

전력 설비 다변량 시계열 이상 진단을 위한 시계열 분해 및 변동성 게이팅 기반 트랜스포머

(A Decomposition and Variability-Aware Gating Transformer for Multivariate Time-Series
Anomaly Diagnosis in Power Facilities)

오승민*, 김진술**

(Seungmin Oh, Jinsul Kim)

요약

전력 설비에서 수집되는 다변량 시계열 데이터는 설비 상태 진단에 중요한 역할을 한다. 그러나 기존 Transformer 기반 모델은 추세와 계절성이 혼재된 신호를 단일 표현 공간에서 학습함으로써 서로 다른 시간적 특성을 충분히 반영하지 못하는 한계를 가진다.

본 논문에서는 시계열 분해와 변동성 기반 게이팅 메커니즘을 통합한 Transformer 구조를 제안한다. 제안 모델은 입력 시계열을 추세 성분과 계절성 성분으로 분리하고, 변수별 시간적 변동성을 기반으로 중요도를 동적으로 조절하며, 이중 경로 인코더를 통해 각 성분을 독립적으로 학습한다.

전력 설비 데이터를 이용한 실험 결과, 제안 모델은 정확도 95.1%를 달성하였으며, 기반 모델인 iTransformer의 성능인 92.5%보다 높은 성능을 보였다. 또한 시계열 분해 또는 게이팅 모듈을 제거할 경우 정확도가 각각 85.2%와 84.6%로 감소하여, 두 구성 요소가 성능 향상에 핵심적인 역할을 함을 확인하였다.

■ 중심어 : 다변량 시계열 ; 이상 진단 ; 트랜스포머 ; 시계열 분해 ; 게이팅 메커니즘

Abstract

Multivariate time-series data collected from power facilities play a critical role in equipment condition diagnosis. However, existing Transformer-based models learn mixed temporal patterns within a unified representation space, which may limit their ability to capture distinct trend and seasonal characteristics.

This paper proposes a Transformer architecture that integrates series decomposition and a variability-aware gating mechanism for multivariate time-series anomaly diagnosis in power facilities. The proposed model separates trend and seasonal components, dynamically adjusts variable importance based on temporal variability, and employs dual-path encoders to independently learn distinct temporal patterns.

Experimental results on real-world power facility datasets demonstrate that the proposed model achieves an accuracy of 95.1%, outperforming the baseline iTransformer model, which achieves 92.5%. When either the decomposition module or the gating mechanism is removed, the accuracy decreases to 85.2% and 84.6%, respectively, indicating that both components play a crucial role in performance improvement.

■ keywords : Multivariate Time Series ; Anomaly Diagnosis ; Transformer ; Series Decomposition ; Gating Mechanism

I. 서론

전력 설비 현장의 스마트화로 인해 전력량, 전압, 온도 등 복합 물리 데이터가 다변량 시계열

* 정회원, 전남대학교 지능전자컴퓨터공학과

** 정회원, 전남대학교 지능전자컴퓨터공학과

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-지역지능화혁신인재양성사업의 지원을 받아 수행된 연구임 (IITP-2026-RS-2022-00156287)

이 논문은 2026년 박사 학위논문 일부 수정·보완한 것임

접수일자 : 2026년 02월 20일

수정일자 : 2026년 03월 10일

게재확정일 : 2026년 03월 13일

교신저자 : 김진술 e-mail : jsworld@jnu.ac.kr

(Multivariate Time-Series, MTS) 형태로 실시간 수집되고 있다. 이러한 데이터는 설비 상태 감시 및 예지보전(Predictive Maintenance)의 핵심 자산으로 활용되며, 특히 고장 발생을 사전에 탐지하기 위한 이상 진단 기술은 전력 설비 운영의 경제성과 안전성을 결정짓는 중요한 요소이다[1-4].

최근 Transformer 기반 모델은 자기 주의(Self-Attention) 메커니즘을 통해 변수 간 상관관계를 효과적으로 학습하며 우수한 성능을 보이고 있다. 특히 시계열 데이터를 변수 단위로 투영하는 역전된(Inverted) 아키텍처가 제안되었으며, 이는 각 변수를 독립적인 토큰으로 처리함으로써 기존 타임스텝 중심 모델의 구조적 한계를 보완하였다.

그러나 이러한 역전형 모델 역시 입력 신호에 내재된 추세(Trend)와 계절성(Seasonality)이라는 상이한 시간적 특성이 단일 잠재 공간에서 중첩되어 학습된다는 한계를 지닌다. 이는 복합적인 변동성을 가진 전력 설비 데이터에서 성분별 세부 동특성을 충분히 분리하지 못하게 하며, 모든 변수를 동일한 가중치로 처리하는 구조는 특정 이상 징후와 직결된 핵심 변수에 대한 모델의 탐지 민감도를 저하시킬 수 있다.

