

Denoising Auto Encoder 기법을 활용한 진동 데이터 전처리 및 성능비교

장준교* · 노천명** · 김성수*** · 이순섭**** · 이재철****†

* 경상대학교 해양시스템공학과 석사과정, ** 경상대학교 해양시스템공학과 박사과정,

*** (주)아디아랩 책임연구원, **** 경상대학교 해양시스템공학과 교수

Vibration Data Denoising and Performance Comparison Using Denoising Auto Encoder Method

Jun-gyo Jang* · Chun-myoung Noh** · Sung-soo Kim*** · Soon-sup Lee**** · Jae-chul Lee****†

* MS Candidate, Dept. of Ocean System Engineering, Gyeongsang Nat'l Univ.

** PhD Candidate, Dept. of Ocean System Engineering, Gyeongsang Nat'l Univ.

*** Principal researcher, Adia Lab inc.

**** Professor, Dept. of Naval Architecture and Ocean Engineering, Gyeongsang Nat'l Univ.

요 약 : 기계 장비의 진동 데이터는 필연적으로 노이즈를 포함하고 있다. 이러한 노이즈는 기계 장비의 유지보수를 진행하는데 악영향을 끼친다. 그에 따라 데이터의 노이즈를 얼마나 효과적으로 제거해주냐에 따라 학습 모델의 성능을 좌우한다. 본 논문에서는 시계열 데이터를 전처리 함에 있어 특성추출 과정을 포함하지 않는 Denoising Auto Encoder 기법을 활용하여 데이터의 노이즈를 제거했다. 또한 기계 신호 처리에 널리 사용되는 Wavelet Transform과 성능 비교를 진행했다. 성능비교는 고장 탐지율을 계산하여 진행했으며 보다 정확한 비교를 위해 분류 성능 평가기준 중 하나인 F-1 Score를 계산하여 성능 비교를 진행했다. 고장을 탐지하는 과정에서는 One-Class SVM 기법을 활용하여 고장 데이터를 탐지했다. 성능 비교 결과 고장 진단율과 오차율 측면에서 Denoising Auto Encoder 기법이 Wavelet Transform 기법에 비해 보다 좋은 성능을 나타냈다.

핵심용어 : 디노이징 오토인코더, 건전성 예측 및 관리기술, 웨이블릿 변환, 단일 클래스 서포트 벡터 머신, 노이즈 제거

Abstract : *Vibration data of mechanical equipment inevitably have noise. This noise adversely affects the maintenance of mechanical equipment. Accordingly, the performance of a learning model depends on how effectively the noise of the data is removed. In this study, the noise of the data was removed using the Denoising Auto Encoder (DAE) technique which does not include the characteristic extraction process in preprocessing time series data. In addition, the performance was compared with that of the Wavelet Transform, which is widely used for machine signal processing. The performance comparison was conducted by calculating the failure detection rate. For a more accurate comparison, a classification performance evaluation criterion, the F-1 Score, was calculated. Failure data were detected using the One-Class SVM technique. The performance comparison, revealed that the DAE technique performed better than the Wavelet Transform technique in terms of failure diagnosis and error rate.*

Key Words : *Denoising Auto Encoder, Prognostics Health Management (PHM), Wavelet Transform, One-Class Support Vector Machine, Noise Reduction*

* First Author : wkd4258@gnu.ac.kr, 055-772-9190

† Corresponding Author : j.c.lee@gnu.ac.kr, 055-772-9195

1. 서론

1.1 연구의 배경 및 필요성

전 세계적으로 4차 산업혁명이 화두로 떠오름에 따라 빅데이터, 사물인터넷 및 3D 프린팅 기술 등 다양한 기술이 요구되며 그에 따른 연구가 활발히 진행되고 있다. 또한, 기계설비의 자동화 / 자율화가 진행 중이며 그로 인해 기계 장비의 효율적인 관리 및 유지보수에 관한 연구 또한, 활발히 진행되고 있다(Lee et al., 2019; Jang and Lee, 2020).

현재 건전성 예측관리(Prognostics Health Management; PHM)의 연구가 항공, 제조, 자동차, 철도, 중공업업을 포함한 다양한 분야에서 활발히 진행되고 있다. 건전성 예측관리(PHM)는 장비의 최신 정보를 기반으로 미래 상태(잔존유효수명)를 예측할 뿐만 아니라 실제 운영 조건에서 시스템의 실시간 건강 진단을 가능하게 하는 새로운 엔지니어링 접근 방식이다. 건전성 예측관리(PHM) 기반의 고장 진단 및 잔존유효수명 예측은 장비의 정비 패러다임을 변화시켜 생산성 향상, 장비 수리 비용 절감 등 큰 장점을 보유한다(Lee et al., 2019; Tsui et al., 2015). 기계 장비의 건전성을 예측 및 관리하는데 있어 필요한 과정은 여러 가지 이지만 그 중 데이터 전처리 부분은 많은 시간을 필요로 한다. 그 예로 Forbes에서 연구자들을 대상으로 연구과정 중 가장 많은 시간이 소요되는 과정을 묻는 질문에서 Fig. 1과 같이 약 80%의 연구자들이 데이터를 계측 및 전처리하고 편성을 하는데 가장 많은 시간을 쓴다고 답하였다. 이러한 데이터 전처리 과정은 기계 장비의 유지보수 즉, 고장 진단과 수명을 예측하는 과정에 있어 큰 영향을 끼친다.

기존의 기계 장비의 신호 처리에는 Wavelet Transform 기법이 널리 사용되고 있다. Wavelet Transform은 Fourier Transform과 달리 여러 Wavelet Function을 활용하여 신호를 변환시키는 급변하는 신호의 변환에 사용하기 유리하다는 점 등 여러 장점들이 있다. 하지만 데이터가 편향되거나 분포가 일정하지 않은 경우 데이터의 손실을 야기한다(Yang et al., 2021).

본 연구에서는 영상의 노이즈 제거에 많이 사용되고 있는 Denoising Auto Encoder(이하 DAE) 기법을 활용하여 진동 데이터의 노이즈를 제거하고자 한다. DAE 기법은 특성추출 과정이 포함되지 않는 딥러닝 기법 중 하나로 본 기법을 활용하여 기계 장비 진동 데이터의 전처리 과정을 수행했으며 기계신호 전처리에 널리 사용되는 Wavelet Transform을 비교 기준으로 선정하여 전처리 성능 비교를 진행했다. 전처리 기법의 성능 비교는 고장 탐지율을 비교하여 진행했으며 고장 탐지를 위해 Anomaly Detection에 사용되는 One-Class

SVM(Support Vector Machine) 기법을 활용했으며, One-Class SVM(Support Vector Machine)의 성능 지표는 분류 성능 평가 지표(F-1 Score)를 사용하여 두 가지 전처리 기법에 대한 성능을 비교했다.

