

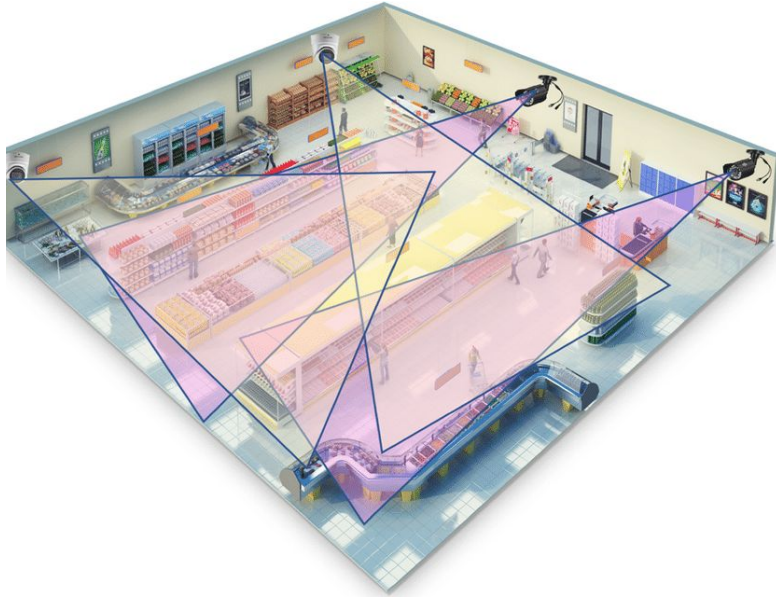


Team: Datadark

Hoang Hung
Tran Bao Nguyen
Dang Quoc Minh
Nguyen Dinh Quan
Hoang Thanh Tu

1. Introduction and Problem Statement

Abstraction



*What problems can be treated
with data of CCTV?*

These years, CCTV has become a crucial part and being applied in millions stores.

CCTV provides groceries enormous data

Only some malls and big stores 've already analyze those data

Most groceries use CCTV for manual tracking by hand



Only minor amount of data used

Problem 1: Unsuitable Items arrangement

The need to bend down to reach products on lower shelves poses challenges for a significant portion of the population. For example, it causes many annoying results

Accessibility Barriers: The current shelf layout makes it difficult for individuals with mobility challenges to reach and retrieve products comfortably. This lack of accessibility may deter potential customers and diminish the overall shopping experience.



Customer Safety: The act of bending down increases the risk of accidents, such as trips or falls, particularly among older customers.

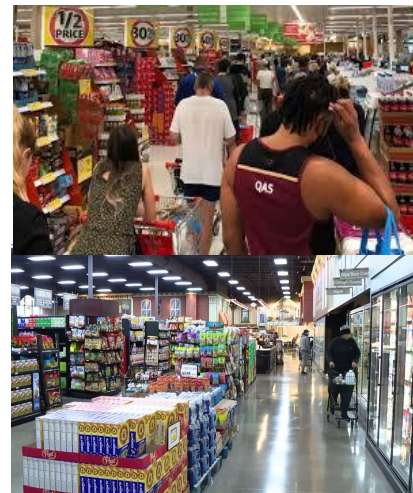


Problem 2: Unequal distribution

The imbalance manifests as crowded shelves in certain areas and empty or underutilized shelves in others, contributing to a suboptimal shopping experience for customers and posing challenges for inventory management.

Operational Inefficiency: Unequal distribution complicates inventory management processes, as some empty places require a lot of money to rent while not generating profit

Customer Frustration: Unequal distribution of products results in crowded shelves, making it challenging for customers to navigate and find desired items.

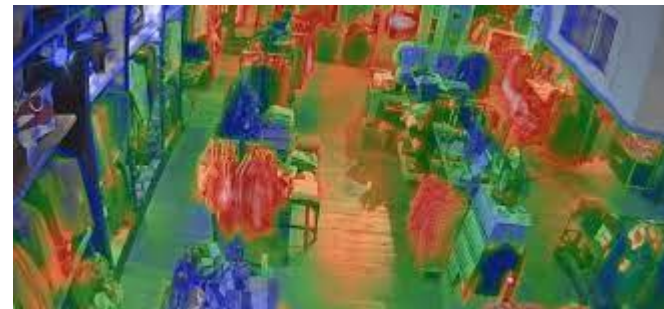


*Those two problems are common in many supermarkets and stores but not many people recognize it. Therefore, our MVP (**Gro**cify) will pioneer to provide solutions for the 2 above problems*

