A data-driven prognostics for an assembly machine for automatic transmissions

Prof. Hyunbo Cho, Ph.D.

hcho@postech.ac.kr

Industrial & Management Engineering
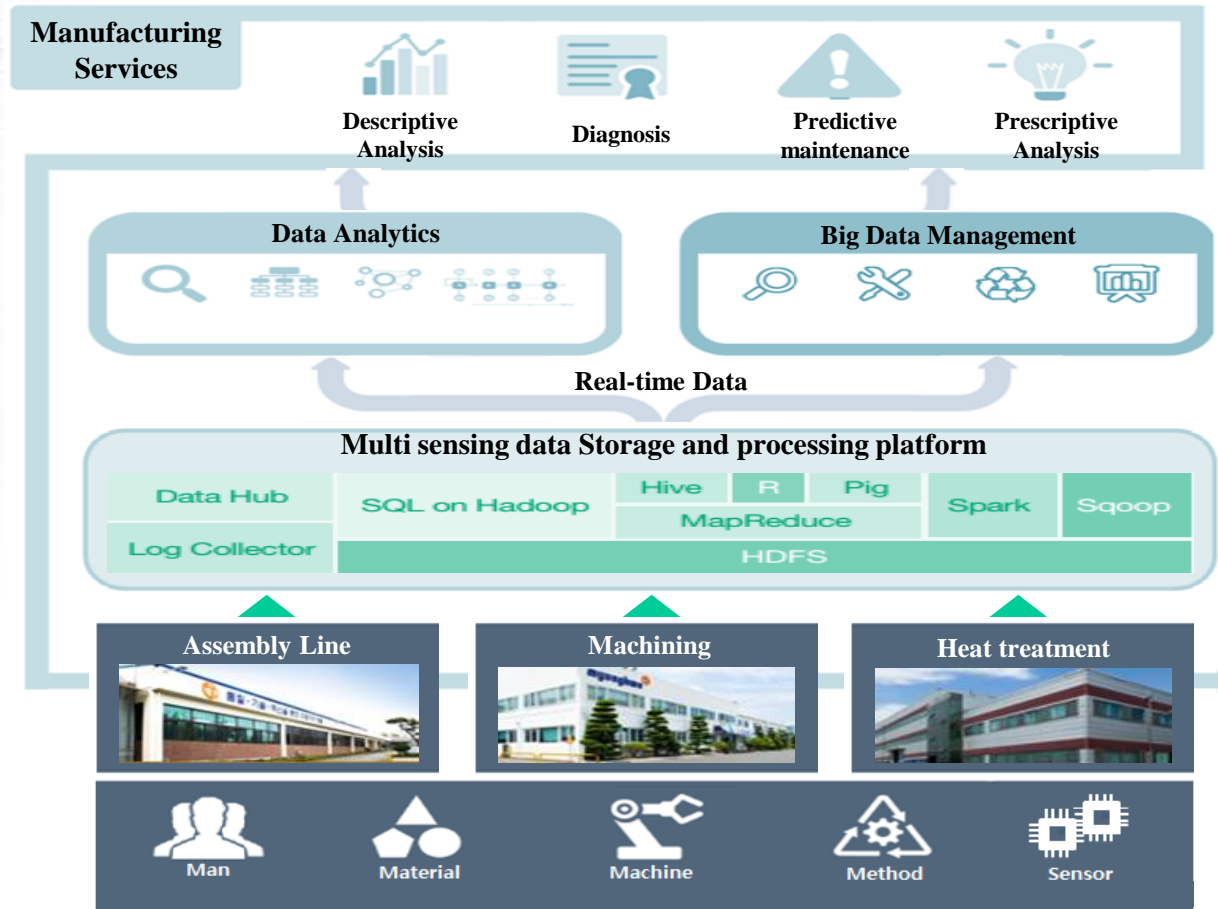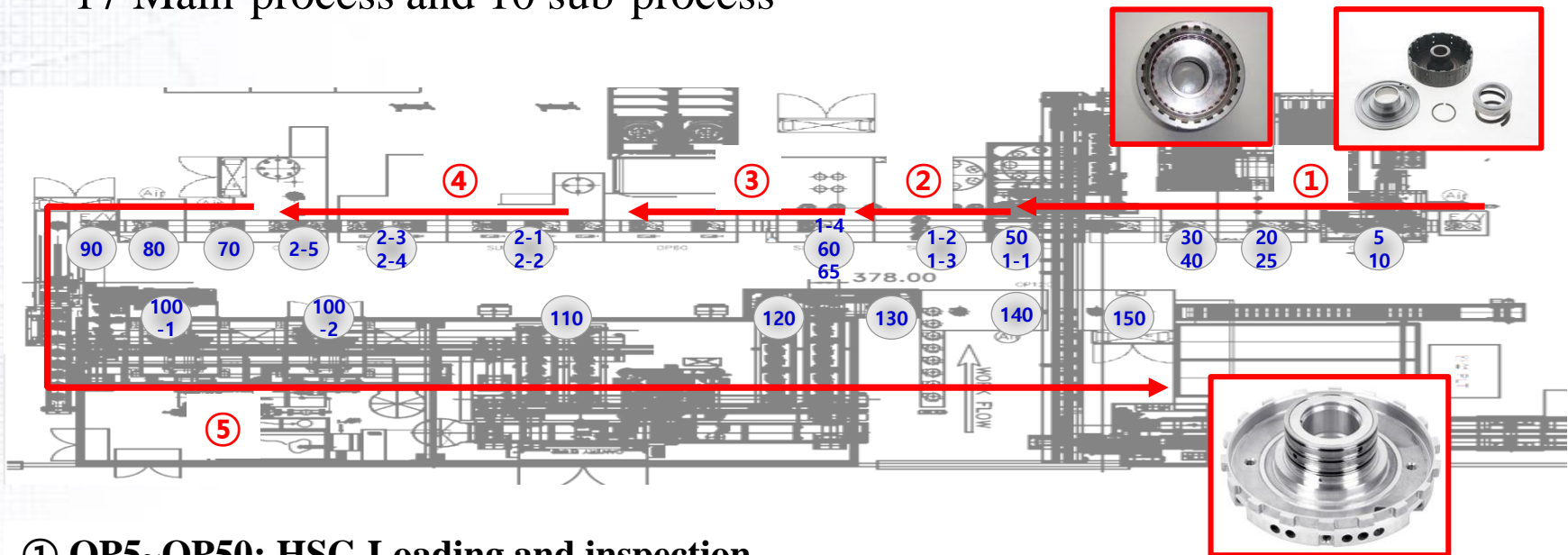
POSTECH

