

Techtonic 2018

-
Thu . Nov 15

-
SAMSUNG SDS Tower
West Campus B1F
Magellan Hall /Pascal Hall



이상감지가 쉬워진다

나는 어느 설비가
고장날 지 알고 있다!

삼성SDS 김재훈 프로



- 이상감지? 수명예측?
- 데이터 준비, 시행착오를 줄이자!
- 본격 이상감지
- 수명예측
- 이상감지 모니터링 시스템

이상감지가 쉬워진다!

이상감지? 수명예측?

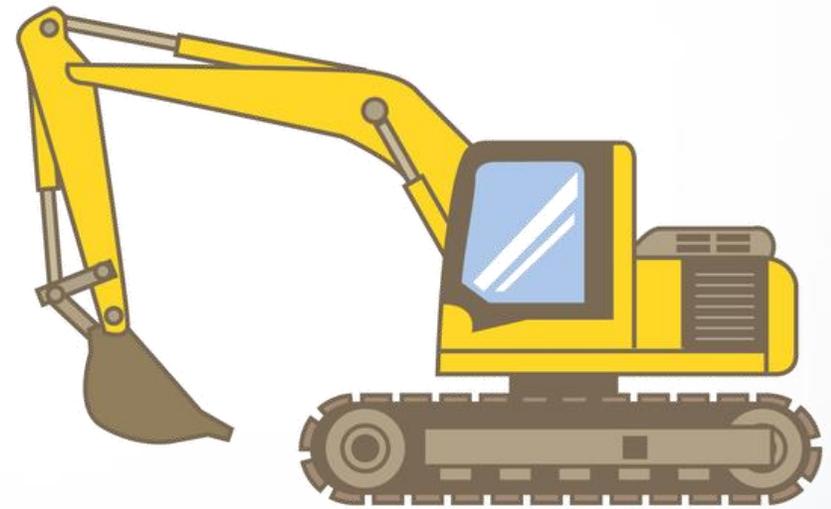
예지보전 (Preventive maintenance)

고장 없이, 효율적으로 기기를 사용하기 위한 기술

이상감지 기술: 시스템 고장 예방



수명예측 기술: 시스템 최대 활용, 가치 발굴



이상감지 사례 #1: 어린이집 아이

내 아이가 달라졌어요!

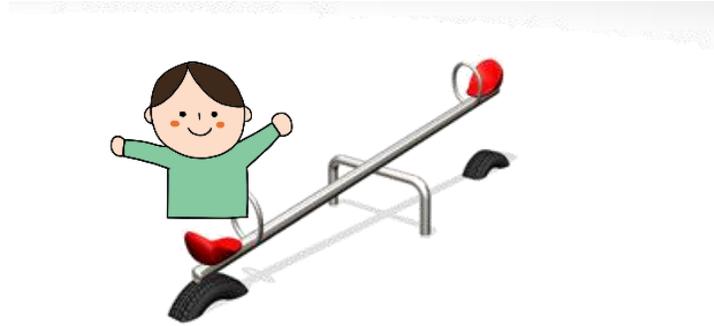


아이의 이상상태를 감지하는 문제
아이  의 행동 패턴을 관찰하자.

시간	장소
9시~12시	도서관
12시~1시	점심식사
1시~3시	시소
3시~5시	미끄럼틀

이상감지 사례 #1: 어린이집 아이

점심밥도 안먹고 시소를 탄다?

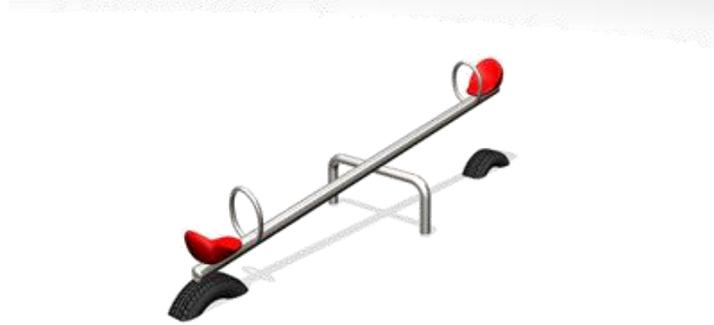


아이의 이상상태를 감지하는 문제
아이  의 행동 패턴을 관찰하자.

시간	장소
9시~12시	도서관
12시~1시	시소 이상!
1시~3시	시소
3시~5시	미끄럼틀

이상감지 사례 #1: 어린이집 아이

활동 영역이 빈도수가 낮은 영역이면 이상으로 진단 가능하다.



아이의 이상상태를 감지하는 문제
아이  의 행동 패턴을 관찰하자.

시간	장소
9시~12시	도서관
12시~1시	점심식사
1시~3시	시소
3시~5시	커피집 이상!

이상감지 사례 #2: 신용카드 결제

결제 장소가 빈도수가 낮다면 이상으로 진단 가능하다.



- 한국에서 쇼핑하던 사람이 갑자기 미국에서 쇼핑?
- 결제한 장소와 시간을 기록해두고 빈도를 계산
- 위치기반 이상감지로도 대략적인 감지 가능

수명예측 사례: 나는 몇살까지 일할 수 있을까?

다양한 정보 활용하여 정확도 극대화



초년생



시간이 지나고



...



사전정보 정년: 60세

이후, 근무실적, 진료기록 등등
참고할 수 있는 정보들을 활용하자.

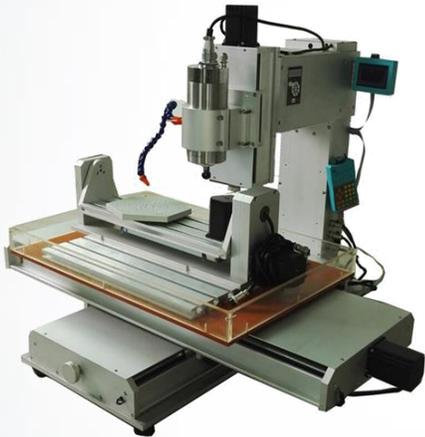
병력이 있다: 정년 ↓

운동을 열심히 한다: 정년 ↑

야근을 자주한다: 정년 ↓

설비에 대한 이상감지, 수명예측

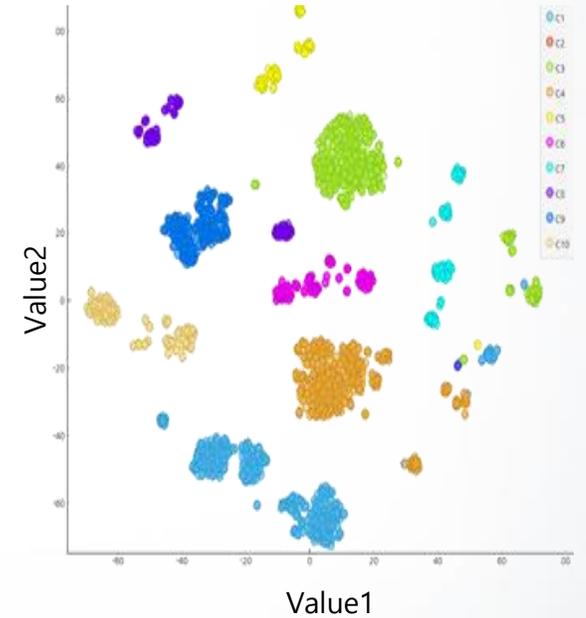
다차원 좌표공간의 위치로 이상감지 및 수명예측을 수행



▶
전처리

시간 (초)	팁 회전속도 (RPM)	온도 (섭씨)	팁에 가해지는 로드 (N*m)
0	75	40	115
1	-	41	116
2	74	40	-
3	75	40	115
...
48	74	41	115

▶
전처리 후 데이터 예시



▶
다차원 공간 plot

이상감지가 쉬워진다!

데이터 준비, 시행착오를 줄이자!

통합데이터베이스 생성

검사이력을 하나의 데이터베이스에 통합

A_1

파라미터	Value
P1	30
P2	0
P3	-20
P3	0
...	...

A_2

파라미터	Value
P1	30
P2	0
P3	-20
P3	0
...	...

⋮

A_4

파라미터	Value
P1	30
P2	0
P3	-20
P3	0
...	...



시간	P1	P2	P3	P4	P5	P6
A_1	+1	-1	0	0		
A_2	1	0	+1	0		
A_3	0	-1	-1	0		
A_4	+1	-1	+1	0		
...		

데이터 준비를 시작해볼까?

전처리 단계에서 만나는 다양한 문제점

저장 데이터의
관리 미흡

결측치

데이터
저장 형태

데이터
세부지식 부재

✔ 통일된 로그기록
포맷 부재

✔ 센서의 오작동 및
기타 오류로 인한 결측치

✔ 상이한 데이터소스
통합 필요

✔ 데이터의 의미
파악 어려움

제일 난감한 상황

임의적인 로그기록 방식, 통일된 포맷 부재

로그 #1

Model : G+
Serial : 4U80RS
DATETIME ≡ 2018-11-08
16:00:00
Lookup : PASS
breakup test : 10.0

로그 #2

Model : G+
Serial : .3125679r29r
DATETIME : 2018-11-08
16:00:00
Look : PASS
Pason test : 10.0

✔ 띄어쓰기 개수

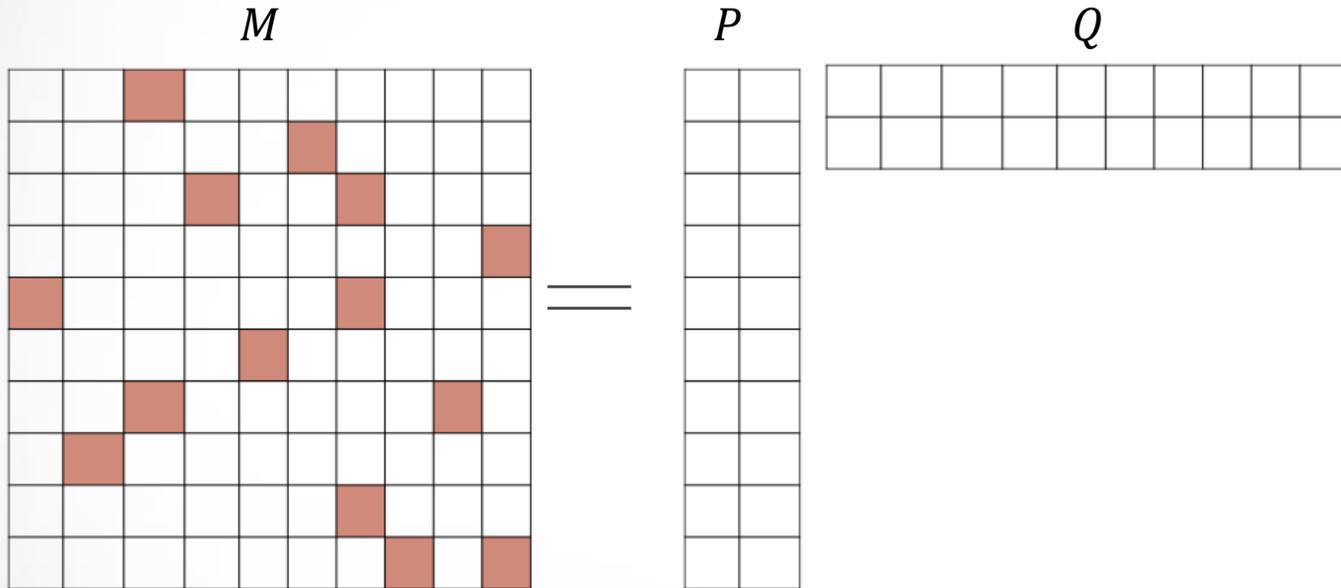
✔ 구분자

✔ 첨자

✔ 같은 센서, 다른 센서명

결측치 문제

행렬 완성 알고리즘으로 빈값 추정



- Low rank matrix factorization 으로 빈 값을 채워넣을 수 있다.

