

碩士學位 請求論文

指導教授 鄭 鍾 筆

소음 환경의 공장 상황에서 베어링 결함
검사를 위한 SSA-SL트랜스포머
프레임워크에 관한 연구

成均館大學校 一般大學院

스마트팩토리 融合學科

李 書 營

碩士學位 請求論文

指導教授 鄭 鍾 筆

소음 환경의 공장 상황에서 베어링 결함
검사를 위한 SSA-SL트랜스포머
프레임워크에 관한 연구

Research on SSA-SL Transformer Framework for
Bearing Fault Diagnosis under Noisy Factory
Environments

成均館大學校 一般大學院

스마트팩토리 融合學科

李 書 營

碩士學位 請求論文

指導教授 鄭 鍾 筆

소음 환경의 공장 상황에서 베어링 결함
검사를 위한 SSA-SL트랜스포머
프레임워크에 관한 연구

Research on SSA-SL Transformer Framework for
Bearing Fault Diagnosis under Noisy Factory
Environments

이 論文을 工學 碩士學位請求論文으로 提出합니다.

2022 年 4 月 日

成均館大學校 一般大學院

스마트팩토리 融合學科

李 書 營

이 論文을 李 書 營의 工學
碩士學位 論文으로 認定함.

2022 年 6 月 日

審査委員長 박 호 건

審査委員 김 재 광

審査委員 정 중 필

목차

제1장 서론	1
1.연구 배경과 목적	1
2.논문의 구성 및 기여	3
제2장 관련연구.	5
1.베어링 결함 감지	5
2.SSA 알고리즘	6
3.LSTM	9
4.바닐라 트랜스포머	10
제3장 SSA-SL트랜스포머	12
1.SSA 변형	14
2.SL 트랜스포머	15
제4장 실험 및 결과	19
1.실험환경	19
2.SSA 노이즈 분해	22
3.Swish 활성화함수	28
4.성능지표	30
5.결과	32
제5장 결론	37
1.연구결과	37
2.향후 연구 계획	39

참 고 문 헌	40
ABSTRACT	45

표목차

표 4-1. SL 트랜스포머와 비교 모델의 Accuracy와 F1-score 표.....	34
표 4-2. SSA 알고리즘 미적용 Accuracy 표.....	35
표 4-3. SSA 알고리즘 미적용 F1-score 표.....	35
표 4-4. SSA 알고리즘 적용한 Accuracy 표.....	36
표 4-5. SSA 알고리즘 적용한 F1-score 표	36

그림목차

그림 2-1. Singular Spectrum Analysis 알고리즘 구조도	8
그림 2-2. 바닐라 트랜스포머 구조도	11
그림 3-1. SSA-SL 트랜스포머 프레임워크	13
그림 3-2. SL 트랜스포머 구조도	17
그림 3-3. 인공지능 모델 학습 과정 블록 다이어그램.....	18
그림 4-1. CWRU 베어링 시뮬레이터	20
그림 4-2. 롤링베어링의 구성요소.....	21
그림 4-3. 데이터 비교 플랏	23
그림 4-4. SSA가 노이즈를 분해하는 과정 (1)	24
그림 4-5. SSA가 노이즈를 분해하는 과정 (2)	26
그림 4-6. SSA가 노이즈를 분해하는 과정 (3)	27
그림 4-7. Activation Function 비교.	29
그림 4-8. SL 트랜스포머의 Accuracy 그래프	34
그림 4-9. SL 트랜스포머의 Loss 그래프	34
그림 4-10. SSA 미적용 Accuracy 그래프.....	35
그림 4-11. SSA 적용 Accuracy 그래프.....	36

논문요약

소음환경의 공장 상황에서 베어링 결함 검사를 위한 SSA-SL트랜스포머 연구

본 연구는 스마트팩토리 관련연구 중 베어링 결함 검사 연구를 다룬다. 베어링 결함 검사 연구는 다음과 같은 한계점이 있었다. 대부분의 베어링 결함 검사 관련 인공지능 모델 연구들은 기존의 모델을 활용하는 데에 그쳤다. 그리고 다양한 전처리 기법에 대한 연구가 부족했다. 특히 노이즈에 관한 연구가 부족했다. 또 실제 공장의 상황을 반영하는 선행연구들이 부족했다. 따라서 본 연구에서는 언급한 한계점들을 극복하는 방향으로 연구가 진행되었다. 최신 동향을 반영하는 트랜스포머(Transformer) 인공지능 모델을 연구에 활용하고 모델을 변형하여 시계열 데이터에 좀 더 맞게 구현하였다. 노이즈에 강한 Singular Spectrum Analysis (SSA) 알고리즘을 재조명하였다. 따라서 SSA-SL트랜스포머 프레임워크를 제안한다. 그리고 베어링 데이터에 가우시안 노이즈를 삽입하여 공장환경을 재현하였다. 해당 프레임워크의 효용성을 입증하기 위해 3가지의 실험을 진행하였으며, 특히 마지막 실험에서는 SSA-SL트랜스포머가 96%의 우수한 Accuracy 성능을 보이기도 하였다.

주제어 : 베어링 결함 감지, Singular Spectrum Analysis, 트랜스포머, 공장 환경 재현

제1장 서론

1.1 연구배경과 목적

스마트팩토리 관련 연구 중에서, 기계의 상태 모니터링에 대한 수요는 점점 늘어나고 있다. 이 이유는 기계 고장으로 인한 다운타임의 증가가 엄청난 경제적 손실로 이어지기 때문이다. 따라서, 기계 시설물이 안정적으로 작동하기 위한 보다 나은 진단 모니터링 방법에 대한 요구가 지속되고 있다. 본 논문에서 설명하는 방법은 주로 딥러닝 방법론을 사용한다. 본 논문은 다양한 기계 및 구성 요소 간의 베어링 결함을 진단하기 위한 딥러닝 접근법을 연구한다[1-5].

공장에서 사용하는 기계시설물의 가장 중요한 부품은 모터다. 전기 장치인 모터의 결함은 공장에 치명적인 손상으로 올 수 있다. 모터의 결함 중 베어링의 결함이 가장 많이 나타난다. 모터 결함은 주로 과전류로 인해 발생한다. 본 논문에서는 이러한 과전류로 인한 베어링 결함의 검사가 주요 관심사다. 가장 간단하게, 베어링은 축을 지지하는 장치다. 베어링은 기계에서 사용되는 동안 회전축과 왕복운동을 지원하는 기계적인 요소로, 정확하고 매끄러운 움직임을 제공한다. 즉, 베어링은 샤프트가 부드럽게 회전하도록 유도하여 하중을 지지하고, 회전축의 위치를 결정하며, 하중이 변화해도 위치를 유지하는 역할을 한다. 따라서 베어링은 축이 있는 기계 설비에서 매우 중요한 요소다. 이러한 이유로 많은 기존 연구에서는 베어링 이상 진단에 초점을 맞췄다 [6-10].

다음은 베어링 결함 검출에 대한 기존의 연구들을 설명한다. 초기에는 베어링 데이터를 기반으로 한 통계적 접근법 혹은 기계학습 방법론이 주로 사용되었다. 결함 분류 프로세스는 Support Vector Machine (SVM) 및 XGBoost와 같은 다양한 기계 학습 방법을 사용하여 해결되었다 [11-13]. 그러나 [14] 연구 이후 딥러닝 방법이 등장함에 따라 베어링 이상 연구에 딥러닝 방법론 접근법이 베어링

결함감지 연구에도 본격적으로 도입되었다. 딥러닝 도입 이후 다음과 같은 연구들이 주로 적용된다. Recurrent Neural Network (RNN) 종류의 인공지능 신경망 모델은 베어링 진동 데이터의 시계열 데이터를 처리하는 주요 방법으로 사용되었다 [15, 16]. 인공지능 모델의 성능을 향상시키기 위해 대부분의 연구는 Convolutional Neural Network (CNN) 모델을 사용하기도 한다. 1차원 데이터를 2차원 데이터로 전처리 작업을 진행 후 CNN모델을 활용한다 [17].

그러나 베어링 결함진단에 대한 연구는 아직 불충분하다는 것을 확인했다. 인공지능 모델에 최신 트렌드를 적용할 필요가 있었다. 그리고 다양한 방법의 전처리 방법론을 적용할 필요가 있었다. 실제 공장 상황을 가정한 연구가 부족했다. 첫째로, 인공지능 모델링의 최신 경향인 트랜스포머 모델을 적용한 베어링 결함 검사 연구가 부족한 것을 발견했다. 텍스트 생성 분야에서 압도적인 성능을 보인 트랜스포머 모델은 현재 이미지 분류와 같은 다양한 분야에서 사용되고 있다 [18]. 이 트랜스포머 모델은 다른 분야에서도 압도적인 성능을 보였다. 본 연구에서도 트랜스포머 모델이 가장 좋은 성능을 낼 수 있다는 것을 가정할 수 있었다. 현재 여러 연구에서 SOTA 성능을 산출한 것으로 알려져 있다 [19, 20]. 둘째로, 베어링 결함 진단 연구에 사용될 진동 데이터는 다양한 방법의 전처리 기법을 사용해 형태 변환을 할 수 있다. 기존의 연구에서는 Fast Furier Transform (FFT)나 Wavelet같은 전처리 기법들을 베어링 결함 감지 연구에 사용하였다 [14, 21]. 그러나 이 전처리 기법들은 여전히 데이터 내부 노이즈가 섞인 상황에서 취약한 성능을 보였고 노이즈에 강건한 전처리기법을 찾을 필요가 있었다. 셋째로, 대부분의 공용 베어링 데이터는 데이터 내부에 잡음이 섞여 있지 않았다. 그러나 실제 공장 환경은 데이터 수집 과정에서 다양한 문제들이 발생한다. 따라서 깨끗한 베어링 데이터 연구는 실제 공장 환경을 에 적합하지 않을 수 있다. 선행 연구 중 공장 환경을 고려해서 데이터에 노이즈를 추가하여 프레임워크의 성능을 평가하는 논문도 있었다. 그러나 잡음이 얼마나 추가되었는지는 밝히지 않았다 [24, 25]. 따라서 본 논문의 또 다른 목적은 언급된 연구의 한계를 보완함도 목적이다.