이에 본 논문에서는 iTransformer의 역전형 토큰 구조를 기반으로, 시계열 분해(Series Decomposition)와 변동성 인식 게이팅(Variability-Aware Gating)을 통합한 동적 적응형 트랜스포머 아키텍처를 제안한다. 제안 모델은 입력 신호를 추세와 계절성 성분으로 명시적으로 분해하여 이중 경로 인코더(Dual-Path Encoder)를 통해 독립적으로 학습하며, 변수별 변동성을 정량화하여 정보 흐름을 동적으로 조절함으로써 진단 성능을 향상시킨다. 실제 대규모 전력 설비 데이터를 활용한 실험 결과, 제안 모델이 최신 기법인 iTransformer 대비 향상된 진단 정확도를 보임을 확인하였다.

본 연구의 주요 기여는 다음과 같다.

첫째, 다변량 시계열의 역전형 임베딩 구조를 기반으로 시간적 위계 특성을 반영한 분해 기반 Transformer 구조를 제안하였다.

둘째, 변수별 변동성을 활용한 게이팅 메커니즘을 통해 데이터의 동적 특성에 적응적인 변수 중요도 학습 방안을 제시하였다.

셋째, 실제 산업 전력 설비 데이터를 활용한 실험을 통해 제안 모델의 유효성을 검증하였다.

II. 관련 연구

다변량 시계열 이상 진단을 위한 연구는 전통적인 통계 기반 방식에서 딥러닝 기반 모델로 빠르게 확장되고 있다. 초기 연구에서는 ARIMA와 같은 자기회귀 모델이나 이동 평균 기반 기법을 활용하여 평균 및 분산 변화에 기반한 이상을 판별하였다. 그러나 이러한 방법은 비선형성과 복합적인 장·단기 패턴을 동시에 가지고 있는 산업 데이터의 특성을 충분히 반영하기에는 한계가 있다.

기반 접근은 크게 CNN, MLP/Linear, 그리고 Transformer 계열 모델로 구분할 수 있다. CNN 기반 모델은 합성곱 연산을 통해 국소적 패턴을 효과적으로 추출하며 단기적 변동이나 스파이크 형태의 이상 탐지에 강점을 가진다. 다중 스케일 합성곱 구조를 활용한 모델도 제안되었으나, 정적인 필터 기반 특성 추출에 의존한다는 한계가 있다[5-7].

MLP 및 Linear 기반 모델은 구조가 단순하고 학습 효율이 높으며 일부 장기 시계열 예측 과제에서 경쟁력 있는 성능을 보였다. 특히 추세와 잔차 성분으로 분해한 뒤 선형 변환을 적용하는 방식은 안정적인 결과를 제공하였다. 그러나 변수 간 복잡한 상호의존성을 충분히 반영하지 못하는 한계가 존재한다[8, 9].

최근에는 Transformer 기반 모델이 다변량 시계열 분석의 주요 흐름으로 자리잡고 있다. 자기 주의(Self-Attention) 메커니즘은 시계열 내 변수 간 관계를 직접적으로 학습할 수 있어 장기

의존성과 상호작용 모델링에 효과적이다[10]. 특히 변수 차원을 토큰 단위로 구성하는 역전형(Inverted) 구조가 제안되었으며, 이는 변수 간 상관관계를 보다 직접적으로 학습할 수 있도록 설계되었다[11-16].

그러나 기존 Transformer 기반 시계열 모델들은 추세와 계절성 성분을 명시적으로 분리하지 않고 단일 표현 공간에서 함께 학습하는 구조적 한계를 가진다. 또한 변수별 중요도를 동적으로 조절하는 설계는 충분히 고려되지 않은 경우가 많다. 이러한 한계는 전력 설비와 같이 복합적인 시간적 특성을 가진 데이터에서 모델의 표현력을 제약할 수 있다.

이에 본 연구는 iTransformer 구조를 기반으로 시계열 분해와 변동성 기반 게이팅을 통합함으로써, 시간적 위계 특성과 변수 중요도를 동시에 반영하는 학습 구조를 제안한다.

III. 제안 모델

본 연구에서는 전력 설비 다변량 시계열 이상 진단을 위해 시계열 분해와 변동성 기반 게이팅 메커니즘을 통합한 Transformer 구조를 제안한다. 제안 모델은 추세와 계절성이 혼재된 산업 시계열의 시간적 위계를 명시적으로 반영하고, 변수별 정보 기여도를 동적으로 조절할 수 있도록 설계되었다.

제안 모델은 세 가지 핵심 구성 요소로 이루어진다. 첫째, 입력 시계열을 장기 추세 성분과 단기 변동 성분으로 분리하는 시계열 분해 모듈이다. 둘째, 변수별 시간적 변동성을 기반으로 정보 중요도를 산출하는 게이팅 메커니즘이다. 셋째, 분리된 두 성분을 독립적으로 학습하는 이중 경로 Transformer 인코더이다. 이러한 구성은 시간적 특성과 변수 중요도를 동시에 고려하는 이상 진단 구조를 형성한다. 전체 구조는 그림 1에 나타내었다.