시간	P1	P2	P3	P4	P5	P6
A_1	+1	-1	NaN	0		
A_2	1	0	+1	0		
A_3	0	NaN	-1	0		
A_4	+1	-1	NaN	0		
...		

$M \approx PQ$ 

시간	P1	P2	P3	P4	P5	P6
A_1	+1	-1	0	0		
A_2	1	0	+1	0		
A_3	0	-1	-1	0		
A_4	+1	-1	+1	0		
...		

상이한 데이터소스 통합

데이터 기록 방식의 차이 존재

시간 (초)	팁 회전속도 (RPM)	온도 (섭씨)	팁에 가해지는 로드 (N*m)
0	75	40	115
1	-	41	116
2	74	40	-
3	75	40	115
...
48	74	41	115



시간 (초)	가동중인가?
0	no
1	yes
2	yes
3	yes
4	yes
5	no

기록번호	가동시작시점 (초)	가동중단시점 (초)
0	1	4
1	10	48



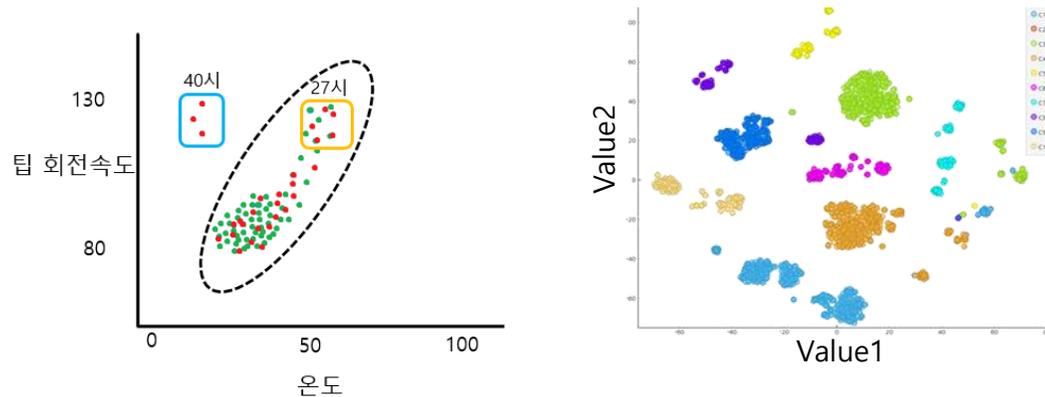
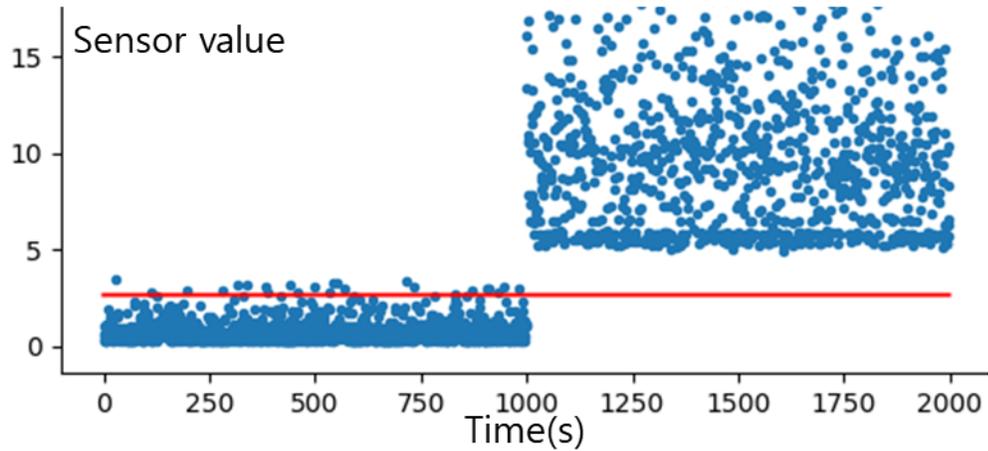
**다른 기록방식
통합작업 필요.**

이상감지가 쉬워진다!

본격 이상감지

이상감지의 목표

다양한 이상상태 정의에 대해 고찰



과거에 발생하지 않고 최근에 발생한 사건

갑자기 빈번하게 발생한 사건

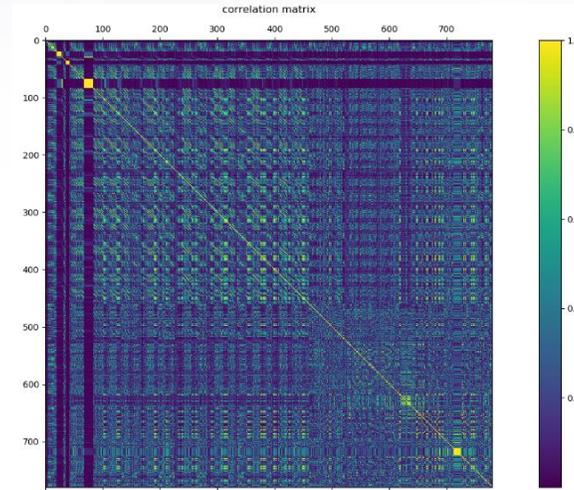
논리적으로 함께 발생할 수 없는 사건

평소의 추이에서 크게 벗어난 극단값

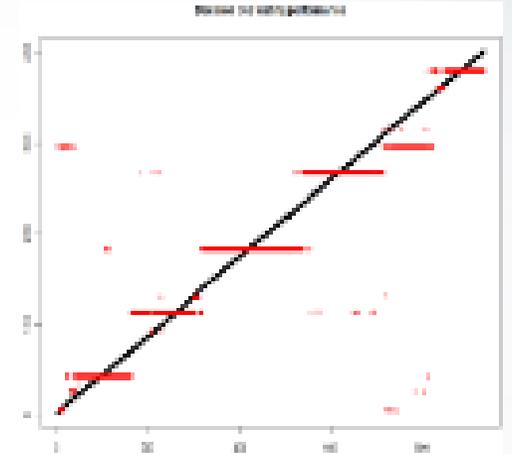
데이터의 모양 파악

EDA (Exploratory Data Analysis) 수행

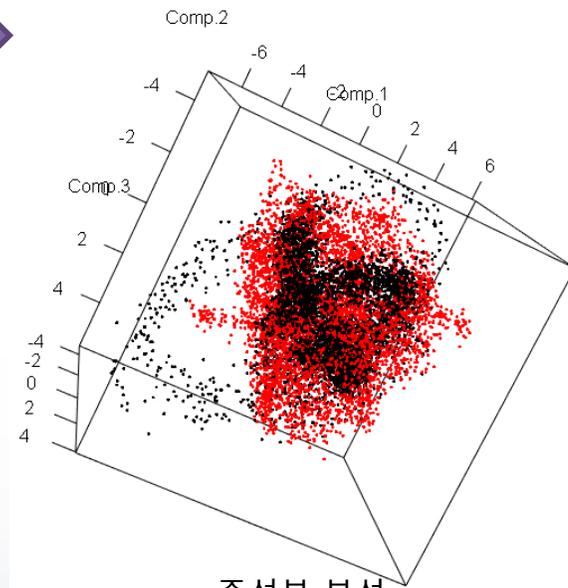
시간 (초)	팁 회전속도 (RPM)	온도 (섭씨)	팁에 가해지는 로드 (N*m)
0	75	40	115
1	-	41	116
2	74	40	-
3	75	40	115
...
48	74	41	115