1.2 논문의 구성 및 기여

다음은 본 논문의 Contributions다. 첫째, SSA-SL 트랜스포머 프레임워크를 제안한다. 첫째, SSA-SL 트랜스포머 프레임워크를 제안한다. SSA는 데이터 전처리 방식이며 SL 트랜스포머는 기존 바닐라 트랜스포머 모델을 변형한 인공지능 모델이다. Singular Spectrum Analysis (SSA) 는 시계열 데이터를 여러 개의 sub-series로 분해 후 데이터를 재구성하는 알고리즘이다. SL 트랜스포머는 Swish-LSTM 트랜스포머의 약자다. 본 연구에서는 Swish 활성화 함수를 사용한 LSTM 블록들을 활용한다. 이 블록들은 본래 바닐라 트랜스포머 내부의 MLP 블록들을 대체한다. 시계열 데이터의 노이즈를 고려하는 알고리즘이란 점에서 SSA가 베어링 결함 진단에 적합한 방법이 될 것이라는 가설을 세울 수 있었다. 그리고 실제로 SL 트랜스포머는 기존 바닐라 트랜스포머보다 성능이 좋게 나타난다. 두번째는, 공장 상황을 가정한다는 것이다. 베어링 이상 진단과 관련된 선행 연구는 이미 95%이상의 Accuracy 성능을 보였다. 따라서 트랜스포머를 활용한 연구는 이미 높은 성능의 정확도를 예상할 수 있었다. 그러나 이 수치는 현실적이지 않다고 여겨졌다. 공장에는 다양한 변수들이 존재하기 때문이다. 실제 공장에서 센서를 통해 데이터를 수집하면 소음 발생으로 어려움을 겪는다는 선행연구들이 존재한다. 따라서 본 연구에는 데이터에 노이즈를 섞는다. 그 후 소음의 정도를 구분한다. 주요 목표는 논문에서 제안된 프레임워크가 소음에 얼마나 강력한지를 확인하는 것이다. 엔지니어와 후속 연구자들이 보다 실질적인 연구 결과를 얻을 수 있으리라 기대한다. Contributions의 요약은 다음과 같다.

1. SSA-SL 트랜스포머 프레임워크를 제안한다. 본 연구에서는 노이즈에 강건한 전처리 기법인 SSA 전처리 기술을 재조명하고 다양한 Metrics를 활용해 SL 트랜스포머의 성능을 확인한다.

2. 실험은 현실적인 공장 환경을 가정하여 수행한다. 베어링 데이터에 백색 가우시안 노이즈를 섞어서 실제 공장환경을 재현하여 실험을 수행했다. 엔지니어와 연구자들은 본 연구를 통해 좀 더 실질적인 결과를 살펴볼 수 있으리라 기대한다.

본 논문은 다음과 같이 구성되었다. 2장에서는 선행 베어링 결함 감지 연구, LSTM, SSA 알고리즘 및 트랜스포머 인공지능 모델을 소개한다. 그리고 3장은 제안된 SSA-SL트랜스포머 프레임워크를 설명한다. 4장은 데이터와 실험 환경에 대한 설명, SSA가 소음을 처리하는 방법, Swish 활성화함수의 기능에 대해 설명 후 3개 실험 결과에 대해 설명한다.

제2장 관 련 연 구

본 장에서는 베어링 결함 감지의 선행 연구들, LSTM 인공지능 모델, SSA 알고리즘, 트랜스포머 인공지능 모델에 대해 설명한다.

1. 베어링 결함 감지

베어링 결함 형성 과정은 다음과 같다. 베어링 내부에서 미세한 균열이 발생한 후 내부에도 미세한 균열이 응집되며 표면 손상이 발생한다. 이후에 윤활유 부족으로 인해 베어링 표면이 접촉하거나 금속이 손상되거나 비정상적으로 과도한 외력이 베어링에 가해진다. 대부분의 결함은 바깥쪽 레이스, 안쪽 레이스 또는 공에서 나타난다. 결함은 직경에 따라 분류할 수 있다. 이러한 베어링 결함에 대한 선행 연구는 [6–10]이 있다. 다수의 선행 연구는 오토인코더를 활용한 이상감지 프레임워크 연구로 접근하였다 [26, 27]. 또 다른 다수의 선행 연구에서는 인공지능 모델 성능을 높이기 위해 데이터를 전처리 하는 방법에 대해 고민한다. 1차원 데이터를 2차원 데이터로 변환하는 시도들이 대표적이다. 2D 이미지 데이터로 변환 후 CNN 인공지능 신경망을 사용하여 분류성능을 향상시키는 기존의 연구들이 대표적이다. 이 외에도 베어링 고장 진단 프로세스에서 노이즈를 해결하기 위한 다양한 방법론이 있다 [28, 29].

2. SSA 알고리즘

본 연구는 비교적 주목받지 않았던 SSA 전처리 방법에 대해 재조명하는 것이 주요 목표다. SSA는 시계열 데이터를 간단한 구성요소로 분해하여 재구성하는 알고리즘이다 [24, 25]. [그림 2-1]은 SSA를 잘 나타내는 그림이다. [그림 2-1]의 분해 과정을 거치면서 중간에 SVD 프로세스를 지나고 그 후 재구성 프로세스를 통해 데이터가 재구성되는 구조다 [33].

분해 과정은 [그림 2-1]의 임베딩 (a) 과 SVD (b) 과정을 거친다. 임베딩 프로세스에서 1차원 시계열 데이터는 행켈 행렬로 변환된다. 이 행켈행렬은 꺾적 행렬이라고도 불린다. 길이 N 이 1차원 시계열이 있다고 가정하자. $F = (f_0, \dots, f_N)$ 가 있다고 할 때, 임베딩 프로세스에서 시계열 F 의 지연된 버전의 정의로 나타난다. 행렬은 $(L \leq I/2)$ 에 의존한다. 이는 윈도우 사이즈이며 사용자의 임의로 결정된다. 행켈행렬인 $K = I - L + 1$ 을 정의하면 다음과 같다.

$$X = [X_1, \dots, X_K] = \begin{bmatrix} f_0 & f_1 & f_2 & \cdots & f_{K-1} \\ f_1 & f_2 & f_3 & \cdots & f_K \\ f_2 & f_3 & f_4 & \cdots & f_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{L-1} & f_L & f_{L+1} & \cdots & f_T \end{bmatrix} \quad (1)$$

여기서 대각선 ($I + j = \text{const}$)의 모든 요소들은 대칭적으로 동일하다. 그리고 [그림 2-1] (b)의 SVD 프로세스는 다음과 같다. SVD 프로세스에서 행켈행렬은 2차원 쌍직교 기본 행렬의 합으로 변경된다 (2).

$$X = X_1 + \cdots + X_K \quad (2)$$

기본 행렬 X_i 은 $X_i = s_i \mathbf{U}_i \mathbf{V}_i^T$ 로 정의된다. s_i 는 X 의 i 번째 단일 값이다. \mathbf{U}_i 는

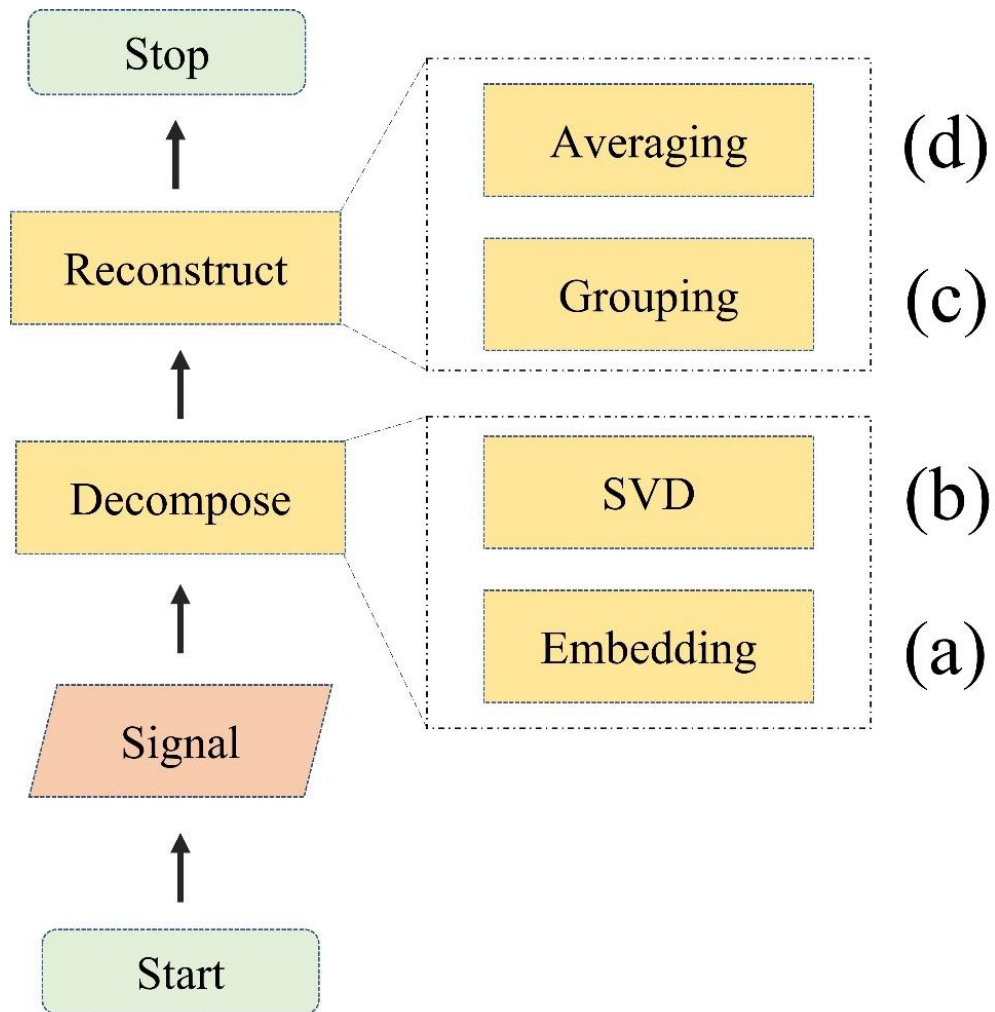
X 의 i 번째 좌벡터 단일 값이다. V_i 는 X 의 우벡터 단일 값이다. s_i, U_i, V_i 들을 SVD의 정삼항이라고 한다. 여기까지 SSA알고리즘의 분해 과정이었다.