본 모델은 iTransformer의 변수 토큰 기반 역전(Inverted) 구조를 기반으로 확장되었다. 시간 축이 아닌 변수 축을 토큰 단위로 구성하는 역전

형 입력 방식을 유지함으로써 변수 간 상호작용 학습 능력을 보존하면서, 시계열 분해와 변동성 기반 게이팅을 통합하여 구조를 고도화하였다. 이를 통해 기존 iTransformer 대비 전력 설비 데이터의 시간적 위계 특성을 보다 효과적으로 반영할 수 있도록 설계하였다.

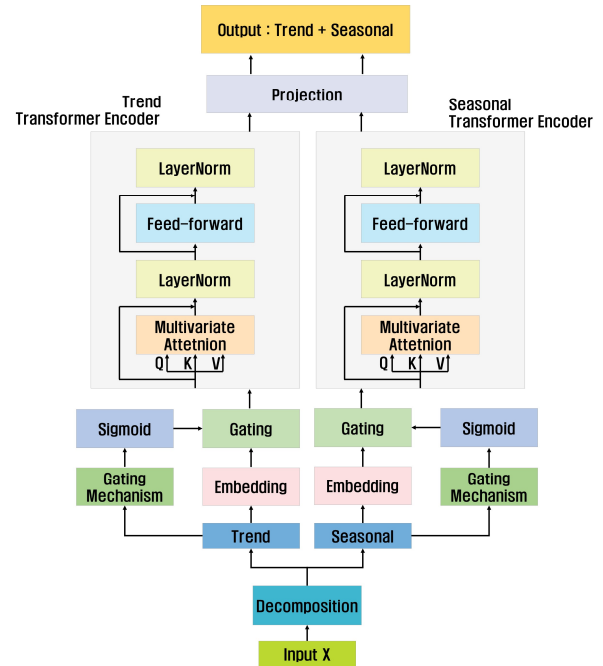


그림 1. 제안된 모델의 전체구성도

1. 시계열 분해 모듈

전력 설비에서 수집되는 다변량 시계열은 장기적 추세와 단기적 변동이 동시에 존재하는 복합 구조를 가진다. 이러한 시간적 위계가 단일 표현 공간에서 함께 학습될 경우, 서로 다른 성분 간 간섭이 발생하여 표현 학습의 효율이 저하될 수 있다.

이를 완화하기 위해 본 연구에서는 이동 평균 기반 분해 기법을 적용하여 입력 시계열을 추세 성분과 계절성 성분으로 분리한다. 추세 성분은 시간 축을 따라 평활화된 신호로 정의되며, 다음과 같이 표현된다.

$$X_T = \text{AvgPool}(X) \tag{1}$$

식 (1)은 입력 시계열 X 에 대해 시간 축 방향으로 이동 평균 연산을 수행하여 장기적 변동 경향을 추출하는 과정을 나타낸다. 즉, 급격한 단기 변동을 완화하고 전반적인 상태 변화 흐름을 반영하는 저주파 성분을 계산한다.

계절성 성분은 원본 시계열에서 추세 성분을 제거하여 계산된다.

$$X_S = X - X_T \quad (2)$$

식 (2)는 입력 신호에서 추세 성분을 차감함으로써 단기 변동 및 반복 패턴을 포함하는 고주파 성분을 분리하는 과정을 의미한다. 이를 통해 장기 추세와 단기 변동이 명확히 구분된 두 개의 표현을 얻을 수 있다.

이와 같은 분해를 통해 저주파 성분(추세)과 고주파 성분(계절성)을 구조적으로 분리할 수 있으며, 분리된 두 성분은 이후 독립적인 학습 경로로 전달되어 시간적 특성에 특화된 표현 학습이 이루어진다. 시계열 분해 모듈의 전체 구조는 그림 2에 나타내었다.

Series Decomposition

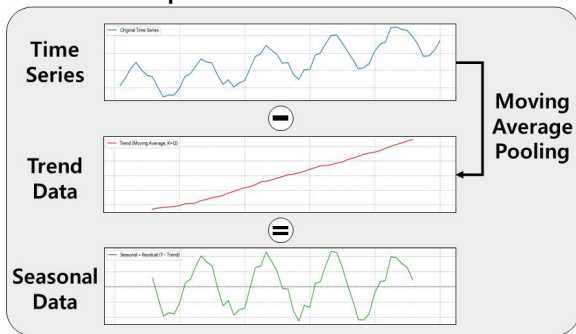


그림 2. 시계열 분해 모듈 구조

2. 변동성 기반 게이팅 메커니즘

다변량 시계열 환경에서는 모든 센서 변수가 동일한 수준의 정보를 제공하지 않는다. 일부 변수는 시간 구간 내에서 거의 변화하지 않는 정적 특성을 보이는 반면, 다른 변수는 고장 상태에서 뚜렷한 변동성을 나타낸다. 모든 변수를 동일한 중요도로 처리하는 경우, 진단과 직접적으로 관

련이 없는 정보까지 모델에 반영되어 표현 학습의 효율이 저하될 수 있다.

