

The background image shows a high-speed train, possibly a KTX, traveling at night. The train is white with blue accents and is moving towards the right. A large, translucent blue wireframe hand is positioned as if it is holding or interacting with the train. The background features a city skyline with illuminated buildings, including a prominent skyscraper. The overall scene is dark with blue and white light effects.

철도차량 유지보수 패러다임 전환 (디지털 처방 정비)

국립 한국교통대학교

철도교통창의부품연구소 **김 철 수**

국내 철도환경에서의 AI 한계 – XAI 필요성

철도환경에서의 머신러닝 vs 딥러닝



- 철도분야와 같이 전문가의 적극적인 개입과 제한적인 데이터셋으로부터 특징추출을 통한 **모델의 일반화 능력 검증** → **XAI / 머신러닝**이 더 적합
- [EU 인공지능법(Artificial intelligence act), '21.04.21] EU는 AI 규제안을 통해 **AI 시스템의 투명성을 의무화**하며, **설명가능한 'XAI'**의 중요성을 부각

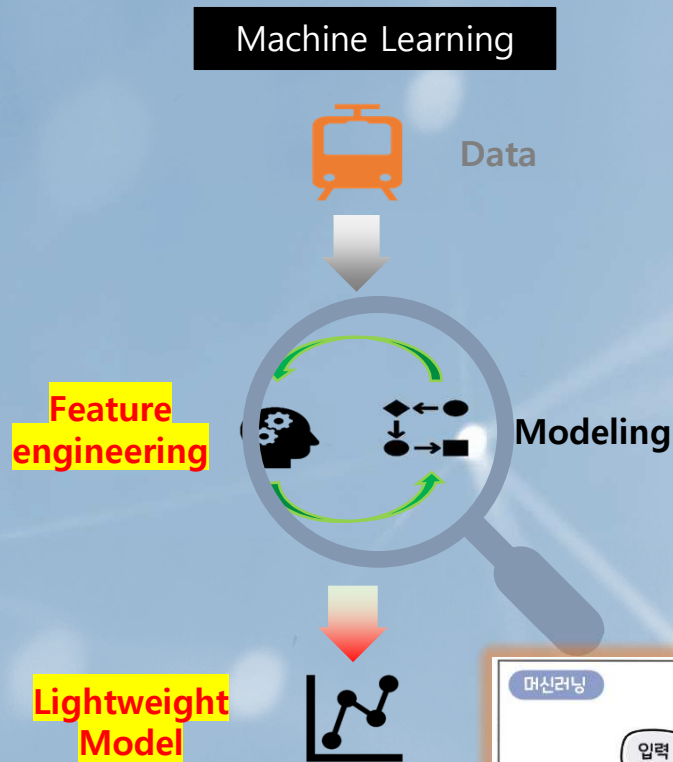
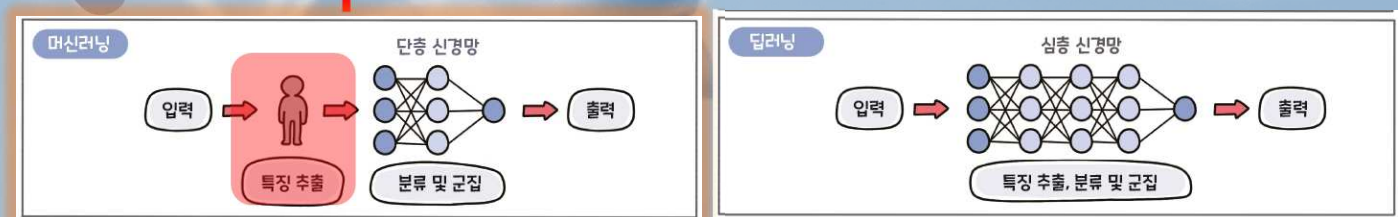


표2_고위험 AI 시스템 분류

유형	세부유형	AI 시스템 유형
유형 I (안전 위험)	① AI 시스템이 안전 규제 대상 제품 또는 제품의 구성요소인 경우	장난감, 기계(machinery), 레크레이션용/개인용 소형 선박(recreational craft and personal watercraft), 리프트, 폭발위험환경에서 사용되는 보호장비, 무선장비, 압력장비, 케이블카, 개인보호장비(Personal Protection equipment), 가스연료 연소기기, 의료기기, 체외 진단 의료기기
	② AI 시스템이 제품자체 또는 제품 구성요소이며 그 제품이 시장 출시/서비스 전에 제3자의 적합성 평가 대상인 경우	민간항공, 이/상/하운 차량, 농업차량/임업차량, 해양장비, 철도 시스템, 자동차 및 트레일러, 이러한 차량의 부품이나 별도의 기술장치, 무인항공기 및 엔진, 프로펠러, 부품 및 장비



해외 사례 - 독일 DB

ETR, "Ausfall prognose als Baustein der Condition Based Maintenance im Instandhaltungssystem eines EMU Nov. 2018, www.eurailpress.de/etr

독일어→영문번역: Failure forecast as a component of Condition Based Maintenance in the maintenance system of an EMU

