Junction 2021 - CGI & KONE Challenge

EVEN SMARTER MAINTENANCE SERVICES

KONE

A global leader in the elevator and escalator industry

Our mission at KONE

To improve the flow of urban life

A typical day at KONE

*) Based on the total amount divided by the number of average working days.
**) Based on an estimate.

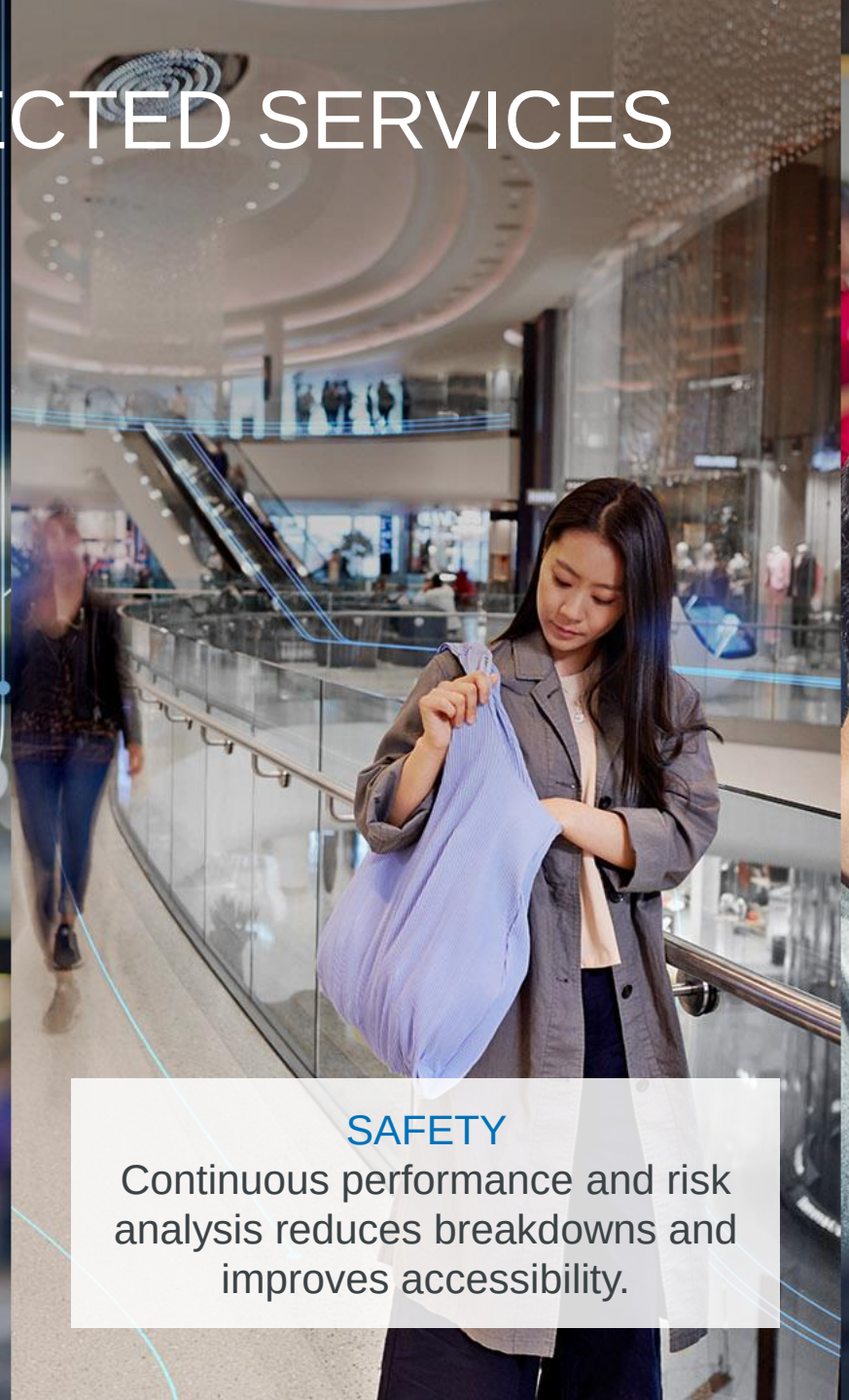
24/7 Connected Services

THE BUSINESS BACKGROUND OF THE CHALLENGE

KONE 24/7 CONNECTED SERVICES

INTELLIGENCE

AI-based predictive maintenance increases the value of your building and prolongs equipment lifetime.