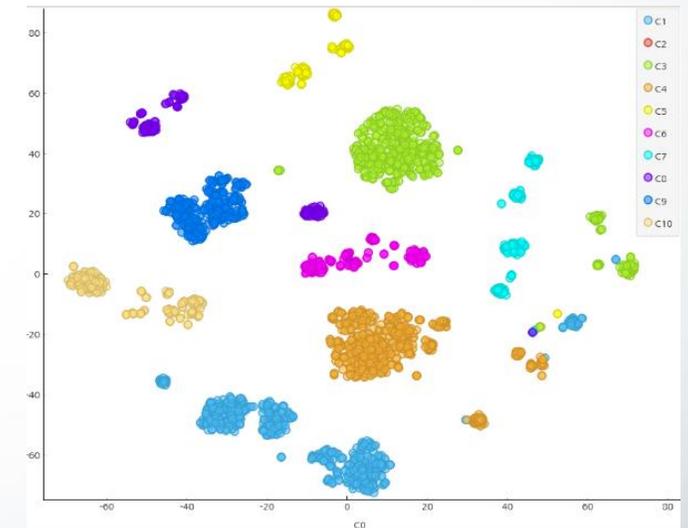
상관관계 분석



잔차그래프



주성분 분석



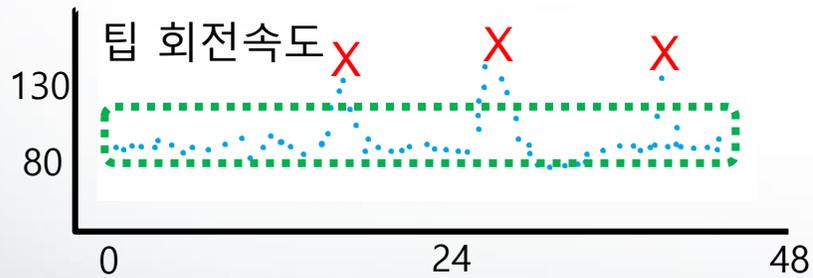
t-SNE 분석

다변량 분석 기반 이상감지 & 모니터링

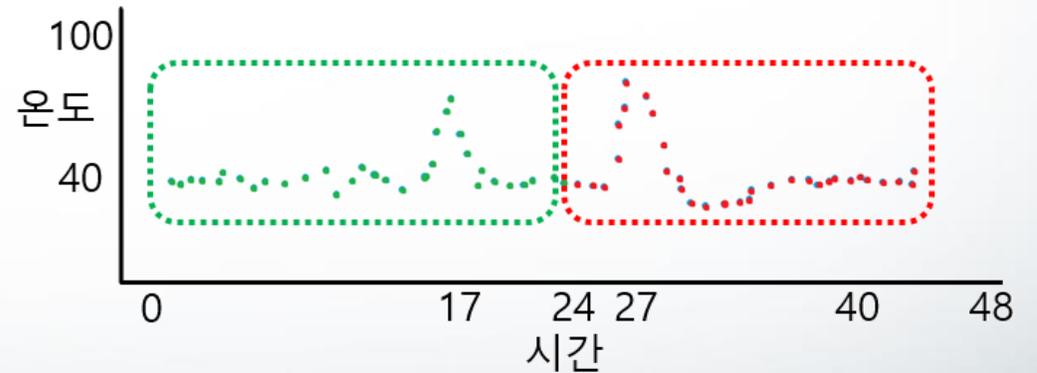
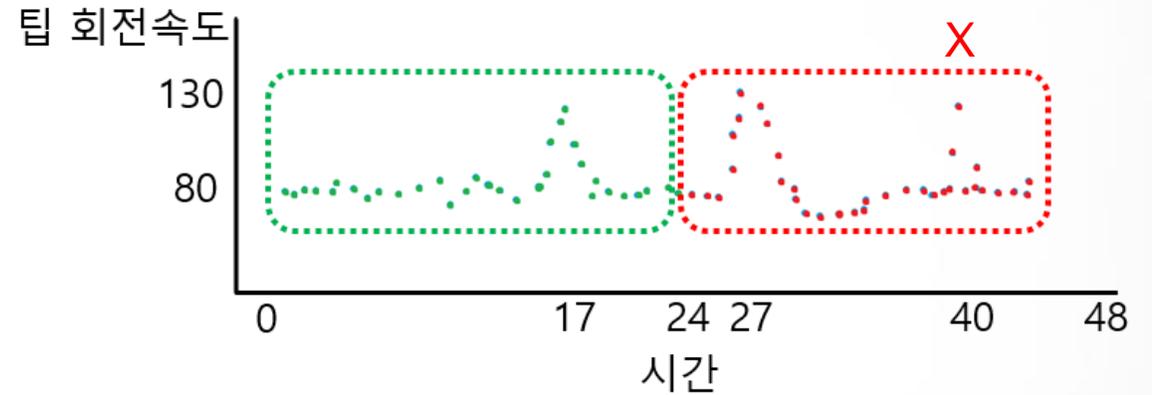
다변량 분석으로 Robust한 분석 수행

단변량 관리한계 기반 분석

시간 (초)	팁 회전속도 (RPM)	온도 (섭씨)	팁에 가해지는 로드 (N*m)
0	75	40	115
1	75	41	116
...
48	74	41	115

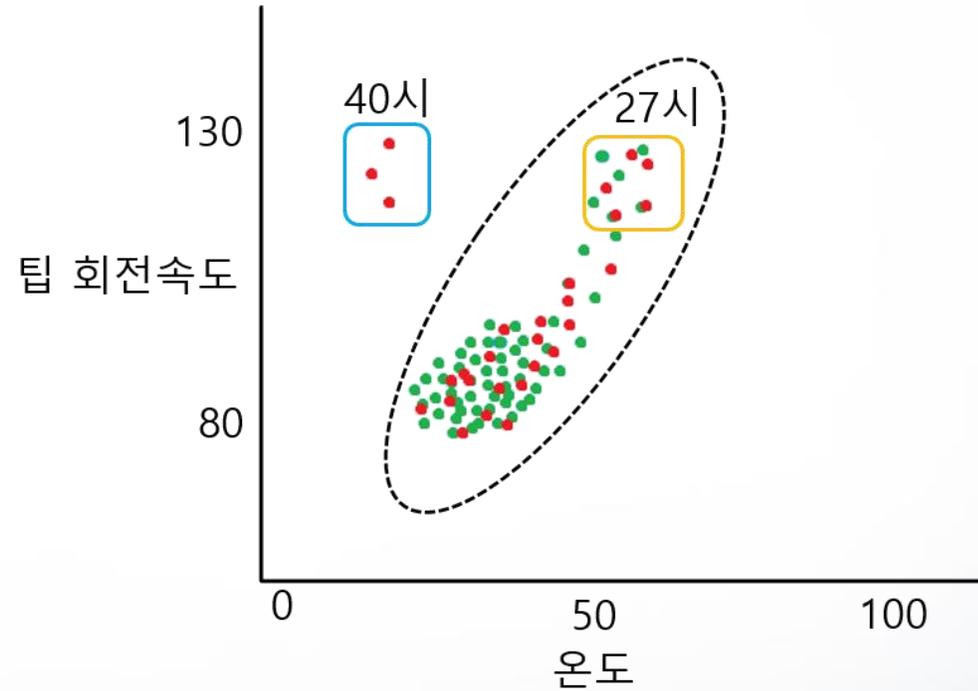
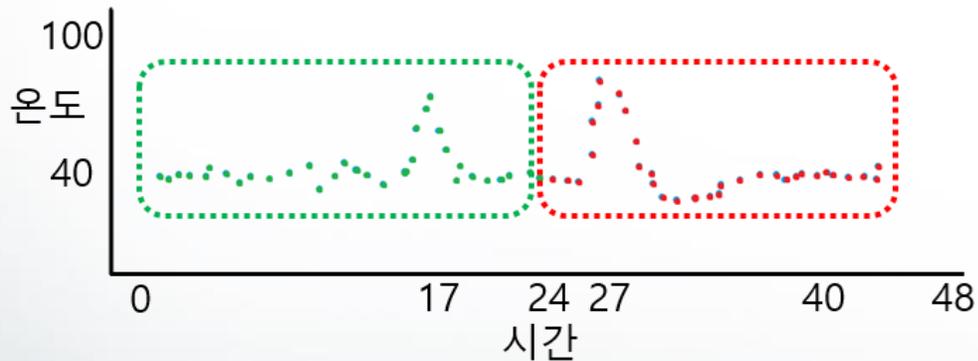
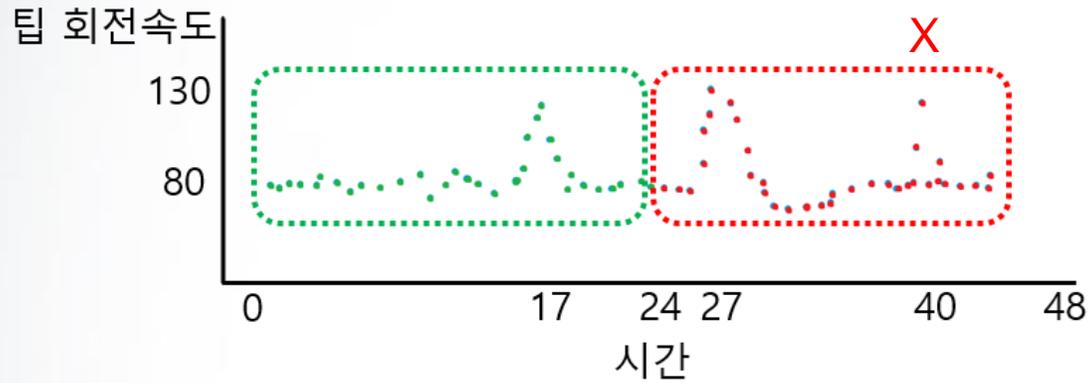


다변량 기반 분석



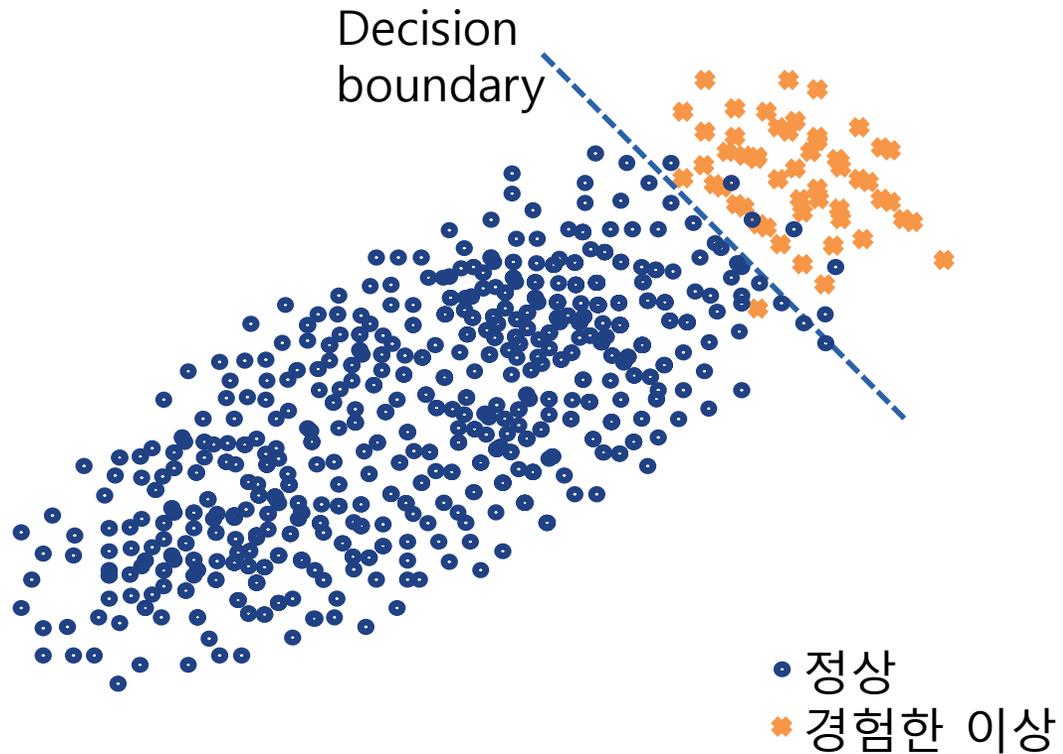
다변량 분석 기반 이상감지 & 모니터링

다변량 분석으로 Robust한 분석 수행



이상감지, 지도학습으로 될까?