이를 반영하기 위해 본 연구에서는 변수별 시간적 변동성을 기반으로 정보 중요도를 산출하는 게이팅 메커니즘을 도입한다. 먼저 분해된 각 성분 X_S 와 X_T 에 대해 시간 축 방향의 표준편차를 계산하여 변수별 변동성 벡터를 구한다.

$$v_S = \sigma(X_S), v_T = \sigma(X_T) \quad (3)$$

식 (3)은 각 변수의 시간 구간 내 변동 폭을 정량화하는 과정을 나타낸다. 여기서 $\sigma(\cdot)$ 는 시간 축 기준 표준편차 연산이며, 값이 클수록 해당 변수의 시간적 변화가 크다는 것을 의미한다.

계산된 변동성 벡터는 두 개의 독립적인 다층 퍼셉트론을 통과하여 게이트 값으로 변환된다.

$$g_S = \text{Sigmoid}(\text{MLP}_S(v_S)) \quad (4)$$

$$g_T = \text{Sigmoid}(\text{MLP}_T(v_T)) \quad (5)$$

식 (4)와 식 (5)는 각각 계절성 및 추세 성분에 대해 변수별 가중치를 생성하는 과정을 의미한다. 시그모이드 함수는 각 변수의 중요도를 0과 1 사이의 값으로 정규화하며, 이는 상대적 기여도를 나타낸다.

이후 게이트 값은 임베딩된 특징 표현에 요소별 곱셈으로 적용된다.

$$E'_S = E_S \odot g_S \quad (6)$$

$$E'_T = E_T \odot g_T \quad (7)$$

식 (6)과 식 (7)은 변수 차원에서 가중치를 적용하여 특징 표현을 재조정하는 과정을 나타낸다. 이를 통해 변동성이 높은 변수는 강조되고, 변화가 미미한 변수는 억제된다.

결과적으로 변동성 기반 게이팅 메커니즘은 변수별 동적 특성을 반영하여 이상 상태와 밀접한 변수에 모델의 표현 학습을 집중시키는 역할을 수행한다. 전체 구조는 그림 3에 나타내었다.

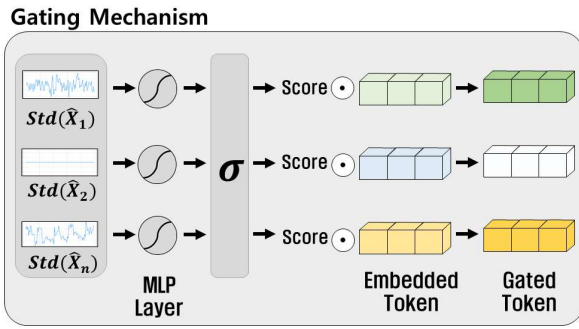


그림 3. 변동성 기반 게이팅 메커니즘 구조

3. 이중 경로 Transformer 인코더

시계열 분해와 변동성 기반 게이팅을 거친 계절성 성분과 추세 성분은 각각 독립적인 Transformer 인코더에 입력된다. 본 연구에서는 두 성분이 서로 다른 시간적 특성을 가진다는 점을 고려하여, 동일한 구조의 인코더를 병렬로 구성하는 이중 경로 구조를 채택하였다.

계절성 경로는 단기적 변동과 주기적 패턴을 중심으로 학습을 수행하며, 고주파 성분에 포함된 국소적 변화와 반복 구조를 포착하는 데 초점을 둔다. 반면, 추세 경로는 장기적 변화와 전반적인 상태 이동을 반영하는 저주파 성분을 학습하여 설비의 점진적 성능 변화나 장기적 상태 전이를 모델링한다.

각 인코더는 다중 헤드 자기 주의(Self-Attention) 연산과 위치 정보가 반영된 피드포워드 네트워크로 구성된다. 자기 주의 연산은 다음과 같이 정의된다.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (8)$$

식 (8)은 쿼리 Q , 키 K , 벨류 V 는간의 유사도를 기반으로 가중합을 계산하는 과정을 나타낸다. 여기서 d 는 스케일링을 위한 차원 계수이며, 이를 통해 변수 간 상호작용이 동적으로 모델링된다. 본 연구에서는 변수 차원을 토큰 단위로 구성하는 역전형 입력 구조를 유지함으로써, 변수 간 상관관계를 직접적으로 학습하도록 설계하였다.

각 경로에서의 인코더 출력은 다음과 같이 표현된다.

$$O_S = \text{Encoder}_S(E'_S) \quad (9)$$

$$O_T = \text{Encoder}_T(E'_T) \quad (10)$$

식 (9)와 식 (10)은 게이팅이 적용된 계절성 및 추세 특징 표현이 각각 독립적인 Transformer 인코더를 통과하여 정제된 표현으로 변환되는 과정을 나타낸다. 두 인코더는 구조적으로 동일하지만 서로 다른 성분을 입력으로 받아 각 시간 스케일에 특화된 표현을 학습한다.

