

# 제5회 K-인공지능 제조데이터 분석 경진대회 보고서

프로젝트명	AI 기반 지능형 금속가공 품질 향상 시스템
팀명	Hunters42
내용요약	<p><b>1. 분석 목표</b> 소성가공 압출공정의 센서 데이터를 기반으로, 불량 발생을 사전에 탐지하는 AI 품질보증 조기경보 모델을 구축하였다.</p> <p><b>2. 모델 구조</b></p> <ul style="list-style-type: none"><li>정적 템형 양상블(RandomForest + LightGBM)과 시계열 모델(TCN)을 결합한 투 트랙 하이브리드 구조</li><li>Soft 양상블: 두 트랙 확률 가중 평균 → 정확도 중심</li><li>Cascade 양상블: 단계적 판정 → 오탐 최소화 중심</li></ul> <p><b>3. 주요 성능 결과</b></p> <ul style="list-style-type: none"><li>Soft 양상블: PR-AUC 0.9667, ROC-AUC 0.9983</li><li>Cascade 양상블: PR-AUC 0.8538, ROC-AUC 0.9070</li><li>리드타임 중앙값 <math>\geq 1</math>분, FP/day <math>\leq 1</math> 유지</li><li>운영정책: A(F1 최대), B(재현율 <math>\geq 0.9</math>), C(오탐 <math>\leq 1/일</math>)</li></ul> <p><b>4. 해석 및 검증(XAI)</b></p> <ul style="list-style-type: none"><li>SHAP 기반 로컬 해석으로 알람별 주요 인자 식별</li><li>PDP · ICE로 상위 변수의 비선형 임계 구간 검증</li><li>TCN 기여도 열지도로 시간 · 센서별 이상 패턴 시각화</li></ul> <p><b>5. 최종 결론</b></p> <p>Soft 양상블을 기본으로, Cascade 임계 제어를 결합한 하이브리드 운영 구조가 성능(탐지력)과 안정성(오탐 최소화)을 동시 달성하였습니다. 이는 제조 AI의 실질적 현장 적용성과 확장 가능성을 입증합니다.</p>

상기 본인(팀)은 위의 내용과 같이 제5회 K-인공지능 제조데이터 분석 경진대회  
결과 보고서를 제출합니다.

2025년 11월 03일

팀장 : 황재성   
팀원 : 박성호   
팀원 : 이유진 

한국과학기술원장 귀중

# 목 차

<b>1. 문제정의</b> .....	<b>4</b>
<b>1.1 공정 개요 및 현황 분석</b> .....	<b>4</b>
1.1.1 공정 개요 및 목적 .....	4
1.1.2 설비 구조 및 데이터 수집 체계 .....	4
1.1.3 주요 공정 변수 정의 .....	4
<b>1.2 현황 분석 및 문제 정의</b> .....	<b>4</b>
1.2.1 현행 검사 방식 및 데이터 품질의 한계 .....	4
1.2.2 데이터 품질 및 활용성의 한계 .....	5
1.2.3 품질 불량에 따른 산업적 리스크 및 시스템화 미흡 .....	5
<b>1.3 해결하고자 하는 문제 및 분석 목적</b> .....	<b>5</b>
<b>1.4 기대 효과</b> .....	<b>5</b>
<b>2. 제조데이터 정의 및 처리과정</b> .....	<b>6</b>
<b>2.1 데이터 수집 및 구성 현황</b> .....	<b>6</b>
2.1.1 데이터 정의 및 수집 현황 .....	6
2.1.2 데이터 수집 개요 .....	6
<b>2.2 데이터 품질 평가 및 확보 전략</b> .....	<b>6</b>
2.2.1 데이터 품질지수 평가 .....	6
2.2.2 결측치 및 이상치 처리 결과 .....	6
<b>2.3 데이터 탐색(EDA) 및 피처 엔지니어링</b> .....	<b>7</b>
2.3.1 변수 분포 및 상관관계 분석 .....	7
2.3.2 데이터 정제 및 유효 변수 선택 .....	8
2.3.3 전처리 전후 품질 개선 결과 요약 .....	8
<b>3. 분석모델 개발</b> .....	<b>9</b>
<b>3.1 모델 개요</b> .....	<b>9</b>
3.1.1 하이브리드 조기 경보 체계 구성 .....	9
3.1.2 후처리 및 경보 필터링 .....	9
<b>3.2 모델 학습 및 검증</b> .....	<b>9</b>
3.2.1 데이터 분할 및 스케일링 .....	9
3.2.2 학습 최적화(TCN) .....	9
3.2.3 양상별 가중치 최적화 .....	10
<b>3.3 모델 평가</b> .....	<b>10</b>
3.3.1 성능 평가 지표 .....	10
3.3.2 최종 성능 결과 .....	14

<b>4. 분석결과 및 시사점 .....</b>	<b>16</b>
<b>4.1 모델 해석 방법론 .....</b>	<b>15</b>
4.1.1 정적 텁형 트랙(랜덤포레스트+LGBM) .....	16
4.1.2 정적 텁형 트랙 - 개별 예측 해석(로컬) .....	16
4.1.3 시계열 트랙(TCN) - 시간축 해석 .....	16
4.1.4 시퀀스 수준 해석 - 프로토타입 · 반사실 .....	16
4.1.5 알람 단위 설명 카드(운영 리포트용) .....	16
<b>4.2 주요 인자 도출 및 기여도 분석 .....</b>	<b>16</b>
4.2.1 전역 중요도 기반의 핵심 인자 도출 .....	16
4.2.2 알람 시점별 로컬 기여도 분석 (운영 근거 확보) .....	17
4.2.3 오류 및 신뢰성 분석을 통한 안정성 확보 .....	17
4.2.4 모드별 성능 비교 및 검증 .....	17
<b>4.3 분석 결과 요약 및 시사점 .....</b>	<b>18</b>
4.3.1 데이터 기반 의사결정 체계 확립 .....	18
4.3.2 운영 안정성 및 선제적 대응 강화 .....	18
4.3.3 기술적 확장성 확보 .....	18
<b>5. 시스템 설계 및 구현 .....</b>	<b>19</b>
<b>5.1 Forge AI 플랫폼 개요 .....</b>	<b>19</b>
<b>5.2 아키텍처 구성 .....</b>	<b>19</b>
<b>5.3 사용자 흐름 .....</b>	<b>19</b>
<b>5.4 주요 페이지 구성 .....</b>	<b>19</b>
<b>6. 중소제조기업에 미치는 파급효과 .....</b>	<b>20</b>
<b>6.1 기술적 파급효과 .....</b>	<b>20</b>
6.1.1 품질 검사 자동화 및 디지털 전환 가속화 .....	20
6.1.2 범용성 확보 및 AI 리포트 자동 생성 고도화 .....	20
<b>6.2 경제적 파급효과 .....</b>	<b>20</b>
6.2.1 불량 비용 최소화를 통한 수익성 개선 .....	20
<b>6.3 시스템 확장 전략 및 운영 효율성 .....</b>	<b>20</b>
6.3.1 향후 시스템 확장 계획 .....	21
<b>경진대회 만족도 조사 완료 페이지 캡쳐 화면 .....</b>	<b>22</b>

# 1. 문제 정의

## 1.1 공정 개요 및 현황 분석

### 1.1.1 공정 개요 및 목적

소성가공 압출공정은 금속 소재에 외력을 가하여 영구 변형(plastic deformation)을 유도함으로써 원하는 형상을 성형하는 대표적인 제조기술입니다. 이 공정은 일정한 단면 형상을 가진 제품을 연속적으로 생산하는 데 적합하며, 자동차-항공-건축 등 다양한 산업 분야에서 꼭넓게 활용되고 있습니다.

소성가공의 핵심은 소재의 내부 미세조직을 제어하여 강도-연성-치수 정밀도를 동시에 확보하는 것으로, 공정 중 온도-압력-속도 등 주요 인자의 정밀 제어가 제품 품질을 결정짓는 핵심 요소로 작용합니다.

### 1.1.2 설비 구조 및 데이터 수집 체계

본 프로젝트의 대상은 자동차 제동 및 구동계 부품에 사용되는 스틸 튜브 압출성형 설비입니다. 해당 설비는 고온-고압 조건에서 금속 소재를 다이(die)를 통해 밀어내어 일정한 단면의 관형 제품을 생산하는 구조입니다.

이 설비의 각종 센서 데이터는 PLC(Programmable Logic Controller)를 통해 5초 간격으로 자동 수집됩니다. 5초 단위의 샘플링 주기는 공정의 열적-기계적 변화를 안정적으로 반영할 수 있는 해상도를 제공합니다.

### 1.1.3 주요 공정 변수 정의

공정 데이터는 온도, 압력, 속도, 부하 등 주요 변수로 구성됩니다. 전체 데이터셋은 약 17,000행 X 20열로 구성된 정형 시계열(tabular time-series) 형태입니다.

