

# 건설기계 동력계통 고장진단을 위한 AI 예측모델 연구

성지원<sup>1</sup>, 김한영<sup>1</sup>, 조한결<sup>1</sup>, 조현수<sup>1</sup>, 오세철<sup>1\*</sup>, 박성철<sup>2\*</sup>

<sup>1</sup>부산대학교 기계공학부 학부생

<sup>1\*</sup>부산대학교 기계공학부 교수

<sup>2\*</sup>볼보그룹코리아(주)

[siw00310@pusan.ac.kr](mailto:siw00310@pusan.ac.kr), [hanyoungk09@pusan.ac.kr](mailto:hanyoungk09@pusan.ac.kr), [0707chg@pusan.ac.kr](mailto:0707chg@pusan.ac.kr), [vlzkc477@pusan.ac.kr](mailto:vlzkc477@pusan.ac.kr)

\*교신저자, E-mail : [sechuloh@pusan.ac.kr](mailto:sechuloh@pusan.ac.kr) , [seongcheol.cr.park@volvo.com](mailto:seongcheol.cr.park@volvo.com)

## Study of an AI-Based Predictive Model for Fault Diagnosis in Construction Machinery Power Systems

Ji-Won Seong<sup>1</sup>, Han-Young Kim<sup>1</sup>, Han-Gyeol Cho<sup>1</sup>, Hyun-Soo Jo<sup>1</sup>, Se-Chul Oh<sup>\*1</sup>, Seong-Cheol Park<sup>\*2</sup>

<sup>1</sup>Dept. of Mechanical Engineering, Pusan National University

<sup>1\*</sup>Dept. of Mechanical Engineering, Pusan National University

<sup>2\*</sup>Volvo Group Korea

### 요 약

본 논문은 엔진 정지(Engine Shutdown)와 직결되는 냉각 계통을 대상으로, 딥러닝 기반 시계열 예측을 통해 이상 시점을 현재 시점 기준으로 정량적으로 추정하는 기법을 제안한다. 이를 통해 기존의 이상 탐지 중심 예지 정비에서 정비 시점 예측 중심의 접근으로 확장 가능성을 제시한다.

### 1. 서론

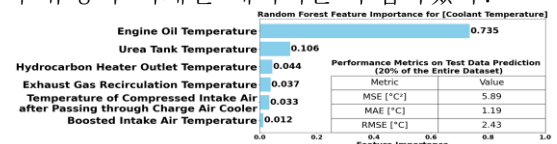
최근 건설 장비 산업에서는 IoT 기술을 접목하여 주요 장비 상태를 실시간으로 확인하고 고장을 사전에 예측하려는 노력이 이루어지고 있다. 이 과정에서 예지 정비(Predictive Maintenance)는 데이터 기반 분석을 통해 정비가 필요한 시점을 합리적으로 판단할 수 있는 기술로 주목받고 있다.[1] 특히 동력 계통 중 엔진의 경우 냉각 시스템의 안정성이 중요하며, 냉각이 원활하지 않으면 부품 손상에 따른 고장으로 이어져 공사 지연, 유지비 증가 등의 손실을 초래할 수 있다. 한편, 기존의 예지 정비 연구들은 대부분 이상 탐지에 따른 경고 발신에 초점을 두고 있어, 실제로 고장 발생 시점을 예측하는 접근은 부족한 실정이며, 정비 시점을 구체적으로 제시할 수 있는 모델의 필요성이 제기되고 있다.[2] 이에 본 연구에서는 자연어 처리 분야에서 활용되는 시계열 예측 기법의 구조를 응용하여, 냉각수 온도의 향후 변화를 다중 시점에 걸쳐 예측하고, 이를 통해 이상 발생 가능성이 높은 시점을 역으로 도출하는 방식을 제안한다.

### 2. 본론

#### 2.1 데이터 구성

데이터는 볼보코리아(Volvo Korea)에서 제공된 건설 기계 엔진 내 센서 수치들로, 정상 운전 상태 및 이상 시점에 도달하는 상태(104℃)로 구분된다. 실험은 외기 냉각이 가능한 상태와 이를 차단한 두 가지

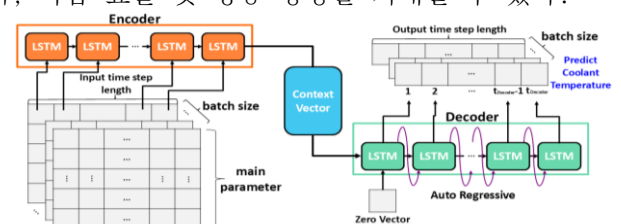
환경에서 수행되었으며, 엔진 부하와 팬 속도를 독립 변수로 설정하였다. 따라서 한 변수를 고정한 상태에서 다른 변수를 단계적으로 변화시키며 장비를 운행하고, 해당 과정에서 발생하는 데이터를 취합하여 총 15 가지 유형의 시계열 데이터를 수집하였다.