Introduction to Project

Goal — Development of artificial intelligence-based diagnosis / prediction system for the **Seamless Manufacturing**

Layout for Assembly Line
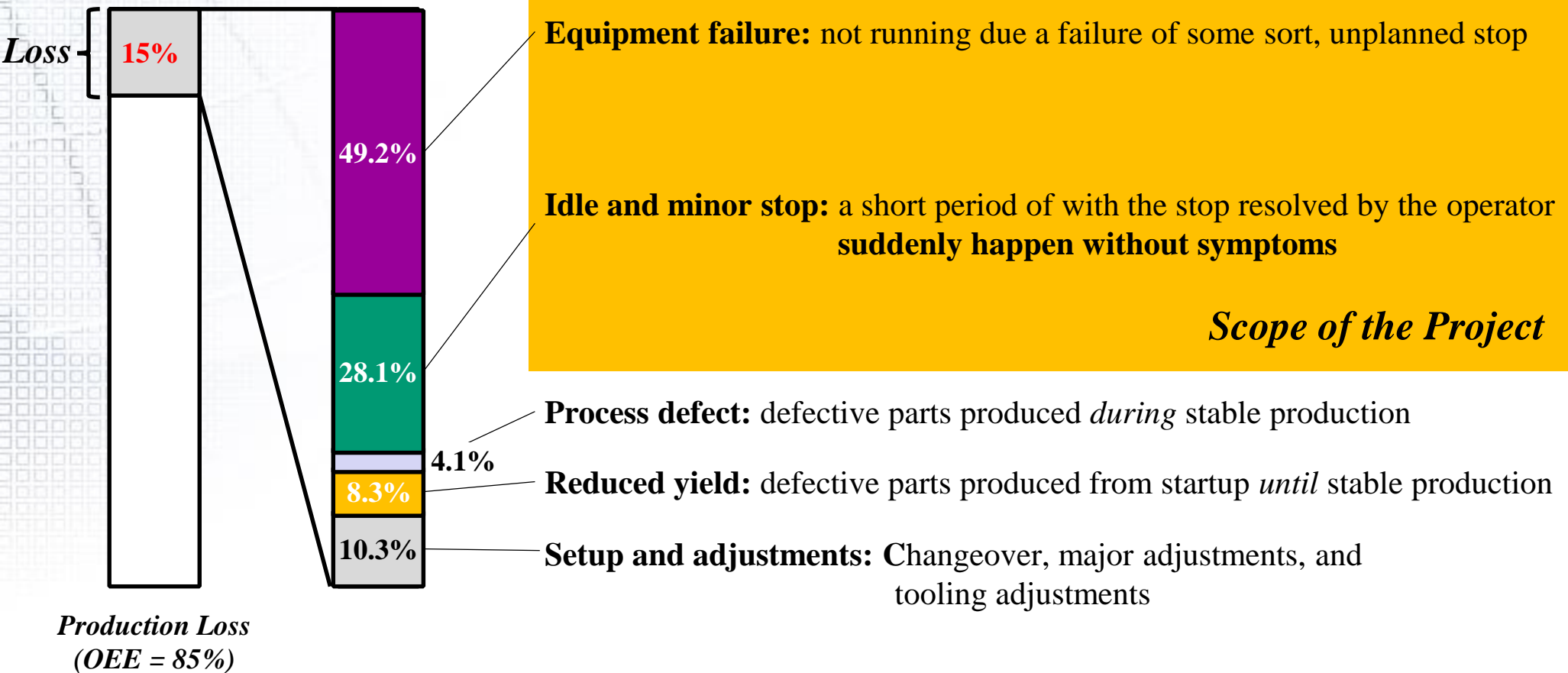
❑ Automatic Assembly Line for Transmission Drum

- 17 Main-process and 10 sub-process

① OP5~OP50: HSG Loading and inspection
② SUB1-1~SUB1-4: Piston, Cover, HSG Assembly
③ OP60~OP65: Fastening bolt and allowance inspection
④ SUB2-1~SUB2-5: SEAL Ring, O-RING assembly and operation check
⑤ OP70~OP150: Attaching barcode, Oil leak inspection, Oil remove, Package

Losses in Assembly Line

Loss — 15%

Production Loss
(OEE = 85%)

49.2%

28.1%

4.1%

8.3%

10.3%

Equipment failure: not running due a failure of some sort, unplanned stop

Idle and minor stop: a short period of with the stop resolved by the operator **suddenly happen without symptoms**

Scope of the Project

Process defect: defective parts produced *during* stable production

Reduced yield: defective parts produced from startup *until* stable production

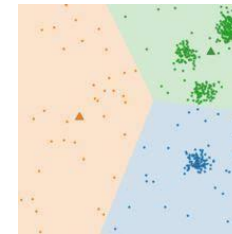**Setup and adjustments:** Changeover, major adjustments, and tooling adjustments