요즘 핫한 AI 지도학습, 여기에도 적용 가능하겠군요?

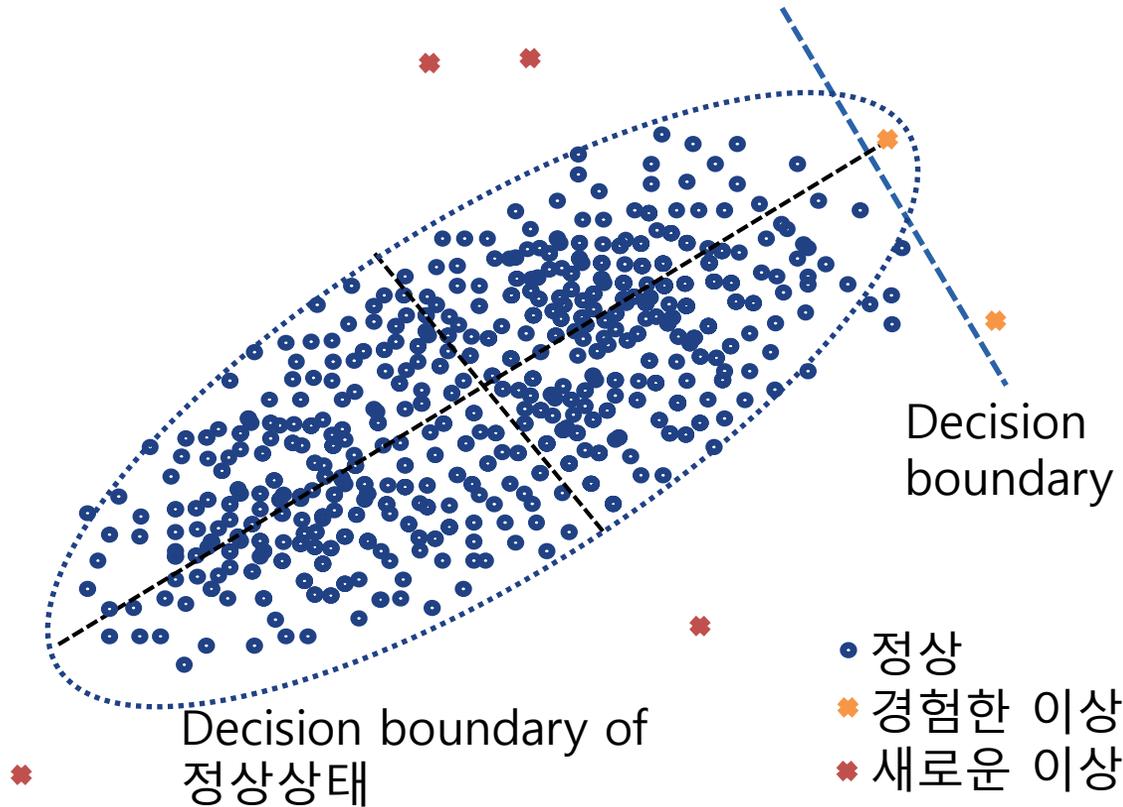


정상과 이상 라벨을 달자.

분류문제로 보고 지도학습!

이상감지, 지도학습으로 될까?

요즘 핫한 AI 지도학습, 여기에도 적용 가능하겠군요?



실제 상황은 이렇습니다.

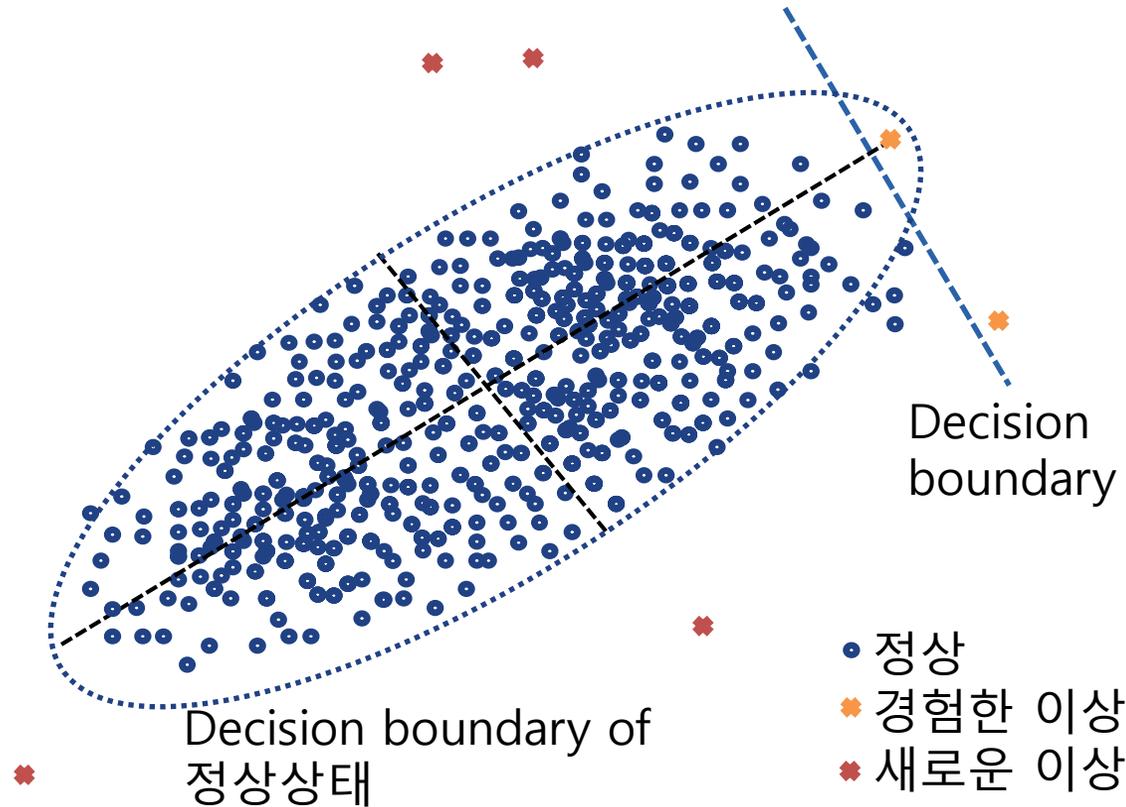
이상상태의 데이터가 극히 적을 뿐더러, 트레이닝 데이터에 없는 새로운 이상상태가 출현한다.

해결책은?