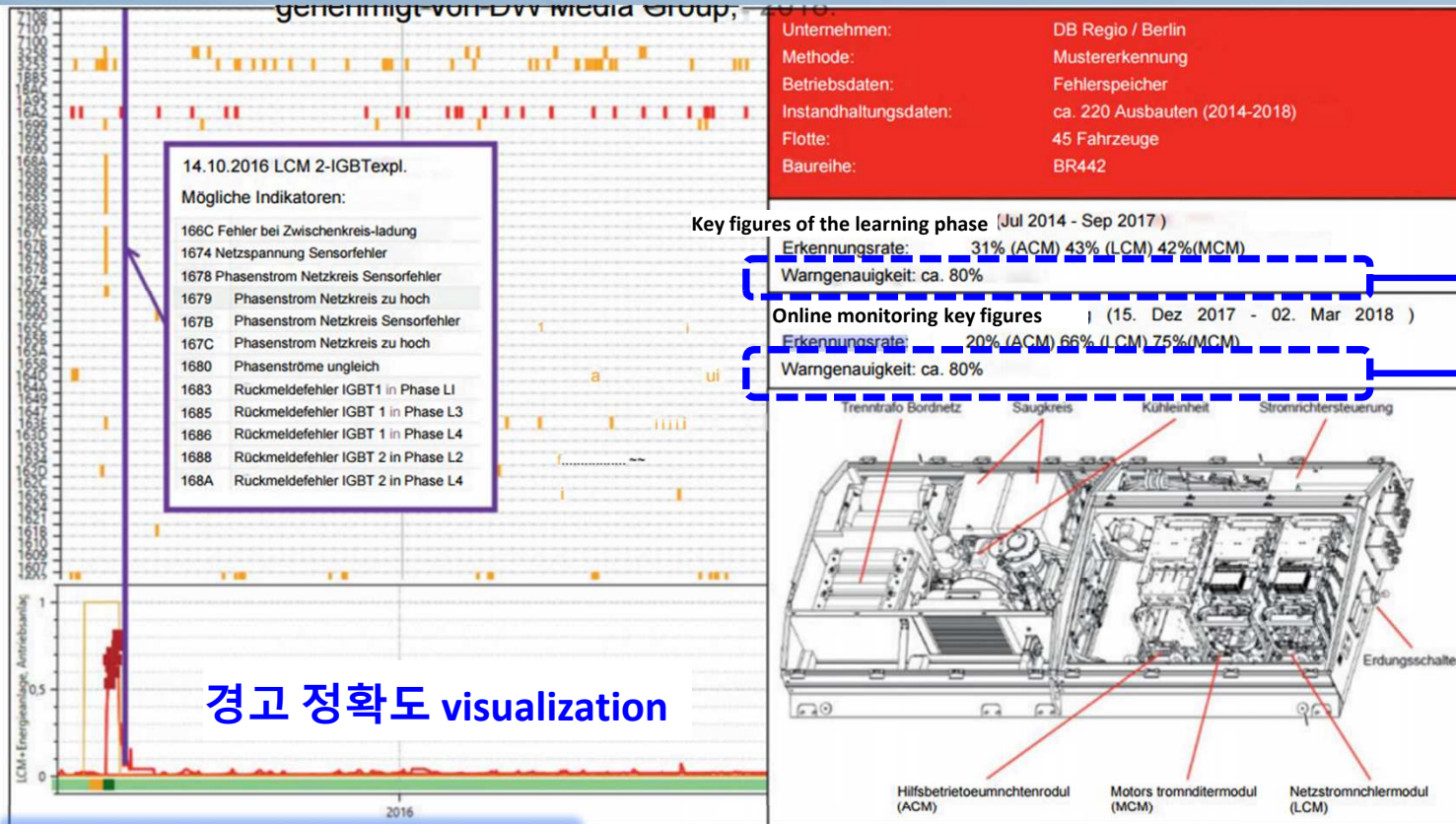


BILD 3:

The graphic on the left shows the converter failure forecast using the logistic regression; the right graphic contains the key figures from this project (retrospective and prospective)

Kennzahlen	Kein tatsächlicher Eintritt	tatsächlicher Eintritt	I. Erkennungsrate	Fehler 2. Art, "übersehener Defekt"
Kein prognostizierter Eintritt	#TN	#FN	$\frac{\text{\# prognostizierte Ereignisse}}{\text{\# alle Ereignisse}}$	$\frac{\text{\# FN}}{\text{\# FN} + \text{\# TP}}$
prognostizierter Eintritt	#FP	#TP	II. Genauigkeit	IV. Fehler 1. Art, "blinder Alarm"
			$\frac{\text{\# erfolgreiche Warnungen}}{\text{\# alle Warnungen}}$	$\frac{\text{\# FP}}{\text{\# FP} + \text{\# TN}}$

BILD 4: Wahrheitsmatrix als Grundlage für weitere Kennzahlen

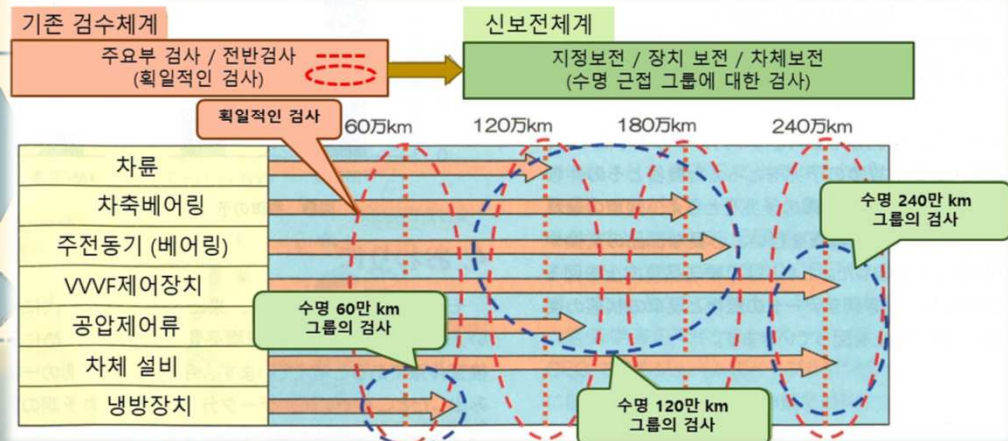
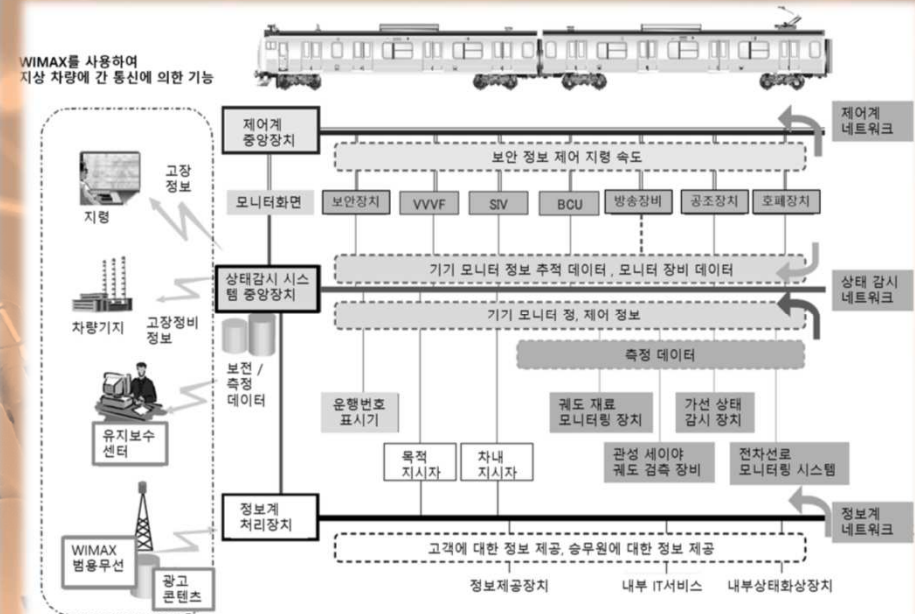
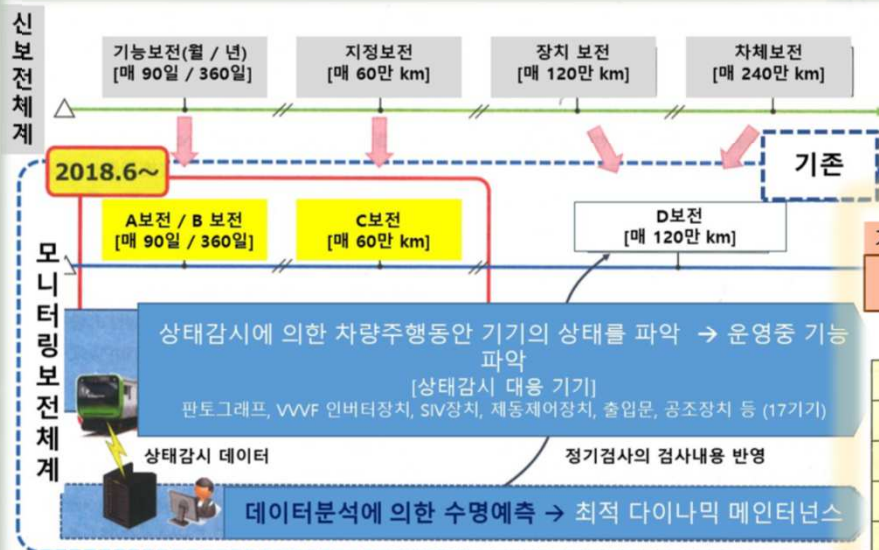
경고 정확도: 80%