SAFETY

Continuous performance and risk analysis reduces breakdowns and improves accessibility.

TRANSPARENCY

Equipment status and service reports, based on real usage data, keep you in the know.

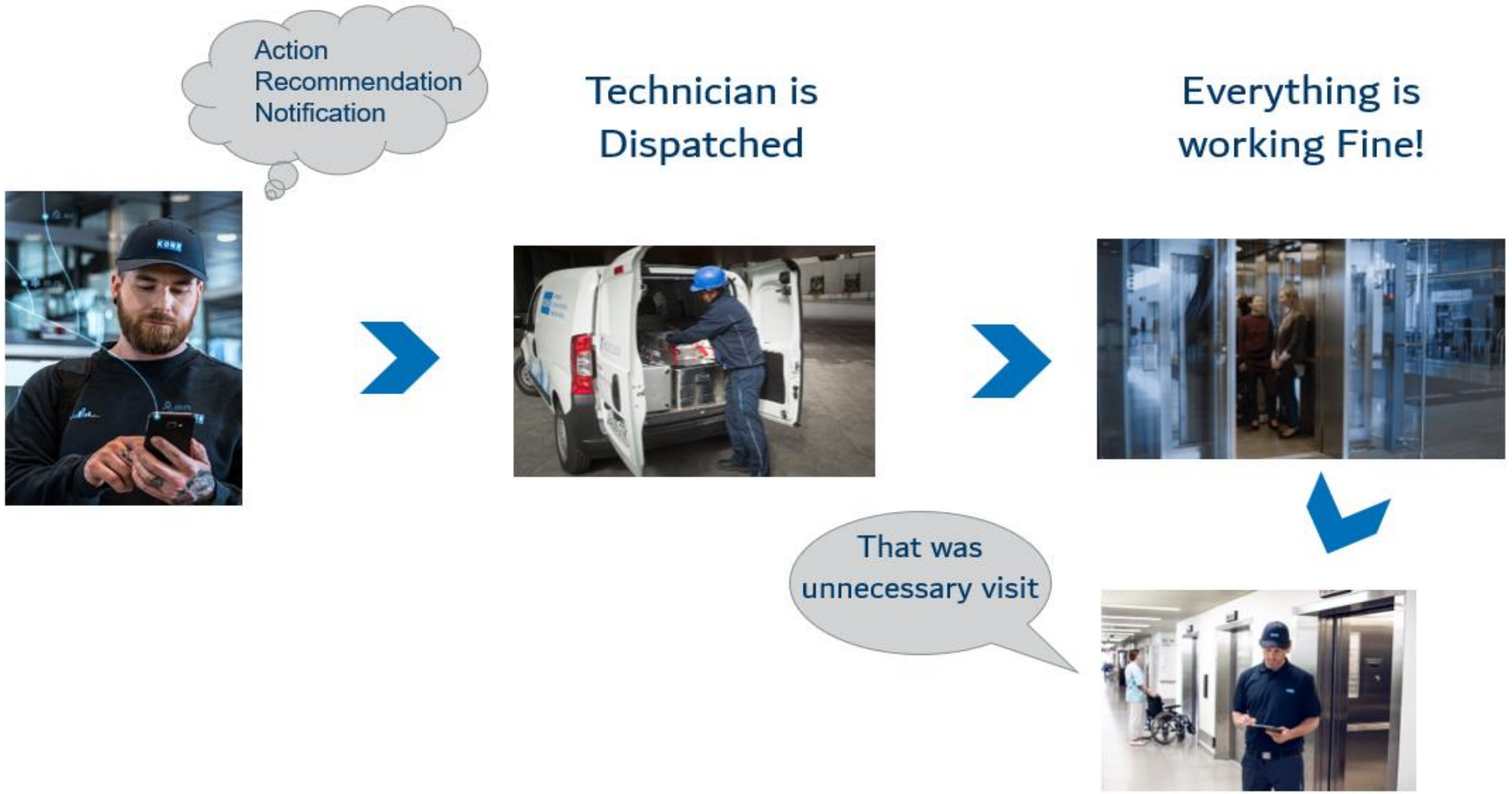
24/7 Connected Services

Equipment with 24/7 Connected Services send data about their operation and condition to the cloud.