(그림 1) 수십 개의 파라미터 중 중요도 상위 일부

#### 2.2 랜덤포레스트(Random Forest) 변수 중요도 추출

학습에 앞서, 냉각수 온도와 관련성이 높은 변수를 추출하기 위해 랜덤포레스트 회귀 모델을 도입하여 변수 중요도를 산출하였으며, 그림 1은 그 결과를 나타낸다. 해당 모델은 트리 기반의 구조로, 변수 간의 비선형적 상호작용을 반영하기에 중요도 분석에 유용하게 활용된다.[3] 이를 통해 이후 시계열 예측 모델의 입력 차원을 상관성이 높은 변수들로 구성함으로써, 학습 효율 및 성능 향상을 기대할 수 있다.



(그림 2) Seq2Seq 기반 LSTM 구조

## 2.3 모델 구조

예측 모델은 순환신경망 기법 중 하나인 LSTM(Long Short-Term Memory)을 기반으로 하며, 이는 시계열 데이터의 장기 의존성을 효과적으로 학습할 수 있다. 본 연구에서는 미래 시점의 온도 수치를 다중 타임스텝에 걸쳐 예측하고, 이를 통해 이상 발생 시점을 역으로 도출하기 위해 Seq2Seq(Sequence-to-Sequence) 구조를 적용하였다. 해당 구조는 기계 번역 모델에서 입력과 출력 시퀀스의 길이가 다를 수 있다는 점에서 착안된 방식이다. 따라서 제안하는 모델은 LSTM 기반의 인코더와 디코더로 구성되며, Auto-Regressive 방식을 도입하여 모델이 이전 시점의 출력을 다음 입력으로 활용하도록 한다.[4] 이를 통해 정답 데이터가 없는 실제 예측 상황을 고려하여 학습을 수행하도록 하였다. 그림 2는 해당 구조를 나타낸다.

## 2.4 Optuna 기반 하이퍼파라미터 튜닝

하이퍼파라미터 튜닝은 Optuna 라이브러리를 활용하여 Bayesian Optimization 기반의 조합 탐색 방식으로 수행하였다.[5] 총 30 회의 trial을 진행하였으며, 튜닝 대상은 입력 시퀀스 길이, 예측 시퀀스 길이, 은닉 벡터(hidden size) 크기, LSTM 층 수로, 모델 구조에 관련된 항목들로 설정하였다. 또한 일반화 성능을 반영하기 위해 검증 손실(validation loss)을 성능 지표로 사용하였고, 각 에폭(epoch)에서의 검증 손실이 기존 trial들의 동일 에폭 기준 중앙값보다 높을 경우, 해당 trial을 조기 종료하여 탐색 효율을 향상시켰다.

<표 1> 30 회 Optuna trial 중 검증 손실 상위 5 개 결과

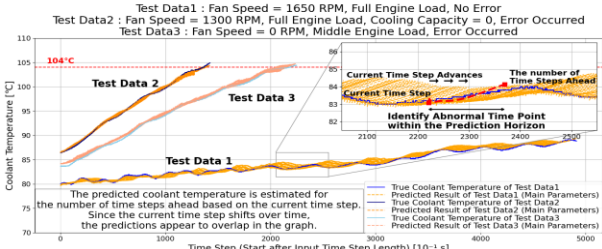
Trial	Trial state	Validation loss	Input time steps	Output time steps	Hidden vector size	Number of LSTM layers
24	COMPLETE	0.140	400	150	96	1
22	COMPLETE	0.157	400	150	32	1
14	COMPLETE	0.162	600	150	96	1
16	COMPLETE	0.168	500	150	64	1
30	COMPLETE	0.260	600	300	96	1

※ To ensure consistency across trials, the number of epochs (25), batch size (64), and learning rate (0.005) were fixed.