독일 DB Regio / Berlin (2014 - 2018)
Fleet: 45 vehicles
Model series: BR442

해외 사례 - 동일본 철도

JR East : INTEROS

- 차량 주행 / 입출고시에 17 종의 상태진단
- 일체형 기계류 및 전장품 정비 주기를 그룹핑
- 일상검사를 제한검사 LI化



해외 사례 - 이탈리아

철도회사 (Trenitalia)의 예지정비기반 시스템 구축사례



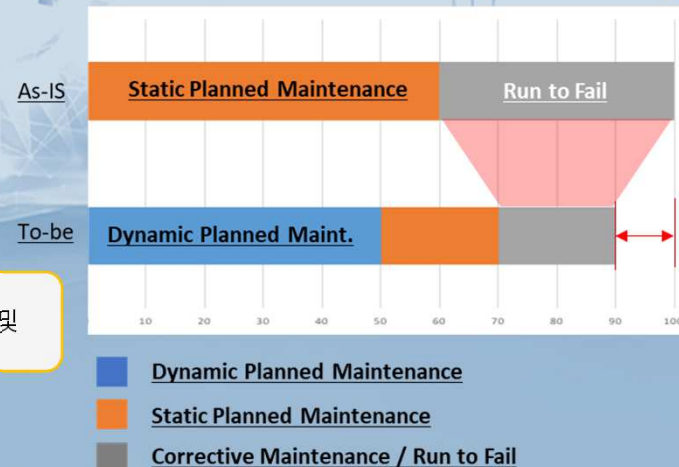
- 이탈리아 국영 철도 업체
- 임직원 : 32,000명
- 연 매출 약 7조(원)
- 연 6천만 명 수송 (중장거리 노선)

- 보유 차량 자산
 - 전기열차 2,000량
 - 기관차 2,000량
 - 객차 30,000량

❖ 예지정비 시스템 DMMS 구축(Dynamic Maintenance Management System)

- 예지정비 대상설비
- 1) 배터리 : 전압, 전류, 온도, 충전량 등의 데이터를 이용하여 정상/비정상 등의 상태로 구분하는 알고리즘을 적용
 - 2) 브레이크 : 마찰에 의해 발산되는 에너지량에 의해 브레이크 수명이 결정된다는 것을 밝혀냄 → 예상수명예측
 - 3) 팬터그래프 : 팬터그래프의 상승에 요구되는 필요시간을 추정하고, 정상과 비정상 패턴을 감지하는 알고리즘을

적용



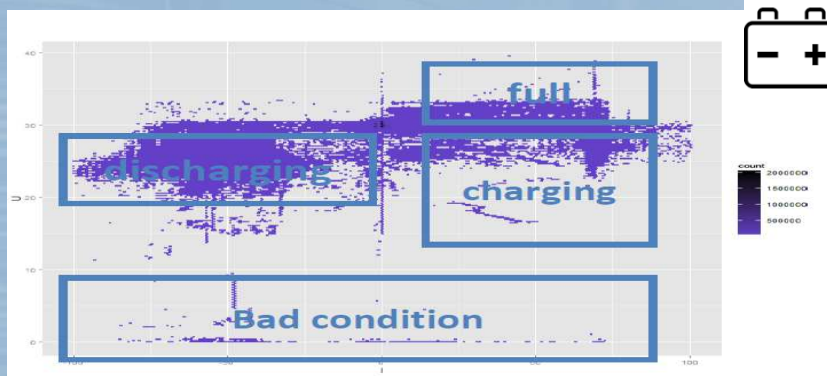
매년 1,700억(원)
정비비용 절감
(총 정비 비용의
8~10% 감소 효과)