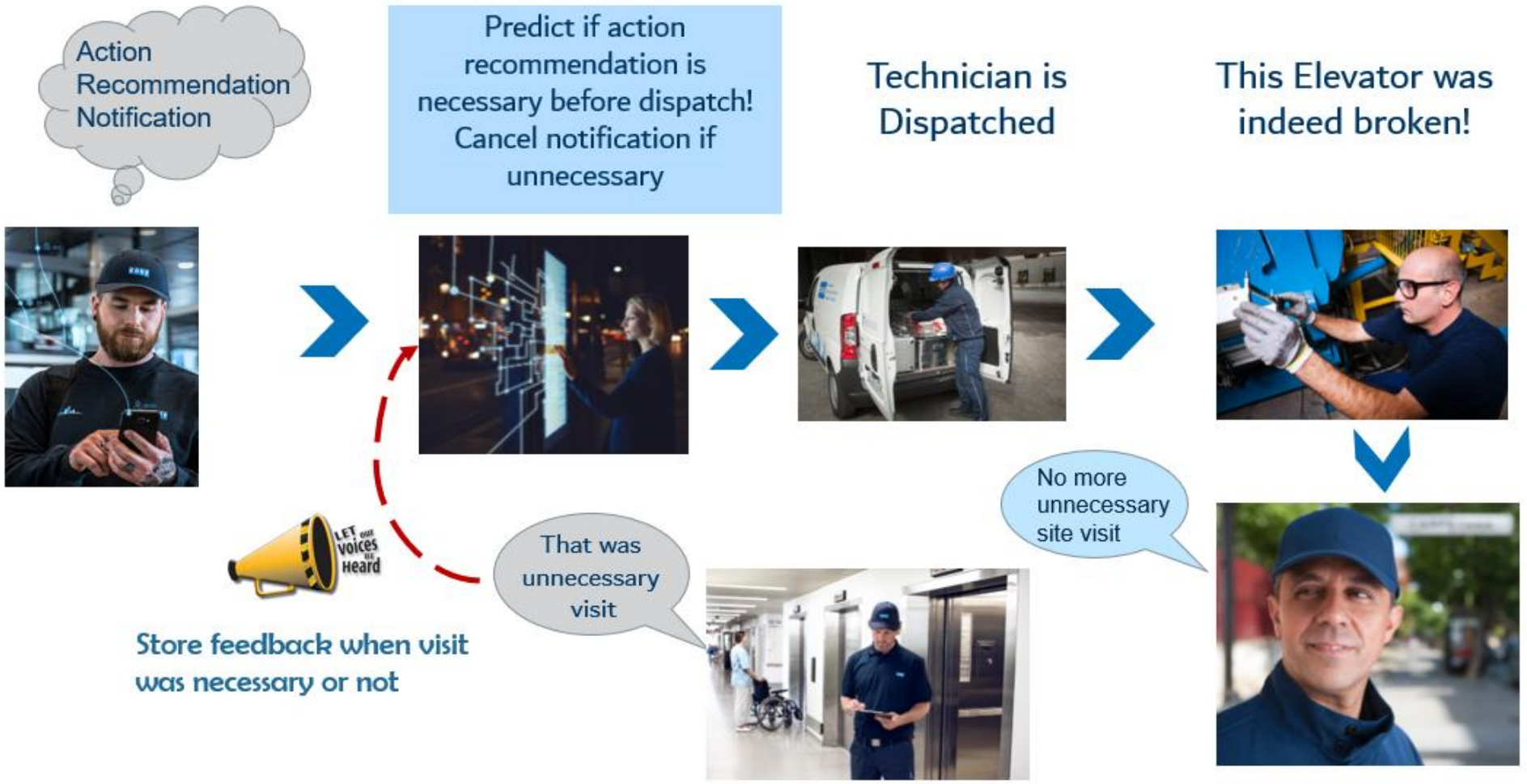
Based on the data from the equipment various algorithms generate service action recommendations that are sent to the technician.