분해 후에는 재구성 절차가 진행된다. 재구성에서는 그룹화 (c) 와 애버리징 (d) 절차가 수행된다. 그룹화에서는 SVD의 결과물을 사용한다. d 개의 정삼항에서 r 개를 선택한다. R 은 SSA알고리즘의 파라미터다. $X_I = X_{i1} + X_{i2} + \dots + X_{ir}$ 에서 X_I 는 신호 F 와 관련이 있다. 그리고 남은 $(d - r)$ 정삼항은 에러 ε 를 대표한다.

최종적으로 애버리징 과정 (d)에서는 (c)에서 r 그룹이 선택된다. $I = X_{i1} + X_{i2} + \dots + X_{ir}$ 는 시계열 데이터 y_{i1}, y_{i2}, y_{ir} 의 행켈화 과정에 의해 변형된다. 만약 k 번째에서 재구성된 시계열 데이터가 $i + j = k + 1$ 일 경우, k 번째 요소는 Z 행렬을 애버리징 하는 과정에서 구할 수 있다. $H()$ 를 행켈화 라고 하자, 즉 행렬 Z 를 행켈화 하는 과정을 $H(Z)$ 라 한다. 결론적으로 다음과 같은 식으로 마무리 지을 수 있다.

$$y = H(X_{i1}) + H(X_{i2}) + \dots + H(X_{ir}) + e \quad (3)$$

공식 (3)에서 y 는 SSA의 목표인 재구성된 데이터를 의미한다. SSA 알고리즘은 다음과 같은 선행 연구들이 있었다. SSA는 LSTM 모델을 사용하여 시계열 예측을 위한 전처리 알고리즘으로 사용되었다. 또한 유사한 연구가 기계 결함 탐지 예측 모델 전처리 방법론으로도 사용되었다 [34, 35].



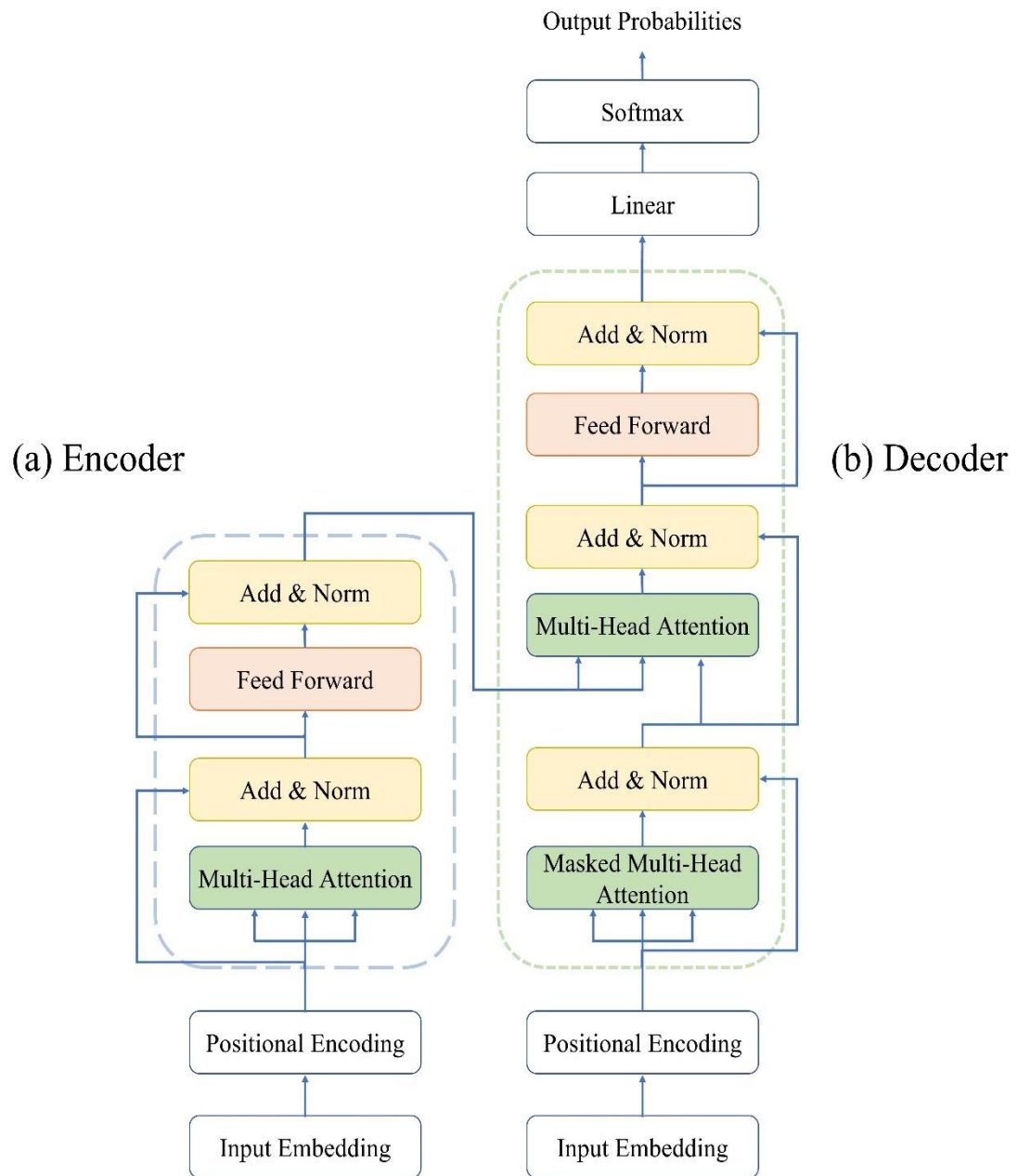
[그림 2-1] Singular Spectrum Analysis 알고리즘의 구조도

3. LSTM

이전에는 순환 신경망 인공지능 모델 (RNN) 이 시계열 데이터와 관련된 연구로 사용되었다. 그러나 RNN에는 한계가 있었다. 따라서 개발된 장기 메모리 신경망 (LSTM)은 RNN의 단점을 보완하기 위해 개발되었다. LSTM은 RNN의 기울기 손실 문제를 해결한다. LSTM은 먼 과거의 이벤트를 학습 가중치로 적용한다. LSTM의 장점은 시계열 데이터 처리 능력이 우수하다는 것이다 [36]. LSTM 시계열 데이터를 분류하는 방법으로 LSTM 모델을 사용한 사전 연구들이 있었다 [37]. 또한 본 연구에서는 LSTM 모델을 사용하여 다변량 시계열 데이터와 함께 사용하여 높은 성능을 보여준다 [38].

4. 바닐라 트랜스포머

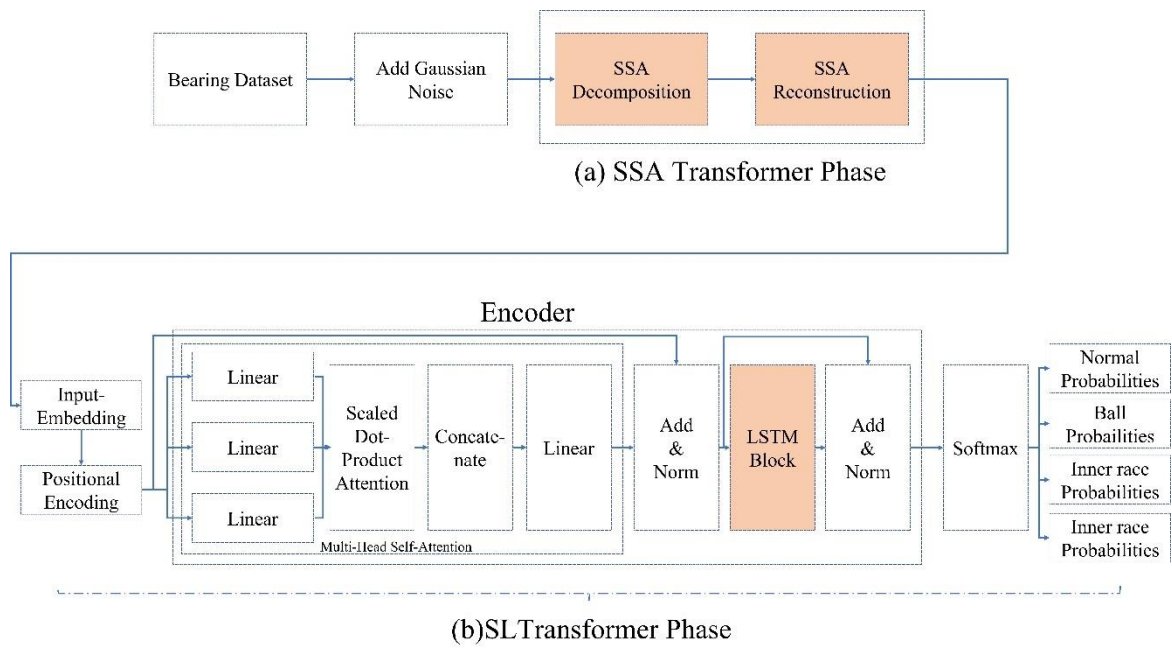
바닐라 트랜스포머는 self-attention의 중요성을 강조하면서 유명해졌다 [18]. 그리고 이 모델은 텍스트 생성 분야에서 곧바로 State-of-the-art (SOTA)의 성능을 점령한다. 이후 텍스트 분류, 비전 분류 등 다양한 분야에서 바닐라 트랜스포머를 변형한 형태들이 활용되어 왔다. 트랜스포머 모델의 가장 큰 특징은 기존 RNN 구조 인공지능 모델 아키텍처에서 반복 특성을 제거하고 self-attention module을 도입했다는 것이다. 트랜스포머 모델은 주로 사용 목적에 따라 인코더와 디코더로 나뉜다. [그림 2-2] (a)는 인코더를 나타내고 [그림 2-2] (b)는 디코더를 나타낸다. 인코더는 주어진 텍스트를 일종의 가중치로 변환하는 역할을 수행한다. 이 경우, 텍스트의 표현을 임베딩 혹은 특징이라는 표현으로 서술된다. 주어진 입력에 대해, 입력은 양방향 self-attention을 적용하여 인코딩된다. 또한 디코더는 인코더에서 표현을 디코딩하는 역할을 한다. 디코더는 마스크 된 self-attention을 사용하여 단방향 auto-regressive 학습을 수행한다. 본 논문에서는 트랜스포머의 여러 변형 형태 중 대표적인, Bidirectional Encoder Representations (BERT)를 수정하여 활용한다 [39]. BERT는 바닐라 트랜스포머에서 Encoder를 활용한 대표적인 모델로 텍스트 분류작업을 수행한다. BERT는 Rectified Linear Unit (ReLU) 활성화 함수 대신 Gaussian Error Linear Unit (GeLU)를 활용한다 [40, 41]. 그 후로, 트랜스포머 인코더를 사용하는 접근법이 다양한 분야에서 연구된다.