정상이 아닌건 이상이야

정상이 아닌 것 = 이상

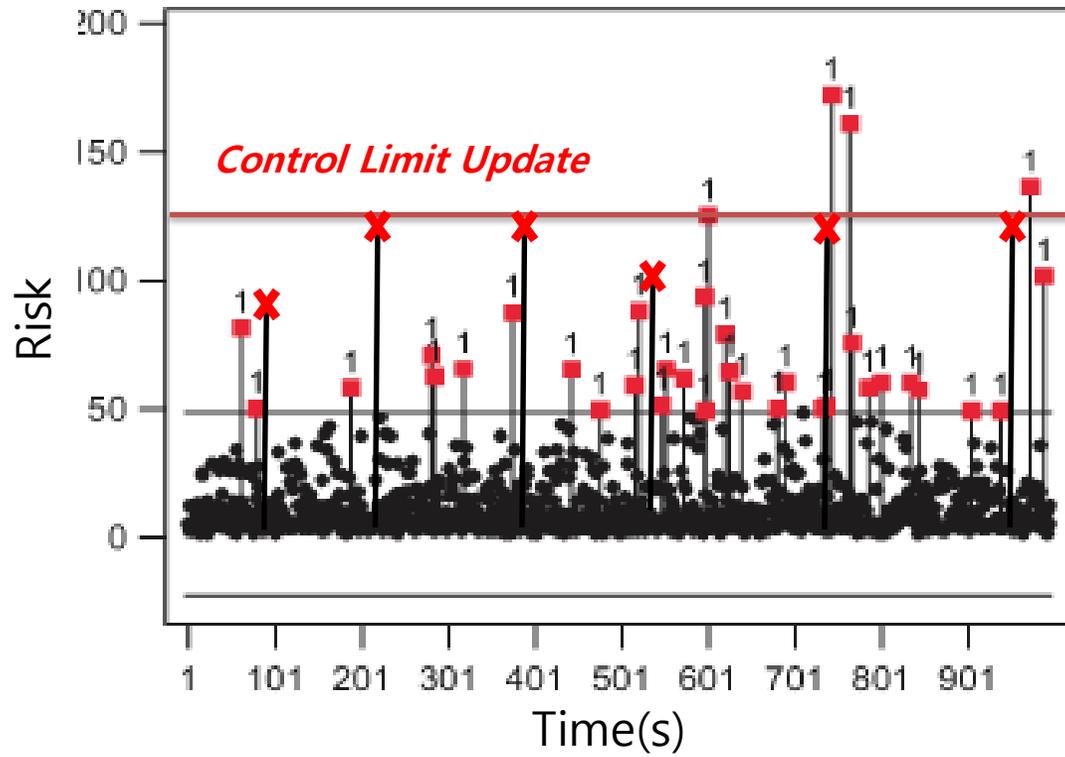
정상을 학습하여 이상을 구분



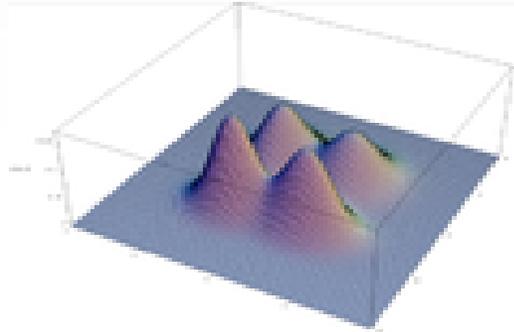
거리를 정규화하면 정상/이상을 구분하는 척도로 활용 가능

- Mahalanobis 거리
- Autoencoder 복원오차
- MSET 알고리즘

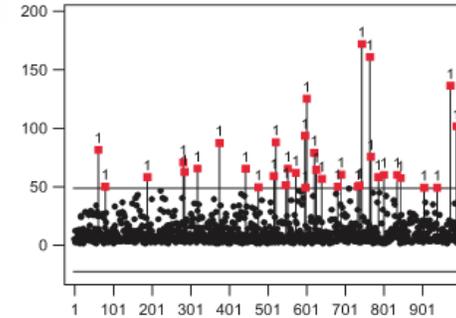
거짓경보 리스크



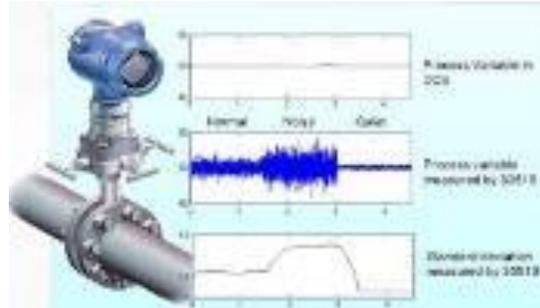
다양한 거짓경보 이유들



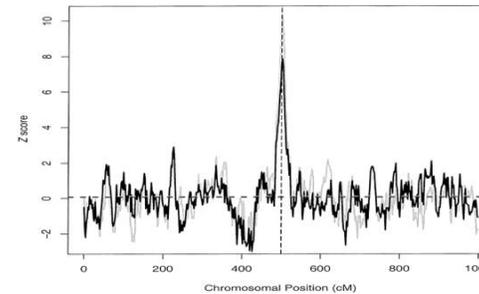
Nonnormal distribution



Inaccurate control limit



Change of process condition

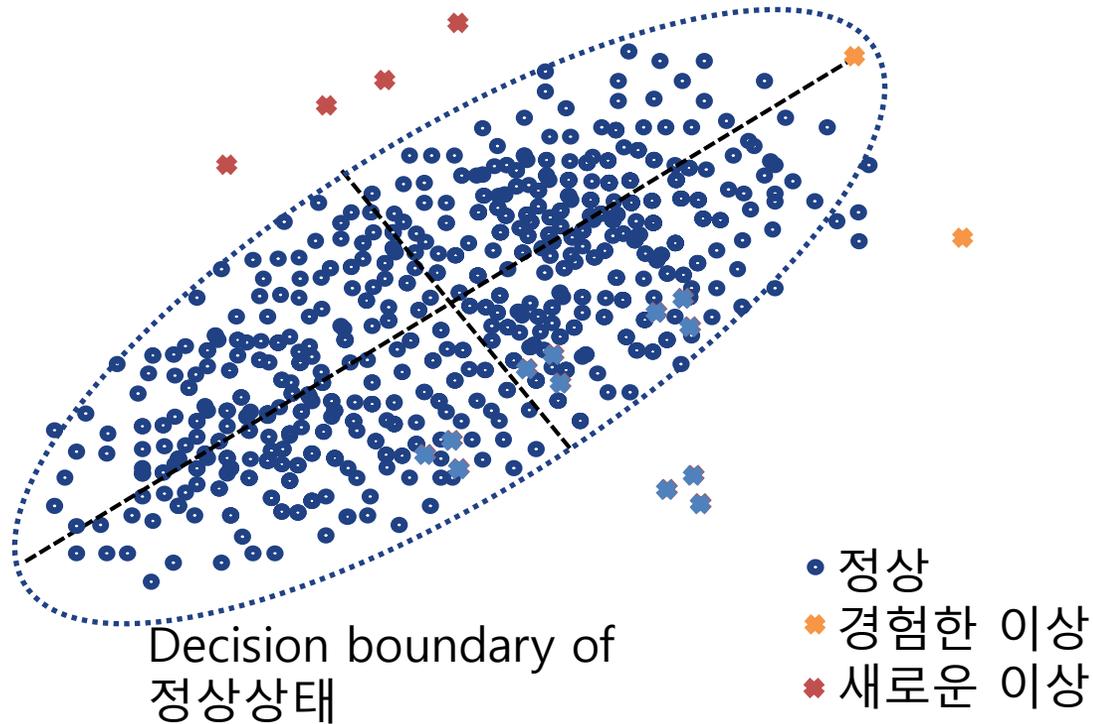


Incorrect statistic

- ✓ Type I error(α) 와 Type II error(β) 는 서로 trade off 관계 : 민감하게 문제를 찾아낼수록 false alarm 은 상승
- ✓ 대부분의 실제 모니터링 현장에서는 매우 높은 수준의 confidence threshold 채택 (매우 둔감)
- ✓ 일관적인 process 이상에 대해 잡아내는 것이 더 중요함

거짓경보 제어 알고리즘

정상 도중 잠깐의 이상을 정상으로 판정

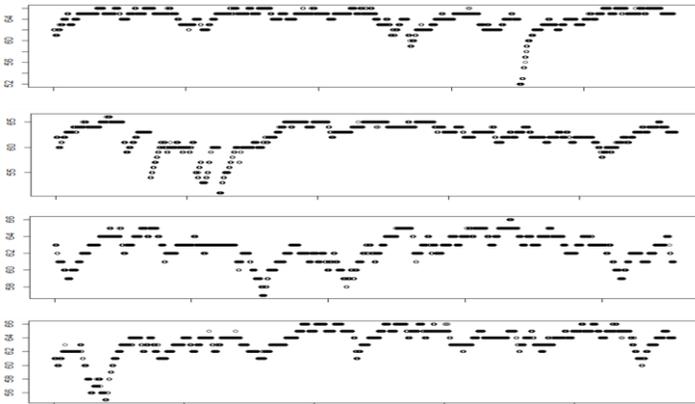
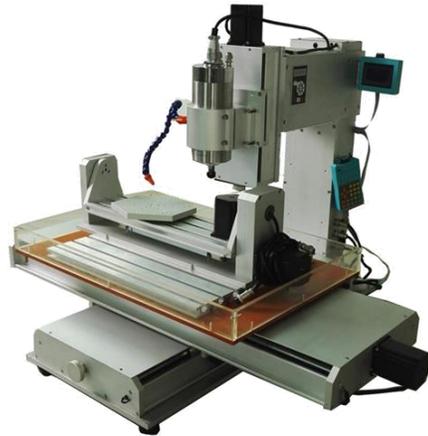


관찰된 때의 센서값만을 참조하지 말고, 이전 값도 참조하여 정상과 이상을 구분하도록 한다.

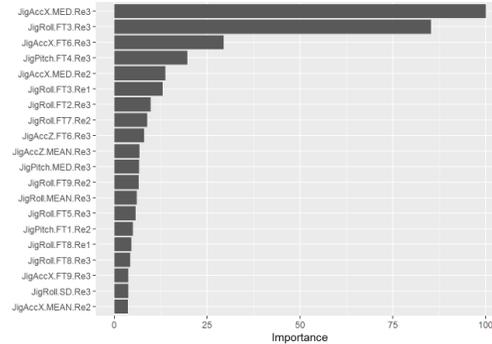
정상 테두리 바깥에 있을지라도 비슷한 시기의 값들이 압도적으로 정상 테두리 안에 있다면 정상으로 판정!

업종별 주요 분석사례

제조, IoT, 헬스 분야 등 다양한 레퍼런스 확보



RF-based Variable Importance



- ✓ CNC로그이벤트
- ✓ 공정 제조설비
- ✓ 상세 부품 (베어링 등)
- ✓ 의료설비 (MRI 등)

이상감지가 쉬워진다!

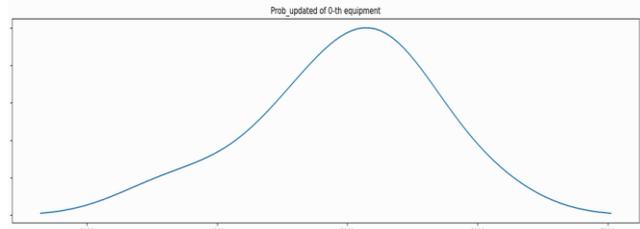
수명예측

수명/신뢰성 분석 로직

쌓인 데이터 이용 대략적인 수명 추정 후
베イズ 정리 이용 수명정보 및 추천 점검시점 실시간 업데이트



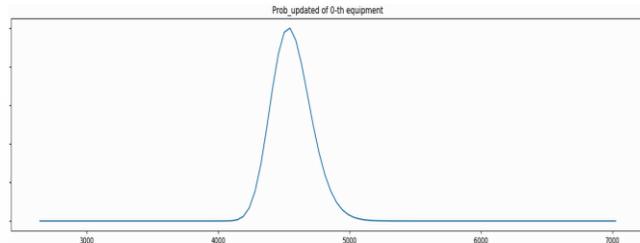
고장이벤트
데이터베이스



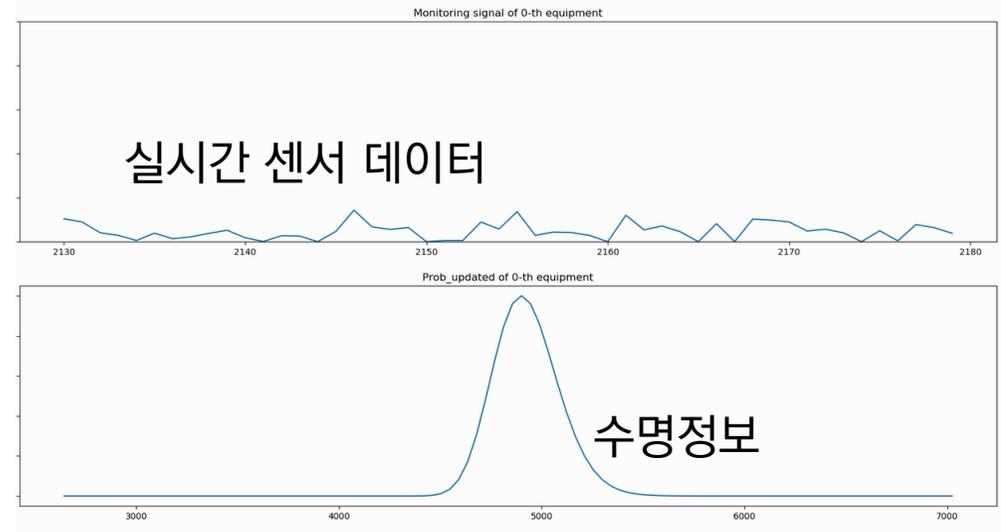
$P(T)$ 학습 (신뢰성 분석)