이중 경로 구조는 단일 인코더 방식과 비교하여 두 가지 측면에서 차별성을 가진다. 첫째, 시간적 위계가 다른 성분을 분리하여 학습함으로써 혼합 신호 처리로 인한 표현 왜곡을 완화한다. 둘째, 계절성 및 추세 정보가 독립적으로 정제된 후 결합되므로, 보다 풍부하고 상보적인 특징 조합이 가능하다.

결과적으로 이중 경로 Transformer 인코더는 단기적 변동과 장기적 변화를 동시에 고려할 수 있는 구조를 제공하며, 전력 설비 다변량 시계열 이상 진단에서 보다 정밀한 특징 학습을 가능하게 한다.

4. 분류 출력 구조

계절성 경로와 추세 경로에서 각각 추출된 특징 표현은 최종적으로 결합되어 이상 진단을 위한 분류 계층에 전달된다. 두 인코더의 출력은 특징 차원을 기준으로 연결되며, 이는 단기적 변동 정보와 장기적 추세 정보를 동시에 포함하는 통합 표현을 형성한다.

$$O_{fused} = \text{Concat}(O_S, O_T) \quad (11)$$

식 (11)은 두 경로에서 산출된 특징 표현을 결합하여 최종 분류를 위한 통합 표현을 구성하는 과정을 나타낸다.

결합된 특징 표현은 선형 변환을 통해 클래스별 예측 값으로 변환되며, 이후 소프트맥스 함수

를 적용하여 각 설비 상태에 대한 확률 분포를 출력한다.

$$\hat{y} = \text{softmax}(W O_{fused} + b) \quad (12)$$

식 (12)에서 W 와 b 는 각각 학습 가능한 가중치 행렬과 편향 벡터를 의미한다. 소프트맥스 함수는 다중 클래스 분류를 위해 각 클래스에 대한 확률 값을 계산한다. 본 연구에서는 설비 상태를 정상, 주의, 경고의 세 가지 클래스로 정의하였다.

모델 학습은 교차 엔트로피 손실 함수를 기반으로 수행된다.

$$L_{cls} = - \sum_c y_c \log(\hat{y}_c) \quad (13)$$

식 (13)은 예측 확률 \hat{y}_c 과 실제 레이블 y_c 간의 차이를 측정하는 손실 함수로, 모델 파라미터가 실제 클래스 분포에 근접하도록 최적화하는 역할을 한다.

이와 같이 제안 모델은 시계열 분해를 통해 시간적 특성을 구조적으로 분리하고, 변동성 기반 게이팅을 통해 변수 중요도를 동적으로 조절하며, 이중 경로 인코더를 통해 상이한 시간 스케일의 특징을 학습한 후 통합 표현을 기반으로 최종 이상 상태를 판별한다.

IV. 실험 환경

1. 데이터셋 구성

본 연구는 AI허브의 전력 설비 에너지 패턴 및 고장 분석 센서 데이터를 활용하여 제안 모델의 성능을 평가하였다. 해당 데이터는 10종의 전력 설비에서 수집되었으며, 총 11개의 센서 변수로 구성된다. 주요 변수로는 누적전력량, 주파수, 선간전압평균, 전류평균, 상전압평균, 온도, 역률평균, 유효전력평균, 무효전력평균, 전압고조파평균, 전류고조파평균이 포함된다.

데이터는 설비 ID 단위로 그룹화한 후 시간 순서를 유지하여 학습 데이터 70%, 검증 데이터

15%, 테스트 데이터 15%로 분할하였다. 결측 구간은 제거하였다.

모델 입력은 슬라이딩 윈도우를 통해 생성된 시계열 데이터를 기반으로 구성되며, 각 샘플은 11개의 센서 변수로 이루어진 길이 100의 다변량 시계열로 표현된다. 본 연구에서는 이러한 입력 데이터를 기반으로 설비 상태를 정상, 주의, 경고의 세 가지 클래스로 분류하도록 모델을 구성하였다.

또한 시계열 분해 모듈에서는 입력 신호의 장기 추세를 추출하기 위해 이동 평균 기반 분해 방식을 적용하였으며, 이동 평균 윈도우 크기는 25로 설정하였다. 이는 입력 시퀀스 길이(100)의 약 1/4 수준으로 설정하여 단기 변동을 완화하면서 장기적인 상태 변화 추세를 안정적으로 반영하기 위한 설정이다.

추세 성분 추출 방법으로는 1D Convolution 기반 방식과 이동 평균 기반 방식(AvgPool)을 비교 실험하였다. 실험 결과, 이동 평균 기반 방식이 더 안정적인 추세 분리와 높은 분류 성능을 보였으며, 이에 따라 본 연구에서는 이동 평균 기반 시계열 분해 방식을 최종적으로 채택하였다.