변수 구분	변수명	설명	단위
온도	EX1~5.MELT_TEMP	압출 호기별 수지 온도	°C
압력	EX1.MELT_P_PV	수지 압출 압력	MPa
속도	EX1.MD_PV	모터 회전 속도	rpm
부하	EX1.MD_TQ	모터 부하율	%
품질	passorfail	외경-두께 기준 품질 판정	0=양품, 1=불량

## 1.2 현황 분석 및 문제 정의

### 1.2.1 현행 검사 방식 및 데이터 품질의 한계

현재 품질 판정은 대부분 작업자의 육안 검사 및 경험적 기준에 의존하고 있습니다. 이로 인해 객관성과 재현성이 확보되지 못하며, 작업자 숙련도-피로도에 따라 결과가 달라지는 문제가 발생합니다. 정량적 품질 기준이 부재하여 과잉 혹은 과소 판정이 빈번히 발생하며, 이는 불필요한 폐기-재검-클레임으로 이어집니다.

## 1.2.2 데이터 품질 및 활용성의 한계

실제 제조 환경에서는 센서 노이즈, 결측치, 이상치, 드리프트가 빈번히 발생하여 모델의 학습 안정성과 예측 성능을 저하시킬 수 있습니다. 또한, 부서별 데이터 관리가 분리되어 데이터 사일로(silo) 현상이 심화되고 있습니다.

이로 인해 전체 공정의 인과 분석이 불가능하며, 실질적인 AI 학습용 통합 데이터셋 구축이 어렵습니다.

## 1.2.3 품질 불량에 따른 산업적 리스크 및 시스템화 미흡

브레이크 라인 등 안전부품의 품질 불량은 인명사고 및 대규모 리콜로 직결될 수 있습니다. 불량 발생 이후의 사후 대응이 아닌, 사전 예측 기반의 예방 체계 구축이 절실히 합니다.

공정 조건 설정이 작업자의 경험에 의존하고, 데이터 분석 기반의 피드백 루프가 부재합니다. 지속적 개선(MLOps 기반 학습-배포 체계)이 정착되지 못한 상황입니다.

## 1.3 해결하고자 하는 문제 및 분석 목적

본 프로젝트의 목표는 데이터 기반의 AI 품질보증 시스템(MetraForgeAI AI)을 구축하여 소성가공 공정의 품질 관리 전 과정을 자동화-지능화하는 것입니다.

구축 영역	세부 내용
데이터 품질 확보 및 자동 전처리 체계 구축	결측치 및 이상치 자동 탐지, 센서별 시간 정렬, 품질지수 산출, 유효 변수 선별 등을 통한 데이터 신뢰성 확보.
불량 판정 예측 모델 고도화	고성능 모델 적용, 하이퍼파라미터 최적화, 클래스 불균형 문제 해결, SHAP 기반 설명 가능 AI(XAI) 적용.
플랫폼 통합 및 실시간 서비스 구현	FastAPI 기반 모델 배포, React 기반 대시보드 시각화, MLOps 체계 확립을 통한 End-to-End 품질보증 파이프라인 구축.

주요 목적은 소성가공 공정 데이터를 활용하여 제품의 품질을 사전에 예측하고, 실시간 공정 데이터를 기반으로 품질 이상 발생 가능성은 정량적으로 제시하는 AI 모델을 개발하는 것입니다.

## 1.4 기대 효과

본 프로젝트를 통해 구축되는 AI 품질보증 시스템의 기대 효과는 다음과 같습니다.

구분	기대 효과
기술적 측면	데이터 품질 진단-전처리-분석을 통합한 표준 알고리즘 프레임워크 확립.
경제적 측면	AI 기반 품질 예측을 통해 불량률을 현 수준 대비 50% 이상 개선하고 비용 절감.
운영적 측면	실시간 품질 모니터링 및 조기 경보 체계 구축을 통한 생산라인 안정성 강화.
확장성 측면	개발된 분석-예측 기법은 단조, 주조 등 유사 공정으로 이식 가능.
지속성 측면	주기적 재학습과 자동 성능 모니터링 기반 MLOps 체계 확립.

## 2. 제조데이터 정의 및 처리과정

### 2.1 데이터 수집 및 구성 현황

#### 2.1.1 데이터 정의 및 수집 현황

본 프로젝트는 '2. 소성가공 품질보증 AI 데이터셋.csv'를 기반으로 수행되었습니다. 데이터는 소성가공 압출 설비에서 PLC를 통해 5초 간격으로 자동 수집된 정형 시계열 데이터이며, 종속변수는 제품 품질 결과(passorfail 컬럼)입니다.

#### 2.1.2 데이터 수집 개요

항목	내용
데이터 명칭	2. 소성가공 품질보증 AI 데이터셋.csv
데이터 형태	정형 시계열
수집 기간	2020.10.30(00:00:04)~ 2020.10.30(13:53:19)
수집 주기	5초 단위
데이터 규모	약 17,000행 X 20열

### 2.2 데이터 품질 평가 및 확보 전략

#### 2.2.1 데이터 품질지수 평가

데이터 품질 평가는 6개 지수(Completeness, Uniqueness, Validity, Consistency, Accuracy, Integrity)를 기준으로 수행되었습니다. 전처리 이후 모든 품질지수는 100.0% 달성을 확인되었습니다.

품질지수	평가 결과 (전처리 후)
Completeness (결측치 비율)	100.0%
Uniqueness (중복 행 비율)	100.0%
Validity (허용 범위 내 값 비율)	100.0%
Consistency (변수 간 논리 일관성)	100.0%
Accuracy (기준편차 검증)	100.0%
Integrity (참조 관계 보존)	100.0%

#### 2.2.2 결측치 및 이상치 처리 결과

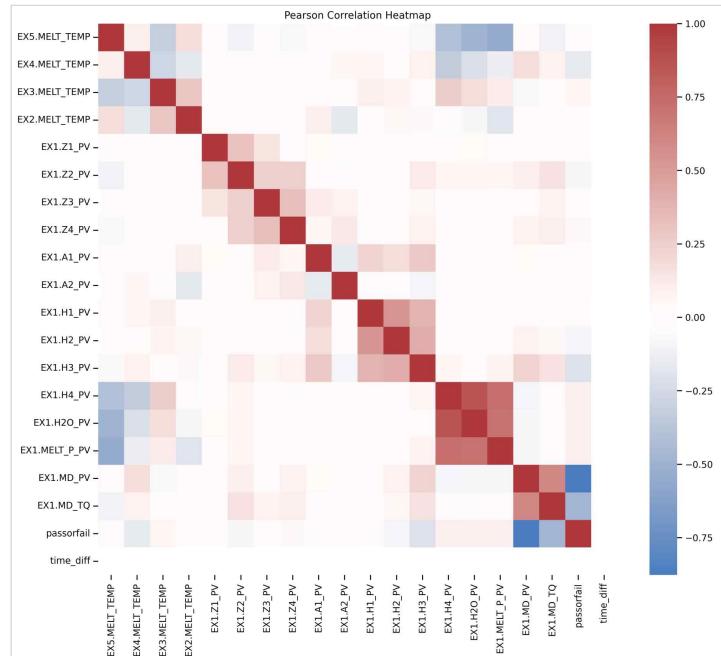
결측률은 초기 데이터의 약 1.3%였으며, 결측은 5초 규칙격자 리샘플 후, 최근 관측값으로 단일 보간(ffill)하였으며 장기 결측은 제외하였다. 이상치는 라벨이 있는 구간에 대해 통계 기반 이상치 탐지를 수행하여 비정상 샘플을 제거하였습니다.

## 2.3 데이터 탐색 (EDA) 및 피처 엔지니어링

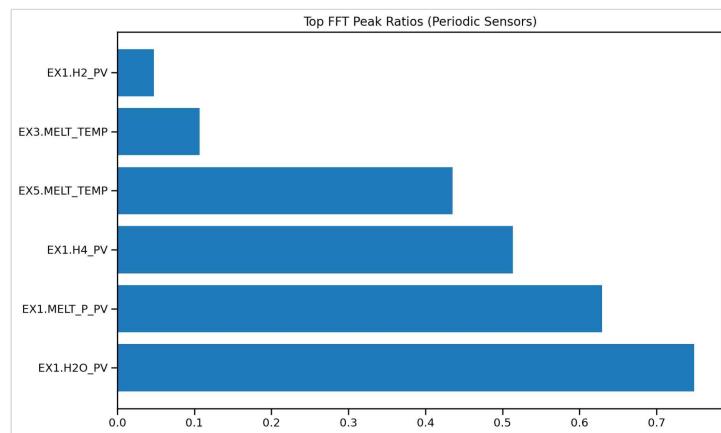
### 2.3.1 변수 분포 및 상관관계 분석

Heatmap(Pearson)으로 변수 간 관계를 확인한 결과, 절댓값  $r \geq 0.8$  이상인 변수쌍은 대표변수로 집약 처리되었으며, 온도와 압력 간 높은 양의 상관이 관찰되었습니다. 불량률은 약 0.64% (양품:불량 약 156:1)로 극심한 클래스 불균형이 존재함을 확인하였습니다.

센서 신호의 시간적 주기성을 검증하기 위해 FFT(Fast Fourier Transform) 분석을 수행하였으며, 주요 주기성 센서의 Peak Ratio는 아래 그림과 같습니다.