[그림 2-2] 바닐라 트랜스포머의 구조도

제3장 SSA-SL트랜스포머

SSA-SL 트랜스포머는 다음과 같은 아이디어에서 탄생한 프레임워크다. 복잡한 공장 상황을 가정하여, 우리는 소음 분해와 재구성에 뛰어난 SSA 알고리즘을 활용한다. 또한 현재 SOTA의 성능을 보여주는 트랜스포머 모델은 시계열 데이터에 적합한 모델로 탈바꿈하였다. 트랜스포머 내부 레이어 중 어텐션 블록에서 나가는 Feed Forward 레이어를 SL 블록이 적합한 것으로 가정하여 설계되었다. [그림 3-1]에서는 본 논문에서 제시한 프레임워크와 전체적인 실험을 참조할 수 있으며, 두 공정으로 구분할 수 있다. 첫째, 노이즈 데이터가 섞인 베어링 데이터로부터 노이즈를 분해하는 SSA 변환 과정이다. 노이즈 제거 작업은 분해 및 재구성 프로세스를 통해 수행된다. 그 다음은 SL 트랜스포머에서 Swish 활성화 함수와 LSTM 인공지능 모델이 트랜스포머 인코더 내부에서 활용된다. 레이어를 통과한 후, 소프트맥스 함수를 통해 데이터의 확률값을 도출한다. 이 확률값을 통해 데이터가 어떤 베어링 결함인지 예측할 수 있다.



[그림 3-1] SSA-SL 트랜스포머 프레임워크

1. SSA 변형

본 논문에서 SSA 알고리즘에 초점을 맞춘 이유는 다음과 같다. SSA는 본 논문의 주요 목적인 공장 환경을 대처하는데 주요한 알고리즘이기 때문이다. 2장에서 설명하듯 SSA는 시계열 데이터를 여러 하위 sub-series로 분해한다. 분해과정에서 다양한 성분들이 추출되고 노이즈가 성분 안에 포함된다고 가정한다. SSA 알고리즘에는 window길이 관련 변수인 L 값과 구성 요소의 수 r 값 두 가지 매개변수가 존재한다. 이 매개변수는 알고리즘 사용자의 재량에 따라 임의로 수정한다. 성분의 수 r 값은 다음과 같은 방법으로 결정된다. 일반적으로는 $\lambda_i/\Gamma(\Gamma = \sum_{i=1}^a \lambda_i)$ 에 대한 기여의 합이 적어도 사전 정의된 임계값이 되도록 d 구성요소 중 r 을 선택하는 것이다. 일반적으로 소음 성분의 기여도는 낮기 때문에 SSA 변환의 목적은 기여도가 낮은 부분을 제거한 후 시계열 데이터를 재구성하는 것이다. 일반적으로 추세와 잡음을 잘 관찰하려면 L 의 큰 값을 선택해야 한다. 그러나 시계열 데이터의 추세가 너무 복잡하면 많은 수의 고유 삼중항만을 고려하여 추출할 수 있다. 따라서 SSA 변환 단계의 목표는 잡음을 제거하고 정상 추세를 복원하는 것으로 볼 수 있다. 주요 변수인 L 을 추출하는 방법은 4장에서 설명한다.

2. SL 트랜스포머

이번 장에서는 SL 트랜스포머가 다음과 같은 가설로 설계되었다는 것을 설명한다. [그림 3-2]에서는 기존 트랜스포머의 Feed Forward가 SL 블록으로 어떻게 대체가 되는지를 보여준다.

먼저 [그림 3-2]를 설명한다. SSA 알고리즘을 거쳐 재구성된 데이터를 입력값으로 받는다. 입력된 시계열 데이터는 Input Embedding과 Positional Encoding을 거치면서 시계열 데이터의 특성을 나타낼 수 있게 쿼리(Q), 키(K), 값(V)으로 편집된다. 그 후 Q, K, V는 멀티헤드-셀프 어텐션을 거친다. 멀티헤드 셀프 어텐션을 거친 결과물은 Add 레이어에서 residual 값을 더하고 Normalization을 진행한다. 그 후 LSTM 블록을 거친다. LSTM은 Normalization을 한 후, Swish 활성화함수를 사용하는 LSTM 모델을 거친 뒤 Dropout, 그리고 LSTM 모델을 한 번 더 거치면서 끝난다. 다시 한 번 Add 계층에서 Residual을 받고 Normalization을 한 뒤, Softmax 함수를 통해 확률 값을 도출한다.

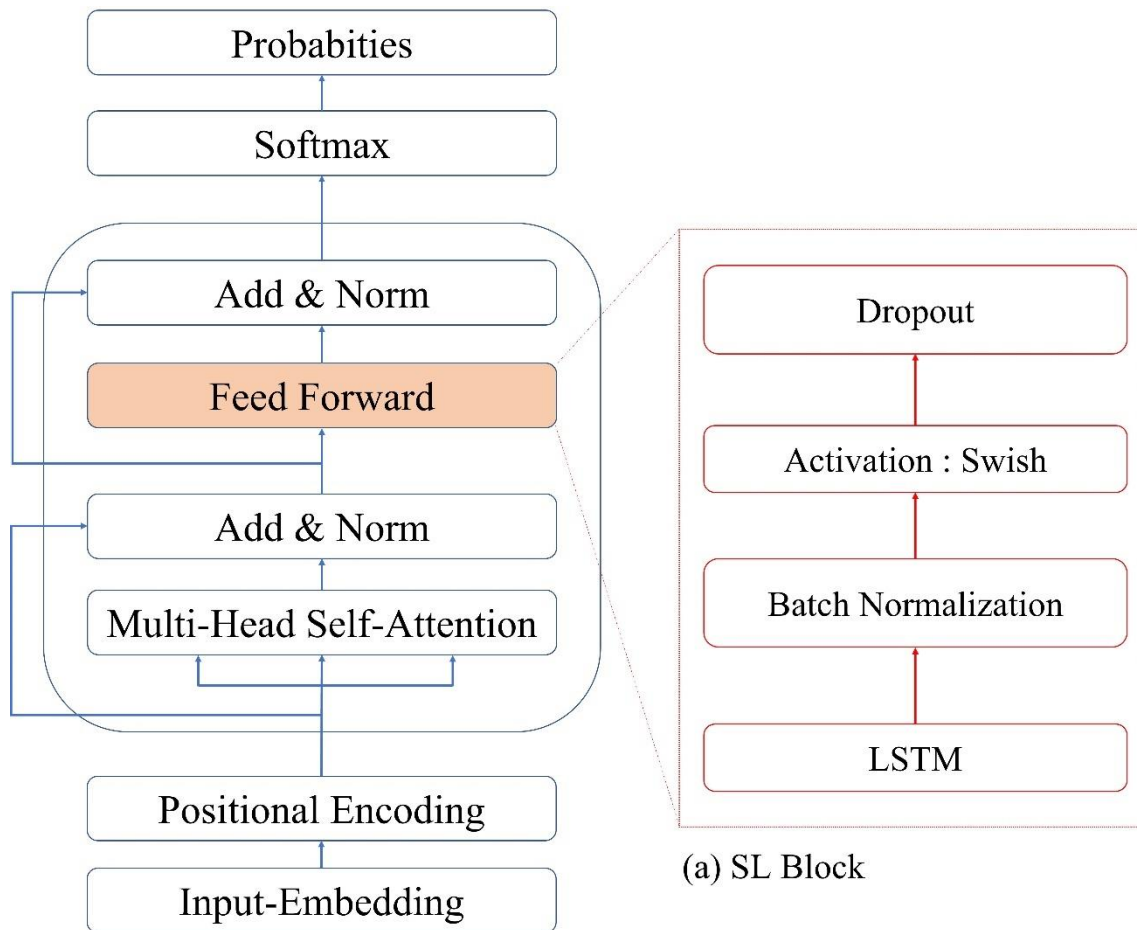
SL 트랜스포머는 Feed Forward 계층에 LSTM을 구성하는 방향을 가정한 방법론이다. 기존 연구에서는 Feed Forward 계층에 MLP를 구성하거나 CNN 등을 사용하는 다양한 접근법들이 있었다. 하지만 LSTM을 직접적으로 집어넣었던 선행 연구는 존재하지 않았다. Feed Forward에 LSTM 블록으로 레이어를 구성함으로써 다음과 같은 장점과 단점을 가정할 수 있었다. LSTM과 달리 트랜스포머의 장점은 입력 데이터들을 병렬처리 할 수 있다는 점이었다. 병렬처리용도의 멀티헤드 셀프 어텐션 아이디어는 고수하였다. Feed Forward를 MLP로 구성한다는 것은 시계열 속성을 잃어버릴 수 있음을 의미하기도 하였다. 따라서 LSTM으로 MLP를 대체하는 블록 레이어를 구성하였다. Recurrent 속성을 부여하여 gradient vanishing현상을 막는 아키텍처를 고안하였다. 그리고 기존의 ReLU보다 더 under bound가 깊은 Swish 활성화함수를 구성하여 gradient vanishing을 더 막을 수 있을

것으로 예상할 수 있었다. Swish는 구글이 제안한 활성화 함수다. Swish는 ReLU 활성화 함수보다 더 부드러운 곡선을 가져서 시계열 데이터에 적합한 가중치를 반영하는 것으로 알려져 있다. 또한 다양한 연구에서 시계열 데이터를 학습시킬 때 더 적합한 인공지능 모델로 수렴시킬 수 있음을 확인할 수 있었다. 다만 MLP보다 파라미터의 양이 늘어서 훈련 연산량이 늘어날 수 있다는 단점도 존재하였다. 그러나 병렬 처리의 철학은 그대로 고수하고 있어서 트랜스포머의 장점을 그대로 살릴 수 있을 것으로 보았다.