고장 장치
센서 데이터베이스



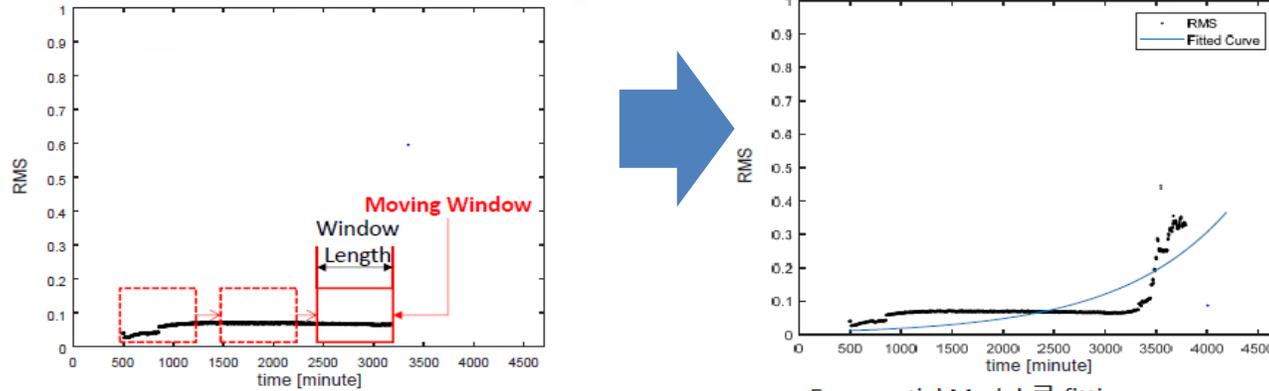
$P(s|T)$ 학습



$$P(T|s) = \frac{P(s|T)P(T)}{P(s)}$$

베어링 수명 예측 사례

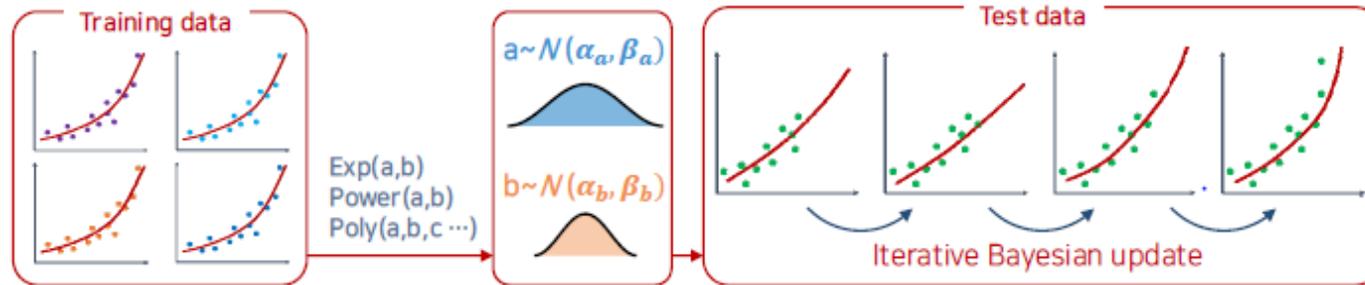
RUL Prediction : Sequential Pattern Prediction with Bayesian Update [with SNU]



• Exponential Model 로 fitting

많은 Training data를 이용하여 Statistical 파라미터 도출

Bayesian update process 로 새로운 데이터에 대한 수명거동 예측

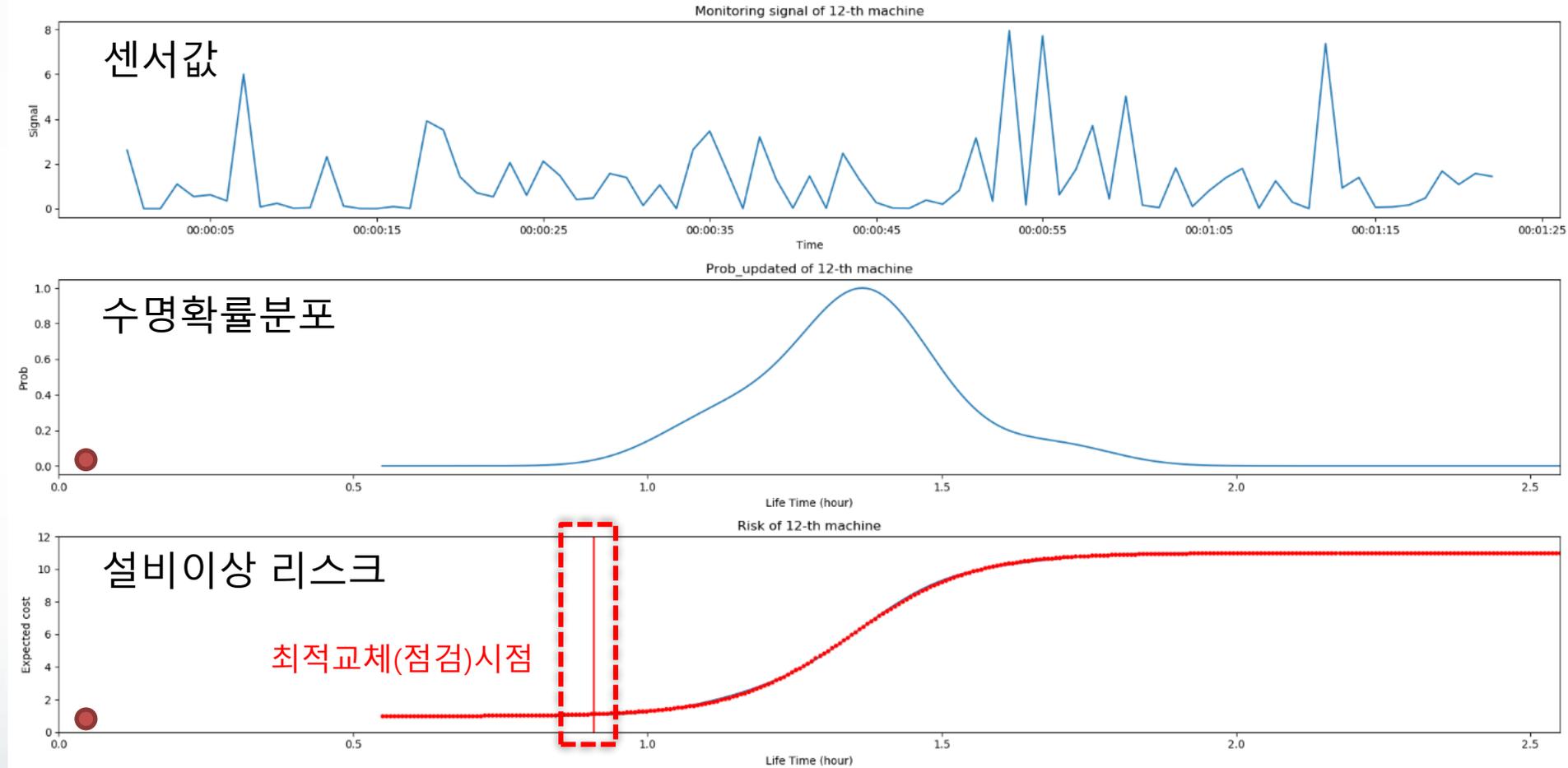


회전체 진동 RMS 진행 패턴 학습 > 초기 수명 모델 설정

Moving window를 통해 관측, 기대하던 모델과 실제값이 상이할 경우, 베이지안 업데이트(ex. 분포 모델, 파라미터 등)

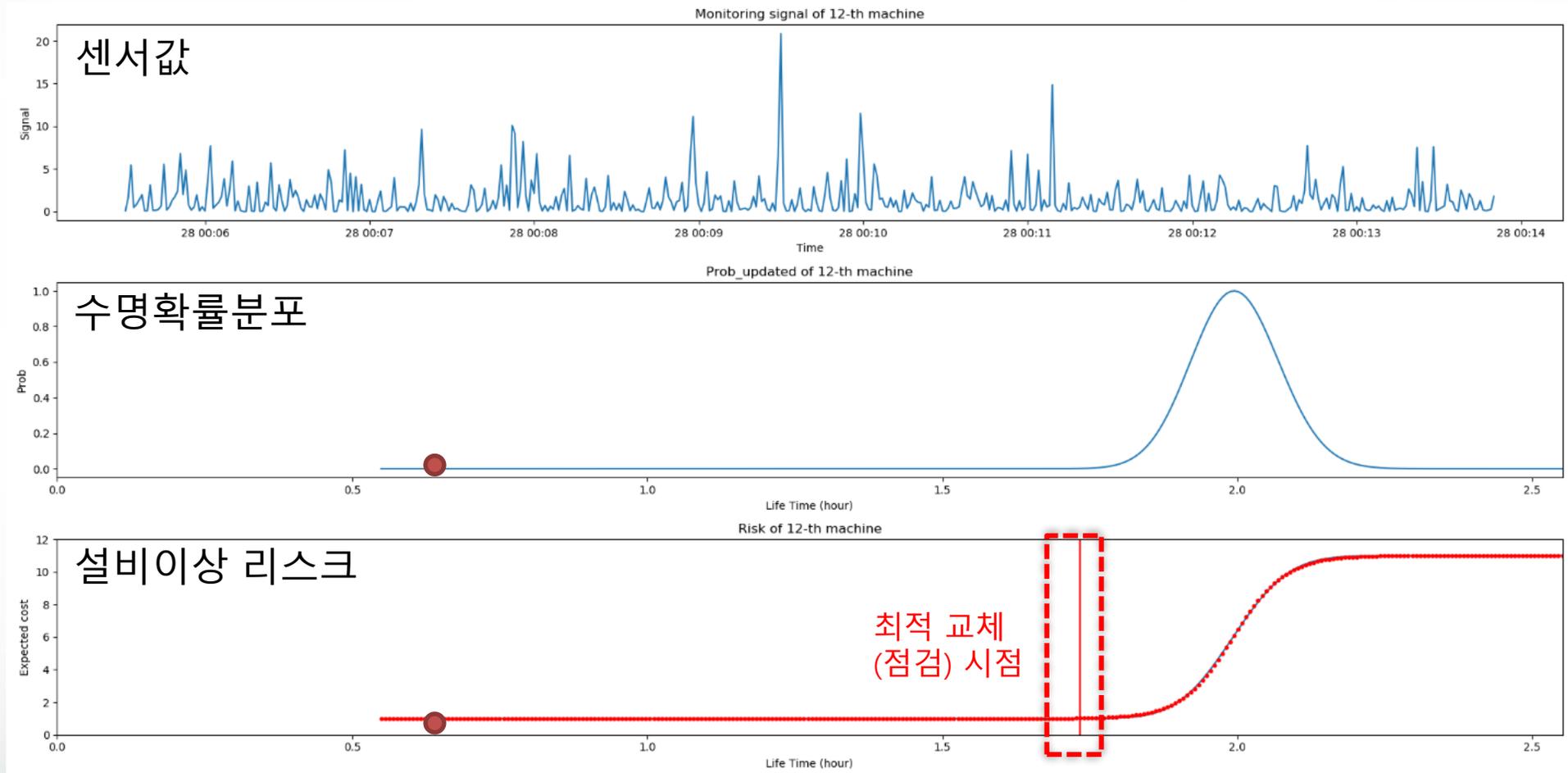
수명예측 실시예

가동 시작시 (5월 28일 자정에 시작, 1시간정도의 수명으로 세팅 후 시뮬레이션) – 센서 데이터가 거의 없다.