(1) Pearson Correlation Heatmap



(2) Top FFT Peak Ratios

[그림 2.1] 변수 분포 및 상관관계 분석 종합

### 2.3.2 데이터 정제 및 유효 변수 선택

데이터 정제는 통계 검정, 차원 축소, 그룹 피처 생성 단계로 구성되었습니다.

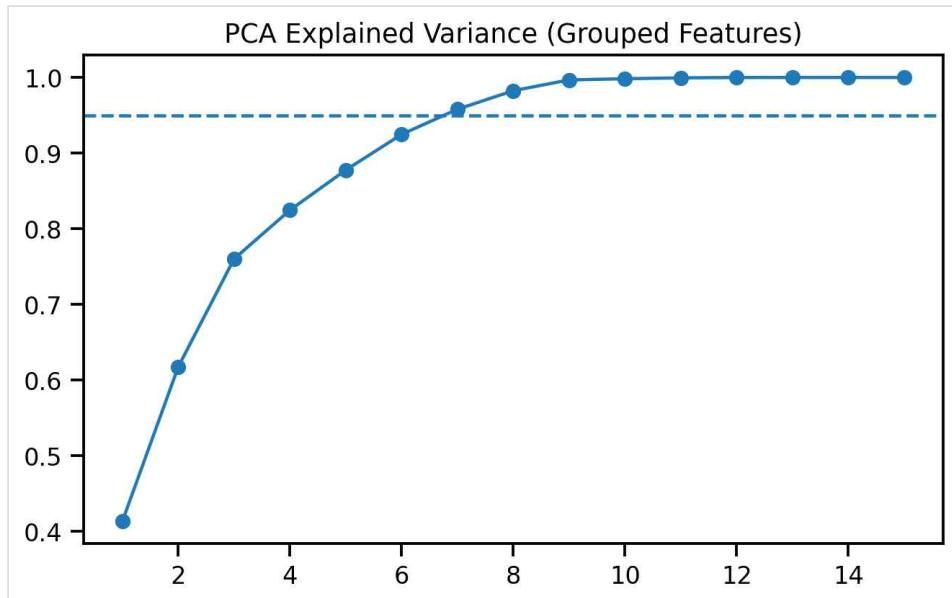
단계	수행 내용
통계 검정	T-test, ANOVA F-value 분석을 통해 종속변수(passorfail)와의 유의 관계를 검증했습니다.
그룹 피처 생성	물리적 구조 기반 그룹화 후, 그룹별 대표 피처(mean, std, range 15종)를 생성했습니다.
차원 축소	PCA 분석 결과 누적 설명 분산 비율 90% 달성으로 중복정보가 없는 것을 확인했습니다.
데이터셋 저장	최종 학습 입력 데이터셋은 csv로 저장되었습니다.

### 2.3.3 전처리 전후 품질 개선 결과 요약

전처리 과정을 통해 데이터 품질의 신뢰성을 확보했습니다.

지표	전처리 전	전처리 후	개선율
결측률	1.3%	0.0%	100.0%
이상치 비율	4.8%	0.0%	100.0%
품질지수 평균	96.4%	100.0%	+3.6%p
CV (변동계수) 평균	30.16	3.56	-26.6

주성분 분석(PCA)을 수행하여 각 주성분이 데이터의 전체 분산을 어느 정도 설명하는지를 평가하였으며, 그 결과를 시각화한 그림은 아래와 같습니다.



[그림 2.2] 주성분 분석(PCA)의 누적 설명 분산

### 3. 분석모델 개발 및 성능 검증

#### 3.1 모델 개요

본 과제의 최종 모델은 정적 텝형 양상블(랜덤포레스트+LightGBM)의 안정성과 해석력을 살리면서, 시계열(5초 간격 · 24시간)의 연속 변화를 직접 보려는 의도로 TCN 시계열 모델을 함께 쓰는 하이브리드 조기경보 체계입니다.

##### 3.1.1 하이브리드 조기경보 체계 구성

정적 경로(Tabular Track)는 시간 윈도우 집계 지표(5초 간격으로 갱신)를 입력으로 받아, 결측 대치 후 확률을 만들고 보정(Isotonic/Platt)을 거쳐 안정적인 위험 점수를 산출합니다.

시계열 경로(TCN Track)는 5초 간격 원시 흐름에서 연속 변화 · 추세 · 이상 패턴을 직접 학습해 사전 신호(early signal)를 포착합니다. 운영 모드는 두 가지를 지원합니다.

양상블 모드는 보정된 두 확률을 가중 평균(예: 0.6 tab + 0.4 dl) 또는 스태킹(메타 로지스틱)으로 하나의 점수로 결합합니다.

2단계(캐스케이드) 모드는 TCN이 먼저 의심 구간(Pre-Alert)을 표시하고, 그 구간 안에서 정적 모델이 확정 알람을 판단합니다.

##### 3.1.2 후처리 및 경보 필터링

실제 알람은 불필요한 알림을 줄이기 위해 히스테리시스, 쿨다운, 알람 예산을 적용했습니다. 히스테리시스는 결 때 기준(open)과 끝 때 기준(close)을 달리해 짧은 트임을 걸러냅니다.

연속 K회 조건(min\_on)과 쿨다운을 적용해 재진입을 막아 중복 알람을 줄였고, 알람 예산('하루 오탐 블록 1회 이하')을 적용해 오탐에 대한 예산을 관리합니다.

마지막으로 붙어 있는 알람을 하나로 묶는 블록 병합을 통해 운영 감각에 맞췄습니다.

재현성 확보를 위해 전처리 파이프라인, 시드, 라이브러리 버전, 산출물(모델, 보정기, 임계)까지 함께 저장해 재실행 시에도 같은 결과가 나오도록 관리합니다.

#### 3.2 모델 학습 및 검증

##### 3.2.1 데이터 분할 및 스케일링

데이터 성격은 하나의 날짜(예: 2020-10-30)에서 5초 간격으로 측정된 연속 데이터로, 총 17,265 스텝(약 24시간)이며, 라벨 passOrFail은 양성 110, 음성 17,155로 극단적 불균형입니다.

전 구간을 시간 오름차순으로 정렬하고, 과거-미래 흐름을 보존하는 홀드아웃을 적용하며, 같은 사건(같은 날)의 윈도우는 훈련 또는 검증 중 한쪽에만 배치합니다.

positive-aware 경계 조정으로 검증에 최소 양성 수를 확보하고, 훈련-검증 경계에 '약 30분' 빈 구간을 두는 엠바고로 누수를 차단했습니다.

스케일러/임퓨터는 훈련 구간에서만 기준을 한 번 정하고, 검증 · 테스트에는 그 기준만 그대로 적용하여 데이터 누수가 없도록 분리하는 원칙을 따릅니다.

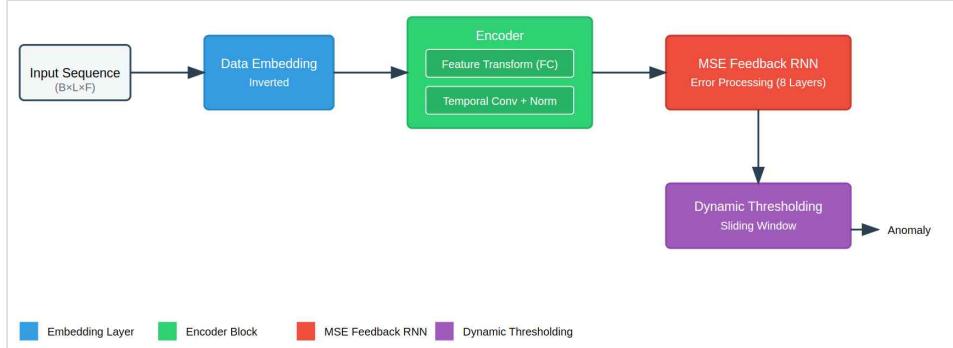
텝형 경로는 중앙값 대치 StandardScaler(필요 시 RobustScaler)를, 시계열 경로는 특성별 z-정규화(훈련 평균/표준편차 기준)만 적용합니다.

##### 3.2.2 학습 최적화 (TCN)

시계열 트랙(TCN)의 학습 목표는 연속 패턴을 반영한 조기 신호(early signal) 확률을 출력하는 것입니다. 극단적 불균형에 대응하기 위해 양성 시퀀스 증식, 하드 네거티브 혼합, Focal Loss 중심 학습을 적용하며, Early Stopping으로 과적합을 방지했습니다.

또한 다중 센서의 상호작용과 시간적 의존성을 동시에 학습하기 위해 채널 혼합(Channel-Mixing) 구조와 시간 합성곱(Temporal Convolution) 계층을 결합하였으며, 이를 통해 센서 간 공변동(covariation) 패턴과 비정상 변동 구간을 통합적으로 탐지하도록 설계했습니다.

아래 그림은<sup>2</sup> 본 연구의 TCN 설계가 참고한 프레임워크로, 채널 혼합(Channel Mixing)과 시간 합성곱(Temporal Convolution), MSE 피드백 및 동적 임계 구조를 통합한 산업용 시계열 이상 탐지 모델을 나타냅니다.