해외 사례 - 이탈리아

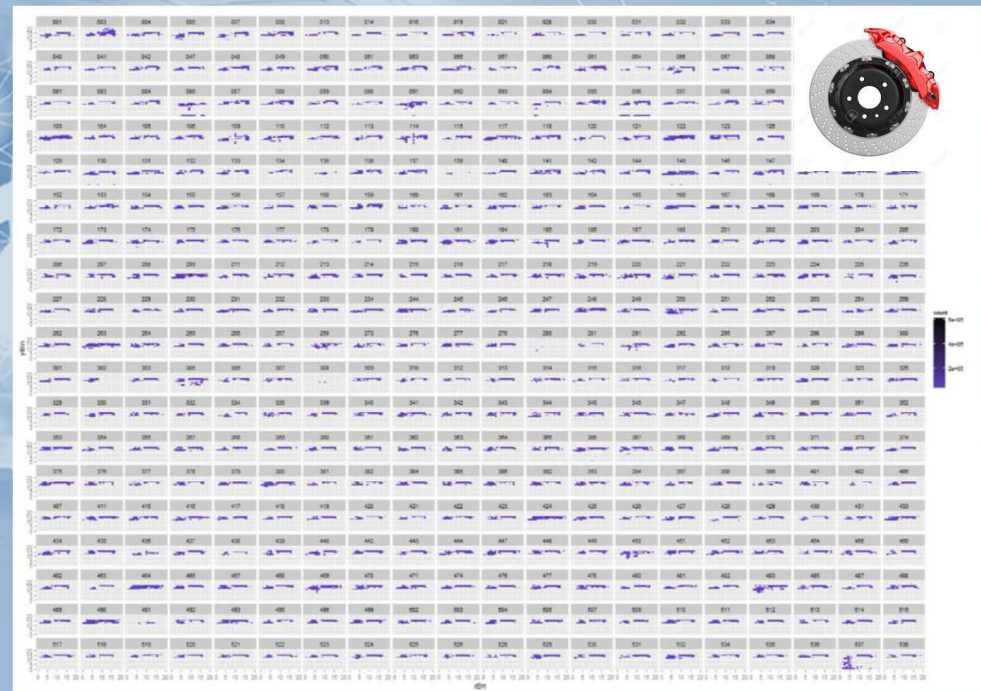
이상패턴분석 및 예측



① 센싱정보를 통한 배터리 상태 군집분석



② Distance based Scoring 기법에 의한 패턴화 및 점수화



③ 현재상태에 대한 모니터링 및 진단



국내 연구사업 개요

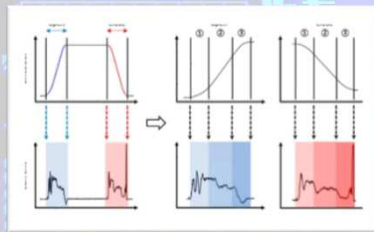
단계	내용	세부 내용	비고
1단계	데이터 수집	·레벨1 데이터:TCMS(진단 1) ·레벨2 데이터:TCMS- Analogue(진단 2) ·레벨3 데이터:열화모델 센싱데이터(진단 3) ·레벨4 데이터:검수데이터 및 이력데이터 ·레벨5 데이터:학률론적 임계값(RUL 평가 기준치)	TCMS와 Edge Server 연계 Edge Server와 Ground Server 연계
2단계	데이터 필터링 및 저장	·데이터 변환 및 데이터 표준화 ·노이즈 제거, 필터링 ·샘플링(열화 자료 등)	Analysis & Storage Server
3단계	데이터 관리	·데이터 유형분석 ·빅데이터 관리 ·BOM 표준화 및 통합(차종 vs.기지)	Analysis & Software Ground Server Server
4단계	고장진단	·장치별 상태기반 고장진단: 식별등급 : RPN(위험도), KPI(가용도, 경제성) ·시계열 데이터 분석: 논리/추세/전용 알고리즘 선정 ·머신러닝/딥러닝 모델 구축	Feature Design, Selection, Extraction Classification Model - Algorithm 선정
5단계	고장예지	·분석 플랫폼 구축 ·검수/정비 관점의 다차원분석 ·노후화에 따른 잔여수명 예측 ·5단계 Threshold/limit	Algorithm 적용 예지 정확도 개선 유지보수지령 및 경고값 등 생성
6단계	최적 유지보수 결정	·예지정비 모델 분석/관리 ·유지보수 주기연장/단축 ·예지/유지보수 전략	
7단계	가시화	·Condition Monitoring ·Life Indicator ·고장원인지표 및 정비항목 디지털화, 고장통계자료 ·모바일 유지보수 효율화 정보 ·다양한 상태 가시화: 기지 차종 편성 차량	UI 개발

Feature Engineering

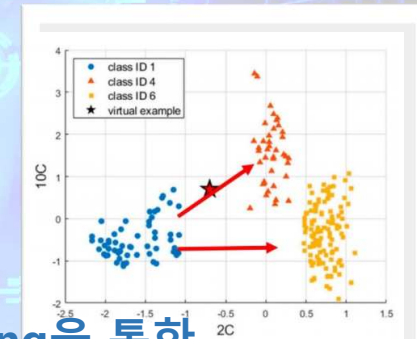
머신러닝 알고리즘을 작동하기 위해 데이터에 대한 도메인 지식을 활용하여 특징(Feature)를 만들어내는 과정

<출처: Wikipedia>

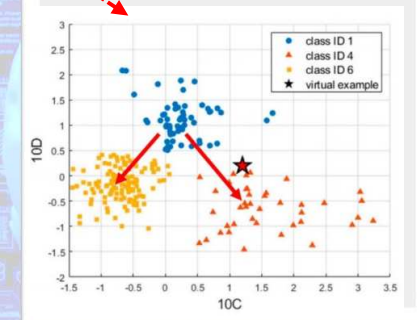
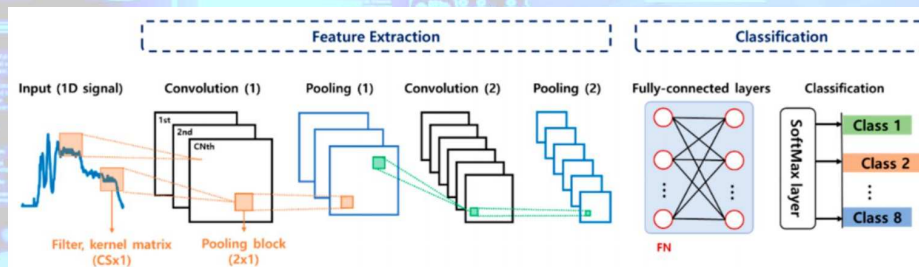
시계열 데이터(온도, 진동, 전류, etc) vs. Feature(통계적 파라미터, 정상성, etc)