Efforts for Avoiding Equipment Failure

❑ Predictive maintenance for motors and machine tools

Servo motor

Vibration →
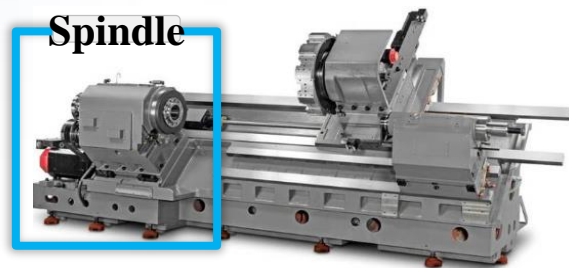
Spindle

Vibration →
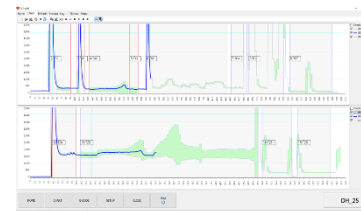
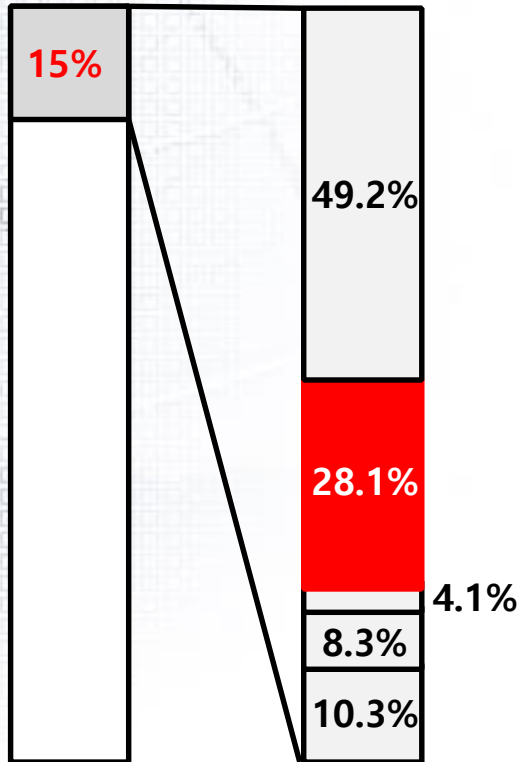## Theses are not Today's Topic

Goal: Reduce Minor Stops


15%

49.2%

28.1%

4.1%

8.3%

10.3%

Definition: Unlike equipment failures, machine is stopped by the machine itself or operators for a time shorter than threshold due to transient errors

Cases of Minor Stops (difficult to measure or expensive)
- Screw jam
- Breakdown of proximity sensors
- Misaligned sensors
- Disconnection of harness

Problems on worksite
(1) Stops without symptoms → Cannot predict when it will happen
(2) Cannot detect them all automatically → Operators cannot recognize
(3) Cannot find reasons without experts → Need time to fix
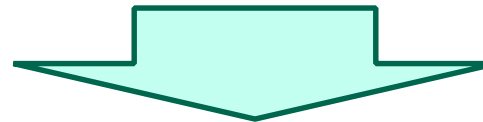
Data-driven Minor Stops detect and Reasoning

POSTECH

Data Exploration

❑ Summary of Data

- 06/2017~07/2017 Data (> 800,000 rows)

- 400 kinds of sensor data

- 3 types of Minor stops on data

SerialNo	Cycle Time	Minor Stop	...	Torque	ScrewLoading	Air Pressure	Temperature	Proximity sensor_1	...
A0001	201	449		6.2	1	12.5	24.3	1	
A0001	201	449		6.2	1	12.5	24.3	1	
A0002	28.9	0		6.2	1	12.4	24.3	1	
A002	28.9	0		6.3	1	12.5	24.3	0	
...				...	...	...	...		

< Example of data>

Problems on Data

SerialNo	Cycle Time	Minor Stop	…	Torque	ScrewLoading	Air Pressure	Temperature	Proximity sensor_1	…
A0001	201	449		6.2	1	12.5	24.3	1	
A0001	201	449		6.2	1	12.5	24.3	1	
A0002	28.9	0		6.2	1	12.4	24.3	1	
A0002	28.9	0							
…									
A1210	**177**	**0**							
A1210	**177**	**0**		6.3				0	
A1210	**177**	**0**		6.2				1	
A1210	**177**	**0**		6.3	1	12.5	24.3	0	
A1211	28.9	0		6.2	1	12.4	24.3	1	
A1211	28.9	0		6.3	1	12.5	24.3	0	