Field technician checks the equipment and gives feedback on whether the action recommendations were accurate.

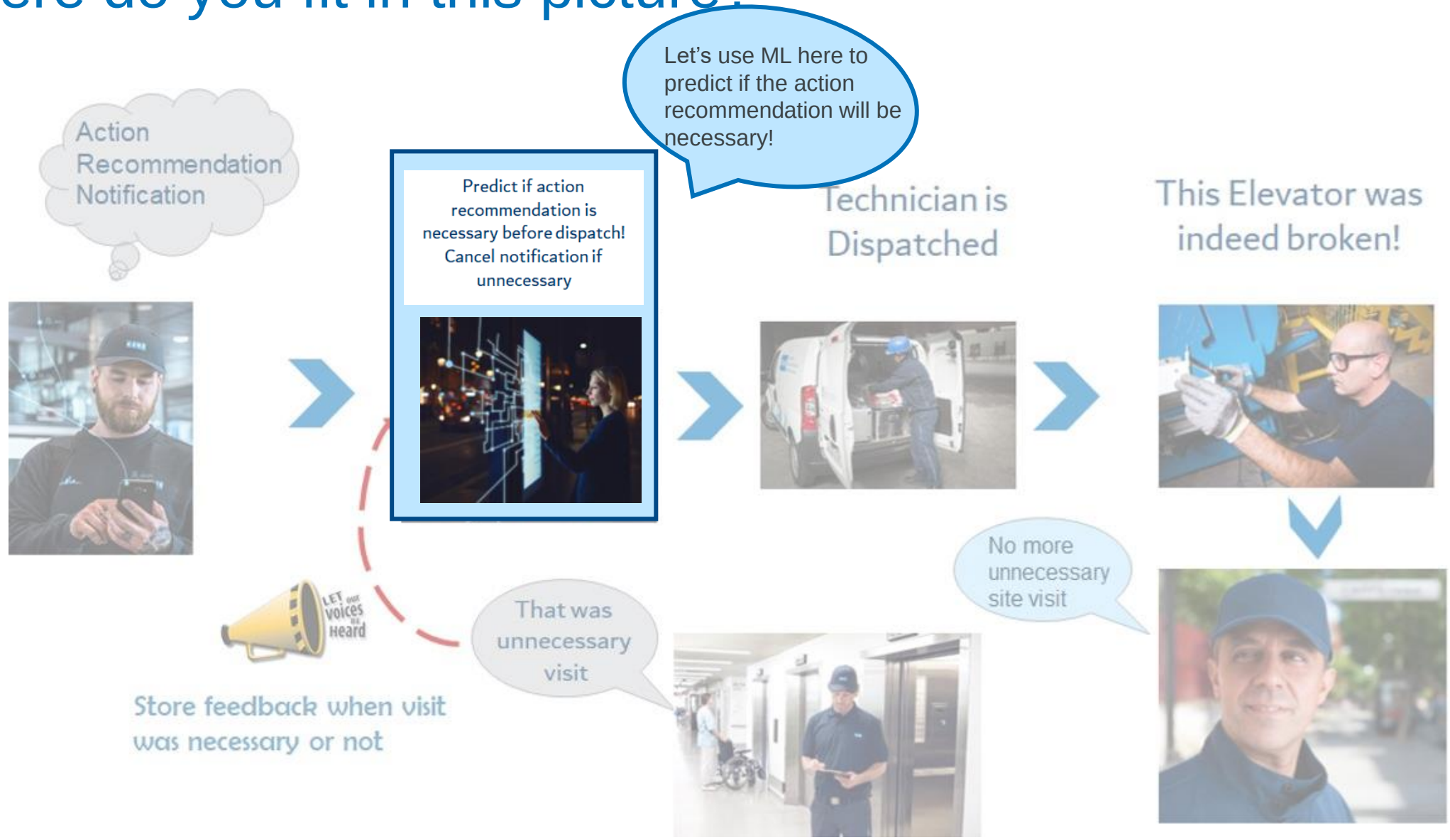
Challenge Introduction:
A short journey into technician's life



How can we improve the work of the technician with Machine Learning?