Insight: Heatmap and analytical data

Providing a heatmap and statistical data can be an effective solution to address the two problems mentioned

Heatmap: Implementing a heatmap that visually represents the product distribution across shelves can quickly convey information. The goods arrangement map visualizes places where mostly people bend down, while the distribution heatmap highlights crowded and deserted places



Statistical data: Alongside the heatmap, providing statistical data such as customer bend-down rates, product turnover rates, sales data, and popularity metrics for each product category can offer a comprehensive understanding of demand patterns. This data can guide decisions on shelf placement and stocking levels.

2. Solution Overview

Solution Overview

Product

“Grocity” is a responsive web application providing shop owners visualization and statistic of customer movement with data from video source (webcam, CCTV, etc).

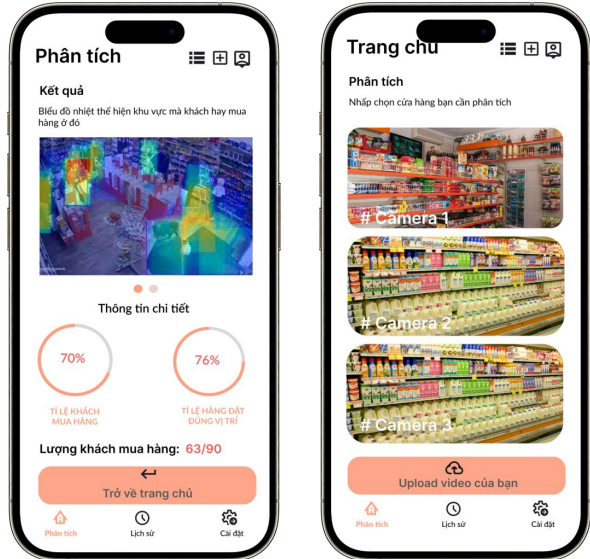
The solution applies computer vision deep learning in detection and classification on top of open source YOLO V8 model (from Ultralytics Developer).

Unique features

DETECTION	People tracking
	Keypoint estimator
VISUALIZATION	Distribution Heatmap
	Pose Frequency Heatmap
STATISTIC	People counter
	Customer pose comparison