- About 32% of data longer than Cycle time of the line (29s)
- Don't have error codes of minor stops
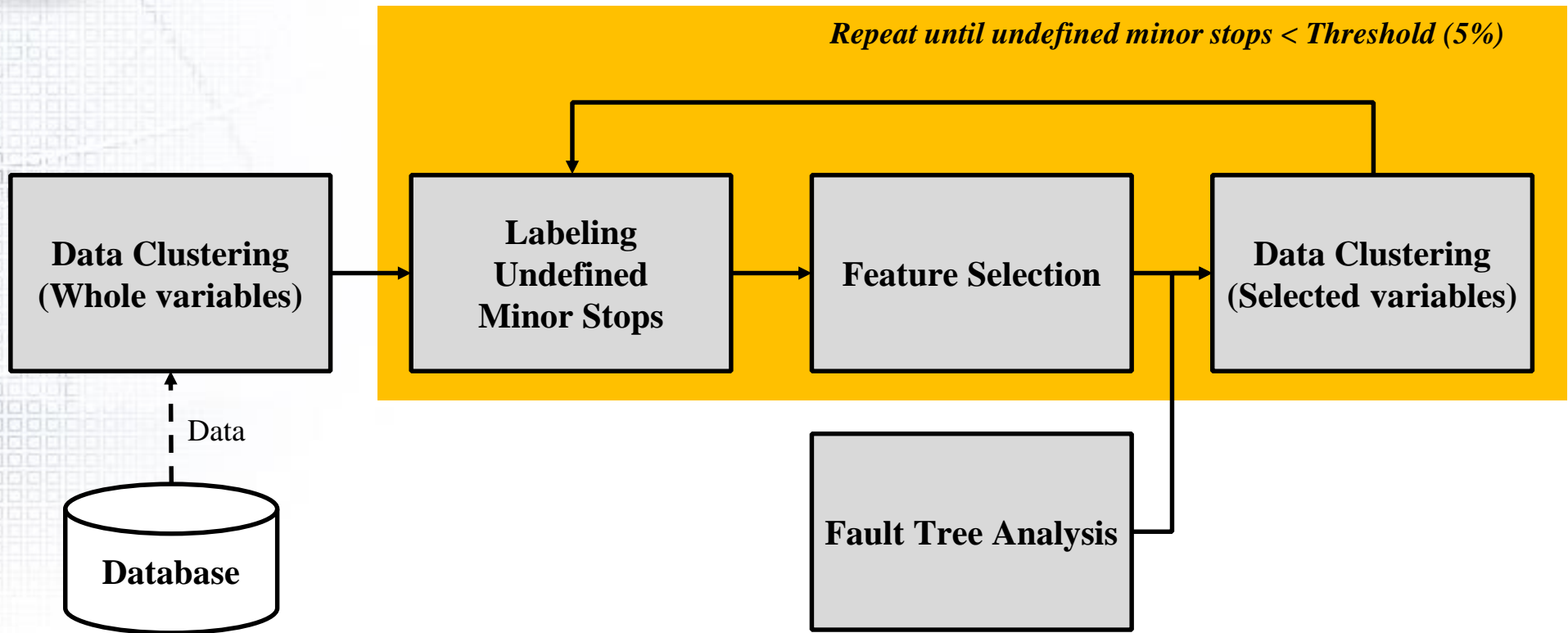
Low Reliability of the Data

Define that processing time **more than 30s is Minor Stop**
Divide undefined minor stop data from normal condition data based on clustering

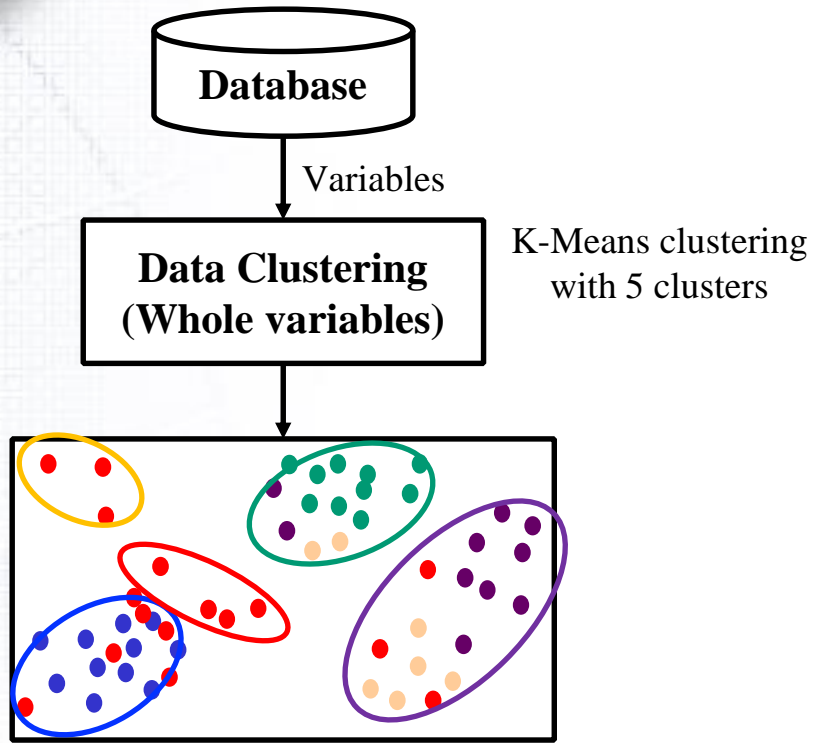*Data-driven diagnosis model for detecting and reasoning minor stops*