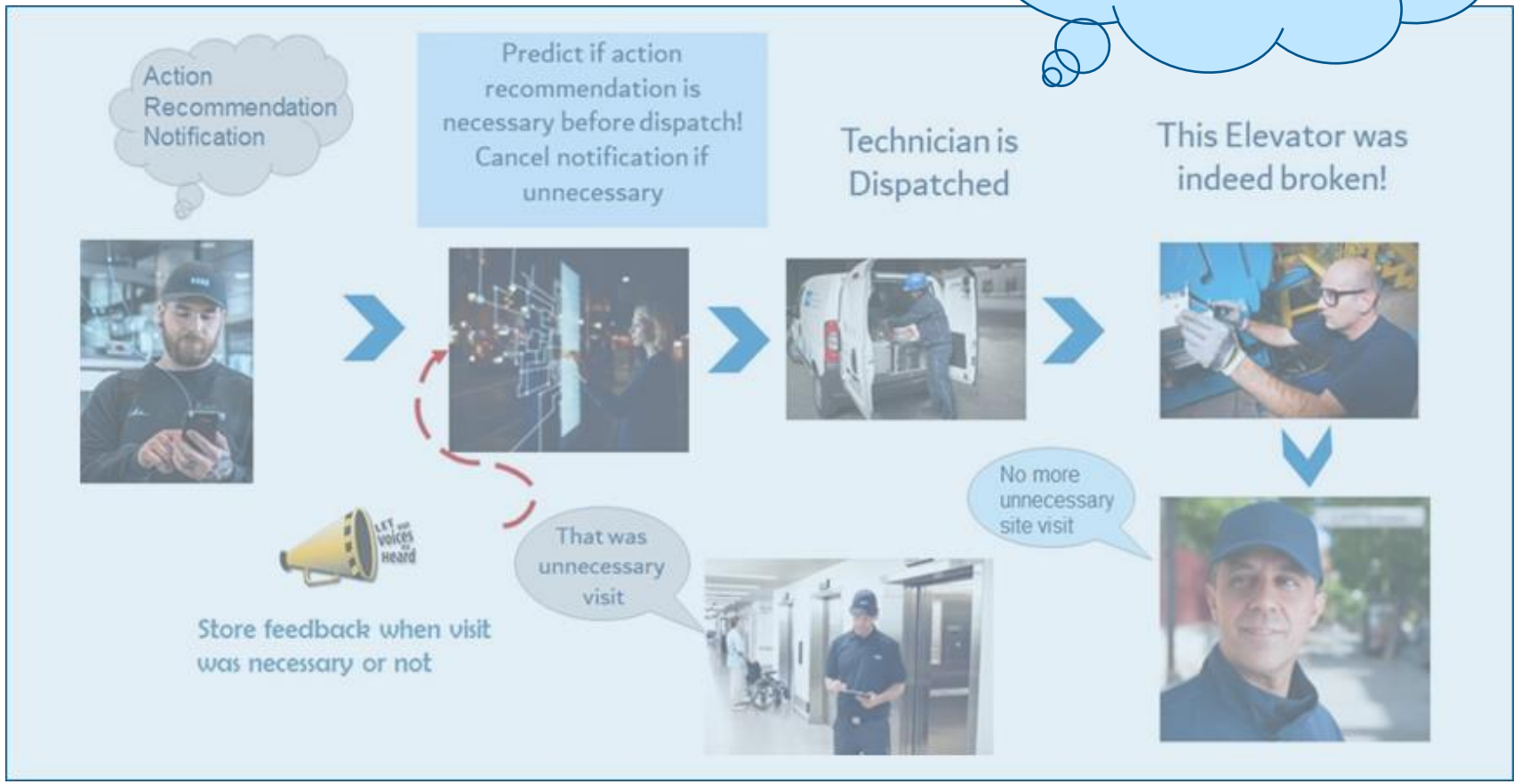


Where do you fit in this picture?