Solution Overview

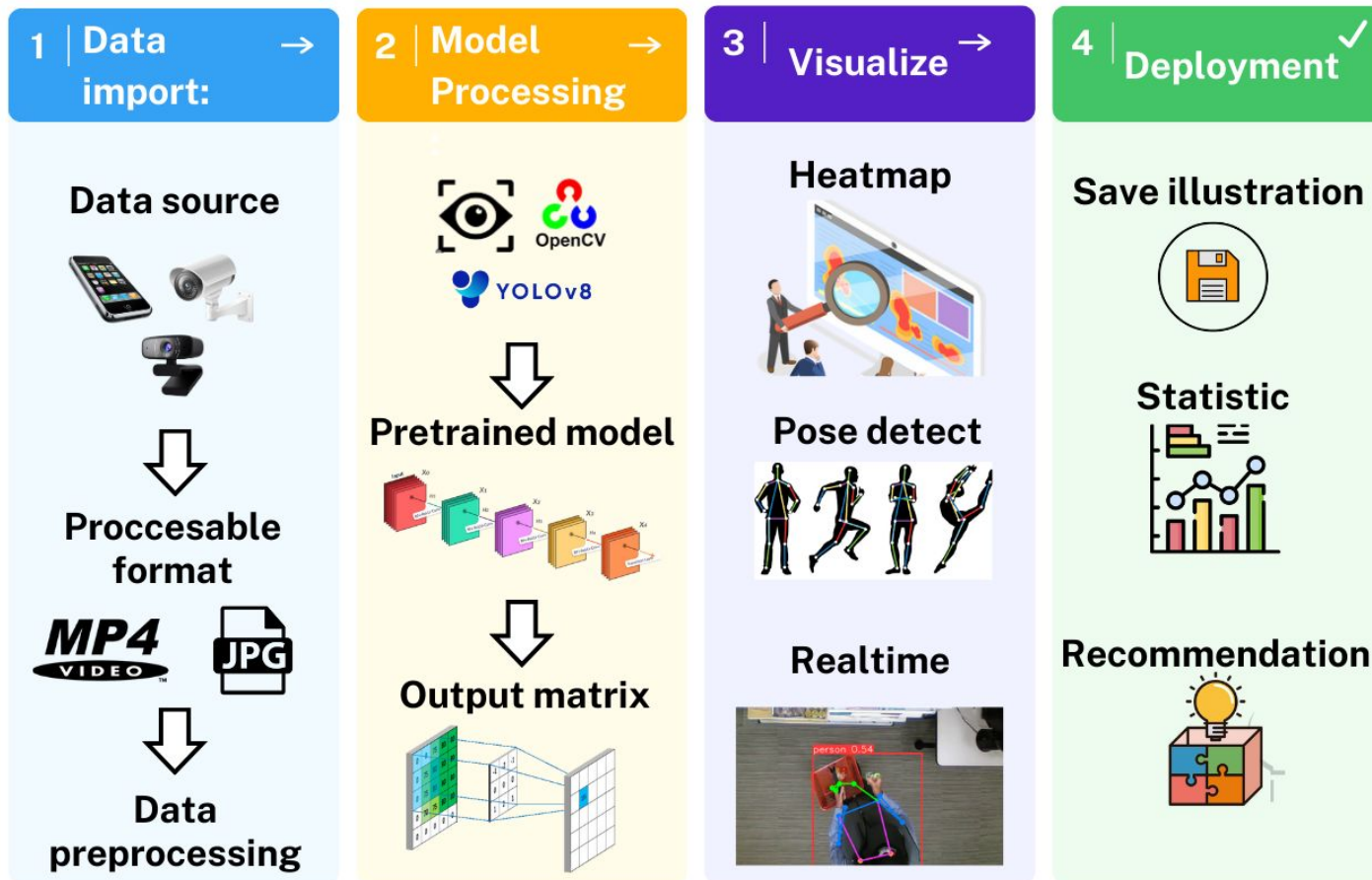
Usually, shop owners only consider on where their customer shopping. A bunches of solution to track customer's heatmap have been established. However, we believe that *“How customer shopping is absolutely much that important”*.



Accordingly, our product **Grocify** is proposed to provide shop owners not only position but also movement tracking of their customer in the stores.

We believe the solution plays a crucial role helping them in optimizing good/products placement.

Overview



3. MVP Introduction

Product concept

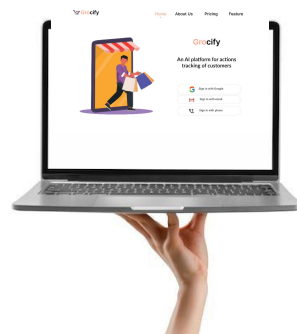
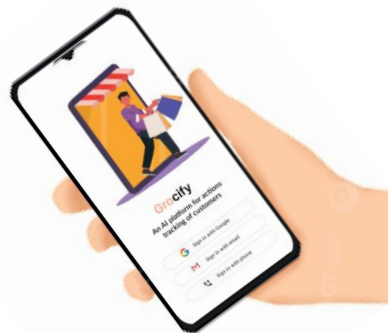
Target users Supermarket managers and grocery owners, who want to know the distribution of customers and get better goods arrangement.

Functional benefits Statisticizing details about the items that **do not have a good arrangement** that people bend down when purchasing items, helping increase customer satisfaction

Analyzing clearly the **distribution of buyers** in specific periods of time by plotting heatmaps, helping optimize the market space

Connect to the camera's storage and generate reports at the scheduled time (daily, every night, twice a week, etc), helping reduce **operational time**