Feature ID	1	2	3	4
Feature Name	Mean	Root amplitude	Root meansquares	Standard deviation
Equation	$\frac{\sum x_i}{N}$	$\left(\frac{\sum x_i^2}{N} \right)^{1/2}$	$\sqrt{\frac{\sum x_i^2}{N}}$	$\sqrt{\frac{\sum (x_i - \bar{x})^2}{N-1}}$
Feature ID	5	6	7	8
Feature Name	Peak	Skewness	Kurtosis	Crest factor
Equation	$\max(x)$	$\frac{\sum (x_i - \bar{x})^3}{(N-1) \left(\sqrt{\sum (x_i - \bar{x})^2} \right)^3}$	$\frac{\sum (x_i - \bar{x})^4}{(N-1) \left(\sqrt{\sum (x_i - \bar{x})^2} \right)^4}$	$\frac{\max(x)}{\sigma}$
Feature ID	9	10	11	12
Feature Name	Clearance factor	Shape factor	Impulse factor	Peak-to-peak
Equation	$\frac{\max(x)}{\min(x)}$	$\frac{\sum x_i^3}{\sum x_i^2}$	$\frac{\sum x_i^4}{\sum x_i^3}$	$\max(x) - \min(x)$
Feature ID	13			
Feature Name	Root sum of squares			
Equation	$\sqrt{\sum x_i^2}$			



시계열데이터(아날로그 이산신호)로부터 Feature Engineering을 통한
Feature selection / extraction / construction / classification



특징선택: 다수 데이터로부터 필수 패턴의 최소 특징을 찾는 것(신호 선별, 순차적 특징 선택, 정규화)

특징 추출: 다양한 시계열 데이터에 존재하는 중복성을 제거함으로써, 학습 동안 일반화를 촉진하도록 주요 특징을 추출함

일반화: 학습모델이 특정한 예제에 오버피팅되는 것을 방지(예:특징을 너무 많이 사용하면 오버피팅 발생)

적용 알고리즘

유사사례(국내/국외)

머신러닝 알고리즘중 정확도/F1스코어값 높은 순위 선택

	모델	설명
Decision Trees (의사결정트리)	Fine Tree (조밀트리)	클래스 간에 세밀한 구분이 가능한 결정 트리 (최대 분할 수 = 100)
	Medium Tree (중간트리)	클래스 간에 보다 미세한 구분을 위한 중간 수준의 유연성을 가지는 결정 트리 (최대 분할 수 = 20)
	Coarse Tree (성긴트리)	클래스 사이를 대략적으로 구분하기 위한 단순한 결정 트리 (최대 분할 수 = 4)
	Optimizable Tree (최적화가능트리)	하이퍼파라미터를 최적화하는 결정 트리 분류기
Discriminant Analysis (판별분석)	Linear Discriminant (선형판별)	클래스 사이에 선형 경계를 생성하는 판별 분류기
	Quadratic Discriminant (2차판별)	클래스 사이에 비선형 경계(타원, 포물선 또는 쌍곡선)를 만드는 판별 분류기
	Optimizable Discriminant (최적화가능판별)	하이퍼파라미터를 최적화하는 판별 분류기
Neural Network Classifiers (신경망분류기)	Narrow Neural Network (좁은신경망)	첫번째 계층의 크기가 10인 하나의 완전 연결 계층을 갖는 신경망 분류기
	Medium Neural Network (중간신경망)	첫번째 계층의 크기가 25인 하나의 완전 연결 계층을 갖는 신경망 분류기
	Wide Neural Network (넓은신경망)	첫번째 계층의 크기가 100인 하나의 완전 연결 계층을 갖는 신경망 분류기
	Bilayered Neural Network (이중계층신경망)	분류를 위한 마지막 완전 연결 계층을 제외하고 두 개의 완전 연결 계층을 갖는 신경망 분류기
	Trilayered Neural Network (삼중계층신경망)	분류를 위한 마지막 완전 연결 계층을 제외하고 세 개의 완전 연결 계층을 갖는 신경망 분류기
	Optimizable Neural Network (최적화가능신경망)	하이퍼파라미터를 최적화하는 신경망 분류기
Naive Bayes Classifiers (나이브베이지 분류기)	Gaussian Naive Bayes (가우스나이브베이지)	숫자형 예측 변수에는 가우스 분포 사용, 카테고리형 예측 변수에는 MVMN 분포 사용
	Kernel Naive Bayes (커널나이브베이지)	숫자형 예측 변수에는 커널 분포 사용, 카테고리형 예측 변수에는 MVMN 분포 사용
	Optimizable Naive Bayes (최적화가능나이브 베이지)	하이퍼파라미터를 최적화하는 나이브 베이지 분류기

Nearest Neighbor Classifiers (최근접이웃분류기)	Fine KNN (조밀KNN)	유클리드 거리 측정법을 사용하며 클래스 사이를 조밀하게 구분하는 KNN (이웃 수 = 1)
	Medium KNN (중간KNN)	유클리드 거리 측정법을 사용하며 조밀 KNN보다 더 적은 수의 클래스로 구분하는 KNN (이웃 수 = 10)
	Coarse KNN (성긴KNN)	유클리드 거리 측정법을 사용하며 클래스 사이를 대략적으로 구분하는 KNN (이웃 수 = 100)
	Cosine KNN (코사인KNN)	코사인 거리 측정법을 사용하는 KNN (이웃 수 = 10)
	Cubic KNN (3차KNN)	3차 거리 측정법을 사용하는 KNN (이웃 수 = 10)
	Weighted KNN (가중KNN)	거리 가중치를 사용하는 KNN (이웃 수 = 10)
	Optimizable KNN (최적화가능KNN)	하이퍼파라미터를 최적화하는 KNN
Ensemble Classifiers (앙상블분류기)	Boosted Trees (부스팅트리)	AdaBoost + 중간 결정 트리 앙상블
	Bagged Trees (배깅트리)	Random forest Bag + 조밀 결정 트리 앙상블
	Subspace Discriminant (부분공간판별)	Subspace + 판별 분류기 앙상블
	Subspace KNN (부분공간KNN)	Subspace + KNN 앙상블
	RUSBoosted Trees (RUSBoosted 트리)	RUSBoost + 결정 트리 앙상블
	Optimizable Ensemble (최적화가능앙상블)	하이퍼파라미터 최적화 + 결정 트리 앙상블
Efficiently Trained Linear Classifiers (고효율훈련선형분류기)	Efficient Logistic Regression (고효율로지스틱회귀)	모델 훈련 시간을 줄이는 기법을 사용하는 선형 로지스틱 회귀 분류기
	Efficient Linear SVM (고효율선형SVM)	모델 훈련 시간을 줄이는 기법을 사용하는 선형 SVM
Support Vector Machines (서포트벡터머신)	Linear SVM (선형SVM)	선형 커널을 사용하여 클래스 사이를 단순하게 선형으로 구분하는 SVM
	Fine Gaussian SVM (조밀가우스SVM)	클래스 사이를 조밀하게 구분하는 SVM
	Medium Gaussian SVM (중간가우스SVM)	조밀 가우스 SVM보다 더 적은 수의 클래스로 구분하는 SVM
	Coarse Gaussian SCM (성긴가우스SVM)	클래스 사이를 대략적으로 구분하는 SVM