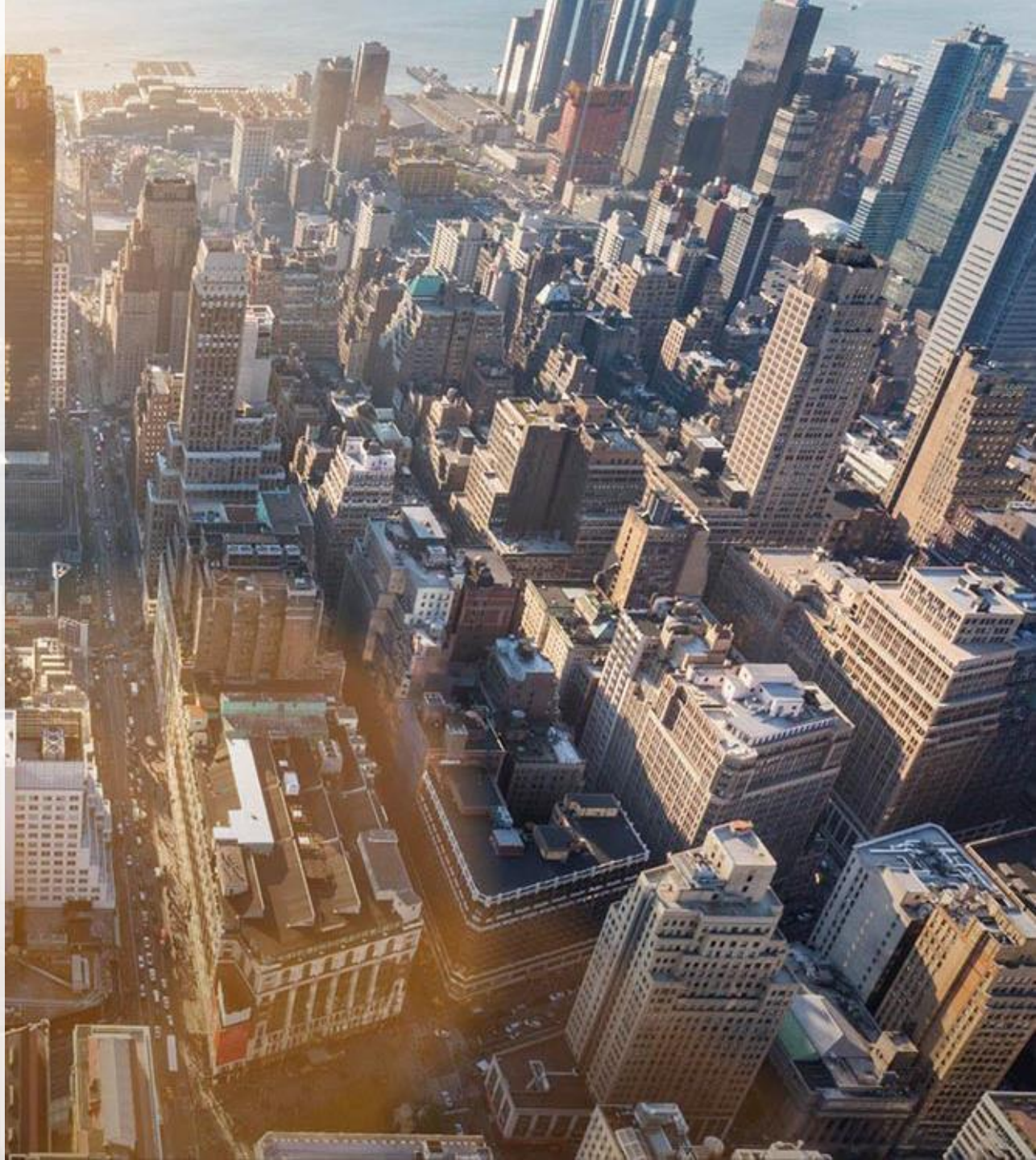
Where do you fit in this picture?

Or can you rethink this entire process to reach our goal?