본 논문은 Fig. 2의 순서로 진행했으며, 1장은 연구의 배경 및 필요성과 관련 연구 현황을 정리했다. 2장은 데이터 셋 및 적용 기법으로 연구에 사용된 데이터 셋 그리고 전처리 기법과 Anomaly Detection에 사용된 기법에 대한 전반적인 내용을 기술했다. 3장은 사례연구로 데이터 전처리 결과 및 전처리 된 데이터를 활용한 Anomaly Detection 결과와 분류 성능 평가지표에 따른 성능 비교 결과를 기술했으며, 4장은 결론 및 향후 연구 계획에 대하여 기술했다.

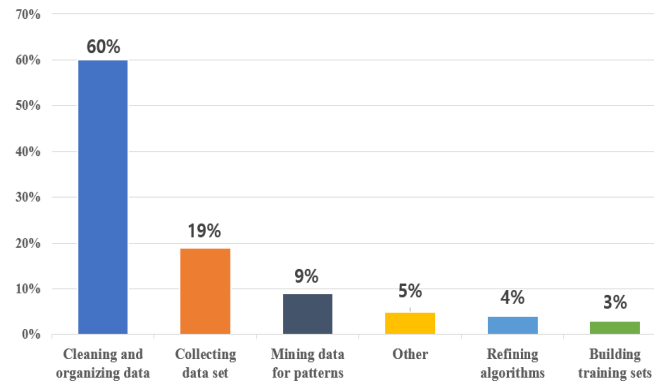


Fig. 1. Importance of data preprocessing.
(from 2016 Data Science Report, CrowdFlower)

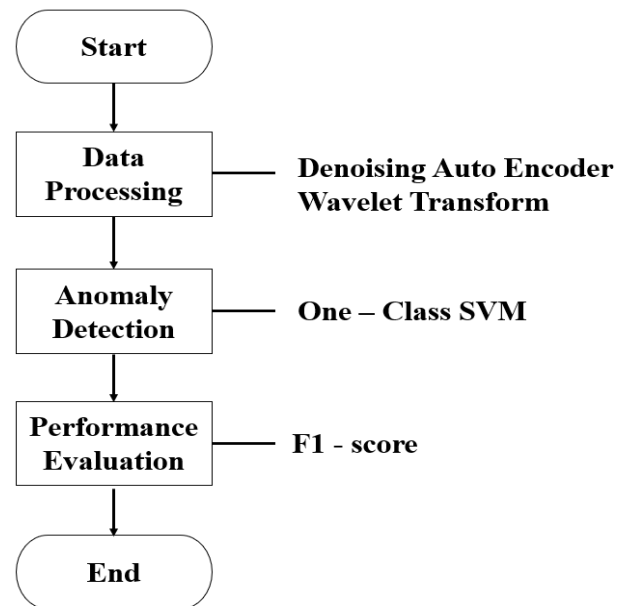


Fig. 2. Flow Chart of Research.

1.2 관련 연구 현황

진동 데이터는 계측 과정 내 필연적으로 노이즈를 포함하여 계측된다. 그에 따라 진동 데이터를 활용한 여러 연구에서 진동 데이터의 노이즈를 제거하기 위해 다양한 기법들을 활용한 것을 볼 수 있다. Lu, Li는 진동 신호가 노이즈에 의해 간섭을 받는 문제를 해결하고자 데이터 주파수 분석을 진행하고자 했고 그에 따라 Wavelet Transform 기법을 활용하여 진동 데이터의 노이즈를 감소시켰다(Lu and Li., 2021). Jia, Li, Dong, He는 Ensemble Empirical Mode Decomposition 기법을 회색 이론과 접목시켰으며 그에 따라 진동 신호의 새로운 노이즈 제거 방법을 제안했다(Jia et al., 2021). Yang, Bo, Yawu, Xi, Dongdong, Yalan은 진동 데이터를 계측하여 수력 발전 장치의 추세 예측을 하고자 했으며 연구 초기에 Wavelet Transform 기법을 활용하여 진동 데이터의 노이즈를 제거했다(Yang et al., 2021).

Wavelet Transform 기법은 기계 신호처리에 널리 사용되어 여러 분야에서 많은 연구가 진행되었다. A. Santhana Raj, N. Murali는 Wavelet Transform 기법을 활용하여 베어링 데이터의 노이즈를 제거했고(Raj and Murali, 2013), Bo-Yeon Hwang, Jae-Hoon Jung, Jang-Myung Lee는 Wavelet Transform 기법을 활용하여 음원 신호의 잡음을 제거함으로써 위치 추정 성능을 향상시켰다(Hwang et al., 2015).

DAE 기법은 기본적인 Auto Encoder의 복원 능력을 더 강화하기 위해 학습 방법을 조금 변형시켜 노이즈를 제거한다. Lovedeep Gondara는 DAE 기법을 활용하여 X-rays, Magnetic Resonance Imaging(MRI), Computer Tomography(CT) 등을 포함한 여러 의학 이미지의 노이즈를 제거했고, Jae-Yeon Jang, Byeong-Wook Min, Chang-Wook Kim는 DAE 기법을 활용하여 반도체 공정 이상탐지에 활용하여 SN 관리도를 제안했다(Jang et al., 2016).

One-Class SVM 기법은 학습된 데이터를 포함하는 최적의 Support Vector를 구하고 이 영역 밖의 데이터들은 Outlier로 간주하는 방식으로 Anomaly Detection, 이미지 검색 문서/텍스트 분류 등에 사용된다. Yeo-jeong Yoon, You-jin Jung은 One-Class SVM 기법을 활용하여 사이버 공격에 의한 시스템 이상상태를 탐지하는 방안을 제안했다(Yoon and Jung, 2019). Yeongjin Cho, Yonghyeok Ji, Gihye Nam, Seongyong Jeong, Hyeongcheol Lee은 One-Class SVM 기법을 활용하여 하이브리드 자동차 엔진 클러치 접합 / 해지 Anomaly Detection를 진행했다(Cho et al., 2021).