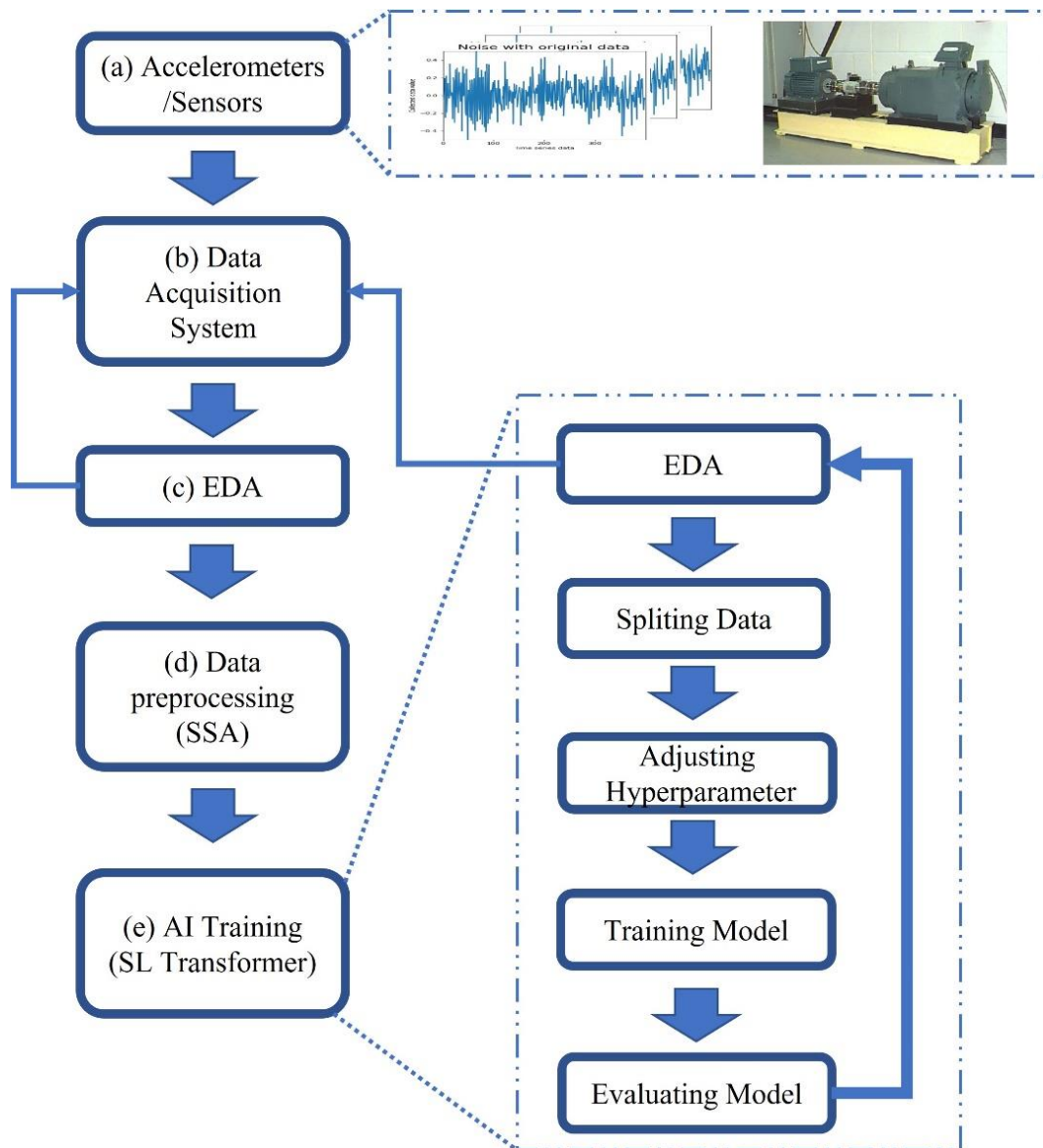
마지막으로 [그림 3-2]의 (a) SL 블록을 구성한 방식에 대한 설명이다. 해당 블록 구조는 [42]에서 설명한다. [42]에서는 Batch normalization과 dropout 층을 구성하는 방법론에 대해 설명한다. 모든 매개변수 값에 대해 네트워크가 항상 원하는 분포로 활성화를 생성하도록 보장하는 것이 목표다. 따라서 [42]에서 제안하는 구조는 Fully Connected, Batch Normalization, Activation, Dropout 순서다. 본 연구에서는 해당 방식을 적용하여 LSTM, Batch Normalization, Activation, Dropout으로 block layer를 구성하였다.

3장의 끝에서, 4장으로 넘어가기 전에 [그림 3-3]을 통해 전체 프로세스를 설명한다. [그림 3-3]에서 각 단계마다 중요한 과정이 있음을 알 수 있으며, (a)는 센서로부터 수신된 데이터를 나타낸다. 데이터의 경우, 가속도 (b)는 데이터 획득 시스템을 나타낸다. (a)와 (b)의 내용은 4.1장에 나타난다. CWRU 공공 데이터 수집 프로세스에 대한 자세한 설명이 나타난다. (c)는 탐색적 데이터 분석(EDA)에 대한 블록이다. EDA는 주로 데이터 결측 값이 없는지 또는 통계적으로 문제가 없는지 등을 확인한다. CWRU 데이터는 기존 연구에서 침도나 왜도에 문제가 없어 인공지능이 학습하기에 적합한 데이터로 알려져 있다. 결측 데이터가 너무 많거나 (c)에서 의도적인 조작을 의심하는 경향이 있을 경우 (b)단계로 되돌아가 데이터를 다시 수집한다. (d) SSA를 통한 전처리 과정이다. 이 과정에서 SSA 알고리즘을 적용하기 위한 최적의 L 값을 구한다. 이는 4.2장에 소개되었다. 그리고 (e)에는 인공지능이 학습하는 과정을 설명한다. 전처리된 데이터는 다시

EDA를 거치고 그 후 데이터는 훈련, 테스트 그리고 검증 셋으로 나뉜다. 모델 학습을 위한 하이퍼파라미터는 분할된 데이터를 기반으로 선택된다. 그리고 연구에서 선택한 메트릭에 따라 점수가 매겨진다. 성능이 가설과 다르거나 사전 연구와 결과가 크게 다르다면 EDA 프로세스로 돌아가 다시 데이터가 불량인지부터 확인한다. 실험의 주요 목표는 이 과정을 반복하여 가설로 설정한 메트릭 점수를 확인할 수 있는지를 확인하는 것이다.



[그림 3-2] SL 트랜스포머 구조도



[그림 3-3] 인공지능 모델 학습 과정 블록 다이어그램.

제4장 실험 및 결과

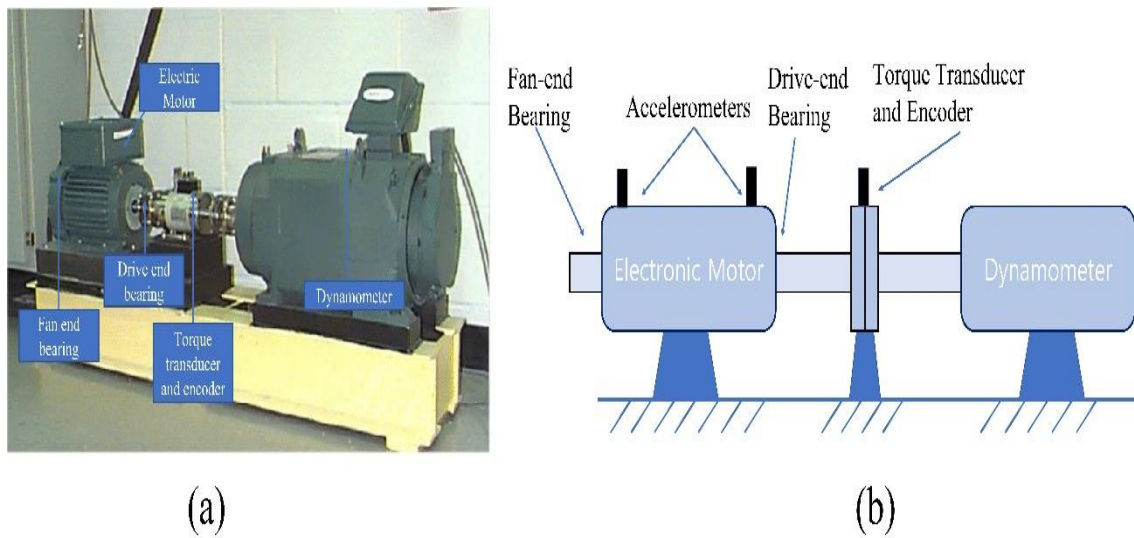
1. 실험 환경

본 논문에서 사용된 데이터는 케이스 웨스턴 리저브 대학교 (CWRU) 베어링 데이터 세트라고 한다. CWRU 데이터 세트는 정상 베어링과 결함 베어링을 나누는 지도 학습 기반의 데이터다. 이 데이터에는 베어링 내부 볼 결함, 내부 레이스 결함, 외부 레이스 결함 총 세가지의 결함이 존재한다. 베어링 진동 데이터는 0-3 마력 (모터속도 1797~1720 RPM)의 부하를 통해 측정되었다. 불량 테스트는 [그림 4-1] (a)에 나타난 기계 설비를 통해 측정되었다. [그림 4-1] (a)와 (b)를 보면 2개의 베어링 (왼쪽), 2hp 모터(왼쪽), 토크 트랜스듀어 및 인코더(가운데) 및 제어 전자기기(그림에 표시되지 않음)으로 구성된다. 그리고 [그림 4-1] (b)는 (a)의 단면도를 나타낸다.