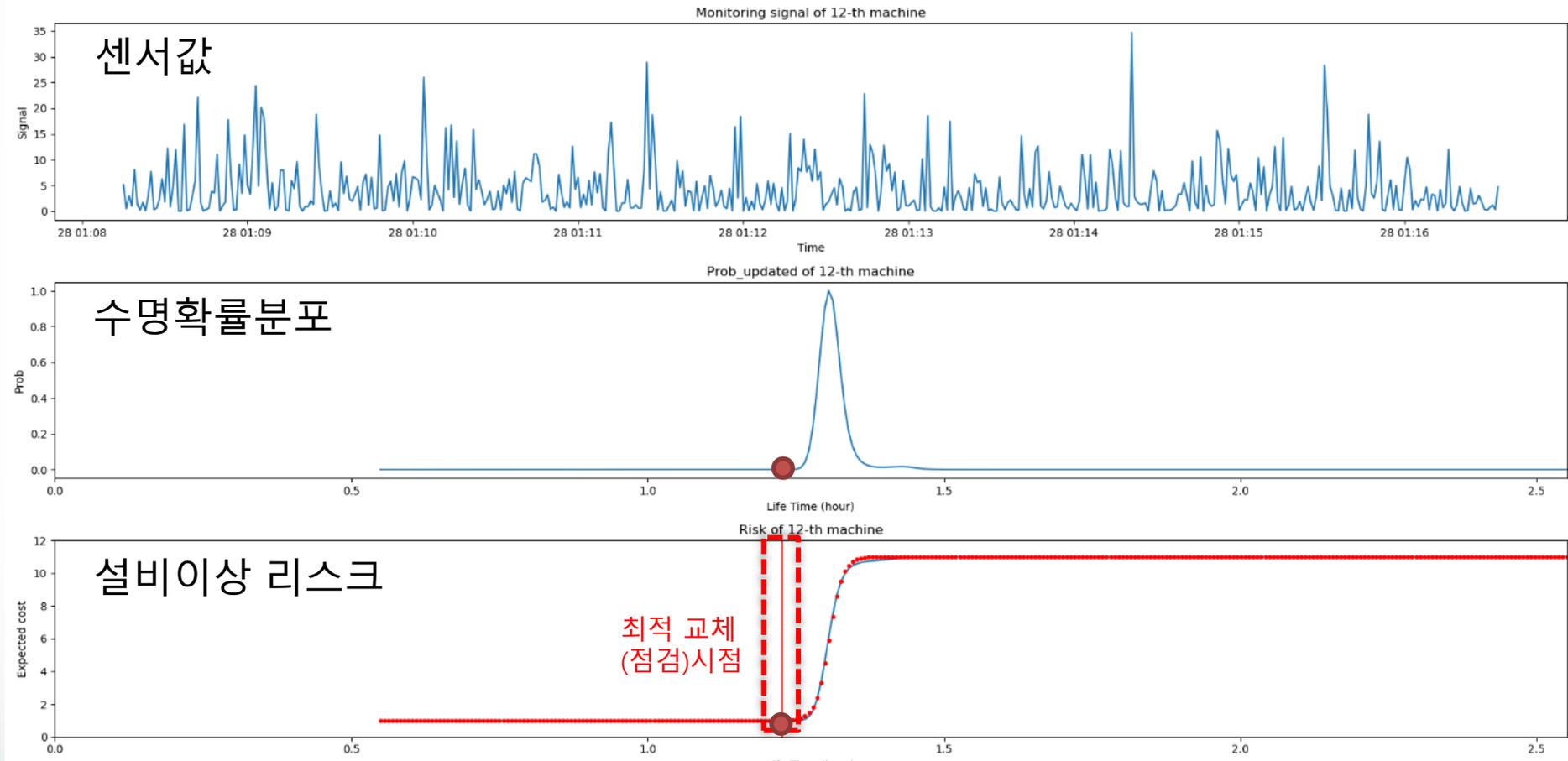
수명예측 실시예

가동 중 - 베이지안 업데이트 적용



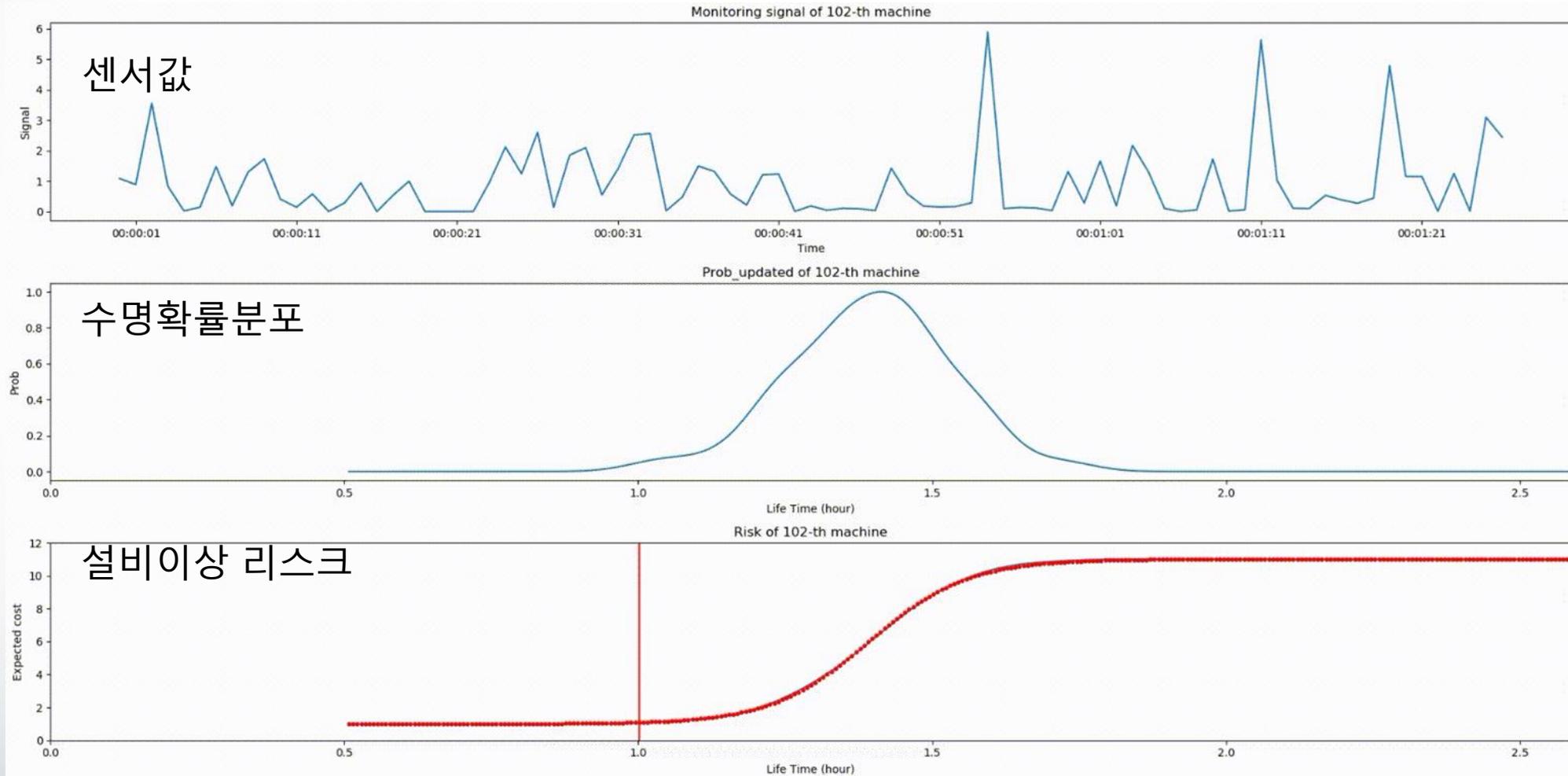
수명예측 실시예

설비 중단 임박



수명예측 실시예

전체 학습 동영상- 베이지안 업데이트를 통한 수명분포의 수렴

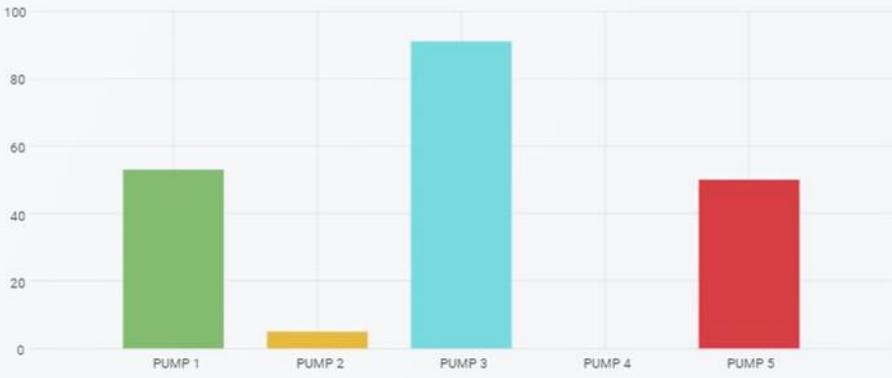


이상감지가 쉬워진다!