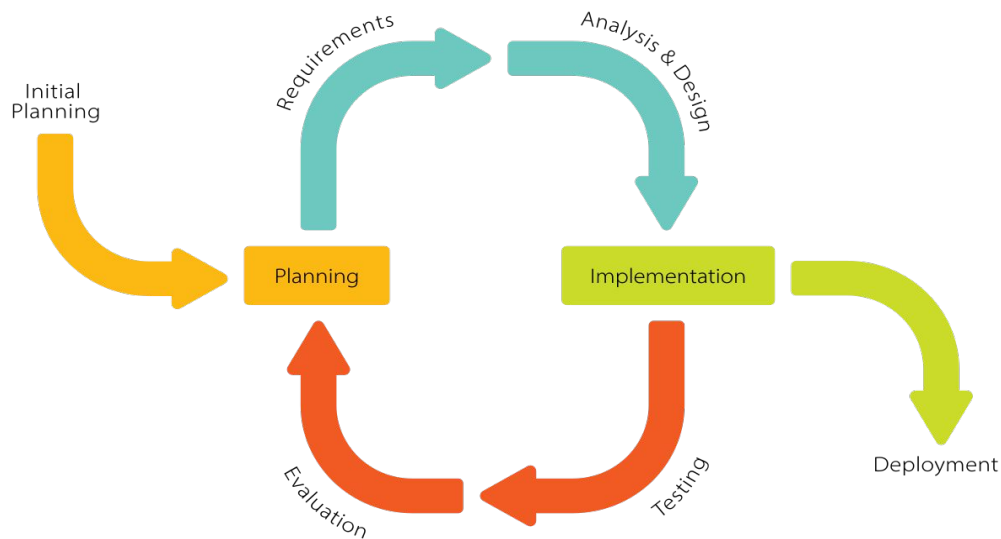
Platform **Web** (suitable for office users) and **App** (suitable for owners away from the store)



4. Methodologies, architecture design and tech stack

Methodology

Grocify is built on the architecture of incremental development:



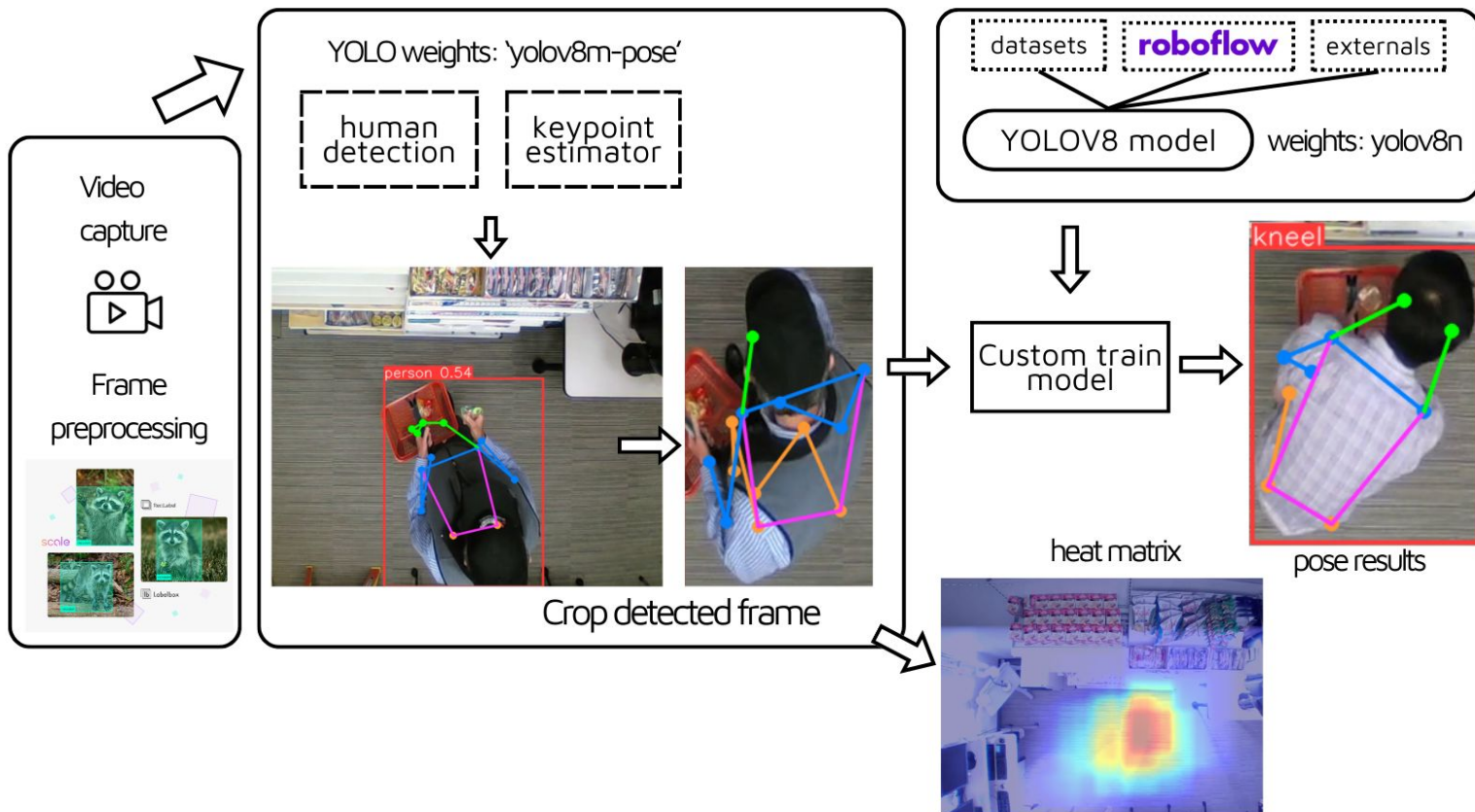
Project is not within budget

Reduce the cost of accommodating changing user requirements

Easier to get customer feedback on the development work that has been done.

Easier to get customer feedback on the development work that has been done.

AI architecture



Tech stack: Video/frame processing

OPENCV is chosen for processing video before pass it to YOLOv8 model



- Wide range of functions and algorithms
- Open-source
- Flexibility
- Speed
- Ease of use

Tech stack: Real-time object detection model

YOLOv8 is chosen for detecting person in our MVP



- Support for converting model to others format
=> Deploy model in real life application
- Updated and maintained frequently by Ultralytics
- Open-source
- Speed
- Accuracy
- Ease of use

Tech stack: Developing MVP

Our MVP mainly focused on **rapid development, the flexibility of the NoSQL**

database, and asynchronous I/O. Therefore, we will choose the **MERN** stack for the whole development.



- **“Type-safety”** and strict typing
- Maintainability and readability
- Catch errors at **compile time**



- **Scalability** and ease of deployment
- **Light-weight**, which can be adapted for 1-month project



- Supports **statistical** modeling
- Sketch **interactive** visualizations for web-app MVP.



- **Versatile** API frameworks
- Supports **middlewares** and **conventions**

Tech stack: Data Storage and Processing



- Do **not** require **Relationships** and **References** between items
- Data structure can **change quickly** and **frequently** during development
- **Faster query** time compared to SQL, embedding Atlas Search as a full-text search



- Support workflow orchestration due to the scheduling platform of DAGs



- Cloud-based database use to store images and video with fast query time

Tech stack: Deployment



- Package our application and its dependencies into an image through the Dockerfile.
- Support CI/CD, helping define staging and production processes.



- Simplifies the deployment and scaling of web app applications
- Provide analytics, number of access, access region