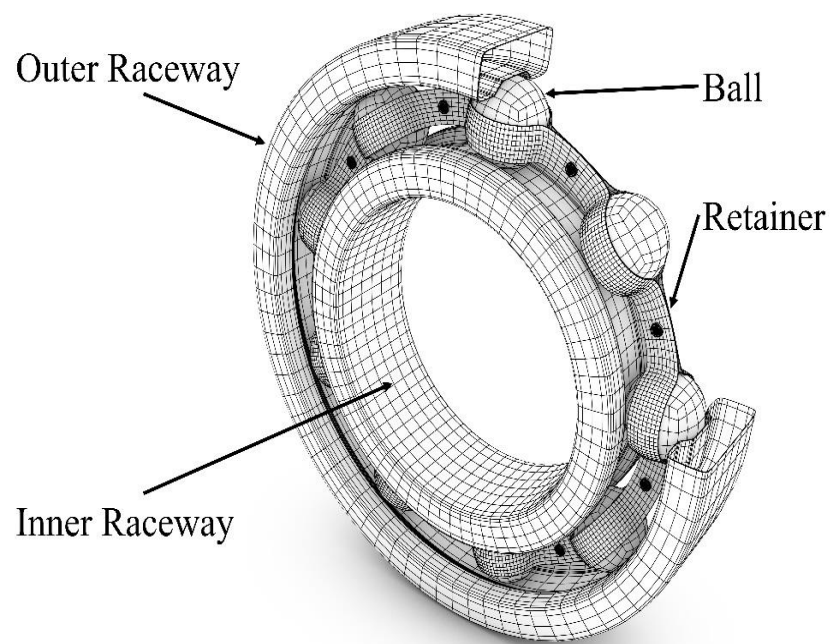
[그림 4-2]는 SKF6205-HC5C3의 롤링베어링 외관을 보여준다. 대부분의 결함은 외부 레이스 결함, 내부레이스 결함 그리고 공에서 나타난다. 데이터 셋에서는 4244개의 정상 데이터 사례, 4860개의 볼 결함 사례, 4862개의 내부 레이스 결함 그리고 8529개의 외부 레이스 결함이 포함되어 있다. 데이터 구조를 가정하면서, 우리는 데이터 증강 방법을 적용하는 것을 고려했다. 그러나 이 논문은 공장 환경을 성공적으로 재현하는 것을 목표로 하였다. 실험은 클래스 불균형 상황도 가정하여 진행되었기 때문에 데이터 증강은 시도하지 않았다.

실험에 사용된 하드웨어 성능은 다음과 같다. Intel core I7-9750h CPU @ 2.60Hz 2.59 GHz 2.59 GHz 32.0 Gb ram, Nvidia GeForce GTX in 1650. 같은 실험은 구글 코랩 프로를 사용해서 수행되기도 하였다. Google Colap Pro: GPU(T4 or P100). 두 환경 모두 근사한 성능을 보였다. 다만 인공지능 신경망을 학습할

때는 GPU 성능에 따라 신경망의 수렴 정도가 달라질 수 있음은 유의해야한다.



[그림 4-1] (a) CWRU의 베어링 시뮬레이터 (b) 단면도.



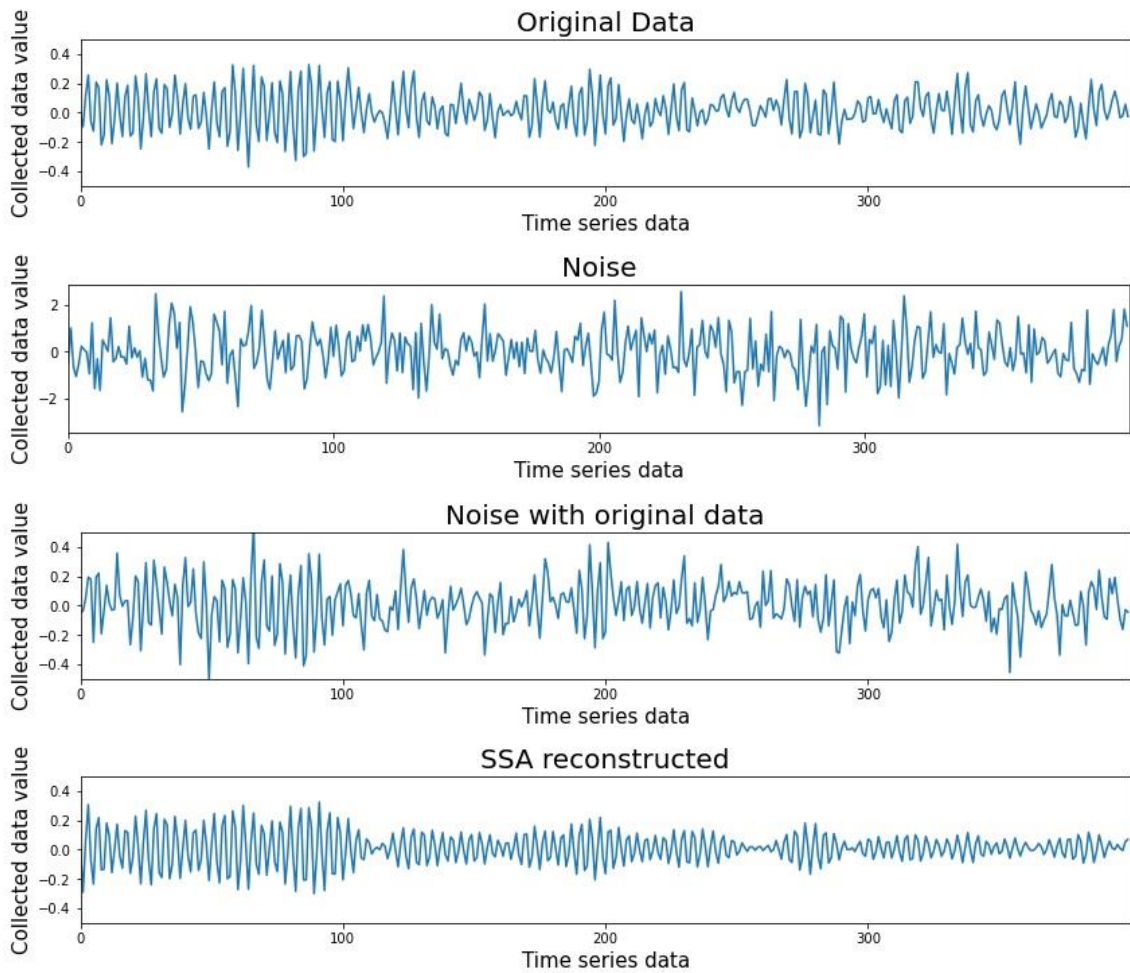
[그림 4-2] 롤링베어링의 구성 요소

2. 노이즈 분해

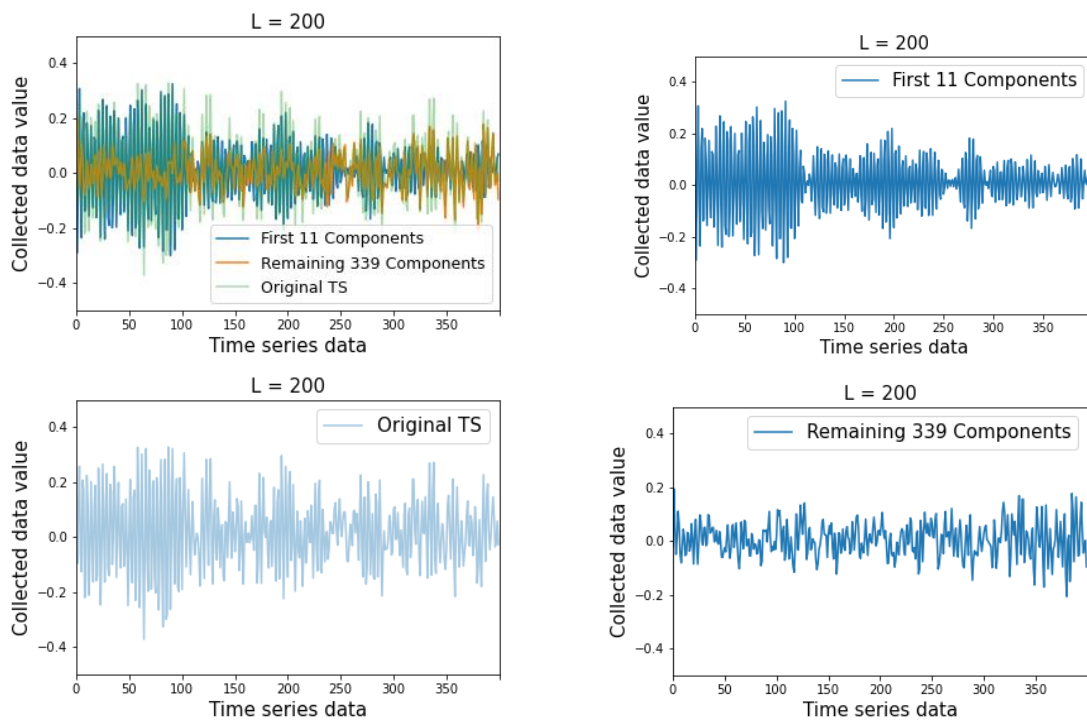
[그림 4-3]의 네 그래프는 각각 원본 데이터, 노이즈, 원본데이터가 포함된 노이즈 그리고 SSA 알고리즘으로 재구성된 플롯을 보여준다. 원본 데이터는 내부레이스 결함 데이터 중 하나의 사례를 가져왔다. 세 번째 그래프는 이 두 요소를 결합하고, 네번째 그래프는 세번째 그래프의 데이터를 SSA 알고리즘을 통해 재구성한 것이다.