전처리 후 생성된 전체 학습 샘플 수는 15,803,429개이며, 클래스 분포는 정상 30.4%, 주의 27.5%, 경고 42.1%이다. 경고 클래스의 비중이 가장 높은 불균형 구조를 가지므로, 평가에는 매크로 F1 점수를 포함한 불균형 대응 지표를 활용하였다.

본 데이터셋에서 경고 상태의 비중이 높은 현상은 데이터 구성 방식과 데이터 생성 과정의 영향으로 해석될 수 있다. 해당 데이터는 실제 전력 설비에서 수집된 센서 데이터를 기반으로 공개된 데이터셋으로, 설비 상태 변화가 포함된 구간들이 일정 비율 포함되어 있다. 또한 본 연구에서는 슬라이딩 윈도우 기반 데이터 생성 방식을 적용하였으며, 경고 상태가 일정 시간 동안 지속되는 구간에서는 다수의 학습 샘플이 생성될 수 있다. 이러한 특성으로 인해 일반적인 산업 이상 탐지

데이터셋과 달리 경고 클래스의 비중이 상대적으로 높게 나타난 것으로 판단된다.

또한 변수별 변동성 수준에는 뚜렷한 차이가 존재한다. 그림 4는 경고 상태 샘플의 변수별 시계열 예시를 나타낸 것이다. 일부 변수는 시간 구간 동안 거의 일정한 값을 유지하는 반면, 특정 변수는 큰 변동 폭을 보인다. 이러한 특성은 변수 중요도를 동적으로 조절할 필요성을 시사하며, 변동성이 큰 변수에 대한 정보 반영이 중요함을 보여준다.

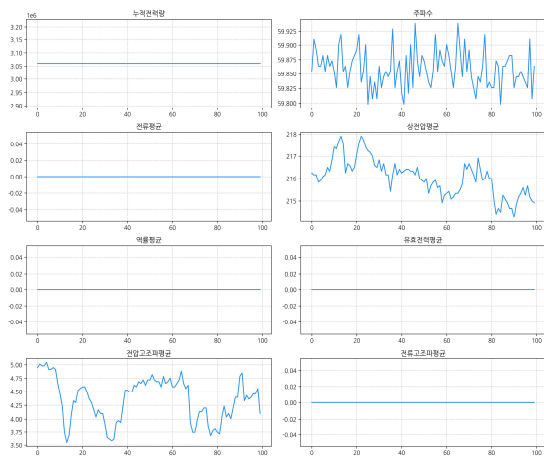


그림 4. 경고 상태에서의 변수별 시계열 변동성 예시

2. 학습 설정

제안 모델은 iTransformer 기반 구조를 따르며, 임베딩 차원은 512, 피드포워드 은닉 차원은 2,048로 설정하였다. 각 경로의 인코더는 6개의 레이어로 구성하였다.

모델 학습에는 Adam 옵티마이저를 사용하였으며, 초기 학습률은 0.0001로 설정하였다. 배치 크기는 1,024로 설정하였고, 최대 50 에포크 동안 학습을 수행하였다. 손실 함수는 다중 클래스 분류를 위한 교차 엔트로피를 사용하였다. 검증 성능이 일정 에포크 동안 개선되지 않을 경우 조기 종료를 적용하여 과적합을 방지하였다.

모든 실험은 NVIDIA H100 GPU 환경에서 수행되었으며, PyTorch 프레임워크를 기반으로 구현하였다.

V. 실험 결과 및 분석

1. 분류 성능 비교

제안 모델의 이상 진단 성능을 평가하기 위해 CNN 기반 모델, MLP/Linear 기반 모델, 그리고 Transformer 기반 모델과 비교 실험을 수행하였다. 평가지표로는 정확도, 매크로 F1 점수, 가중 F1 점수를 사용하였다. 매크로 F1 점수는 클래스 불균형 환경에서 각 클래스의 성능을 동일 비중으로 평가하기 위해 활용되었다.

표 1은 제안 모델과 비교 모델들의 분류 성능을 나타낸다. 제안 모델은 정확도 95.1%, 매크로 F1 점수 94.7%, 가중 F1 점수 95.1%를 기록하여 모든 지표에서 가장 우수한 성능을 보였다. 이는 기반 모델인 iTransformer의 정확도 93.0% 및 매크로 F1 점수 92.4% 대비 각각 2.1%p 및 2.3%p 향상된 결과이다.

또한 LightTS, PatchTST, TimeXer 등 다른 Transformer 기반 모델들과 비교해도 일관된 성능 우위를 보였다. 특히 매크로 F1 점수의 향상은 클래스 불균형 환경에서도 안정적인 판별 능력을 유지함을 의미한다.

이러한 결과는 시계열 분해와 변동성 기반 계이팅을 결합한 구조가 전력 설비 다변량 시계열 이상 진단에서 효과적으로 작동함을 보여준다.