진단 알고리즘 성능 검증

- 오차행렬(Confusion Matrix) : 예측의 다양한 결과와 분류 문제의 결과를 테이블 레이아웃으로 표시

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

- True Positive(TP) : 긍정 예측을 성공한 횟수
- False Positive(FP) : 긍정 예측을 실패한 횟수
- True Negative(TN) : 부정 예측을 성공한 횟수
- False Negative(FN) : 부정 예측을 실패한 횟수

- 정확도(Accuracy) : 실제 데이터와 예측 데이터가 얼마나 같은지를 평가하는 지표

$$Accuracy = \frac{\text{실제와 예측 결과가 동일한 건수}}{\text{전체 예측 건수}} = \frac{TP + TN}{TP + TN + FP + FN}$$

- 정밀도(Precision) : 긍정 값을 올바르게 분류하는 모델의 능력을 계산

$$Precision = \frac{\text{긍정 예측 성공 건수}}{\text{긍정으로 예측한 건수}} = \frac{TP}{TP + FP}$$

- 재현율(Recall) : 실제 긍정인 값을 잘 맞추는지 평가, 민감도(Sensitivity)라고도 함

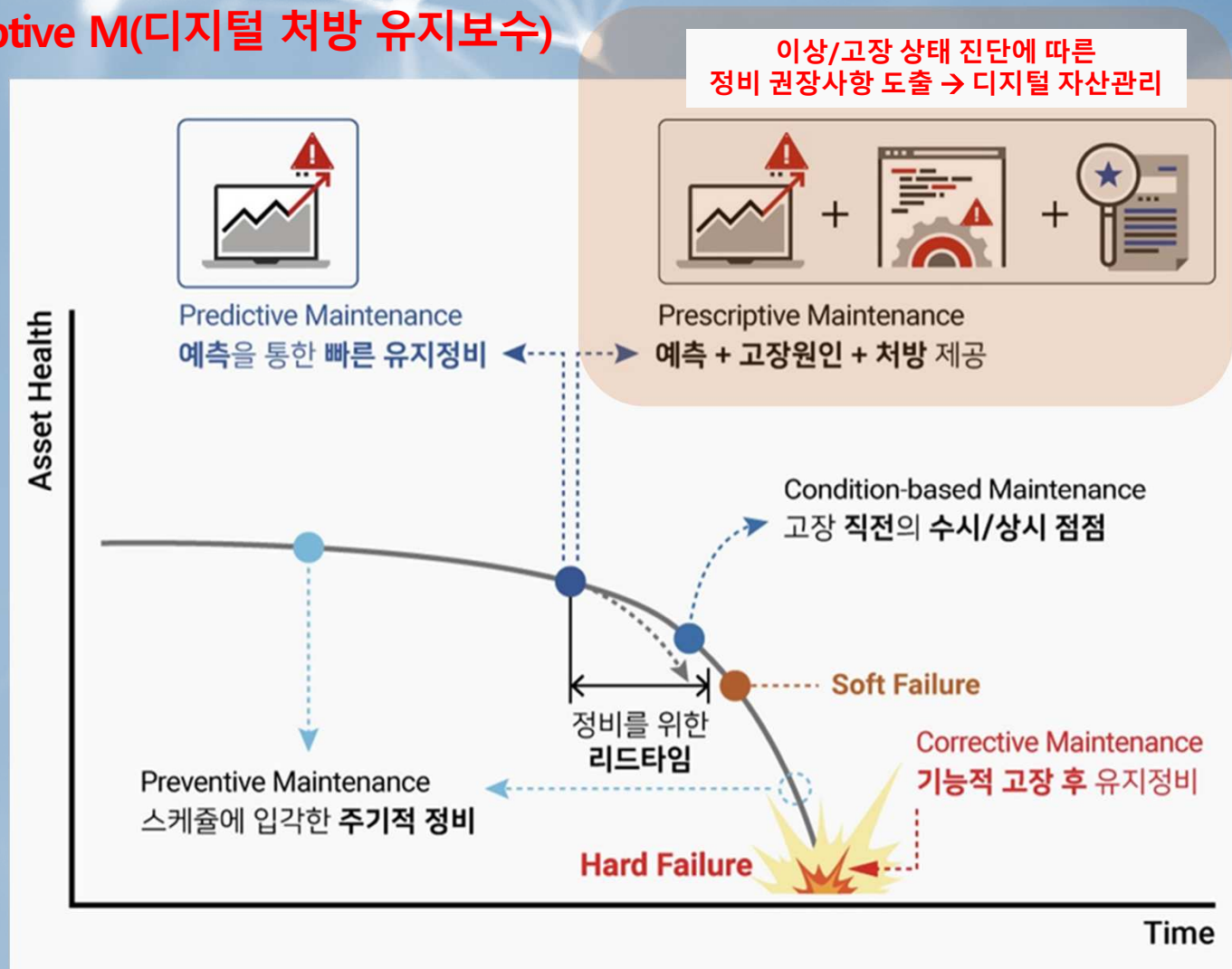
$$Accuracy = \frac{\text{긍정 예측 성공 건수}}{\text{실제 긍정 건수}} = \frac{TP}{TP + FN}$$