Modeling Approach

Data Clustering with Whole Data and Labeling undefined Minor Stops

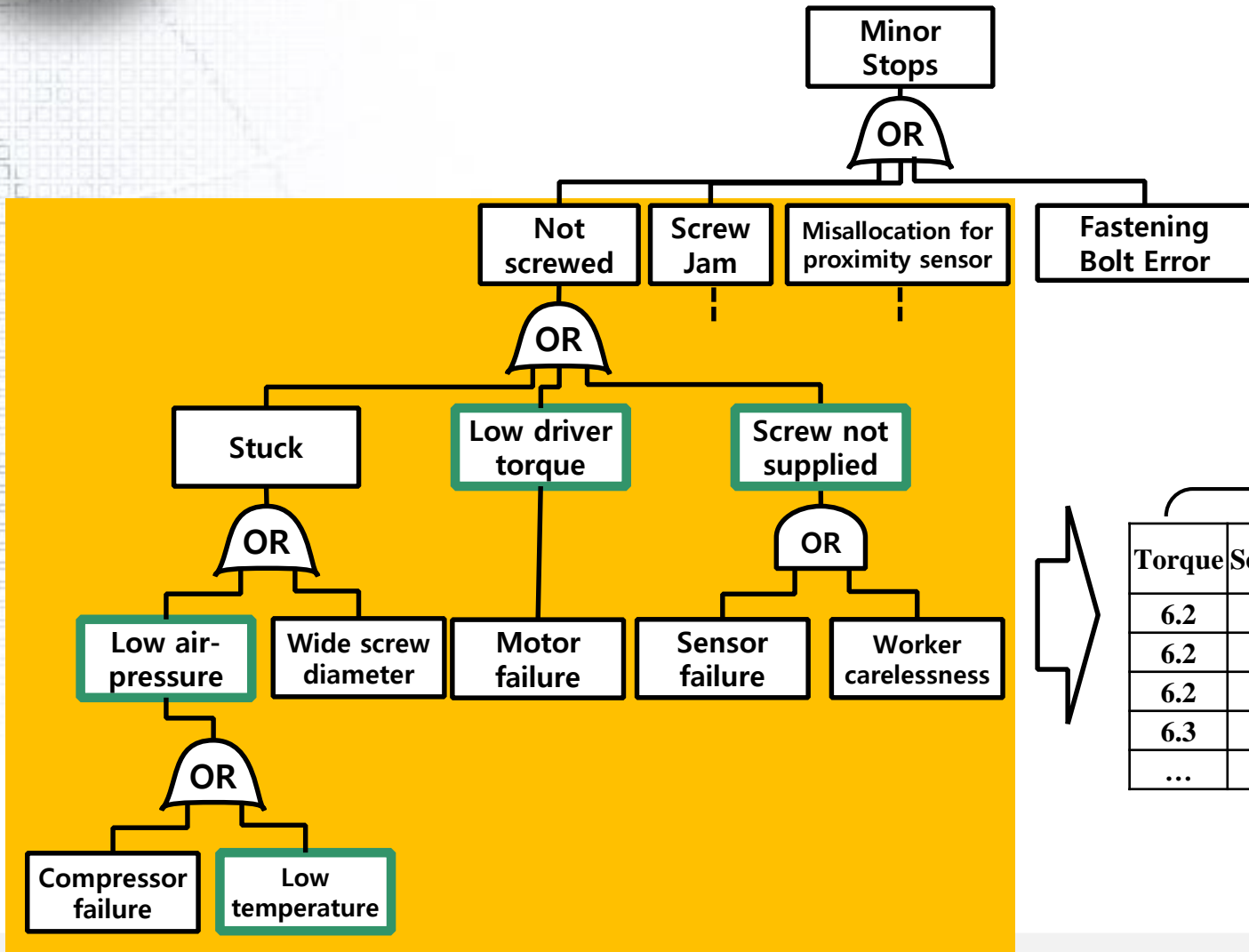**Database**

Variables

Data Clustering (Whole variables)