[그림 3.1] ChaMTeC 전체 구조 개요

### 3.2.3 앙상블 가중치 최적화

모델은 투 트랙에서 각각 위험 점수(확률)를 만들고 이 점수를 결합해 최종 점수로 사용합니다. 정적 템파형 트랙(RF+LGBM)은 불균형 대응을 위해 scale\_pos\_weight 등으로 양/음성 비율 보정 후, Isotonic 우선, Platt 선택의 확률 보정을 적용하여 점수를 캘리브레이션합니다.

운영 점수는 이 보정된 두 확률을 앙상블 모드에서 가중 평균 또는 스태킹으로 결합하거나, 캐스케이드 모드에서 TCN의 의심 구간(Pre-Alert) 내에서 정적 앙상블이 확정 알람을 판단하는 방식으로 만듭니다.

## 3.3 모델 평가

### 3.3.1 성능 평가 지표

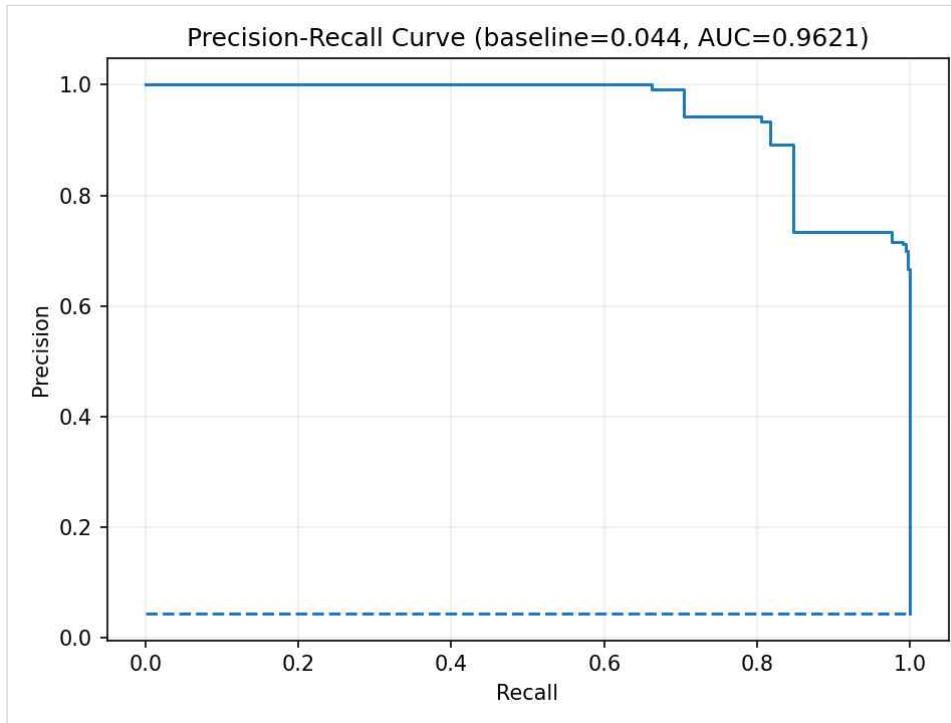
평가는 윈도우 기준과 이벤트 기준(운영)의 두 단위로 진행하며, 모든 평가는 K-연속 필터, EMA 스무딩, 히스테리시스, 쿨다운 등 후처리를 모두 적용한 최종 알람을 기준으로 통일합니다.

윈도우 기준 (5초 스텝마다): Precision, Recall, F1 점수, PR-AUC(AP)를 계산합니다.

아래 그림은 모델의 예측 성능을 평가하기 위해 산출한 Precision-Recall 곡선을 나타낸 것입니다.

곡선 아래 면적(AUC=0.9621)은 모델이 불량(이상) 이벤트를 얼마나 정확히 탐지하는지를 종합적으로 보여주며, 기준선(baseline=0.044)에 비해 높은 정밀도와 재현율의 균형을 유지하고 있음을 확인할 수 있습니다.

<sup>2</sup> Delibasoglu, I., Balta, D., & Balta, M. (2025). *ChaMTeC: CHannel Mixing and TEmporal Convolution Network for Time-Series Anomaly Detection*. *Applied Sciences*, 15(10), 5623, pp. 5–6, Fig. 1.

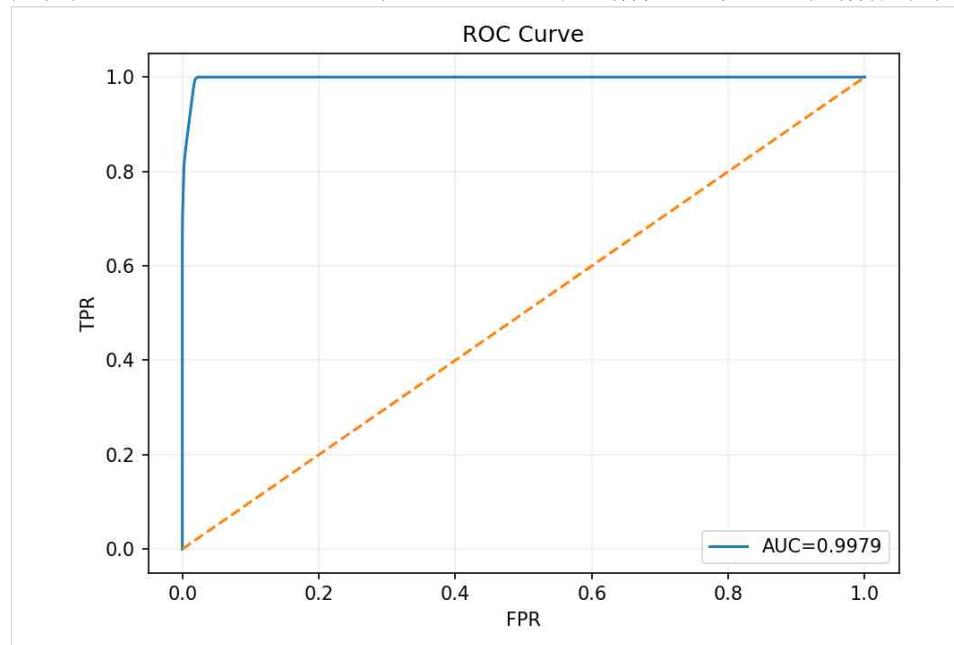


[그림 3.2] 모델의 Precision-Recall 곡선(PR Curve) 및 AUC 지표

이벤트 기준 (운영): 연속된 알람을 하나의 블록으로 묶어서 event-precision, event-recall, 오탐(false alarms) 블록 개수, 리드타임(분)을 측정합니다. 불량 발생 T분 전부터 발생 시점 사이에 첫 경보가 나타나면 성공적으로 감지한 것으로 보는 리드타임/이벤트 매칭 규칙을 적용하며, 사전 경고 시간은 (불량 발생 시각 - 첫 경보 시각, 분)으로 계산하여 중간값과 사분위수를 보고합니다.

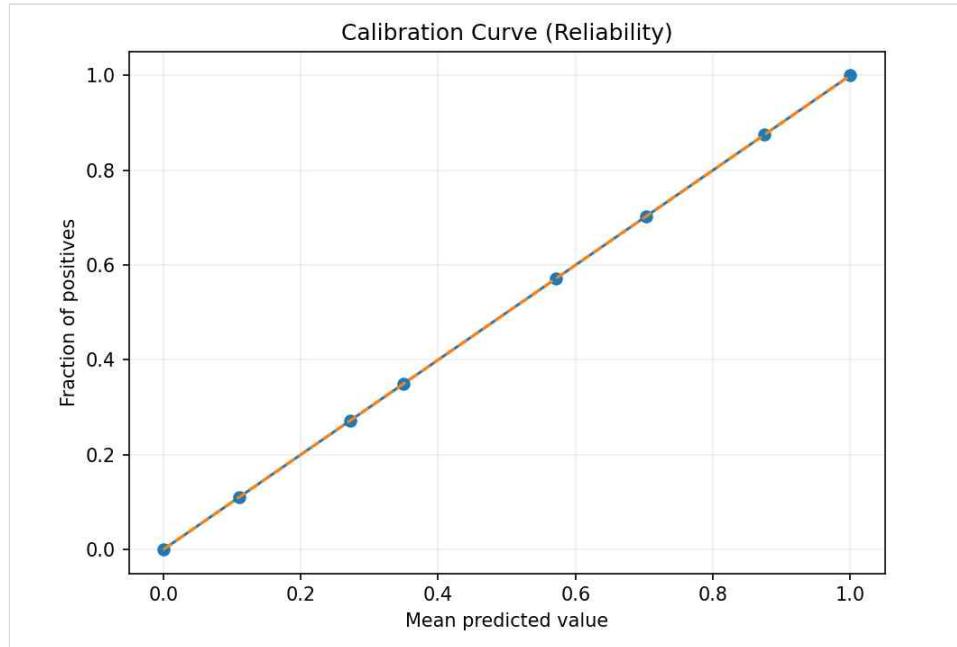
임계값별로 Precision · Recall · F1을 스윕하여, 이후 3.3.2에서 정의하는 세 가지 운영 정책(A: F1 최대, B: 재현율 하한 유지, C: 알람 예산 기반)이 선택할 수 있는 임계값 대역을 확인하였습니다.