2. 데이터 셋 및 적용 기법

2.1 Data Set

연구를 위해 활용된 데이터 셋은 한국지능정보사회진흥원의 사업결과로써, 대전광역시 도시철도공사에서 2020년 구축한 데이터로 도시철도 역사 내 기계설비(전동기)의 상태 모니터링 및 고장 진단을 위해 계측된 데이터이다. 데이터는 크게 진동 데이터와 전류 데이터로 나뉘어져 있으며 전동기의 전압에 따라 2.2KW~55KW까지 다양하게 구축되어 있다. 또한 회전체 불평형, 벨트 느슨함, 축정렬 불량, 베어링 고장을 포함한 4 가지의 고장 유형으로 분류가 되어 있으며 각각의 경우의 정상 데이터를 포함하고 있다. 본 연구에서는 데이터 구축량이 가장 많은 2.2KW 전동기의 정상 데이터와 벨트 느슨함 유형의 고장 데이터를 활용하여 연구를 진행했다. 전체 정상 데이터에서 100,000개의 정상 데이터와 고장 데이터의 구역을 크게 세 구역으로 나누어 각 구역마다 100,000개씩 총 300,000개의 고장 데이터를 선별하여 연구에 활용했다. 고장 데이터를 3가지 구역으로 나누는 기준은 한 가지 유형의 고장 데이터 즉, 벨트 느슨함 데이터에서 크게 세 가지의 다른 특징과 분포를 가지고 있어 세 구역으로 나누어 연구를 진행한다. 전체 데이터에서 각각의 고장 구역의 위치는 Fig. 3과 같다.

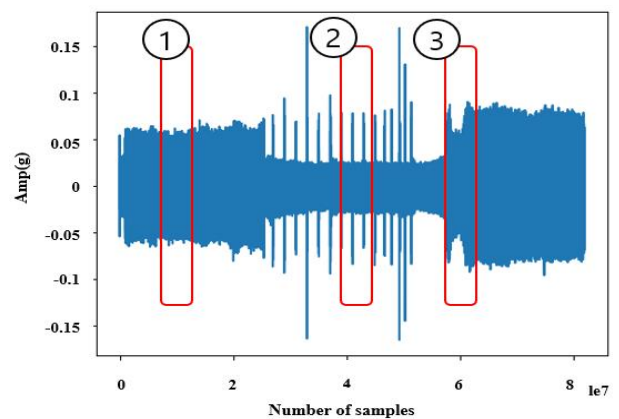


Fig. 3. Entire Failure Data.

본격적인 연구에 앞서 학습의 안정화를 위해서 데이터를 0과 1 사이로 정규화 하여 진행했다. 정규화 된 정상 데이터는 Fig. 4와 같으며 정규화 된 고장 데이터는 Case 1~3 순으로 Fig. 5~Fig. 7과 같다. 고장 데이터 Case 2 즉, Fig. 6은 사진과 같이 정상데이터 범위를 일부분 포함하고 있다(Fig. 6 중 빨간색 영역). 이와 같이 확실한 고장 데이터의 빈도수가 낮아 고장 데이터 Case 2의 고장 탐지율이 다른 고장 데이터 Case 1, 3 비해 낮을 것으로 예상된다.

Denoising Auto Encoder 기법을 활용한 진동 데이터 전처리 및 성능비교

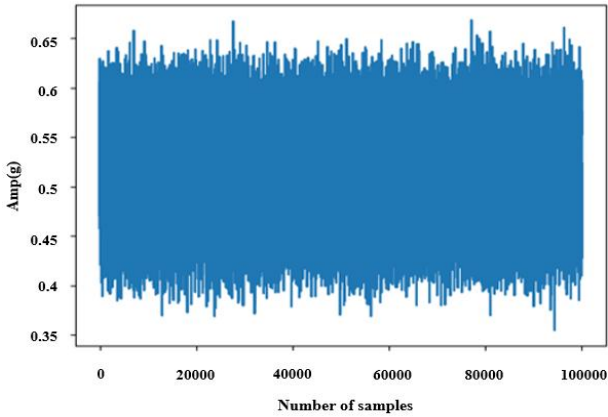


Fig. 4. Normal Data.

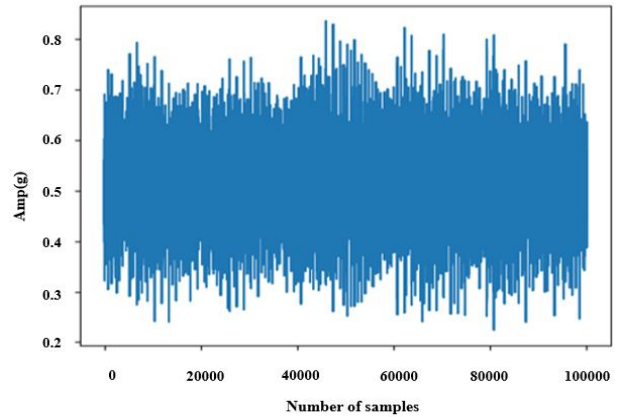


Fig. 7. Failure Case (3).

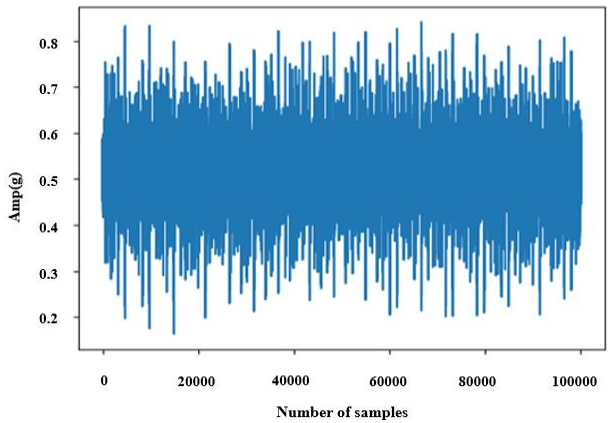


Fig. 5. Failure Case (1).