K-Means clustering with 5 clusters

Labeling as minor stops

- Normal condition
- undefined minor stops
- Labeled minor stop #1
- Labeled minor stop #1
- Labeled minor stop #1

POSTECH
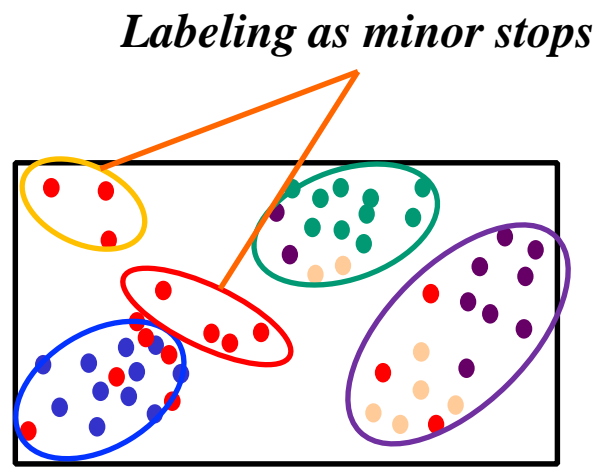
Fault Tree Analysis

Develop fault tree with worksite experts to find meaningful variables

Fixed variables

Torque	ScrewLoading	Air Pressure	Temperature	...
6.2	1	12.5	24.3	
6.2	1	12.5	24.3	
6.2	1	12.4	24.3	
6.3	1	12.5	24.3	
...	...	...	...	

Feature Extraction

- **Variables from FTA are subjective**
- **Select meaningful and objective variables using ANOVA**

Fault Code	X0B	X0F	X10	X11	X12	X13	X14	...
0	11.7	4.3	0	0	1	1	0	
0	11.6	4.4	0	0	1	1	0	
0	11.6	4.3	0	0	1	1	0	
0	11.7	4.4	0	0	1	1	0	
321	11.6	3.9	1	1	1	1	0	
331	11.7	3.9	1	1	1	1	0	
0	11.7	4.4	1	1	1	1	0	
0	11.6	4.3	1	1	1	1	0	
0	11.7	4.4	1	1	1	1	0	
181	11.6	4.3	1	0	1	1	0	
0	11.7	4.4	1	1	1	1	0	
0	11.6	4.3	1	1	1	1	0	
0	11.6	4.4	1	1	1	1	0	
0	11.6	4.3	1	1	1	1	0	
0	11.6	4.4	1	1	1	1	0	
...	...	...	...	...	...	...	...	

Variables	p-value
X0B	0.00
X0F	0.01
X10	0.00
X12	0.13
X13	0.35
X14	0.00
...	...

Variables	p-value
X0B	0.00
X0F	0.01
X10	0.00
X12	0.13
X13	0.35
X14	0.00
...	...

Data Clustering with Selected Variables and Labeling undefined Minor Stops
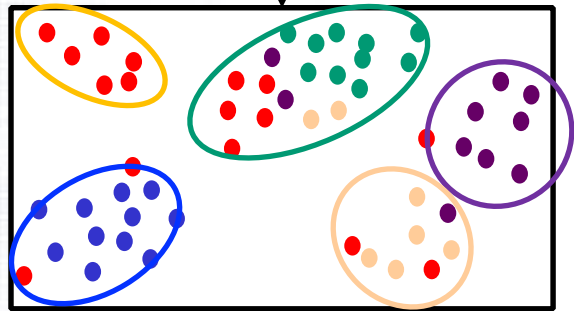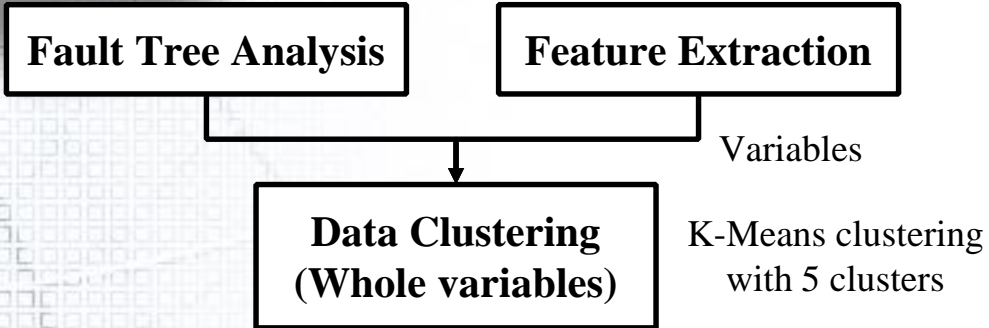
Fault Tree Analysis