[그림 4-3]의 하단에 있는 SSA 알고리즘으로 재구성된 데이터를 보면 노이즈가 제거되어 원래의 데이터가 잘 구성되었음을 확인할 수 있다. 다만 좀 더 자세한 과정을 살펴볼 필요가 있었다. [그림 4-4]는 3.1장에 기술된 SSA알고리즘 작동 원리에 기초한 소음 제거 프로세스를 시각화 하여 보여준다.

[그림 4-4]의 오른쪽 상단 그림은 처음 11개의 주요 구성요소를 보여준다. SSA알고리즘은 신호 데이터를 다양한 시계열 데이터로 먼저 분해한 후 분해된 데이터 중 가장 중요한 인자를 선택해 사용자의 재량으로 결정할 수 있다. 3.1장에서 설명한 바와 같이 SSA 알고리즘 그룹화 프로세스에서 하위 시리즈 요소의 90%를 선택하는 것이 중요하다는 것이 관찰되었다. 왼쪽 하단의 표기법은 하위 데이터에 남아있는 데이터를 나타낸다. SSA 알고리즘의 재구성 프로세스에서 이러한 잔류 데이터는 노이즈로 결정되며 사용되지 않는다. SSA 알고리즘은 다른 전처리 방법에 비해 노이즈를 제거하는데 이점이 있었다. 다만 잔류 데이터들을 제거하고 재구성하는 과정에서 과도한 데이터 재구성이 있기도 하였다. SSA 알고리즘을 이용한 세번째 실험에서 흥미로운 결과가 발견되기도 하였다. 정확도 성능은 잡음이 없는 첫번째 실험보다 잡음이 많은 세 번째 실험에서 더 높게 측정되었다. 이는 SSA 알고리즘이 패턴을 더 명확하게 만든다는 또 하나의 가설을 세울 수 있었다.



[그림 2-3] 데이터 비교 플랏, 위에서부터 원본 데이터, 노이즈, 노이즈와 원본 데이터를 섞은 데이터 그리고 SSA 알고리즘을 통해 재구성된 데이터

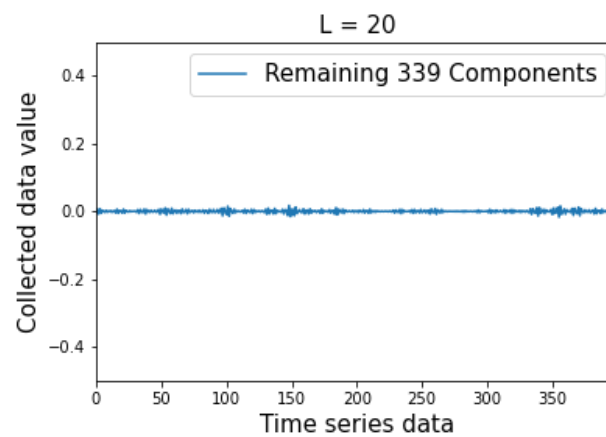
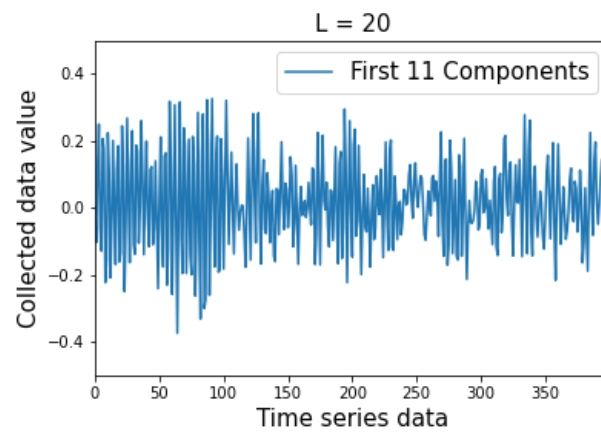
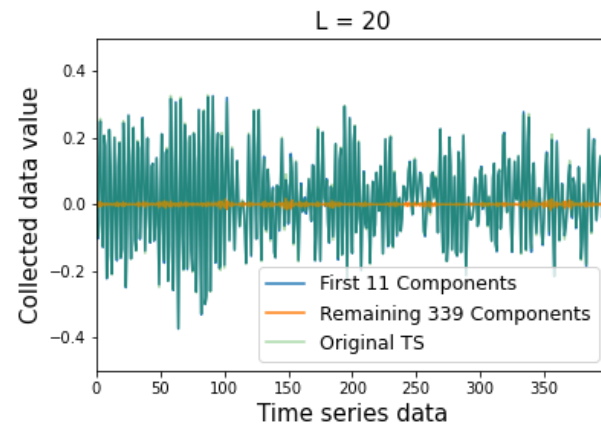


[그림 4-4] SSA가 노이즈가 섞인 데이터를 분해하는 과정을 보인다. 왼쪽 상단은 모든 요소가 섞인 플랏 그래프, 오른쪽 상단은 주요 11개 요소를 시각화 한 플랏 그래프, 왼쪽 하단은 원본 시계열 데이터, 우측 하단은 남은 339 요소를 시각화 한 플랏 그래프.

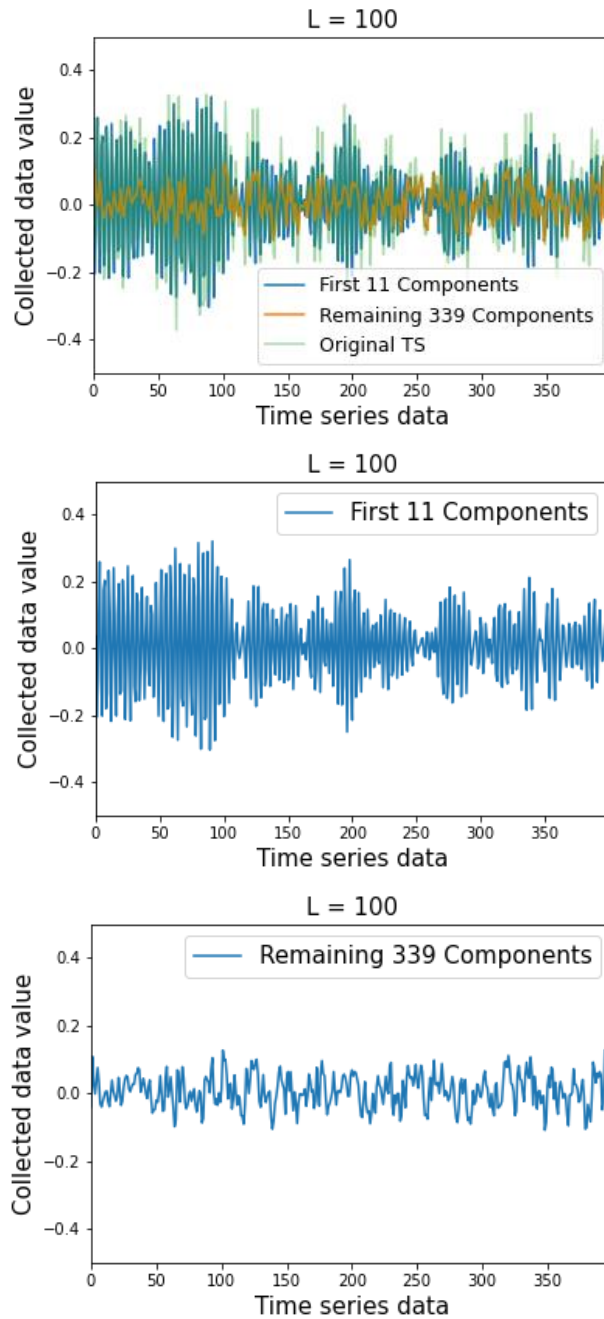
[그림 4-4]에서는 SSA의 파라미터 중 L 의 값이 200으로 설정 되어있다. 이 매개변수는 SSA를 사용할 때 매우 중요한 변수다. 다음은 적절한 L 값을 추정하는 과정을 시각화 한 자료다.

[그림 4-5]의 L 수치는 20이다. [그림 4-5]의 맨 아래 그림을 보면 L 값이 너무 작아 노이즈가 잘 나타나지 않음을 확인할 수 있다. [그림 4-6]은 L 값이 100인 상황을 보여준다. L 이 20인 경우보다 더 노이즈를 잘 분해했음을 알 수 있다.

상기 방법을 거치지 않아도 $L=T/2$ 로 사용되는 방법이 가장 널리 알려져 있다. 그러나 [17]는 SSA를 사용할 때 L 값이 $T/2$ 가 무조건 적합한 파라미터 값으로 보장될 수는 없다고 말한다. 따라서 L 값 이하의 값 중 가능한 경우를 고려하여 추세선과 기타 성분을 잘 분해하는 최적의 L 값을 찾는, 즉 최적화를 하는 L 값을 찾아야한다.



[그림 4-5] L값이 20인 경우.



[그림 4-6] L 값이 100인 경우.