철도차량유지보수 패러다임 변환

CBM(상태진단유지보수)

PreDictive M(예지유지보수)

● Digital Prescriptive M(디지털 처방 유지보수)



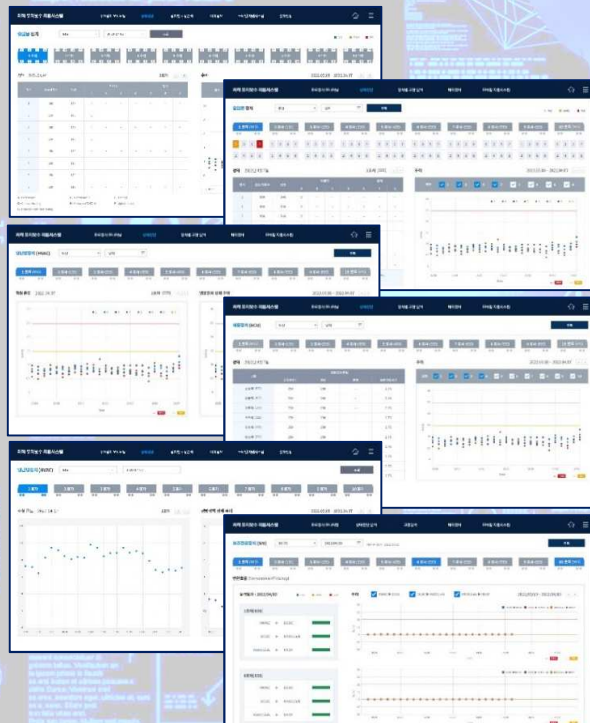
철도차량유지보수 패러다임 변환

예지정비 Prediction	디지털 처방/지능 정비 Prescription
자산이 곧 고장날 것으로 예측	작업 주문서 자동 작성 및 기술자 파견 권장
부품의 신뢰할 수 있는 수명 예측	수명이 다하기 전에 부품을 교체를 권장
자산의 잔여 내용연수 예측	이전 자산이 고장났을 때 복구하는 것보다 폐기하는 것이 좋은지 권장
생애 수명 유지보수 비용 예측	최적의 수명 유지보수 비용이 포함된 유지보수 계획 추천
자산에 대한 고장의 근본 원인 예측	근본 원인을 기준으로 자산을 복구하는 최선의 해결 방법 권장
각 작업 영역의 부품 수요 예측	각 작업 영역의 부품 재고 수준 권장
작업명령 턴타임 예측	마감일을 놓치지 않도록 작업 순서를 상향 조정하는 것이 좋은지 권장
수리 비용 예측	자산을 수리하는 대신 교체하는 것이 좋은지 권장
보증 오류 예측	잘못된 이의제기인지 여부를 판단하기 위해 사례를 공개하는 것이 좋은지 권장

CBM과 디지털 전환 자산관리를 위한 제언

디지털 처방 유지보수 관점에서 지상 상태진단 및 유지보수 지원 시스템

- 진단 알고리즘을 통해 주요 장치의 하위 부품에 대한 상태를 진단
- Regression 분석을 통하여 유지보수 시기 결정
- 이상 상태 발견 시, 정비원에게 고장코드에 따른 조치방법, 정비매뉴얼, 도면 등을 제공(필요 시 다자간 원격지원)



[주요 장치별 상태 진단 및 상태 추이 분석]