Judging Criteria

HOW THE CHALLENGE IS JUDGED



Judging Criteria

- This challenge awards a winner in two distinct categories:
  - Best Performing and Most Explainable Model
    - This category is judged based on your model's performance and how explainable the predictions are. Please elaborate on the latter part in your submission.
  - Out of the Box Thinking
    - You can win in this category by rethinking the whole service process or by creating a novel ML solution to the service action filtering problem.
- Judging in both categories as well as awarding the 1st prize is subjective and not based on any predefined formula

Dataset Introduction

WHAT DATA IS PROVIDED FOR YOU

Dataset Introduction

train.csv

Train ML model

Historical action recommendation data with technician feedback (ground truth)

test.csv

Generate prediction

Historical action recommendation data without technician feedback from other groups equipment

submission_example.csv

Submit prediction

train.csv

115475 entries (rows) + a header row
13 columns: 12 features + 1 ground truth

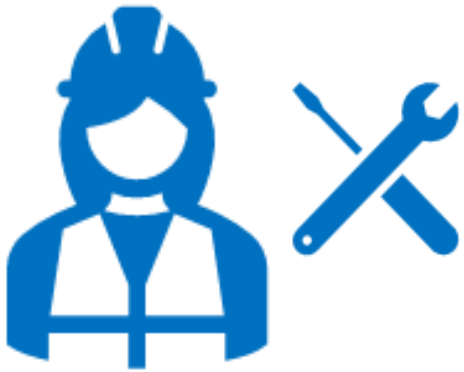
Equipment features

features	definition
equipment_id	Unique equipment identifier
equipment_area	Equipment that is geographically close to each other, belongs to the same equipment_area
equipment_category	Equipment category groups equipment with similar technical features together
usage_type	Equipment belonging to the same usage_type share a similar usage pattern.
speed_category	An ordinal variable, equipment belonging to the same speed_category have similar rated speed and equipment in speed_category 2 move faster than those in speed_category 1 etc
load_category	An ordinal variable, equipment belonging to the same load_category can carry similar rated load and equipment in load_category 2 carry more load than those in load_category 1 etc.
floors_category	An ordinal variable, equipment belonging to the same floor_category have similar number of floors and equipment in floor_category 2 have more floors than those in floor_category 1 etc

train.csv

115475 entries (rows) + a header row
13 columns: 12 features + 1 ground truth

Maintenance events



features	definition
case_id	Unique identifier of technician visiting the equipment. One case can contain multiple service action recommendations.
completion_date	Date when technician visited equipment
action_recommendation_id	Unique identifier of which service action recommendation was sent to the technician
action_recommendation_type	Type of service action recommendation. Different types of action recommendations are handled with separate processes.
action_recommendation_category	Category of service action recommendation based on what kind of monitoring hardware is installed on the equipment

train.csv

115475 entries (rows) + a header row
13 columns: 12 features + 1 ground truth

Technician feedback

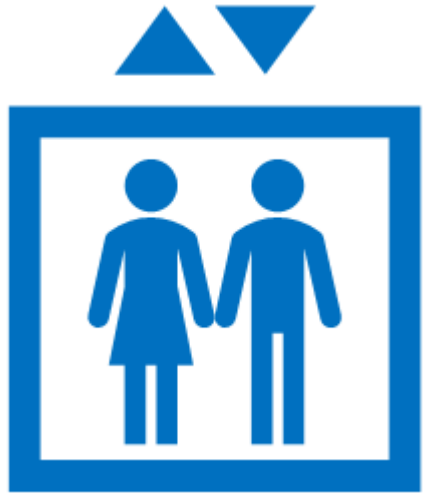
column	definition
feedback	Technician feedback whether service action recommendation was accurate (1) or unnecessary (0).

test.csv

29428 entries (rows)+ a header row
12 columns: 12 features

features
equipment_id
equipment_area
equipment_category
speed_category
load_category
usage_type
floors_category
case_id
completion_date
action_recommendation_id
action_recommendation_type
action_recommendation_category

Distinct equipment_id



submission_example.csv

Submission criteria:

- Submission should have exactly 29428 entries + a header row (example has only 10 entries)
- The file should have exactly 3 columns:
  case_id
  action_recommendation_id
  feedback (contains your binary predictions: 1 for accurate, 0 for unnecessary)

⚠ N.B: Do not change the datatype of the columns or the column names. Invalid submission csv may lead to disqualification.