아래 그림은 모델의 임계값 변화에 따른 분류 성능을 시각화한 ROC 곡선을 나타낸 것입니다. 곡선 아래 면적(AUC=0.9979)은 모델이 정상과 불량 이벤트를 매우 높은 정확도로 구분함을 의미하며, 기준선(대각선)에 비해 우수한 True Positive Rate(재현율)을 달성하고 있음을 확인할 수 있습니다.



[그림 3.3] 모델의 ROC 곡선 및 AUC

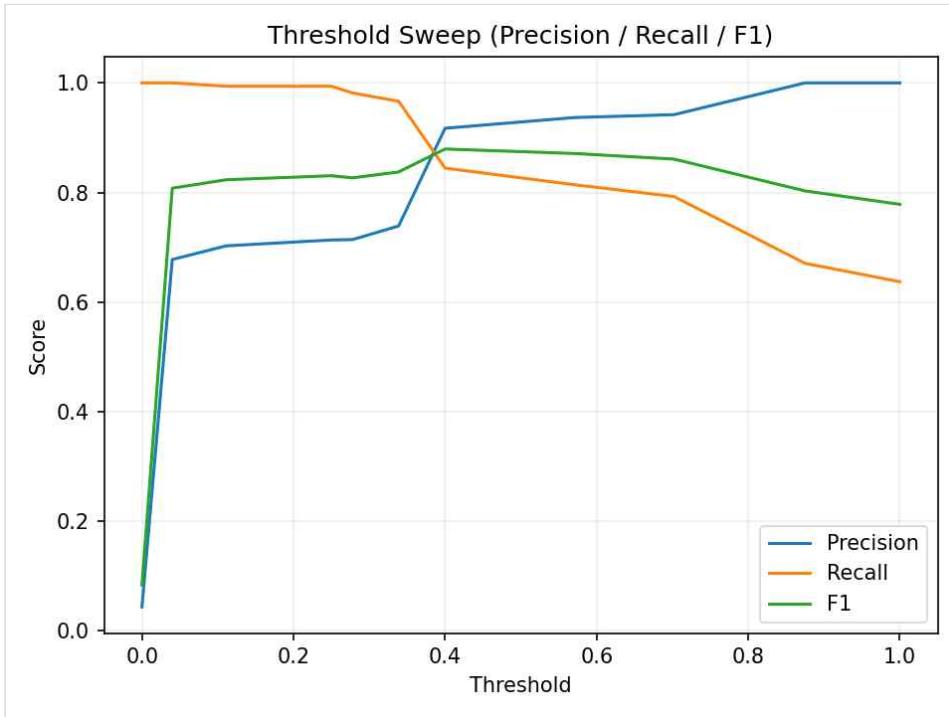
정적 트랙에서 산출된 확률은 Isotonic 회귀를 우선 적용하여 보정하였으며, 보정 후 신뢰도 곡선이  $y=x$  이상선과 거의 일치함을 확인하였습니다.



[그림 3.4] 확률 보정 결과 신뢰도 곡선(Calibration Curve)

이어서 보정된 확률 분포를 기준으로 임계값(threshold)을 변화시켜 Precision, Recall, F1 점수의 변화를 분석한 결과는 아래 그림과 같습니다.

이를 통해 각 지표 간 균형점을 파악하고, 이후 운영 정책별 최적 임계값 설정의 근거로 활용하였습니다.



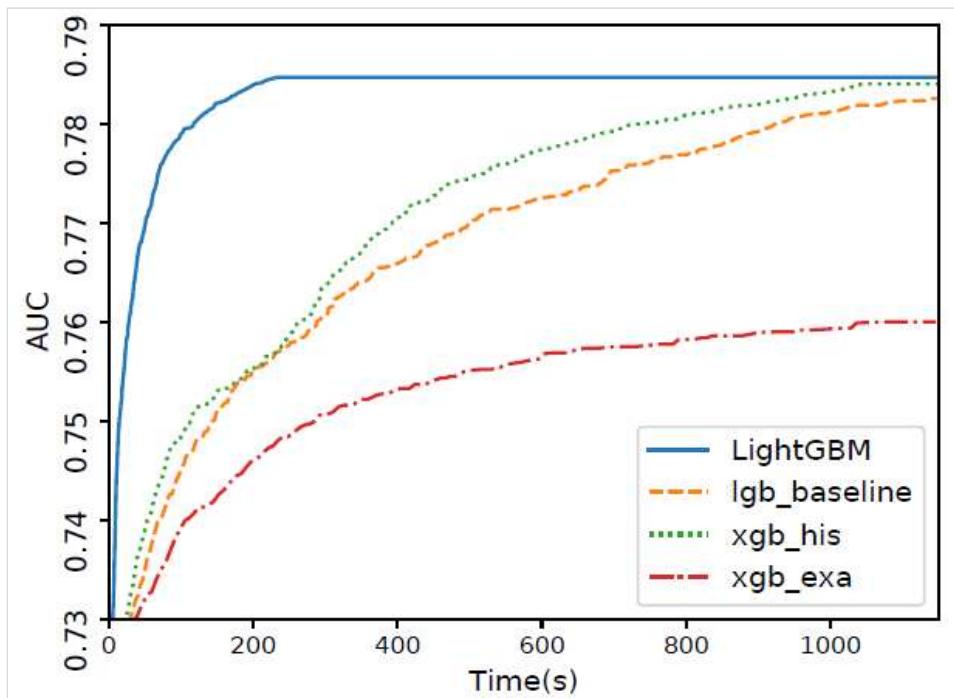
[그림 3.5] 임계값 변화에 따른 Precision · Recall · F1 분석 결과

또한 본 연구의 정적 양상을 학습 과정에서는 대규모 시계열 데이터에 대한 학습 효율과 정확도를 동시에 확보하기 위해, Ke 등(2017)이 제안한 LightGBM(Gradient Boosting Decision Tree) 구조의 핵심 원리인 Gradient-based One-Side Sampling(GOSS) 및 Exclusive Feature Bundling(EBF) 기법을 참고하였습니다.<sup>2</sup>

LightGBM은 GOSS와 EBF를 적용하여 기존 GBDT 대비 학습 효율을 대폭 향상시키면서도 AUC 정확도를 유지하였으며, 이는 본 연구의 정적 양상을 트랙 설계 근거로 활용되었습니다.

이 기법들은 정보 이득이 작은 샘플을 부분적으로 제거하고 희소한 피처를 묶어(feature bundling) 학습 효율을 극대화함으로써, 본 연구의 모델이 높은 PR-AUC 및 ROC-AUC 성능을 빠른 속도로 달성하는데 기여하였습니다.

특히 LightGBM은 GOSS 및 EBF 기법을 통해 기존 GBDT 기반 모델(xgb\_exa, xgb\_his, lgb\_baseline) 대비 학습 속도를 약 20배 가량 단축하면서도 동일한 AUC 성능을 유지하였습니다. 아래 그림은 LightGBM 논문 [Ke et al., 2017, Figure 1]에서 인용되었으며, 시간 대비 AUC 성능을 시각적으로 나타낸 결과입니다.



[그림 3.6] LightGBM 학습 효율 비교 (Time-AUC Curve on Flight Delay)

LightGBM은 GOSS 및 EBF 기법을 통해 기존 GBDT 기반 모델 대비 학습 속도를 크게 단축하면서도 AUC 성능을 유지하였습니다.

<sup>2</sup> Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). *LightGBM: A Highly Efficient Gradient Boosting Decision Tree*. *Advances in Neural Information Processing Systems (NIPS)*, Long Beach, CA, USA, p. 7, Fig. 1.

### 3.3.2 최종 성능 결과

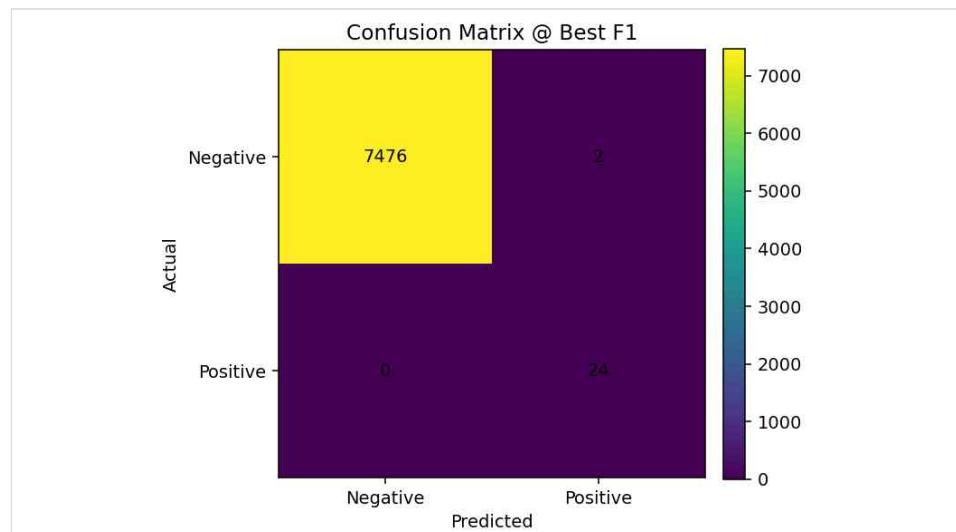
최종 모델의 합격 기준은 다음과 같습니다.