## 2.5 모델 학습 및 테스트 데이터를 활용한 검증

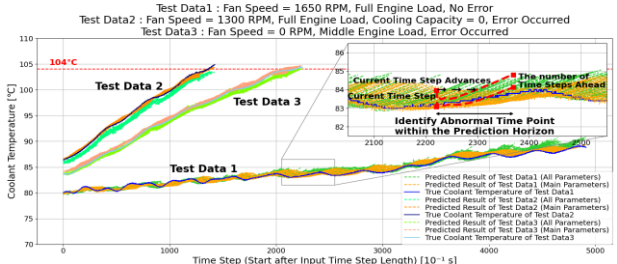
학습 결과는 표 1에 제시된 바와 같이, 24 번째 trial이 0.14의 가장 낮은 검증 손실을 기록하였다.

Compare Actual and Predicted Coolant Temperature (Used Main Parameters)



(그림 3) 최적 파라미터 조합의 냉각수 온도 예측 결과

Compare Actual and Predicted Coolant Temperature (Used Main and All Parameters)



(그림 4) 전체 파라미터를 도입한 예측 결과 추가 비교

따라서 해당 파라미터 조합으로 학습한 다음 테스트 데이터(전체의 20%)로 검증한 결과, 그림 3과 같이 예측 결과가 정답 값과 유사한 분포를 나타낸 것을 확인할 수 있다. 이를 통해, 현 시점 이후의 온도 분포를 산출한 다음, 이상치 발생 시점을 예측 구간 내에서 역으로 도출할 수 있음을 제안할 수 있다.

<표 2> 테스트 데이터 검증 성능 평가

Test Data	Fan speed 1650 rpm Full engine load No Error			Fan speed 1300 rpm Full engine load Cooling capacity 0 Error occurred			Fan speed 0 rpm Middle engine load Error occurred		
	main	all	Improved	main	all	Improved	main	all	Improved
MSE [°C]	0.217	0.346	37.3%	0.091	0.951	90.4%	0.108	0.195	44.6%
MAE [°C]	0.354	0.429	17.5%	0.250	0.939	73.4%	0.266	0.371	28.3%
RMSE [°C]	0.412	0.506	18.6%	0.285	0.954	70.1%	0.305	0.398	23.4%

또한, 입력 차원 축소가 예측 성능에 미치는 영향을 평가하기 위해, 최적의 하이퍼파라미터 조합을 고정한 상태에서 입력 feature 구성을 엔진 전체 파라미터와 그림 1의 랜덤포레스트 기반 주요 파라미터로 구분하여 학습한 다음, 그림 4와 같이 예측 결과를 비교했을 때, 주요 파라미터 만을 사용한 경우가 표 2와 같이 MSE, MAE, RMSE 지표 모두에서 전체 feature 사용 대비 17% 이상의 성능 개선을 보였다.

## 3. 결론

본 논문에서는 건설기계의 예지 정비를 목적으로, Seq2Seq 기반 LSTM 모델을 활용한 이상 시점 예측 기법을 제안한다. 특히, 엔진 정지 현상과 직결되는 냉각 계통에 주목하였으며, 랜덤포레스트 기반 변수 중요도 분석과 Optuna 기반 파라미터 최적화를 적용하여 검증 성능을 보완하였다. 그러나 수집된 데이터가 분 단위로 짧고, 고장 발생 조건을 인위적으로 부여하였기에 실제 운행 환경과의 차이가 있으며, 이로 인해 예측결과가 과대평가되었을 가능성이 있다. 여러 가지 운전 조건 및 장기 운행 데이터를 반영할 경우, 예측 타임스텝 확장 및 복합적인 패턴 학습이 가능할 것으로 기대된다. 또한, 다양한 구조의 예측 모델을 적용하여 성능을 비교함으로써, 최적의 모델 구조를 탐색할 필요가 있다.

## 후기

본 연구는 부산대학교 LINC 3.0 사업단과 볼보코리아(Volvo Korea)의 지원으로 산학협력 기반 캡스톤 디자인 프로그램의 일환으로 수행되었음.

## 참고문헌

- [1] M. Nunes et al., "Challenges of predictive maintenance for Industry 4.0: A review," J. Manuf. Syst., vol. 63, pp. 403–419, Jul. 2022.
- [2] T. Zonta et al., "Predictive maintenance in the Industry 4.0: A systematic literature review," Comput. Ind. Eng., vol. 150, Mar. 2020.
- [3] B. H. Menze et al., "A comparison of random forest and its Gini importance with standard chemometric methods," BMC Bioinformatics, vol. 10, art. no. 213, pp. 1–16, 2009.
- [4] I. Sutskever et al., "Sequence to sequence learning with neural networks," in Proc. NeurIPS, 2014, pp. 3104–3112.
- [5] T. Akiba et al., "Optuna: A next-generation hyperparameter optimization framework," in Proc. KDD, 2019, pp. 2623–2631.