표 1. 제안된 모델과 비교 모델의 분류 성능 비교

아키텍처	모델	정확도(%)	Macro F1-Score(%)	Weighted F1-Score(%)
Transformer	제안 모델	95.1	94.7	95.1
MLP / Linear	LightTS	94.0	93.8	94.1
Transformer	iTransformer	93.0	92.4	93.0
Transformer	PatchTST	83.9	83.5	83.9
Transformer	TimeXer	83.8	83.3	83.8
CNN	MICN	75.7	74.2	75.4
MLP / Linear	DLinear	54.5	45.6	49.1
CNN	TimesNet	51.8	49.1	50.8
Transformer	Autoformer	50.4	37.7	42.0

2. 설비 식별 확장 실험

제안 모델의 표현 학습 능력을 추가적으로 검증하기 위해, 상태 분류와 동시에 설비 식별을 수행하는 확장 실험을 수행하였다. 이를 위해 최종 분류 계층을 수정하여 10개 설비에 대한 분류를 추가적으로 학습하였으며, 상태 예측과 설비 식별을 동시에 수행하도록 모델을 구성하였다.

표 2. 제안 모델의 설비 식별 확장 실험 성능 비교

실험 환경	모델	Accuracy	Facility Accuracy
전체 데이터 환경	제안 모델	93.3	81.9
	iTransformer	91.6	81.3
이상 상황 중심 환경	제안 모델	93.3	85.0
	iTransformer	91.6	83.3

표 2는 두 가지 평가 환경에서의 설비 식별 성능을 비교한 결과를 나타낸다. 첫 번째는 정상, 주의, 경고 상태를 모두 포함하여 상태를 예측하는 전체 데이터 환경이며, 두 번째는 모델이 주의 또는 경고로 예측한 이상 상황 구간만을 대상으로 설비 식별 성능을 평가한 이상 상황 중심 환경이다.

전체 데이터 환경에서 제안 모델은 정확도 93.3%, 설비 식별 정확도 81.9%를 기록하였다. 이는 기반 모델인 iTransformer의 설비 식별 정확도 81.3% 대비 향상된 결과이다.

반면, 이상 상황 중심 환경에서는 제안 모델의 설비 식별 정확도가 85.0%로 증가하였다. 이는 iTransformer의 83.3% 대비 높은 수치이며, 이상 상태로 예측된 구간에서 설비별 특성이 더욱 뚜렷하게 반영됨을 보여준다.

이와 같은 결과는 시계열 분해와 변동성 기반 게이팅 구조가 상태 분류뿐 아니라 설비 고유 특성의 표현 학습에도 기여함을 시사한다. 특히 이상 구간에서의 성능 향상은 제안 모델이 고장과 밀접한 변수 패턴을 효과적으로 포착하고 있음을 보여준다.

3. 제거 실험

제안 모델의 각 구성 요소가 분류 성능에 미치는 영향을 분석하기 위해 시계열 분해 모듈과 변동성 기반 게이팅 메커니즘을 단계적으로 제거하는 제거 실험을 수행하였다. 표 3은 각 구성 요소를 제거했을 때의 분류 성능 변화를 비교한 결과를 나타낸다.

표 3. 제안 모델의 제거 실험 성능 비교

모델	구조	정확도(%)
제안 모델	[시계열 분해] + [변동성 기반 게이팅 메커니즘]	95.1
제안 모델 (w/o 변동성 기반 게이팅 메커니즘)	[시계열 분해]	85.2
DART (w/o 시계열 분해)	[변동성 기반 게이팅 메커니즘]	84.6
iTransformer	기반 모델	92.5

시계열 분해 모듈을 제거한 경우 정확도는 85.2%로 감소하였다. 이는 시간적 위계가 분리되지 않을 경우 서로 다른 특성이 혼합되어 표현 효율이 저하됨을 의미한다.

변동성 기반 게이팅을 제거한 경우 정확도는 84.6%로 감소하였다. 이는 변수 중요도를 동적으로 조절하지 않을 경우 진단에 유효한 정보에 대한 집중도가 낮아질 수 있음을 보여준다.

반면, 두 모듈을 모두 적용한 제안 모델은 95.1%의 정확도를 달성하였다. 표 3에서 확인할 수 있듯이, 각 모듈을 개별적으로 제거했을 때와 비교하여 약 10%p 이상의 성능 차이가 발생하였다. 이는 시계열 분해와 변동성 기반 게이팅이 상호 보완적으로 작용하여 이상 진단 성능 향상에 기여함을 시사한다.

결과적으로 제거 실험은 두 구성 요소가 제안 모델의 핵심 구조임을 정량적으로 뒷받침한다.