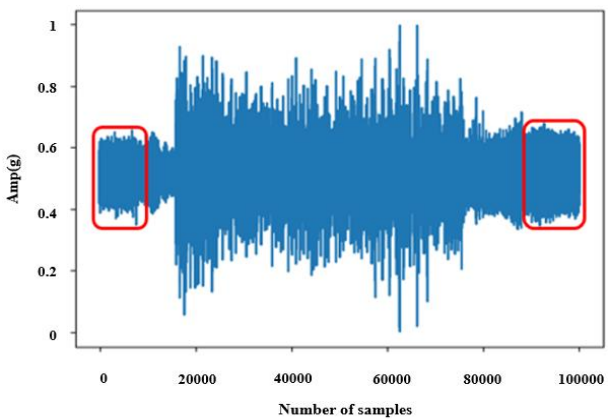


Fig. 6. Failure Case (2).

2.2 Data Preprocessing

필연적으로 모든 데이터는 노이즈를 포함하고 있으며 이 노이즈를 어떻게 효율적으로 제거하는가에 따라 향후 모델의 성능을 좌우한다. 데이터의 노이즈를 제거하는 방법은 특성추출, 차원축소 등 다양한 방법이 있으며 본 연구에서는 사용되는 DAE 기법은 특성추출 과정이 포함되지 않아 데이터의 특성을 간직하는 동시에 노이즈를 제거할 수 있는 장점이 있다.

본 연구에서는 DAE 기법을 통해 데이터 전처리를 진행한 후 기계 신호 처리에 널리 사용되는 Wavelet Transform과 성능 비교를 진행했다.

2.2.1 Wavelet Transform

Wavelet Transform은 기계 신호처리 및 분석에 사용되는 기법으로 1970년대 프랑스의 Morlet에 의해 처음 제안되었으며 다양한 분야에서 활발하게 적용되고 있다(Lee and Kim, 2011).

Wavelet Transform 기반 데이터의 노이즈 제거는 Wavelet Thresholding 기법을 통해 이루어진다. Wavelet Thresholding 기법은 Wavelet Transform을 통해 계산된 Wavelet 계수가 Threshold 값 즉 임계값 보다 작은 경우 0으로 만드는 원리이다. 본 기법은 Fig. 8과 같이 크게 Hard Thresholding, Soft Thresholding으로 구분되며 각각 식(1)과 식(2)와 같이 표현된다.

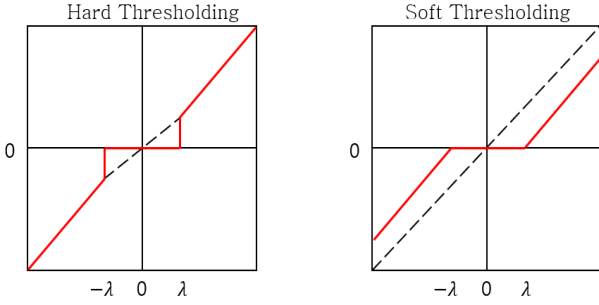


Fig. 8. Hard Thresholding and Soft Thresholding. (Ahn et al., 2015)

$$T_{\lambda}^{hard} = \begin{cases} \mu & \text{if } |\mu| \geq \lambda \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$T_{\lambda}^{soft} = \begin{cases} (\mu - \text{sign}(\mu)\lambda) & \text{if } |\mu| \geq \lambda \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

식(1)과 식(2)에서의 λ 는 임계 기준값이며, μ 는 Wavelet 계수이다. 두 가지의 경우에서 임계 기준값인 λ 는 노이즈 제거 과정에 큰 영향을 끼치는 중요한 요소이다. 임계 기준값 λ 는 식(3)과 같이 결정하게 되며, 식 내의 σ 와 n 은 각각 노이즈를 제거하고자 하는 데이터의 표준편차 값과 샘플 개수를 의미한다.

$$\lambda = \sqrt{2 \ln n} \sigma \quad (3)$$

2.2.2 Denoising Auto Encoder

Auto Encoder는 신경망을 이용하여 입력 값과 같은 출력 값을 재생성 하도록 하는 비지도 학습 알고리즘이다(Lee, 2021). DAE 기법은 Auto Encoder의 이러한 원리를 활용하여 Fig. 9와 같은 구조를 가진다.

DAE 모델의 학습을 위해 Input Layer를 구성할 때 원본 데이터에 임의의 노이즈를 삽입한다. 이때 임의의 노이즈는 원본 데이터의 노이즈가 가지는 특성과 유사한 노이즈를 삽입해야 한다. 그 후, 임의의 노이즈가 추가된 데이터를 입력층 즉, Input Layer로 가져가 Output Layer를 임의의 노이즈가 추가되지 않은 원본 데이터와 가장 유사하게 복원할 수 있는 가중치(W, b)를 학습하는 원리이다(Kim et al., 2014). 모델의 학습이 완료되면 노이즈를 제거하고자 하는 데이터를 학습된 모델에 적용하여 노이즈를 제거하는 원리이다.

기존 전처리 기법들은 전처리 과정 중 데이터의 정보 요약되거나 손실이 발생하는 것과 달리 DAE기법은 이러한 과정을 포함하지 않아 효율적인 노이즈 제거 결과를 추출할 수 있다(Kim et al., 2014).

DAE의 학습 방법을 수식으로 정리하면 다음과 같다.

- (1) 본래의 데이터 x 에 임의의 noise를 추가하여 \tilde{x} 를 만들어 준다.

$$\tilde{x} = x + \text{noise} \quad (4)$$

- (2) Noise를 추가하여 만든 \tilde{x} 를 Input으로 가져가 Encoding 해준다.

$$h = f_{\theta} = s(W\tilde{x} + b) \quad (5)$$

- (3) 마지막으로 Decoding을 해줌으로써 Output 값 x' 을 본래의 데이터 x 와 가장 유사하게 학습한다.

$$x' = g_{\theta'}(h) = s(W'h + b') \quad (6)$$

$$\min_{W, b, W', b'} L(x, x') = \frac{1}{2} \|x - x'\|^2 \quad (7)$$

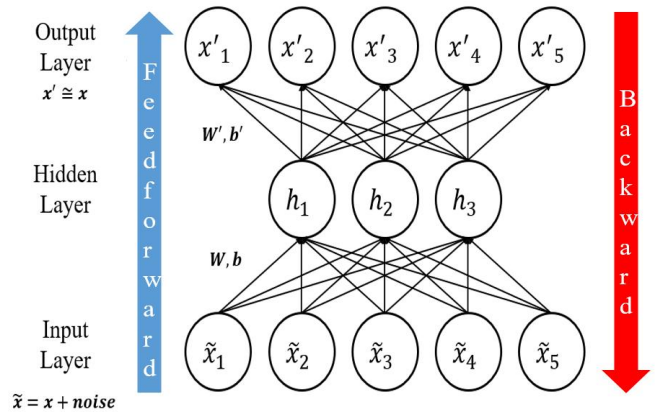


Fig. 9. Denoising Auto Encoder.

