analysis.

These tutorials and resources cater to a wide range of topics in geospatial data analysis, machine learning, and remote sensing. They provide valuable insights and practical guidance for researchers and practitioners interested in leveraging spatial data for various applications and serves as a useful reference for individuals looking to enhance their skills and knowledge in these domains, enabling them to effectively using tools and datasets for research and applications in environmental monitoring, agriculture, and urban planning among other fields.

# Dashboards and Reports

These dashboards summarise the findings and insights from EDA groups and aim to provide comprehensive analyses and predictions related to crop selection and yield in specific regions.

**Dashboard Analysis**

1. **Report - Crop Selection and Yield Prediction by Umuagwo (Link 1)**
   - **Author:** Umuagwo
   - **Link:** Dashboard Link 1
   - **Key Insights:** This dashboard likely focuses on crop selection strategies and yield predictions based on historical data and possibly incorporating machine learning models for predictive analytics.
2. **Report - Crop Selection and Yield Prediction by Umuagwo (Link 2)**
   - **Author:** Umuagwo
   - **Link:** Dashboard Link 2
   - **Key Insights:** This second dashboard authored by Umuagwo may provide further detailed analysis or could be an updated version of the first report, possibly offering refined predictions or additional data insights.
3. **Report - Crop Selection and Yield Prediction by Swathy Ramakrishnan**
   - **Author:** Swathy Ramakrishnan
   - **Link:** No direct link provided in the data.
   - **Key Insights:** This dashboard by Swathy Ramakrishnan adds another perspective, potentially bringing in different methodologies or regional-specific insights into crop selection and yield prediction.

These dashboards collectively represent a robust effort towards enhancing agricultural decision-making through data-driven insights. By leveraging historical data and predictive analytics, these reports aim to assist stakeholders in making informed choices regarding crop selection and yield optimization.

Further collaboration and comparative analysis between these dashboards could provide a broader understanding of effective agricultural strategies and contribute towards sustainable farming practices.

**Recommendations**

- **Interdisciplinary Collaboration:** Encourage collaboration between agricultural experts, data scientists, and policymakers to refine predictive models and ensure practical application.
- **Continuous Updates:** Regular updates to the dashboards with new data inputs and model recalibrations to improve prediction accuracy.
- **User-Friendly Interfaces:** Enhance dashboard accessibility and usability for non-technical users to facilitate widespread adoption and understanding.

In conclusion, the insights derived from these dashboards offer valuable opportunities for improving agricultural productivity and sustainability, paving the way for data-driven decision-making in the agricultural sector.

**References**


- Umuagwo. "Report - Crop Selection and Yield Prediction (Link 1)." Retrieved from Link 1.

- Umuagwo. "Report - Crop Selection and Yield Prediction (Link 2)." Retrieved from Link 2.

This report provides an overview and analysis of three web applications developed by the Omdena Milan Chapter, focusing on urban agriculture in Milan. These applications aim to facilitate the identification of suitable locations for urban farming, explore urban crops suitable for cultivation, and manage pests and diseases affecting urban crops.

# Web Application

1. **Task 1: Identifying Suitable Locations for Urban Farming**
    - **Link:** Task 1 Web App README
    - **Purpose:** This web application is designed to identify potential areas within Milan that are suitable for urban farming. It uses geospatial data, environmental factors, and to select areas with optimal conditions for cultivating crops in an urban setting.
2. **Task 2: Urban Crops**
    - **Link:** Task 2 Web App
    - **Purpose:** The Task 2 web application focuses on providing information about urban crops that are suitable for cultivation in Milan. It includes details such as crop varieties, growing conditions, nutritional values, and cultivation techniques tailored for urban environments.
3. **Task 3: Urban Crops Pests & Diseases**
    - **Link:** Task 3 Web App
    - **Purpose:** This web application is dedicated to managing pests and diseases that commonly affect urban crops in Milan. The main feature offers identification of pests and pathogens for prevention strategies, and treatment options to mitigate agricultural risks.

The web applications developed by the Omdena- Milan Chapter challenge *Identifying potential areas for urban agriculture in Milan, Italy* , have demonstrated a comprehensive approach to supporting urban agriculture initiatives in Milan.

- **Task 1** focused on strategic planning by identifying optimal locations, potentially enhancing food security and sustainability efforts in urban areas.
- **Task 2** provided essential information to urban farmers, promoting the cultivation of diverse crops suited to Milan's specific environmental conditions and community needs.
- **Task 3** addressed critical challenges in urban agriculture, offering solutions to manage pests and diseases effectively, thereby safeguarding crop yields and quality.

**Web app Repository**

- Omdena Milan Chapter. "Identifying Suitable Locations for Urban Farming." Retrieved from Task 1 README.
- Omdena Milan Chapter. "Urban Crops." Retrieved from Task 2 Web App.
- Omdena Milan Chapter. "Urban Crops Pests & Diseases." Retrieved from Task 3 Web App.

This report summarises the functionalities and benefits of the web applications developed by the Omdena Milan Chapter, highlighting their role in advancing urban agriculture practices in Milan.

# Conclusion

This innovative approach successfully integrated advanced technologies within urban agriculture showcased the potential of remote sensing technology and GIS applications played a pivotal role in advancing agricultural productivity while preserving the environment through sustainable practices, contributing to global climate resilience adaptation strategies for urban farming, we aimed at fostering resilient ecosystems, enhancing food security and environmental impact.

Through integrated technological solutions that aim to improve community well-being, part of this project sustainability in agriculture, we aimed to pave the way forward in addressing complex challenges related to agricultural sustainability goals, contributing towards a more sustainable future strategies for urban farming.

The culmination of our efforts over eight weeks resulted in the development and deployment of three distinct web applications designed to visualise and interact with the predictive modelling outcomes from our project aimed at identifying suitable areas for urban agriculture in Milan, Italy. These web applications were crafted to empower stakeholders with actionable insights by providing an intuitive platform where they could explore potential cultivation zones based on advanced data analysis, including soil types, accessibility of water resources, climatic conditions, and more.

# Impact and Future Directions

The potential to use AI tools in advancing agriculture practices in urban areas by leveraging technology to inform decision-making, promote sustainable farming practices, and mitigate challenges associated with urban food production. They empower stakeholders, including farmers, policymakers, and community members, to actively participate in and benefit from urban agriculture initiatives.

- **User Engagement:** Encourage user feedback and community involvement to improve the accuracy and relevance of data within the applications.
- **Continuous Updates:** Regularly update the applications with new data and insights to ensure they remain current and effective in addressing evolving agricultural needs.
- **Integration:** Explore opportunities to integrate these applications with local agricultural networks, educational institutions, and governmental initiatives to maximise impact and scalability.

# Repository

To find datasets, models, Colab notebooks, images, tutorials, and other resources related to the project, please visit the following repositories:

1. **Google Drive Repository:** This repository contains a comprehensive collection of project resources, including datasets, images, and other supporting materials. You can access it here.
2. **GitHub Repository:** This repository hosts the project's code, models, Colab notebooks, and tutorials. It is organised into various folders for easy navigation. You can explore it here.

All the necessary resources for understanding and engaging in the Urban Agriculture in Milan project are available in these repositories. Don't hesitate to contact us if you have any questions.

## Acknowledgments

I want to extend my sincere gratitude to all those who contributed to this project. Your hard work, knowledge, and teamwork have played a key role in reaching our objectives and coming up with effective solutions. We have made great progress in promoting urban agriculture. Thank you all for your hard work and commitment!

### Thank you!

Maria Fisher
Omdena-Milan
river@insiberia.net
Linkedin