VI. 결 론

본 논문에서는 전력 설비 다변량 시계열 이상 진단을 위해 시계열 분해와 변동성 기반 게이팅 메커니즘을 통합한 Transformer 아키텍처를 제안하였다. 제안 모델은 입력 시계열을 추세와 계절성 성분으로 구조적으로 분리하고, 변수별 시

간적 변동성을 기반으로 중요도를 동적으로 조절함으로써 시간적 위계 특성과 변수 중요도를 동시에 반영하도록 설계되었다.

실제 전력 설비 데이터를 활용한 실험 결과, 제안 모델은 정확도 95.1%를 달성하며 기존 Transformer 기반 모델 대비 일관된 성능 향상을 보였다. 또한 제거 실험을 통해 시계열 분해와 변동성 기반 게이팅이 각각 독립적으로, 그리고 결합적으로 분류 성능 향상에 기여함을 확인하였다. 이는 시간적 특성 분리와 동적 변수 중요도 학습이 전력 설비 다변량 시계열 이상 진단에서 효과적인 전략임을 시사한다.

향후 연구에서는 다양한 산업 도메인 데이터에 대한 일반화 성능을 추가로 검증하고, 분해 방식의 적응적 확장과 경량화 설계를 통해 실시간 적용 가능성을 향상시키는 방향으로 연구를 확장할 예정이다.

REFERENCES

- [1] F. Wang, et al., "A survey of deep anomaly detection in multivariate time series: Taxonomy, applications, and directions," *Sensors*, vol. 25, no. 1, article 190, 2025.
- [2] P. Boniol, T. Nguyen, S. Kadam, and T. Palpanas, "Dive into time-series anomaly detection: A decade review," arXiv preprint, arXiv:2412.20512, 2024.
- [3] P. Stahmann and B. Rieger, "A benchmark for real-time anomaly detection algorithms applied in Industry 4.0," *International Conference on Machine Learning, Optimization, and Data Science*, pp. 20-34, Siena, Italy, 2022.
- [4] H.S. Kim and M.H. Lee, "Designing an GRU-based on-farm power management and anomaly detection automation system," *Smart Media Journal*, vol. 13, no.1, pp. 18-12, 2024.
- [5] H. Wu et al., "TimesNet: Temporal 2D-variation modeling for general time series analysis," *International Conference on Learning Representations (ICLR)*, Kigali, Rwanda, 2023.
- [6] H. Wang et al., "MICN: Multi-scale local and global context modeling for long-term series forecasting," *International Conference on Learning Representations (ICLR)*, Kigali, Rwanda, 2023.
- [7] H.S. Jin, K.B. Kim, "TCN-USAD for Anomaly Power Detection," *Smart Media Journal*, vol. 13, no. 7, pp. 9-17, 2024
- [8] A. Zeng et al., "Are transformers effective for time series forecasting?," *AAAI Conference on Artificial Intelligence*, vol. 37, no. 9, pp. 11121-11128, Washington, DC, USA, 2023.
- [9] D. Campos et al., "LightTS: Lightweight time series classification with adaptive ensemble distillation," *Proceedings of the ACM on Management of Data*, vol. 1, no. 2, pp. 1-27, 2023.
- [10] A. Vaswani et al., "Attention Is All You Need," *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998-6008, Long Beach, CA, USA, 2017.
- [11] Y. Liu et al., "iTransformer: Inverted transformers are effective for time series forecasting," *International Conference on Learning Representations (ICLR)*, Vienna, Austria, 2024.
- [12] Y. Nie, "A time series is worth 64 words: Long-term forecasting with transformers," *International Conference on Learning Representations (ICLR)*, Kigali, Rwanda, 2023.
- [13] Y. Wang et al., "TimeXer: Empowering transformers for time series forecasting with exogenous variables," *Advances in Neural Information Processing Systems*, vol. 37, pp. 469-498, Vancouver, Canada, 2024.
- [14] H. Wu et al., "Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting," *Advances in Neural Information Processing Systems*, vol. 34, pp. 22419-22430, Virtual Conference, 2021.
- [15] T. Zhou et al., "FEDformer: Frequency enhanced decomposed transformer for long-term series forecasting," *International Conference on Machine Learning (ICML)*, pp. 27268-27286, Baltimore, MD, USA, 2022.
- [16] J.J. Lee, J.H. Lee, "Time Series Forecasting Method for Mixed Categorical and Continuous Variables," *Smart Media Journal*, vol. 14, no.6, pp. 14-22, 2025

 저 자 소 개

**오승민(정회원)**

2019년 나사렛대학교 디지털콘텐츠학과 학사 졸업.

2021년 전남대학교 ICT융합시스템공학과 석사 졸업.

<주관심분야 : 딥러닝, 인공지능, 시계열 예측, 시계열 이상탐지>

**김진술(정회원)**

2001년 유타대학교 컴퓨터과학 학사 졸업.

2005년 한국과학기술원 정보통신공학 석사 졸업.

2008년 한국과학기술원 정보통신공학 박사 졸업.

<주관심분야 : 클라우드 컴퓨팅, 디지털 미디어, 지능형 네트워크, 딥러닝>