2.3 Anomaly Detection

Anomaly Data는 주어진 데이터 중에서 다른 관측치에 비하여 지나치게 작거나 큰 경우의 데이터를 의미한다(Jun, 2008). 본 데이터에서의 Anomaly Data는 기계 고장 데이터이며 Anomaly Detection을 위해 One-Class SVM 기법을 활용한다. One-Class Classification 기법을 활용하는 이유는 실제 기계 장비의 신뢰성 향상으로 인해 정상 데이터에 비해 고장 데이터의 양이 현저히 적기 때문에, 정상 데이터만으로도 기계의 고장 데이터를 탐지해 낼 수 있는 One-Class Classification 기법을 활용했다.

2.3.1 One-Class SVM(Support Vector Machine)

One-Class SVM은 머신러닝 알고리즘 중 하나로 분류나 회귀, 클러스터링 및 Anomaly Detection의 목적으로 많이 사용된다. 본 연구에서는 One-Class SVM 기법을 Anomaly Detection 즉, 고장 데이터를 탐지하는데 사용했다. One-Class SVM은 학습되는 모든 클래스를 하나의 클래스로 규정하고 학습을 진행한 후 Fig. 10과 같이 Training Data와 Non-Training Data를 구분하는 Hyperplane을 정하여 Anomaly Detection을 진행한다. 본 연구에서는 정상 데이터를 모델에 학습시킨 후 고장 데이터를 탐지했다. 또한, 시계열 데이터의 특성상 선형적으로 분류하는 것이 불가능하기에 Fig. 11과 같은 커널 기법을 이용하여 주어진 데이터를 고차원 특징 공간으로 Mapping 시킨 후 분류를 진행한다. 그 후 다시 2차원 공간으로 매핑 시키게 되면 Fig. 11과 같이 비선형 방식으로 Anomaly Data를 탐지하게 된다(Kim et al., 2019). Fig. 12의 Normal Data는 모델에 학습시킨 정상 데이터이며 Learned Frontier 밖으로 분류가 되는 데이터는 Anomaly Data로서 고장 데이터로 탐지하게 된다.

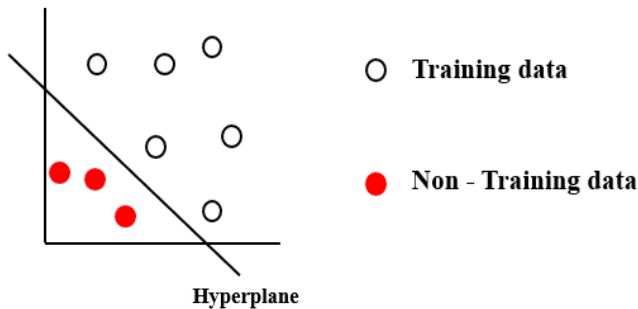


Fig. 10. One - Class SVM.

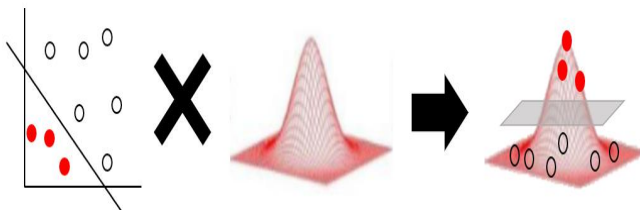


Fig. 11. Kernel Method.

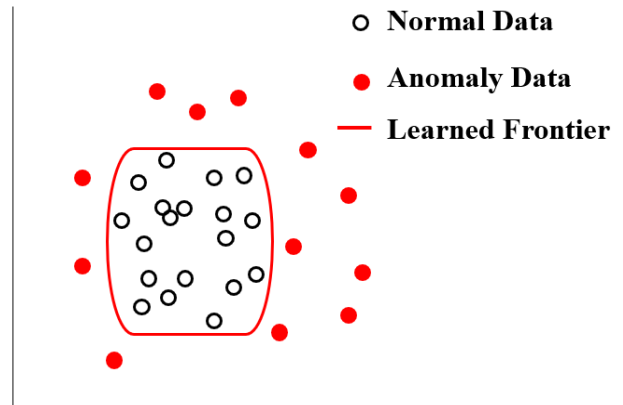


Fig. 12. Nonlinear One - Class SVM.

2.4 분류 성능 평가지표

학습된 모델을 평가하는 요소는 모델의 결과와 실제 결과의 관계를 파악함으로써 모델을 평가할 수 있다. 모델의 성능 평가에 사용되는 지표들은 다양하며, 본 연구에서는 정밀도와 재현율의 조화 평균인 F-1 Score를 활용하여 모델을 평가했다. F-1 Score의 수식은 식(8)과 같으며 Precision과 Recall의 수식은 각각 식(9), 식(10)과 같다.

$$F-1 \text{ Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

위의 식에서 TP는 실제 True인 정답을 True라고 예측한 경우이며 본 연구에서는 정상인 데이터를 모델이 정상으로 예측한 경우를 뜻한다. FP는 실제 False인 정답을 True라고 예측한 경우이며 고장 데이터를 정상 데이터로 예측한 경우이다. FN은 실제 True인 정답을 False라고 예측한 경우이며 정상 데이터를 고장으로 예측한 경우이다.

3. 사례연구

3.1 Result of Data Preprocessing

2장에서 설명한 두 가지의 전처리 기법으로 데이터의 노이즈 제거를 진행했다.

Wavelet Transform의 경우 Matlab의 Wavelet Tool Box를 활용하여 원본 데이터를 진동 데이터의 형상과 가장 유사한

db4 함수를 사용하여 4단계 변환 과정을 거쳤다. 각각의 변환 단계마다 Soft Thresholding을 통해 임계값을 설정하여 노이즈를 제거하며 최종 분해 단계의 데이터를 추출했다.