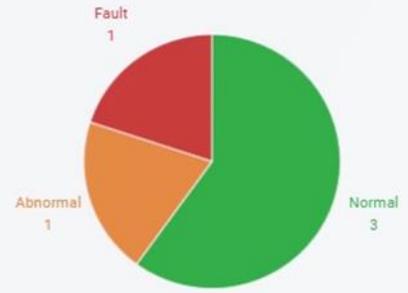
이상감지 모니터링 시스템

Normal Status

Remaining Useful Life %



Overall Status



Anomaly Status



Normal List

pump_name	RUL Rate %
PUMP 1	53
PUMP 3	91
PUMP 5	50

Abnormal List

pump_name	RUL Rate %
PUMP 2	5

Fault List

pump_name	RUL Rate %
PUMP 4	0

Individual Pump Monitoring

Current Status

(Last 1 min)

PUMP 1
Normal

PUMP 2
Abnormal

PUMP 3
Normal

Asset Monitoring

(T-square)



Remaining Gauge %

(Last 1 min)



Breakdown Date

(Predicted)

PUMP 1
2019-04-04

PUMP 2
2018-11-03

PUMP 3
2019-09-25

Normal Status

Overall Scores

Current Status

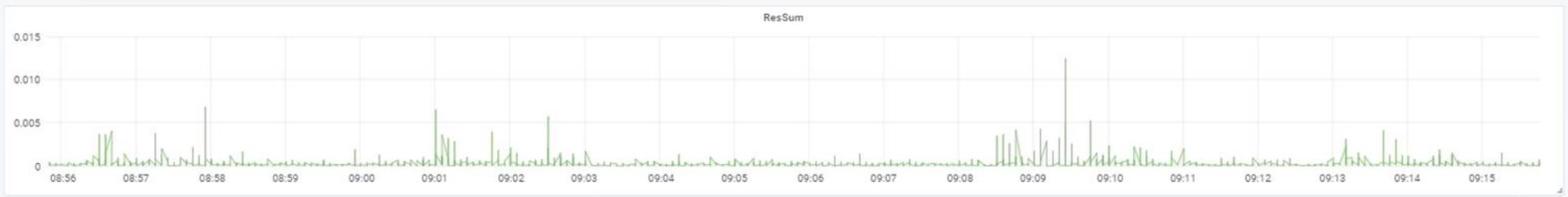
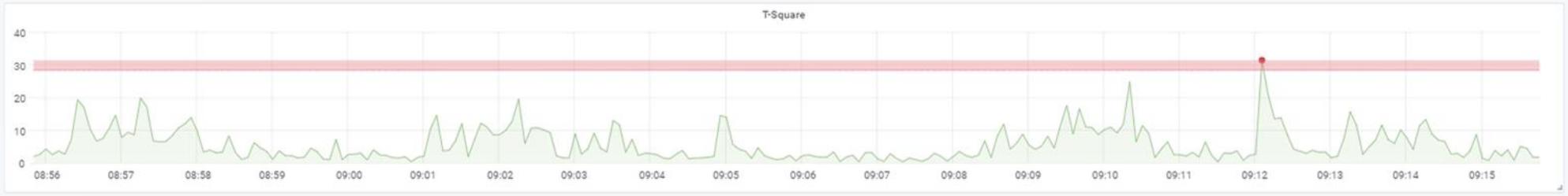
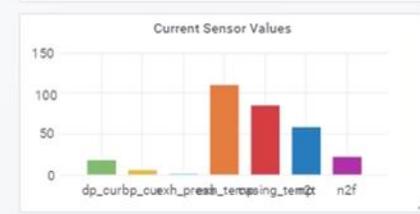
Normal

Remaining Gauge %

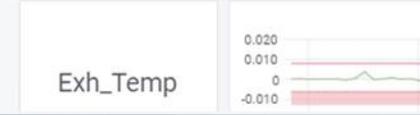
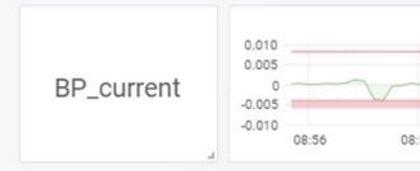
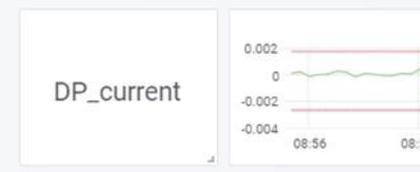
53%

Breakdown Date (predicted)

2019-04-06



Sensor Score



Normal Status

Diagnosis & Interpretation

Monitored Equipment : [Diffusion Pump 1](#)



Time : 2018-10-25 09:04:05 ~ 2018-10-25 09:14:56

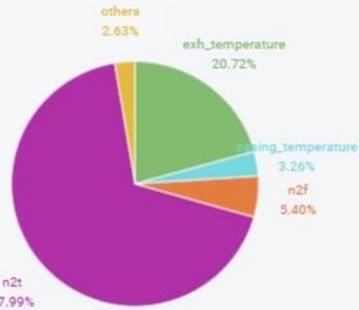
Timely Condition :



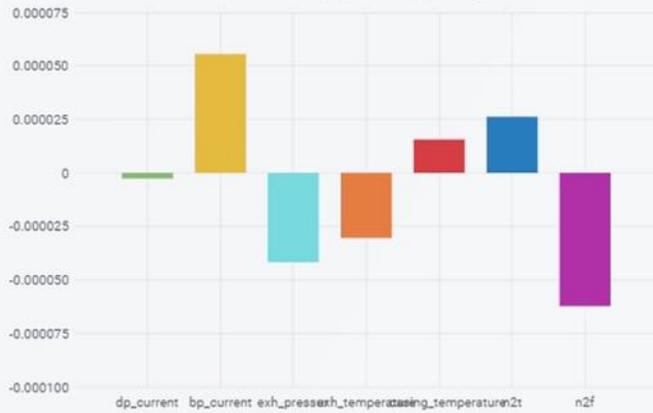
Normal
100.00%

Exploratory Analytics

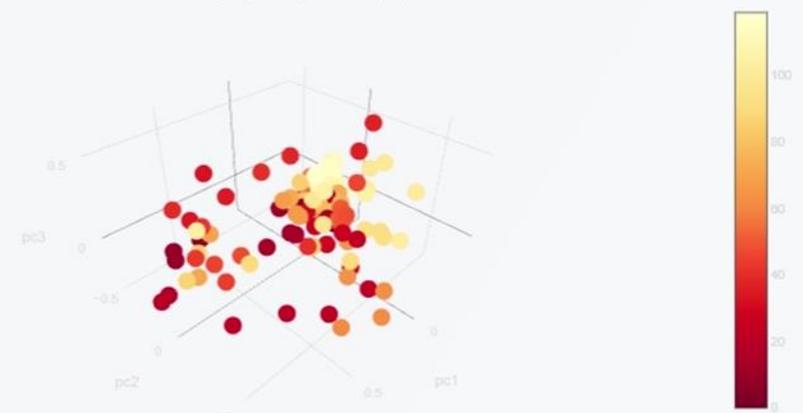
Variable Importance for Anomaly (by Decision Tree)



Sensor Values (Avg within Dragged Area)



PCA (Principal Component Analysis) - 3D



Descriptive Analytics

Correlation Matrix

	dp_cur	bp_cur	e_press	e_temp	c_temp	n2t	n2f
dp_current	1.0000	0.7105	0.2979	0.5045	0.4729	-0.2472	-0.4185
bp_current	0.7105	1.0000	0.1737	0.6784	0.7126	-0.1931	-0.4368

Summary Statistic

	max	min	range	sum	avg	variance	stddev	skewness	kurtosis	row
dp_current	17.68	17.38	0.30	2.11 K	17.58	0.00	0.06	-1.09	1.05	120.00
bp_current	6.12	3.93	2.19	614.28	5.12	0.31	0.55	-1.05	-0.31	120.00

Abnormal Status

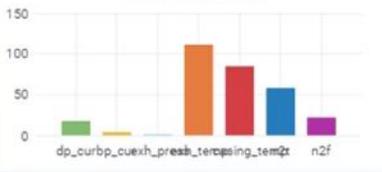
Abnormal



Breakdown Date (predicted)

2018-11-06

Current Sensor Values



▼ Sensor Score

DP_current



BP_current



Exh_Pressur



Exh_Temp



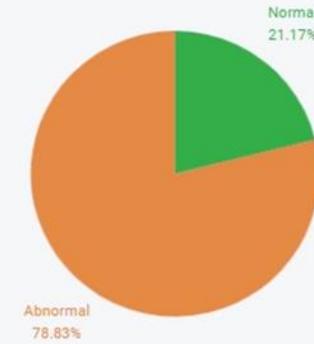
Abnormal Status

Diagnosis & Interpretation

Monitored Equipment : [Diffusion Pump 2](#)

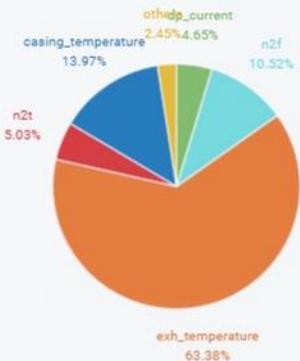
Time : 2018-10-25 09:03:31 ~ 2018-10-25 09:14:57

Timely Condition :

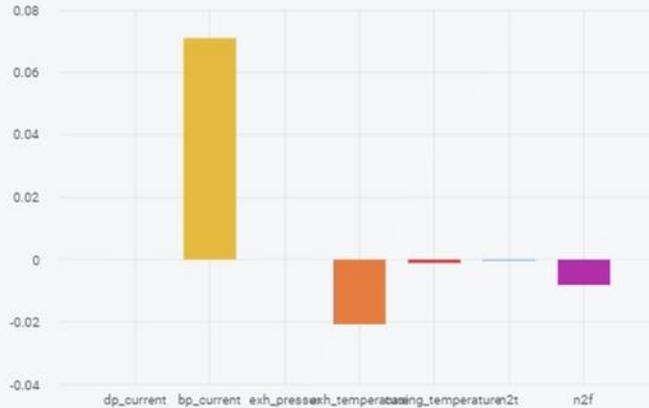


Exploratory Analytics

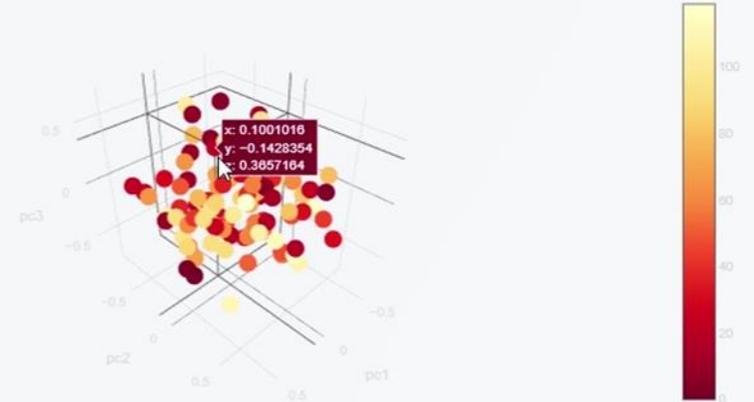
Variable Importance for Anomaly (by Decision Tree)



Sensor Values (Avg within Dragged Area)



PCA (Principal Component Analysis) - 3D



Descriptive Analytics

Correlation Matrix

	dp_cur	bp_cur	e_press	e_temp	c_temp	n2t	n2f
dp_current	1.0000	0.1916	0.0481	0.1257	0.2733	0.0420	-0.0893
bp_current	0.1916	1.0000	-0.0981	0.1628	0.1416	-0.0800	0.0428

Summary Statistic

variable_name	max	min	range	sum	avg	variance	stddev	skewness	kurtosis	nrow
dp_current	17.78	17.28	0.50	2.11 K	17.57	0.01	0.11	-0.31	-0.59	120.00
bp_current	7.05	2.88	4.17	618.81	5.16	1.10	1.05	-0.11	-0.84	120.00

Q & A

Partner

Disrupt

Foresee



Thank you