5. Core functionality and performance metrics

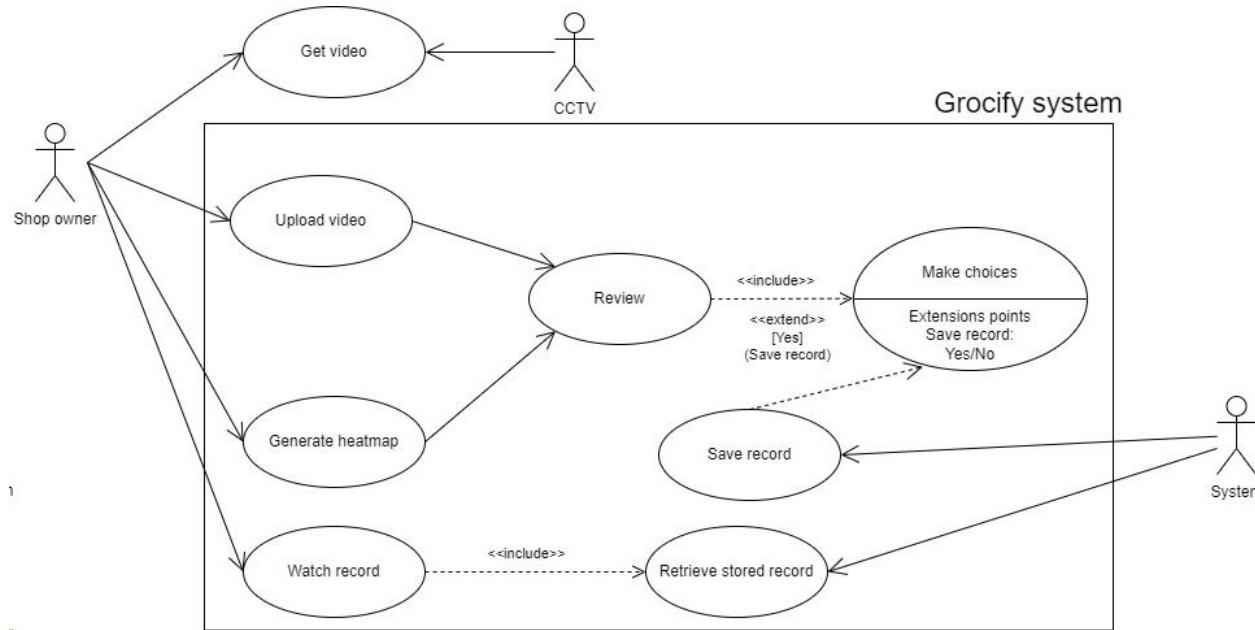
Core Functionality

“Grocify” has following functionalities providing show owners visualization and statistics of their store distribution:

- **Distributed Heatmap Visualization:** describe the frequency that customers appear in the CCTV/Camera.
- **Stand/Kneel Pose Heatmap Visualization:** describe the frequency customers stand/bend down at each shelf in CCTV.
- **Customer & Pose realtime detection:** realtime inference of AI model through CCTV data.
- **Mixed Heatmap:** visualize combination of Distributed and Pose estimation heatmap.
- **Retrieve statistic:** get figures from processed frame/videos (people counter, stand & kneel proportion, return product rate)



USE-CASE DIAGRAM



Due to the use-case diagram, our main flow includes:

1. Shop owner can use video recorded from CCTV or upload new ones
2. The system will return heatmap and analytical data of videos
3. After receiving the heatmap and data, users choose to delete or save record for further review
4. Users can retrieve and watch stored record at anytime, or they can delete helpless records to save memory

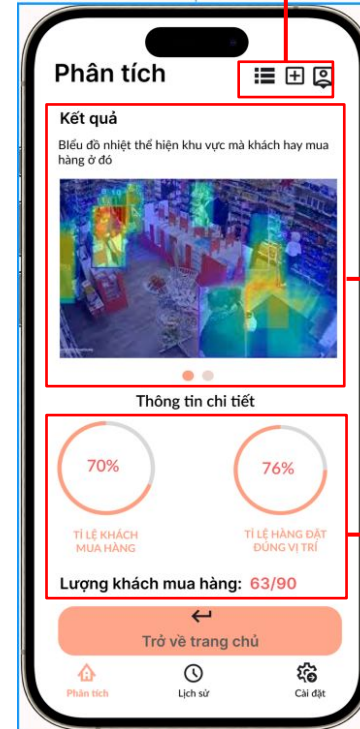
UI explanation

The center of app provide a list of

The navbar helps us to navigate the main page, the history, and setting manager



The header bar provides services, camera adding



A heat map is displayed to represent the locations that do not have suitable goods arrangement

Some information extracted from the video is represented as charts and text

Application

To maintain the convenience and ease of use for **Grocify**, we consider application's response time, with the focusing on working with large size file (as CCTV source video always long, despite its low quality) as well as output image size.

Model

Of course, the accuracy and speed of AI model can not be neglected. As our model is in field of object detection and classification, we will use mAP (mean Average Precision). It is used to evaluate the accuracy of models in identifying and localizing objects within an image.