데이터 전처리 결과는 Fig. 13~Fig. 16과 같다. Fig. 13은 Fig. 4의 Normal Data를 Wavelet Transform기법을 통해 전처리한 결과이며, Fig. 14는 Fig. 5의 고장 데이터 Case 1을 전처리한 결과이다. Fig. 15는 Fig. 6의 고장 데이터 Case 2를 전처리한 결과이며, Fig. 16은 Fig. 7의 고장 데이터 Case 3을 전처리한 결과이다. 4가지 경우 모두 전처리 과정을 거치며 y축 즉, 진동의 범위가 좁아졌으며, Fig. 15의 경우 Wavelet Thresholding을 거치며 상대적으로 큰 진폭의 데이터가 많이 삭제된 결과를 볼 수 있다.

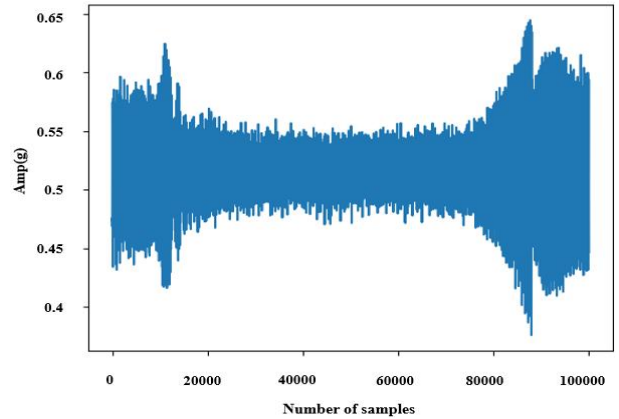


Fig. 15. Wavelet Transform of Failure Data Case (2).

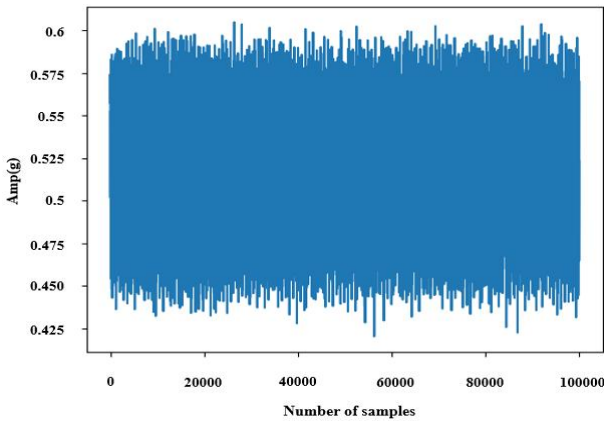


Fig. 13. Wavelet Transform of Normal Data.

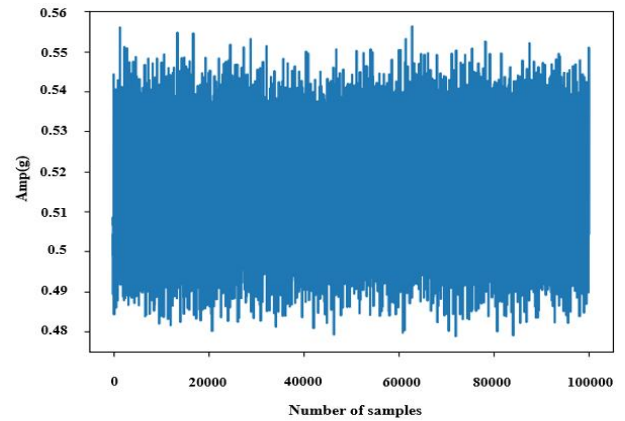


Fig. 16. Wavelet Transform of Failure Data Case (3).

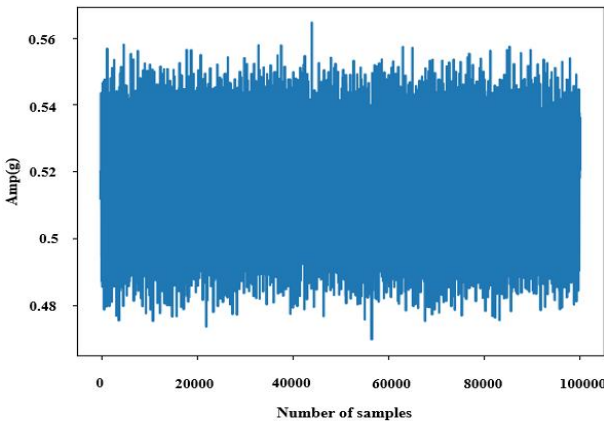


Fig. 14. Wavelet Transform of Failure Data Case (1).

DAE를 활용한 데이터 전처리 수행과정에서 원본 데이터에 추가되는 임의의 노이즈는 진동 데이터 노이즈와 유사한 Gaussian 노이즈를 추가했다. 노이즈가 추가된 Input Data를 Raw Data로 복원하는 모델을 만들어 실험 데이터에 적용시켜 노이즈를 제거했다. 데이터 전처리 결과는 Fig. 17~Fig. 20과 같다. Fig. 17은 Fig. 4의 Normal Data를 DAE기법을 통해 전처리한 결과이며, Fig. 18은 Fig. 5의 고장 데이터 Case 1을 전처리한 결과이다. Fig. 19는 Fig. 6의 고장 데이터 Case 2를 전처리한 결과이며, Fig. 20은 Fig. 7의 고장 데이터 Case 3을 전처리한 결과이다. Wavelet Transform과 마찬가지로 4가지 경우 모두 전처리 과정을 거치며 y축의 범위가 좁아졌으며, Wavelet Transform에 비해 그래프의 형태가 유지되며 전처리된 결과를 확인할 수 있다.

Denoising Auto Encoder 기법을 활용한 진동 데이터 전처리 및 성능비교

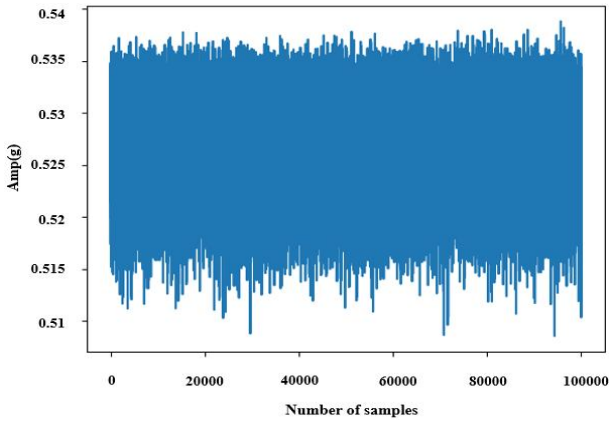


Fig. 17. DAE of Normal Data.