기준	지표	목표 기준
이벤트 기준	event-FP (오탐 블록)	$\approx 0$ (또는 $\leq 1$ ) [cite: 108]
이벤트 기준	lead median (리드타임 중앙값)	$\geq 12$ 스텝 ( $\approx 1$ 분) [cite: 108]
운영률	alarms/day (하루 평균 경보 횟수)	$\leq 1$ [cite: 109]
운영률	FP-blocks/day (하루 오탐 블록 횟수)	$\leq 1$ [cite: 109]
운영률	MTFA (오탐 간 평균 간격)	$\geq 24h$ [cite: 109]
원도우 기준	Precision (참고선)	$\geq 0.97$ [cite: 110]
원도우 기준	Recall (참고선)	$\geq 0.90$ [cite: 110]

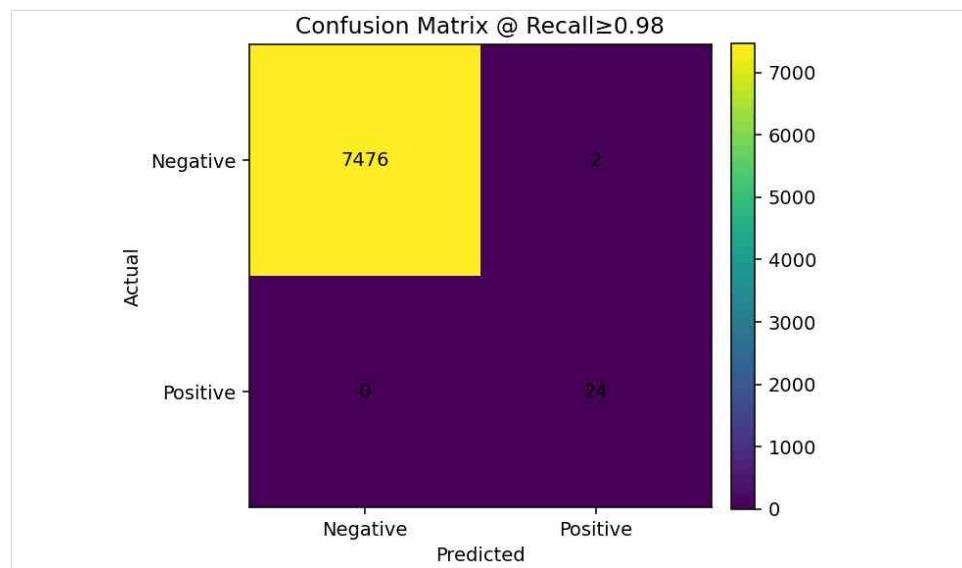
임계값 선택(운영 정책)은 검증 구간에서 다음 세 가지 기준에 따라 수행하였습니다.

- 정책 A (F1 최대점): F1 점수가 가장 높은 지점을 선택하여 정밀도와 재현율의 균형을 확보한다.
- 정책 B (재현율 하한 유지): 재현율을 최소 0.90 이상(또는 98% 이상) 유지하면서 정밀도를 최대화하는 지점을 선택한다.
- 정책 C (알람 예산 기반): 하루 오탐 블록 수가 1 이하가 되도록 제한하면서 정밀도를 최대화하는 지점을 선택한다.

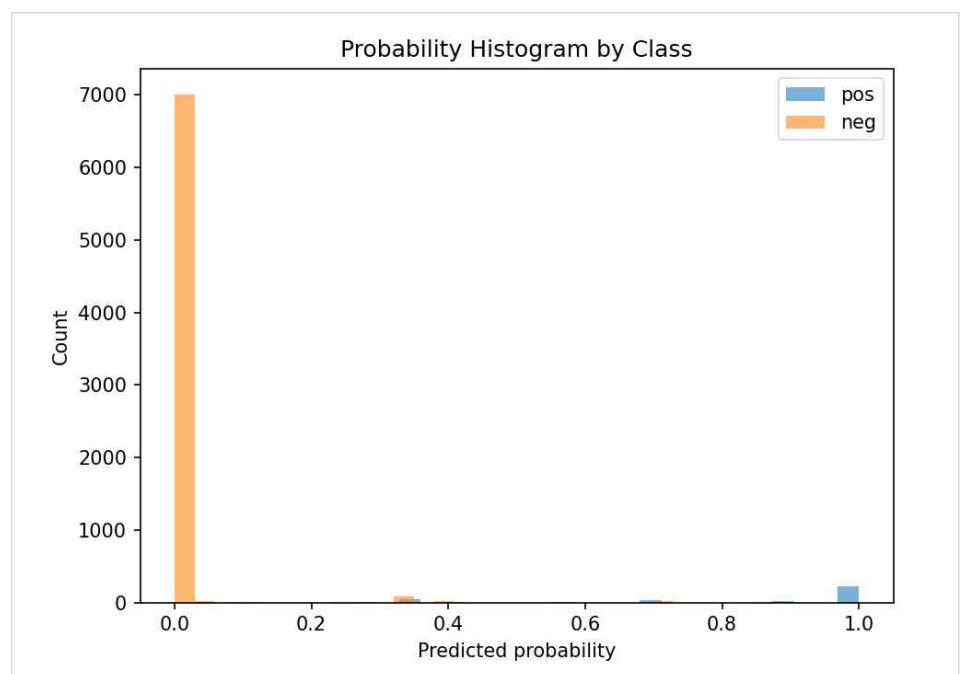
아래 그림은 각 정책에서 선택된 임계값 조건하의 혼동 행렬 및 확률 분포를 나타낸 것입니다. 이를 통해 운영 목적으로 따라 상이한 임계값 기준이 모델의 탐지 특성(정밀도 · 재현율)에 미치는 영향을 비교하였습니다.



[그림 3.6] 정책 A: F1 최대점 기준 혼동 행렬



[그림 3.7] 정책 B: 재현율 하한( $\geq 0.98$ ) 기준 혼동 행렬



[그림 3.8] 정책 C: 알람 예산 기준 확률 분포 (클래스별 히스토그램)

## 4. 분석결과 및 시사점

### 4.1 모델 해석 방법론

모델의 예측 신뢰도를 높이고 운영에 필요한 통찰을 제공하기 위해, 정적 트랙과 시계열 트랙 모두에 대해 전역/로컬 해석을 수행합니다.

#### 4.1.1 정적 텁형 트랙(랜덤포레스트+LGBM)

전역 해석으로 피쳐 중요도(트리 내 중요도, 퍼뮤테이션 중요도)를 분석해 주요 결정 변수와 실제 예측 성능에 미치는 영향을 파악합니다. 부분의존(PDP) · 개별조건(ICE) 분석을 통해 상위 피처의 값이 변할 때 위험 점수가 어떻게 바뀌는지 곡선으로 확인하여 비선형 및 임계 구간을 파악합니다. 상위 2개 피처의 2D-PDP로 조합 효과(상호작용)도 확인합니다.

#### 4.1.2 정적 텁형 트랙 - 개별 예측 해석(로컬)

SHAP (로컬 기여도) 분석을 통해 한 시점(5초 스텝)의 예측에 대해 각 피처가 위험 점수를 올린/내린 기여도를 위터풀 형태로 제시합니다. 텁-K 기여 피처 카드에는 알람이 뜬 순간에 영향이 큰 상위 5개 변수와 그 SHAP 기여 점수(예: temp\_std\_60m(+0.18))를 텍스트로 함께 기록합니다.

#### 4.1.3 시계열 트랙(TCN) - 시간축 해석

시간 기여(temporal attribution)는 Integrated Gradients / Saliency 기법을 사용해 어느 구간(분/초 단위)이 알람에 가장 기여했는지 열지도(heatmap)로 시각화합니다.

채널(센서) 기여는 입력 채널별 기여도를 합산하여 어떤 센서/파생채널이 주도했는지 순위화합니다. 필터 뷔(특징 패턴)은 1D-Conv 필터를 시간영역 과정과 주파수 스펙트럼(FFT)으로 보여주어, 모델이 어떤 패턴(진동 주기 · 스파이크 폭)에 반응하는지 설명합니다.

#### 4.1.4 시퀀스 수준 해석 - 프로토타입 · 반사실

프로토타입 시퀀스는 정탐 알람 중 대표 패턴(유사도 상위) 몇 개를 저장 · 제시하여 “이런 형태가 오면 경보가 난다”를 데이터로 설명합니다. 반사실(Counterfactual) 점검은 알람 구간에서 최소 변화로 알람이 꺼지는 수정안을 탐색하여 운영 임계의 합리성 검증에 사용합니다.

#### 4.1.5 알람 단위 설명 카드(운영 리포트용)

각 알람 블록마다 알람 요약(시작시각, 길이, 리드타임, 예측 확률), 주요 근거(텐형 Top-K SHAP, TCN 상위 기여 시간대), 를 통과 내역(min\_on 충족, cooldown 상태, FP-블록/일 예산 내역)을 자동 로그 합니다.