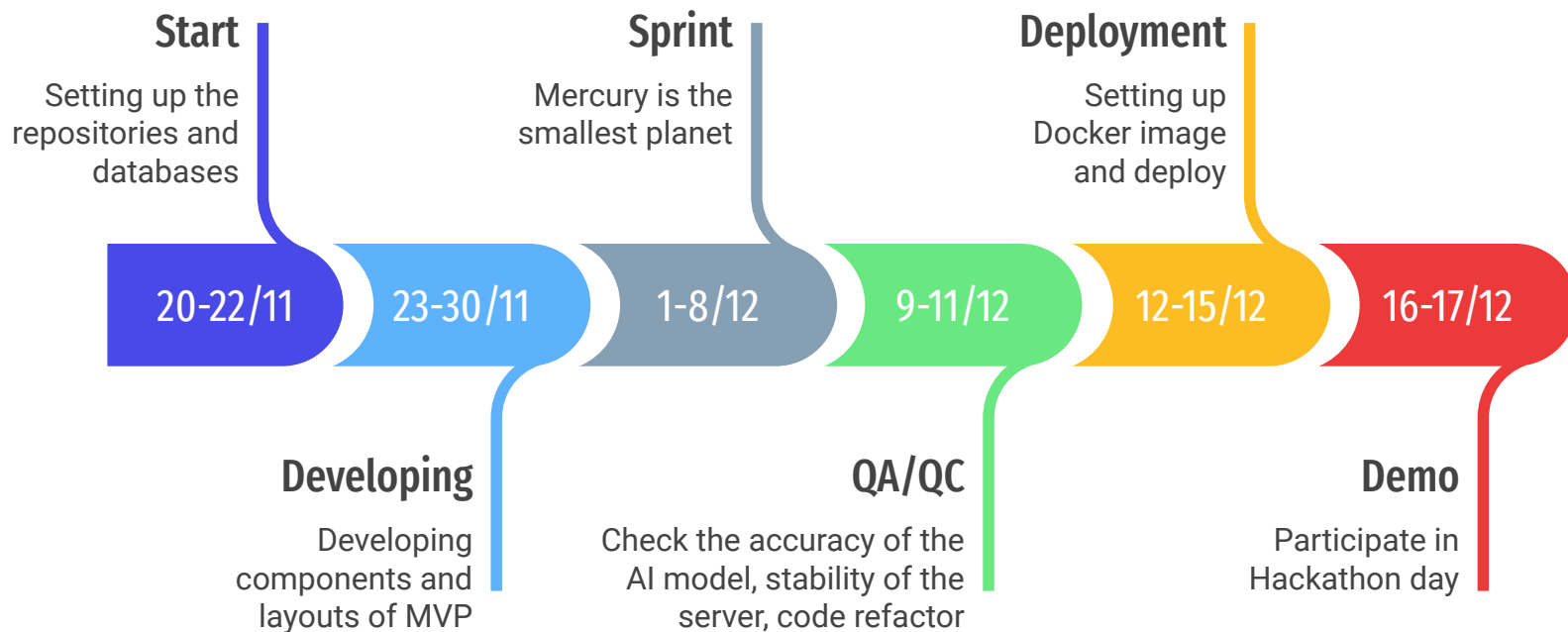
Performance Metrics

Application evaluation		
Processing time/ Video length	Video Length	Time
	Short (<15 mins)	immediate
	Medium (1 hour)	< 1 min
	Long (n hour)	< log2(n) min
Realtime delay	CCTV/Camera	<1500ms
Output storage	Distribution heatmap	< 500KB
	Pose heatmap	< 500KB

Model Evaluation (mAP)		
Person Detection Model	Precision	> 0.95
	Recall	> 0.95
Pose Estimator Model	Precision	> 0.9
	Recall	> 0.9
Whole Model	Runtime	> 45 frames/ second

6. Timeline

Timeline and Roadmap



7. Current limitations and future development

Potential future development

There are some solutions our team suggested for the future enhancement

Initiating new features: Integrating other recognition platform such as shoplifting detection, people shopping habit by analyzing the type of items customer buy



Integrate in public places: Our system can be installed in other public locations. For example, in library, the shelves are high and the amount of books in each shelf is extremely significant, therefore an AI model notify the some shelf need to be rearrangement is important to save the time for lectures and students

Limitations

There are some solutions our team suggested for the future enhancement

Direct connection with CCTV is limited: In order to analyze the outcome of retail, our web-app require access to the storage of the past videos of CCTV instead of connecting directly like other IoT apps. This is because our app does not support the IoT services really much.



CCTV is too expensive: For some middle and small groceries, especially in rural areas. Setting up the whole system may cost a considerable amount of money to operate

Conclusion

Apart from given information, we can conclude about our app as:

Key-point: customers have to bend down for items and some places in shop do not have equal popularity

Value proposition: solution to two critical challenges: unequal product distribution and lost sales opportunities.

Benefits of AI model: improving the image processing stage, helping plotting heating with providing analytics data for retail sellers.

Potential impact: can adding more detection features such as shoplifting, shopping habits, etc in the app.