Feature Extraction

Variables

Data Clustering (Whole variables)

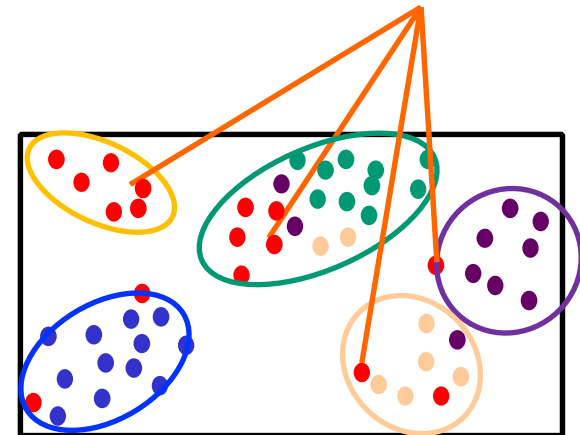K-Means clustering with 5 clusters

Labeling as minor stops

- Normal condition
- undefined minor stops
- Labeled minor stop #1
- Labeled minor stop #1
- Labeled minor stop #1

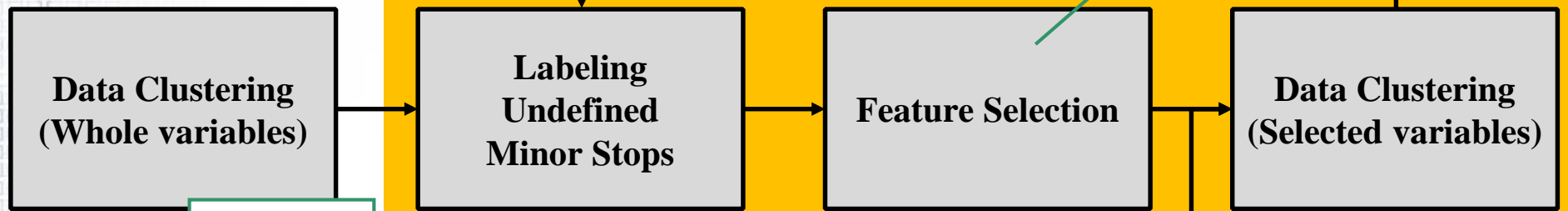$$\frac{Undefined\ minor\ stops}{Normal\ condition\ data + Undefined\ minor\ stops} = 32\% \rightarrow 14.8\%$$

Result

$$\frac{Undefined\ minor\ stops}{Normal\ condition\ data + Undefined\ minor\ stops} = 32\% \rightarrow 14.8\% \rightarrow 5.5\% \rightarrow 4.1\%$$

Feature = Candidate for causing minor stops

Repeat until undefined minor stops < Threshold (5%)

| Data Clustering (Whole variables) | Labeling Undefined Minor Stops | Feature Selection | Data Clustering (Selected variables) |

400 variables

Data

Database

Fault Tree Analysis

43 Fixed variables from worksite experts

❑ Importance of data reliability

   ▪ Developed model without considering cycle time → low recall and precision

   ▪ Sensor data is not always accurate

   ▪ Machines do not make a mistake but humans do

❑ Need more researches for reducing minor stops

   ▪ Many researches for equipment failures but a few for minor stops

   ▪ Minor stops happen more frequently than equipment failures

   ▪ Generally repaired by operators and is not easy to identify and solve problems
     (Office staff couldn't identify minor stops for 2 hours during strike)

❑ Convert tacit knowledge into explicit knowledge

- Reducing feature space of data from Fault Tree Analysis

- Knowledge should be part of the model
  (Data talks. But can operators.)

❑ Increasing productivity is the key because of the regulation

- Korean new labor regulation = no more than 52 hours work in a week

- Cost ↑, Production capacity ↓

- Data analytics will play an important role for productivity