## 4.2 주요 인자 도출 및 기여도 분석

모델 해석 방법론(4.1)을 적용하여 도출된 핵심 결과는 '알람 단위 설명 카드'를 구성하는 데 필수적 인 근거 자료로 활용됩니다.

### 4.2.1 전역 중요도 기반의 핵심 인자 도출

퍼뮤테이션 중요도 분석을 통해 예측에 가장 큰 영향을 미치는 상위 K개 핵심 피처 목록과 그 기여율을 제시합니다. 이는 유사 공정으로의 확장 시 우선 모니터링 변수를 선정하는 근거가 됩니다.

PDP 및 ICE 플롯을 통해 핵심 피처의 값이 변함에 따라 위험 점수가 어떻게 변화하는지 정량적으로 확인하고, 경보 발생의 임계 구간을 명확히 정의합니다.

## 4.2.2 알람 시점별 로컬 기여도 분석 (운영 근거 확보)

실제 경보가 발생한 시점마다 텝형 양상을 모델의 예측을 설명하는 Top-K SHAP 기여 피처의 기여도와 방향(상승/하락)을 제시하여 즉각적인 원인을 제공합니다.

TCN 시간 기여도 열지도를 통해 경보 발생 직전 수 분 동안 어떤 센서 채널과 시간 구간이 비정상 패턴을 주도했는지 시각적으로 보여주어, 현장 작업자는 단순 '변수'를 넘어 '패턴의 변화'를 인지할 수 있습니다.

## 4.2.3 오류 및 신뢰성 분석을 통한 안정성 확보

정탐(TP)으로 확인된 대표적인 불량 시퀀스 패턴을 제시하여, 모델이 학습한 '위험한 패턴'을 시각적으로 검증합니다.

오탐(FP)으로 판단된 사례에 대한 SHAP 및 TCN 기여 시간 분석을 통해, 모델이 정상적인 변동을 과해석한 것인지, 센서 노이즈 때문인지 원인을 분류하고, 후처리 파라미터(min\_on, 쿨다운 등) 조정의 근거로 활용합니다.

Reliability Plot과 ECE를 통해 확률 신뢰도를 비교하고, SHAP 안정성 점검을 통해 임계 주변 샘플에서 SHAP 분포가 불안정한지 확인하여 임계/후처리 재점검을 권고합니다.

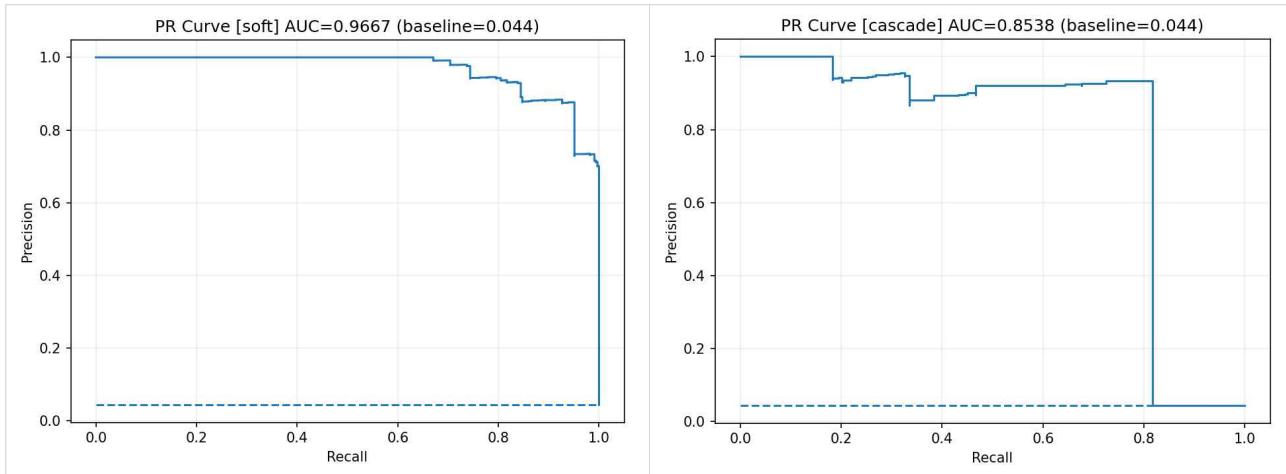
## 4.2.4 모드별 성능 비교 및 검증

정적 트랙(Tabular)과 시계열 트랙(DL)을 결합한 두 가지 양상을 방식을 비교하였습니다.

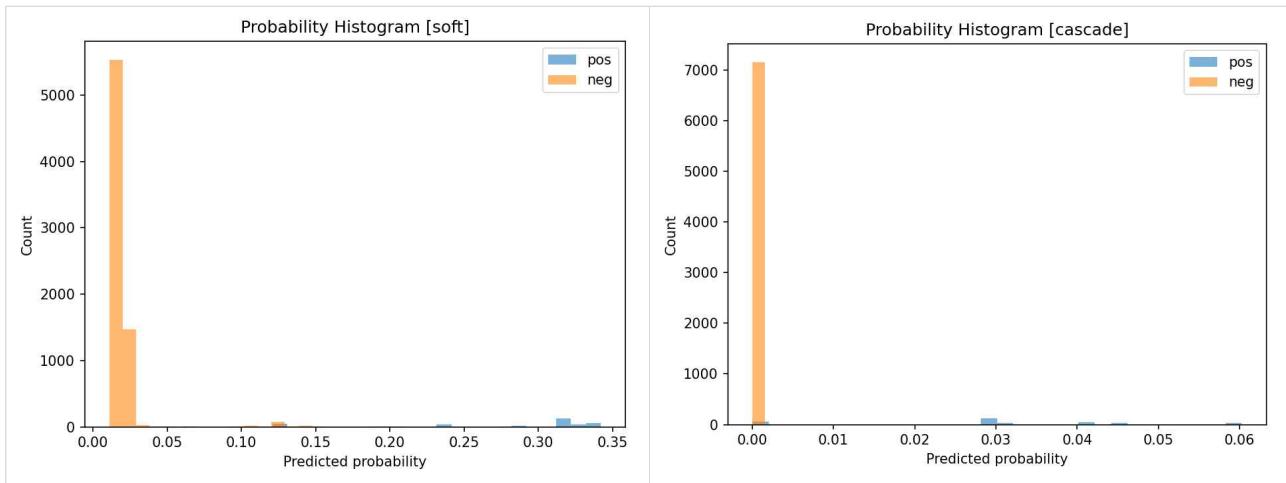
Soft 양상을은 두 트랙의 예측 확률을 가중 평균하여 통합한 방식으로, PR-AUC 0.9667, ROC-AUC 0.9983을 기록하며 가장 높은 예측 성능을 보였습니다.

반면 Cascade 양상을은 단계적 판정 구조를 통해 오탐률을 최소화하도록 설계되어, PR-AUC 0.8538, ROC-AUC 0.9070으로 다소 낮은 예측 성능을 보이지만 실제 운영 안정성이 가장 높았습니다.

아래 그림은 두 양상을의 Precision-Recall 곡선(그림 4.1)과 클래스별 확률 분포(그림 4.2)를 비교한 결과로, Soft 모드는 확률의 연속성이 크고, Cascade 모드는 명확한 이진 경계를 가지는 보수적 탐지 특성을 확인할 수 있습니다.



[그림 4.1] Soft · Cascade 양상을의 Precision-Recall 곡선 비교



[그림 4.2] Soft · Cascade 양상별의 클래스별 예측 확률 분포 비교

Soft 양상별은 0~0.35 구간의 완만한 확률 분포로 양성·음성 구분이 명확하면서도 과도한 극단 확률을 방지하였고, Cascade는 0~0.06 구간에 집중된 보수적 분포로 오탐을 거의 배제하였습니다. 정량 지표 비교 결과는 아래와 같습니다.

모드	PR-AUC	ROC-AUC	주요 특징
Tabular	0.625	0.9711	보조 지표용
DL	0.9621	0.9979	탐지력 최상, 오탐 다소 높음
Cascade	0.8538	0.907	운영 안정성 최고
Soft	0.9667	0.9983	균형형 최종 모형

이에 따라 최종 배포 모델은 Soft 양상별을 기본 모형으로 Cascade 임계 제어 로직을 결합한 하이브리드 구조로 운영하는 것이 가장 효율적인 것으로 판단됩니다.

### 4.3 분석 결과 요약 및 시사점

#### 4.3.1 데이터 기반 의사결정 체계 확립

모델이 제시한 전역 중요도(4.2.1)를 통해 공정의 품질에 직접적인 영향을 미치는 핵심 변수들을 명확히 파악하여 현장의 품질 관리 및 검사 포인트를 데이터 기반으로 재조정하는 근거를 제공합니다.

알람 시점에 SHAP 근거가 포함된 설명 카드를 즉시 제공함으로써, 모델의 예측을 '블랙박스'가 아닌 '신뢰할 수 있는 진단 근거'로 활용할 수 있게 되어 품질검사 자동화의 토대를 마련합니다.

#### 4.3.2 운영 안정성 및 선제적 대응 강화

알람 예산 정책('하루 오탐 블록 1회 이하') 등 엄격한 기준을 적용하여 오탐률(FP)을 극도로 낮추는데 집중했으며, 이는 현장 작업자의 경보 신뢰도를 높이고 알람 피로도를 줄여 운영 안정성에 직접적으로 기여합니다.