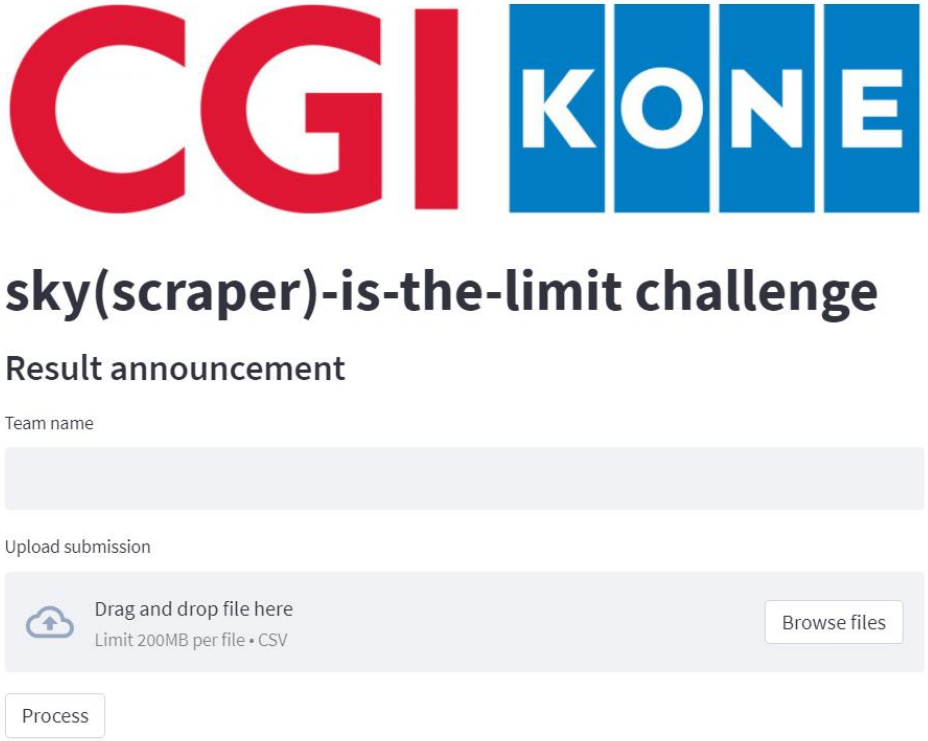
case_id	action_recommendation_id	feedback
9d54504e-c805-4859-b92e-a8df79732	ar00000250	1
554e89db-0d65-44f1-a3cb-d79662cd1	ar00000188	1
6cb91017-5e91-446b-9064-40975833	ar00000005	1
496f3fe9-36a6-4b44-a686-5383584c4	ar00000124	1
f8189755-b79e-4114-8a2b-84e08a22	ar00000250	1
7b7672dd-6e41-49dc-b70e-7e32e5d7	ar00000060	0
f6ce9a79-f7b8-461d-83a7-2a30e841d	ar00000174	1
6bf48aee-a676-4e9d-9ffe-2213c75fe7	ar00000018	1
2b424690-94a9-4b24-b2e7-b772a9f4c	ar00000273	1
e58245ba-a372-4916-b20f-4ee93c0c1	ar00000174	1
328a282a-349e-4db3-a380-41e12679	ar00000188	1
328a282a-349e-4db3-a380-41e12679	ar00000250	1
5d4e7e78-ce86-4bb9-b99d-c1832781	ar00000105	1
70f8070b-d1ed-491b-95fb-fbcb3ff223	ar00000293	1

F2 to evaluate “Best Performing Model”

$$F_2 = (1 + 2^2) \cdot \frac{\text{precision} \cdot \text{recall}}{(2^2 \cdot \text{precision}) + \text{recall}}$$

Cost of missing relevant action recommendation > cost of unnecessary service visits

Overview of streamlit app used by judges to evaluate model performance from submission csv



Come to our booth to get access to the dataset and ask questions


Dedicated to People Flow™
KONE