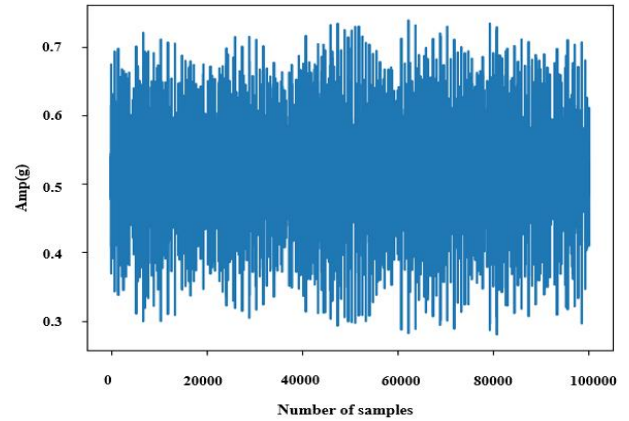


Fig. 20. DAE Failure of Data Case (3).

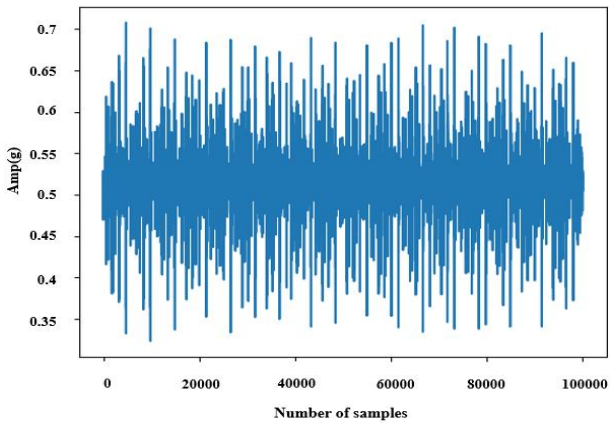


Fig. 18. DAE of Failure Data Case (1).

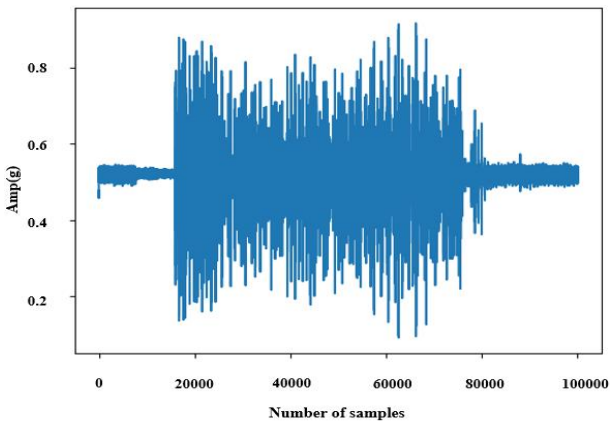


Fig. 19. DAE of Failure Data Case (2).

3.2 Result of Anomaly Detection

전처리가 진행 된 데이터를 활용하여 각각의 전처리 기법마다 Anomaly Detection을 진행했고 결과는 Table 1, Table 2와 같다. 또한, 다음과 같은 탐지 결과를 토대로 F-1 Score를 계산하여 전처리 성능 평가를 진행했으며 결과는 Table 3과 같다.

전처리가 진행 된 정상 데이터를 모델에 먼저 학습한 결과 DAE기법의 경우 Table 1과 같이 전체 100,000건의 전체 데이터 중 93,179건의 데이터를 정상으로 판별하여 약 93.18%의 정상 학습률을 보였다. 이를 통해 고장 데이터를 탐지해본 결과 고장 데이터 Case 1의 경우 100,000건 중 85,860건을 탐지했으며, 고장 데이터 Case 2의 경우 100,000건 중 71,880건을 탐지했고 고장 데이터 Case 3의 경우 100,000건 중 78,965건을 탐지했다.

Wavelet Transform의 경우 Table 2와 같이 91,236건의 데이터를 정상으로 판별하여 약 91.24%의 정상 학습률을 보인다. 이를 통해 고장 데이터를 탐지해본 결과 고장 데이터 Case 1의 경우 100,000건 중 78,784건을 탐지했으며, 고장 데이터 Case 2의 경우 100,000건 중 83,505건을 탐지했고 고장 데이터 Case 3의 경우 100,000건 중 77,146건을 탐지했다.

Anomaly Detection 결과 DAE 기법이 Wavelet Transform 기법보다 약 1.94% 향상된 정상 데이터 학습률을 나타냈다. 또한, 고장 데이터를 탐지하는데 있어 DAE 기법은 정상 데이터 범위를 포함하고 있는 고장 데이터 Case 2의 데이터의 탐지율이 가장 낮게 측정되어 예상과 상충하는 결과를 도출한 반면, Wavelet Transform 기법은 고장 데이터 Case 2의 탐지율이 높게 나와 오차를 포함하고 있다. 고장 데이터 Case1, 3의 탐지율은 DAE 기법이 Wavelet Transform에 비해 약 7.1%, 1.8% 높은 성능을 보인다.

이를 활용하여 계산한 F-1 Score 결과는 Table 3과 같으며 DAE 기법을 활용한 경우 고장 데이터 Case1, 3의 F-1 Score가

Table 1. Result of DAE

(Unit: ea)

DAE		Actual Values			
		Normal	Failure Case (1)	Failure Case (2)	Failure Case (3)
Predictive Values	Normal	93,179	14,140	28,120	21,035
	Failure	6,821	85,860	71,880	78,965

Table 2. Result of Wavelet Transform

(Unit: ea)

Wavelet Transform		Actual Values			
		Normal	Failure Case (1)	Failure Case (2)	Failure Case (3)
Predictive Values	Normal	91,236	21,216	16,495	22,854
	Failure	8,764	78,784	83,505	77,146

Table 3. F-1 Score of Wavelet Transform and DAE

(Unit: ea)

F-1 Score	Failure Case	Wavelet Transform	DAE
	Case (1)	0.8589	0.8989
	Case (2)	0.8784	0.8421
	Case (3)	0.8523	0.8700

Wavelet Transform의 결과에 비해 평균 약 0.03점 높게 계산됐다. 이러한 점을 보아 DAE 기법을 활용한 데이터 전처리 성능이 정상 데이터의 학습률과 고장을 정확하게 진단하는 고장 진단을 또한 향상시킨다는 점을 확인할 수 있다.