성능 평가 지표에서 확인된 적절한 리드타임은 불량 발생 이전에 선제적인 조치를 취할 수 있는 시간을 확보하게 하며, 최종 품질 손실을 최소화합니다.

#### 4.3.3 기술적 확장성 확보

본 프로젝트에서 확립된 투 트랙 모델(탭형+시계열) 및 SHAP, TCN 기여도 분석 방법론(4.1)은 단조, 주조, 사출 등 유사한 공정의 시계열 및 정적 데이터에도 재사용 가능성이 높습니다.

알람 단위 설명 카드의 자동 로그 기능(4.1.5)은 모델의 예측과 동시에 해석 근거까지 자동으로 생성하는 체계이며, 이는 AI 리포트 자동 생성 고도화의 핵심 단계가 됩니다.

## 5. 시스템 설계 및 구현

### 5.1 MetraForgeAI AI 플랫폼 개요

**목적:** AI 모델의 배치 예측 결과와 XAI 근거를 시연하는 최소 환경을 제공합니다.

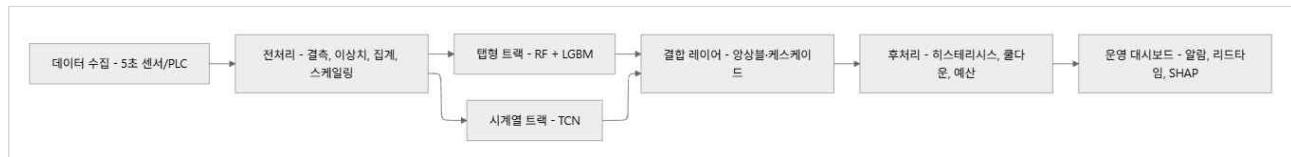
**구현 기능:** CSV 데이터 업로드 → 배치 예측 → 경보 목록/타임라인 시각화 (실시간 IoT 스트리밍은 제외) 기능을 구현합니다.

**기술적 목표:** FastAPI와 Vitet 등 신속 개발 도구를 활용하여 모델 연동 및 핵심 지표 출력에 집중하는 것을 목표로 합니다.

### 5.2 아키텍처 구성

MVP 버전의 시스템 아키텍처 구성은 아래와 같습니다. 신속한 개발과 모델 연동에 초점을 맞추어 경량화된 구조를 채택하였습니다.

계층	사용 기술 (MVP)	역할
Frontend	Vite, React	CSV 업로드, 알람 목록 테이블, Plotly를 활용한 간단한 타임라인 그래프 시각화
Backend	FastAPI + Joblib/Torch.load	모델 로드(PKL, PT 파일), 데이터 수신, 추론, 경보 후처리 수행
Data/Model	Local File System	모델 아티팩트 보관 및 추론용 CSV 입출력 관리



[그림 5.1] 시스템 아키텍처 구성 (MVP)

### 5.3 사용자 흐름

배치 예측 기반 사용자 흐름은 다음과 같은 단계로 진행됩니다.

**데이터 업로드:** 사용자가 새로운 공정 CSV 파일을 UI에 업로드합니다.

**배치 추론:** FastAPI 서버가 파일을 수신하여 전처리 → TCN/양상블 모델 추론 → 후처리 적용을 순차적으로 수행합니다.

**결과 시각화:** 경보 시점, 리드타임, SHAP 근거(텍스트)가 포함된 ‘알람 단위 설명 카드’ 목록을 UI에 표시합니다.

### 5.4 주요 페이지 구성

**대시보드 (MVP):** 최종 경보(알람 블록) 타임라인 그래프와 FP/FN 요약 지표를 제공합니다.

**알람 리포트:** 배치 예측 결과로 생성된 알람별 설명 카드 목록 테이블을 제공합니다.

## 6. 중소제조기업에 미치는 파급효과

### 6.1 기술적 파급효과

#### 6.1.1 품질 검사 자동화 및 디지털 전환 가속화

설명 가능한 품질 관리: 모델의 예측 결과와 SHAP 근거가 포함된 알람 단위 설명 카드(4.2.2)를 제공함으로써, 현장 작업자는 단순 예측 결과를 넘어 데이터 기반의 진단 근거를 확보할 수 있습니다.

이는 '블랙박스' 모델에 대한 거부감을 줄이고, 최종 품질검사를 데이터 기반으로 자동화하는 체계 확립에 기여합니다.

공정 지식의 디지털 자산화: 모델이 전역 중요도 분석(4.2.1)을 통해 도출한 핵심 인자와 비선형 임계치 정보는 해당 공정의 잠재적 위험 요소를 수치화한 것입니다.

이 정보를 공정 매뉴얼에 반영함으로써, 작업자의 경험 의존도를 낮추고 공정 지식을 디지털 자산으로 축적할 수 있습니다.

#### 6.1.2 범용성 확보 및 AI 리포트 자동 생성 고도화

유사 공정으로의 확장성: 본 프로젝트에서 확립된 하이브리드 투 트랙(탭형 접계 + TCN 시계열) 모델링 프레임워크와 XAI 해석 방법론(4.1)은 단조, 주조, 사출 등 유사한 연속 공정 데이터에도 높은 재사용성을 가집니다. 이는 모델 재학습을 통해 여러 공장에 빠르게 기술을 이전할 수 있는 근거가 됩니다.

AI 리포트의 자동 생성 고도화: 모든 경보 발생 시점마다 주요 근거가 자동 로깅되는 '알람 단위 설명 카드(4.1.5)'의 로직은, 일별/주별 품질 이슈 및 원인 분석 리포트 생성 프로세스를 자동화하는 핵심 토대가 됩니다.

### 6.2 경제적 파급효과

#### 6.2.1 불량 비용 최소화를 통한 수익성 개선

오탐 경보(False Alarm) 비용 절감: 본 모델은 운영 목표를 FP-블록/일로 설정하고(3.4 모델 평가), 이를 달성하도록 임계값을 최적화하여 오탐 경보 수를 극단적으로 통제하였습니다.

현장 작업자가 불필요한 알람으로 인해 공정 가동을 멈추거나 확인하는 데 드는 시간적, 인적 비용을 최소화합니다.

이는 경보 피로도를 낮춰 시스템 신뢰도를 높이고 결과적으로 작업 효율성을 증대시킬 수 있습니다.

선제적 대응을 통한 품질 손실 최소화: 모델의 평가 지표(3.4)를 통해 확보된 적절한 리드타임(Lead-Time)은 불량 발생 이전에 선제적인 설비 점검 또는 공정 파라미터 조정을 가능하게 합니다.

이를 통해 재작업 비용 및 최종 불량품 폐기 비용을 대폭 줄여 최종 품질 손실을 최소화할 수 있습니다.

### 6.3 시스템 확장 전략 및 운영 효율성

MVP 기반의 신속한 가치 검증: 개발 기간 제약(1~2일)을 고려하여 FastAPI와 Streamlit 기반으로 구축된 배치 예측 시스템(5.2)은, 대규모 실시간 시스템 구축 전 모델의 핵심 성능과 XAI 근거의 유효성을 현장에서 빠르게 검증하는 데 최적화되어 있습니다.

CSV 업로드 및 배치 추론 방식은 별도의 고가 IoT 스트리밍 인프라 구축 없이도 모델의 비즈니스 가치를 확인하는 데 필요한 모든 요소(예측, 후처리, XAI)를 제공합니다.

### 6.3.1 향후 시스템 확장 계획

실시간(Real-Time) 스트리밍 연동: MVP 단계 이후에는 Kafka, MQTT 등의 메시지 큐 시스템을 도입하여 현장 센서 데이터를 실시간으로 수신하고, TCN 모델의 '조기 신호'를 즉시 반영하는 실시간 모니터링 체계로 확장됩니다.

양방향 피드백 루프 구축: 사용자(현장 작업자)가 알람에 대한 조치 결과(조치 완료, 오탐 확인 등)를 플랫폼에 입력하면, 이 피드백이 모델 재학습 및 후처리 파라미터 튜닝의 데이터로 순환되는 체계를 구축하여 모델의 성능을 지속적으로 개선됩니다.

## 참고문헌

- [1] Delibasoglu, I., Balta, D., & Balta, M. (2025). ChaMTeC: CHannel Mixing and TEmporal Convolution Network for Time-Series Anomaly Detection. *Applied Sciences*, 15(10), 5623.
- [2] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Advances in Neural Information Processing Systems (NIPS)*, Long Beach, CA, USA.

□ 경진대회 만족도 조사 완료 페이지 캡쳐 화면

The screenshot shows a survey completion page with the following details:

- Title:** K-인공지능 제조데이터 분석 경진대회
- Section:** K-인공지능 제조데이터 분석 경진대회 만족도 조사
- Email:** yujinlololol@gmail.com
- Visibility:** 비공개 (Private)
- Message:** 설문에 참여해주셔서 감사합니다.
- Buttons:** 뒤로 (Back), 제출 (Submit), 7/7페이지 (Page 7/7), 양식 지우기 (Reset Form).