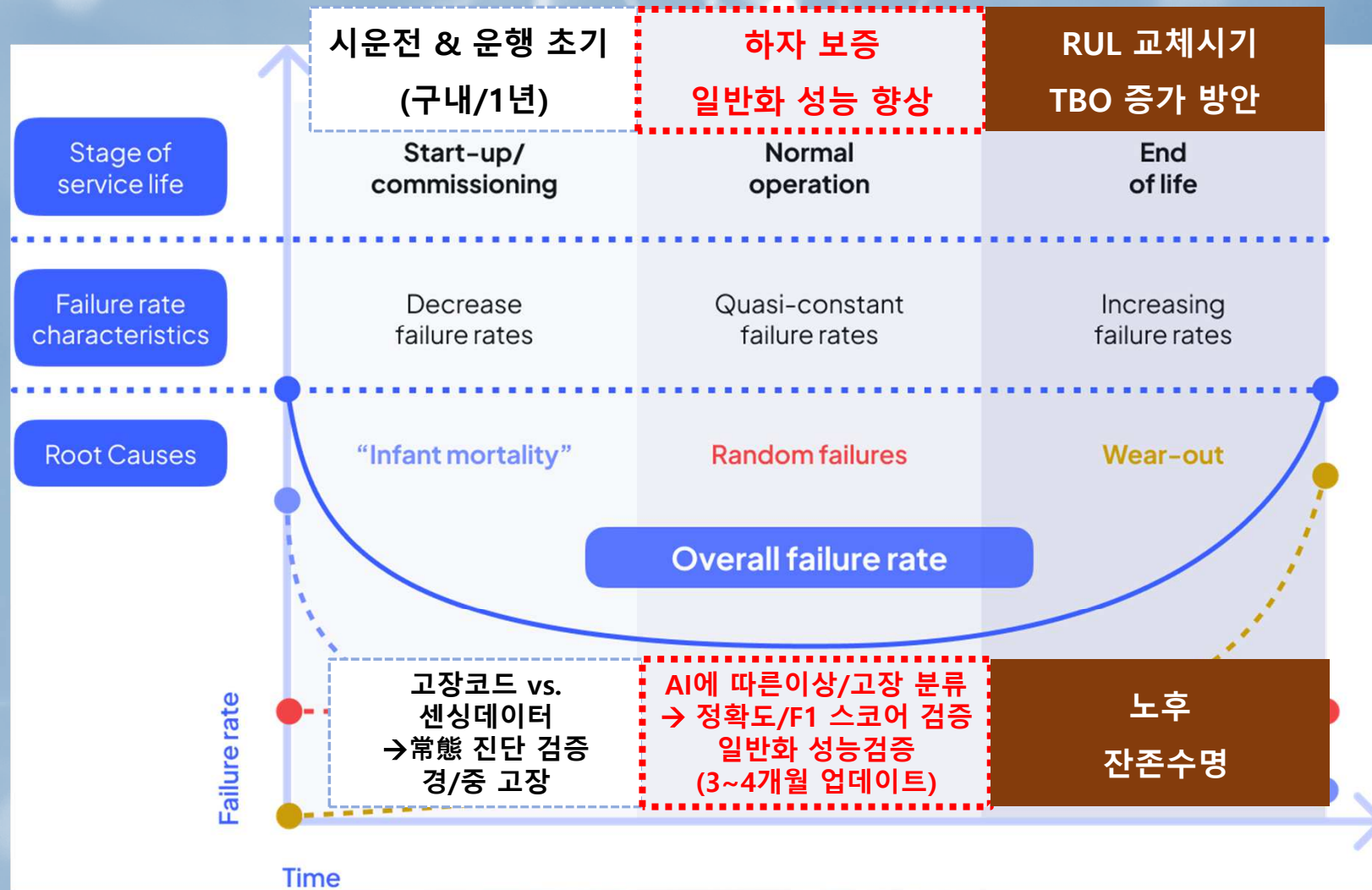


[이상상태 감지 시 조치 방법]



[정비원에게 정비매뉴얼, 도면 등 제공]

CBM과 디지털 전환 자산관리를 위한 제언



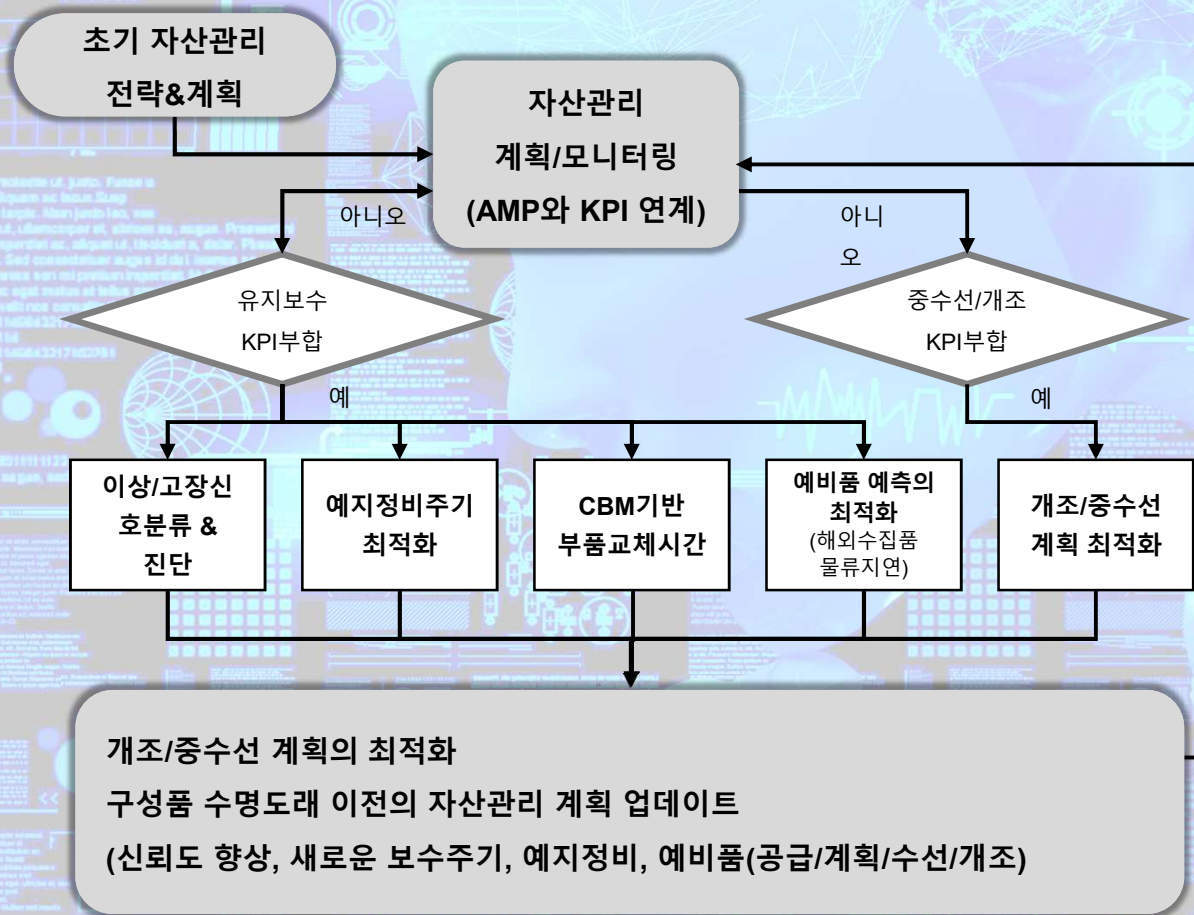
CBM과 디지털 전환 자산관리를 위한 제언

지능형 유지보수관리 시스템 (철도운영자산관리 디지털 전환시스템)

CBM: condition base maintenace

AMP: asset management plan

KPI: key performance index



Code	KPI(key performance index)
C6b	열차 정시 운행률 (Actual Trains Run on Time)
P1a	피크시간 차량가용률
P1b	피크시간 차량운행률
P3a(i)	차량 km당 지연 유발 사고(2분 이상)
P3a(ii)	차량 km당 지연 유발 사고(5분 이상)
P3a(iii)	차량 km당 지연 유발 사고(15분 이상)
P3b(i)	열차 km당 지연 유발 사고(2분 이상)
P3b(ii)	열차 km당 지연 유발 사고(5분 이상)
P3b(iii)	열차 km당 지연 유발 사고(15분 이상)
P5a	Overall Labour Productivity (Passenger Journeys / Staff + Contractor Hours)
P5b	Overall Labour Productivity (Capacity Km / Total Staff + Contractor Hours)
P5c	Overall Labour Productivity (Car Km / Total Staff + Contractor Hours)
F2b	열차 주행 거리 당 유지보수비용
F2c	열차 주행 거리 당 일반관리비용

**본 디지털 콘텐츠는 한국교통대학교에서
전동차 전기식 출입문의 매뉴얼을 기반으로 제작하였습니다**

감사합니다