4. 결론 및 향후연구

본 연구에서는 기존에 영상의 노이즈 제거 과정에 많이 사용되던 DAE 기법을 기계 진동 데이터에 적용했다. 특성 추출 과정을 포함하지 않은 DAE 기법을 활용하여 기계 진동 데이터 전처리를 진행했으며, 성능 검증을 위해 Wavelet Transform과 성능 비교를 진행했다. 성능 비교를 위해 One-Class Classification 기법 중 하나인 One-Class SVM 기법을 활용하여 정상데이터를 학습한 후 고장 데이터를 Anomaly로 판단하여 고장 탐지를 진행했다. 진행 결과 DAE 기법을 활용하여 전처리가 진행 된 데이터의 정상 학습률이 Wavelet Transform에 비해 약 1.94% 높게 나왔으며 고장 탐지결과를 활용하여 계산한 F-1 Score 또한 고장 데이터만을 포함한 경우 0.03점 더 높게 계산이 됐으며, 일부분의 정상 데이터를 포함한 경우에는 더 낮게 나와 정확성은 높고 오차율은 낮은 결과를 나타냈다.

향후 연구에서는 여러 기계 장비의 고장 유형 별 데이터를 활용하여 전처리 기법의 성능 비교를 진행할 계획이며, 딥러닝 기법을 활용한 Anomaly Detection 기법을 활용하여 고장 데이터 즉, Anomaly를 탐지하여 고장 탐지율과 F-1 Score를 비교하여 DAE 기법과 Wavelet Transform 기법의 성능을 좀 더 자세하게 비교하는 연구를 진행할 계획이다.

감사의 글

이 논문은 2021년도 산업통상자원부, 해양수산부 재원으로 한국산업기술진흥원(P0001968, 2021년 산업혁신인재성장 지원사업)과 해양수산과학기술진흥원(스마트 항만-자율운항 선박연계 기술개발)의 지원을 받아 수행된 연구임.

References

[1] Ahn, B. H., H. J. Kim, S. W. Park, Y. S. Kim, and H. K. Choi(2015), Study on Noise Reduction Method of Acoustic Emission Signal for Rotorcraft Gearboxes Condition Monitoring and Diagnosis, Journal of the Korean Society for Noise And Vibration Engineering (KSNVE), pp. 359-363

- [2] Cho, Y. J., Y. H. Ji, G. H. Nam, S. Y. Jeong, and H. C. Lee(2021), PCA-One Class SVM Based Anomaly Detection on TMED Hybrid Vehicle, The Korean Institute of Electrical Engineers, pp. 168-169.
- [3] Hwang, B. Y., J. H. Jung, and J. M. Lee(2015), Advanced Sound Source Localization Study Using De-noising Filter based on the Discrete Wavelet Transform (DWT), Institute of Control, Robotics and Systems. Journal of Institute of Control, Robotics and Systems 21(12), pp. 1185-1192.
- [4] Jang, J. Y., B. W. Min, and C. W. Kim(2016), Semiconductor process anomaly detection based on Denoising Autoencoder, Journal of the spring Academic Conference of Korean Institute Of Industrial Engineers, pp. 1928-1948.
- [5] Jang, J. G. and J. C. Lee(2020), Performance Comparison of Data-Driven Prognostics Techniques by Predicting the Remaining Useful Life of Bearing, Journal of the 26th Winter Academic Conference of the Korean CDE Association, pp. 331-333.
- [6] Jia, Y., G. Li, X. Dong, and K. He(2021), A novel denoising method for vibration signal of hob spindle based on EEMD and grey theory, Measurement, Vol. 169, 105490.
- [7] Jun, S. H.(2008), An Outlier Data Analysis using Support Vector Regression, Journal of Korean Institute of Intelligent Systems, 18(6), pp. 876-880.
- [8] Kim, D. H., S. J. Han, B. K. Jung, S. H. Han, and S. B. Lee(2019), A Machine Learning-Based Method to Predict Engine Power, Journal of the Korean Society of Marine Environment & Safety, Vol. 25, No. 7, pp. 851-857.
- [9] Kim, K. M., D. Y. Kim, and J. H. Lee(2014), Measuring Similarity Between Movies Based on Polarity of Tweets, Journal of Korean Institute of Intelligent Systems, 24(3), pp. 292-297.
- [10] Lee, J. E. and I. S. Kim(2011), A Study on the Fault Detection Technique of the Grid-Connected Photovoltaic System using Wavelet Transformation, The Transactions of Korean Institute of Power Electronics, 16(1), pp. 79-87.
- [11] Lee, J. H.(2021), Experimental Study on Application of an Anomaly Detection Algorithm in Electric Current Datasets Generated from Marine Air Compressor with Time-series Features, Journal of the Korean Society of Marine Environment & Safety, Vol. 27, No. 1, pp. 127-134.
- [12] Lee, J. H., S. Y. Yu, S. C. Sin, D. H. Kang, S. S. Lee, and J. C. Lee(2019), A Study on the Development of Failure Prediction Algorithm for Bearing Using Machine Learning Algorithm, The Korea Marine Engineering Association, 43(6), pp. 455-462.
- [13] Lu, Q. and M. Li(2021), A Method Combining Fractal Analyss and Single Channel ICA for Vibration Noise Reduction, Hindawi Shock and Vibration, Volume 2021, Article ID 5583587, 10 pages.
- [14] Raj, A. S. and N. Murali(2013), Morlet Wavelet UDWT Denoising and EMD based Bearing Fault Diagnosis. ELECTRONICS, Vol. 17, No. 1
- [15] Tsui, K. L., N. Chen, Q. Zhou, Y. Hai, and W. Wang(2015), Prognostics and Health Management: A Review on Data Driven Approaches.
- [16] Yang, B., B. Zheng, Y. Zhang, X. Zhu, D. Zhang, and Y. Jiang(2021), The Vibration Trend Prediction of Hydropower Units Based on Wavelet Threshold Denoising and Bidirectional Long Short-Term Memory Network, 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA).
- [17] Yoon, Y. J., and Y. J. Jung(2019), Detection of System Abnormal State by Cyber Attack. Journal of The Korea Institute of Information Security & Cryptology, Vol. 29, No. 5.

Received : 2021. 11. 08.

Revised : 2021. 12. 14. (1st)

: 2021. 12. 22. (2nd)

Accepted : 2021. 12. 28.