3. Swish 활성화 함수

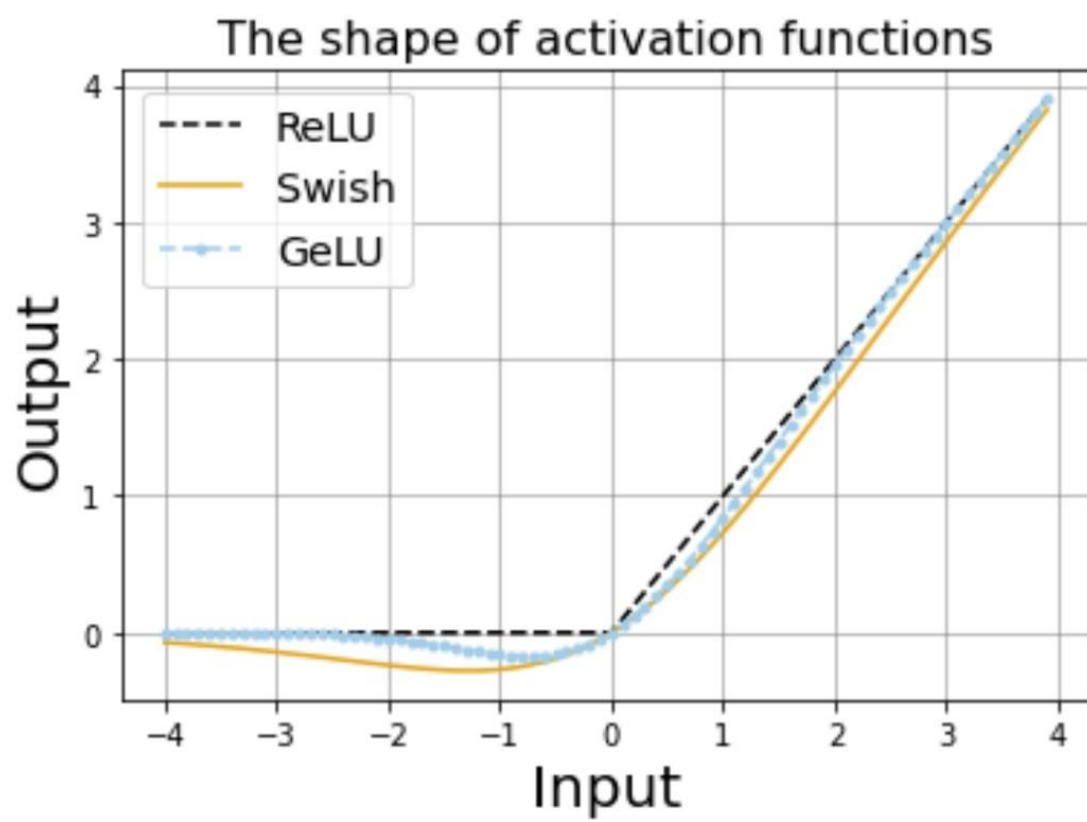
전통적으로, ReLU 활성화 함수는 딥러닝 모델에서 자주 사용되어왔다. 그러나 최근 많은 연구에서 다양한 활성화 함수의 사용에 대해 주목하고 있다. BERT에서 ReLU와 비교하여 GeLU는 부드러운 곡선을 가지고 있어 그라디언트 소실을 줄일 수 있다 [39]. 이는 GeLU가 음수에서도 미분을 할 수 있어 작은 그라디언트를 전달할 수 있다.

이는 [그림 4-7]에서 볼 수 있다. [그림 4-7]에서 검은색 선은 ReLU를, 하늘색선은 GeLU를 나타낸다. GeLU는 ReLU보다 곡선이 매끄럽지만 본 논문에서는 Swish가 더 매끄러워 그라디언트를 더 잘 전달할 수 있다는 추측을 세울 수 있었다. Swish의 공식은 (4)와 같다.

$$f(x) = x \cdot \sigma(x), \text{ where } \sigma(x) = \frac{1}{1+e^{-x}} \quad (4)$$

Swish는 GeLU보다 음수에서 x 까지의 값의 범위가 넓다. 이는 GeLU가 약간 음수 값을 허용하는 것보다 더 넓은 범위를 허용하여 각 노드에서 수신한 그라디언트를 더 잘 표현할 수 있게 한다. 작은 음수가 기울기로 전송되더라도 이전 계층으로 전송될 수 있으므로 학습은 더 성공적이게 된다.

따라서 스위시는 기존 ReLU의 한계를 극복하면서 GeLU보다 더 나은 유형의 활성화 함수가 될 것이라는 가정을 세울 수 있었다.



[그림 4-7] ReLU, GeLU 그리고 Swish 활성화함수 비교 그래프.

4. 성능지표

인공지능 모델 분류 성능은 CWRU 베어링 데이터 세트를 사용하여 측정되었다. 정확도 다음과 같다.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

True Positive (TP)는 참이 참으로 분류됨을 의미한다. True Negative (TN)은 True Negative의 약자로 참이 거짓으로 분류됨을 말한다. False Positive (FP)는 거짓이 참으로 분류됨을 말하고 False Negative는 거짓이 거짓으로 분류됨을 말한다. 정확도는 총 샘플 수와 알고리즘이 올바르게 예측한 것의 비율을 나타낸다. 예를 들어, 알고리즘이 90% 정확하다면 100개의 샘플 중 90개만 정확하게 분류를 한다.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (6)$$

Recall (6)은 인공지능 모델이 참이라 분류한 것 중 실제로 참인 클래스의 비율이다.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (7)$$

Precision (7)은 실제로 참인 것 중 모델이 참으로 예측한 것의 비율이다.

$$F1 - score = 2 \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

F1-score (8)은 조화평균이라고도 불린다. 데이터 레이블이 불균형한 경우 Accuracy보다 모형의 성능을 더 잘 평가하는 것으로 알려져 있다. 모든 실험은 평가 지표에 언급된 것 중 Accuracy와 Accuracy의 단점을 보완하기 위한 F1-score를 사용하였다.

5. 결과

총 3가지 실험을 수행했다. 첫 번째 실험은 CWRU 베어링 데이터 세트의 고장을 진단하는 실험이다. 첫 번째 실험에서는 SL 트랜스포머가 비교 모델 간의 분류를 수행할 때 가장 우수한 성능을 낼 수 있다고 가정할 수 있었다. 두 번째 실험과 세 번째 실험은 SSA 알고리즘의 성능을 평가한다. 두 번째 실험에서는 SSA 알고리즘 없이 데이터를 분류하고 세 번째 실험에서는 SSA 알고리즘을 통해 데이터를 분류한다.

SL 트랜스포머와 성능을 비교하는 대상으로 SVM, CNN, LSTM, Vanilla 트랜스포머를 선정하였다. 나열된 인공지능 모델은 과거부터 지금까지 다양한 영역에서 널리 사용하고 있는 인공지능 모델이다.

첫 번째 실험 결과는 [표 4-1]에 나타난다. 예상한대로, 실험은 SL 트랜스포머가 95%의 정확도와 94%의 F1-score를 기록하여 최고의 성능을 발휘함을 확인하였다. 즉, 바닐라 트랜스포머의 MLP를 LSTM으로 대체하는 방법을 통해 인공지능 모델의 정확도가 더욱 높아졌음을 확인했다. 나아가 Swish 활성화 함수를 사용한 이유가 인공지능 모델의 안정적인 수렴을 의도했었다. [그림 4-8]과 [그림 4-9]를 살펴보면, Accuracy와 Loss 곡선이 안정적인 곡선 형태를 나타내는 것을 확인할 수 있었다. 이는 인공지능 모델의 훈련이 안정적으로 수렴했음을 알 수 있다.

95%의 정확도는 인공지능 모델에게 매우 높은 성적이다. 하지만 이것은 현실적인 수치가 아니라고 판단하였다. CWRU 데이터 세트는 베어링 데이터들 중 노이즈의 양이 없는 깨끗한 데이터로 평가받는다. 따라서 실제 공장 환경을 반영하기 어려우며, 추가 실험이 필요하다는 결론을 내렸다.

두 번째와 세 번째 실험은 베어링 데이터에 삽입된 소음에 관한 것이다. 이 실험의 목적은 SSA 알고리즘이 소음에 얼마나 강한지를 보고자 하는 것이었다.

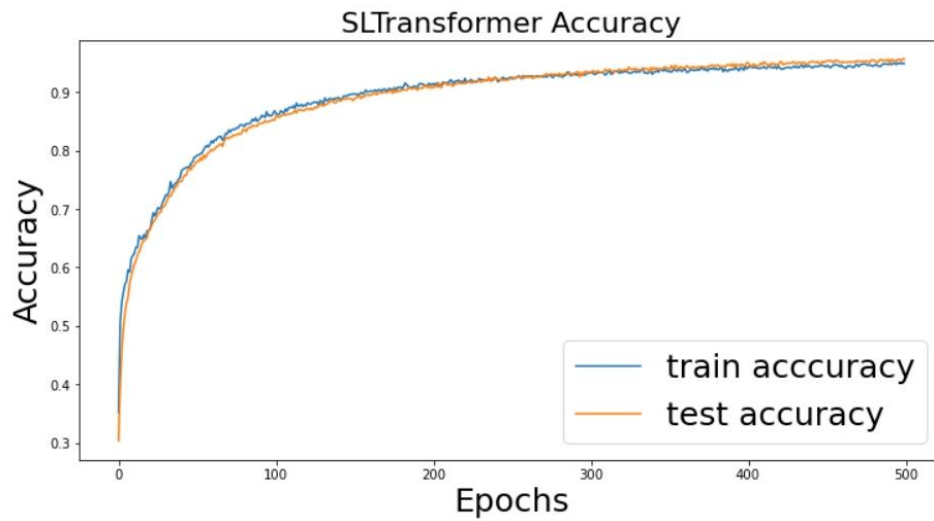
가우시안 백색 노이즈를 CWRU 데이터 세트에 삽입 후 실험을 진행한다. 소음이 수반되는 실험은 상황이 일정하지 않고 다양한 소음이 발생하는 산업 현장에 가깝다. 이 실험에서는 신호 대 잡음비 (SNR) 지수를 사용하여 잡음 변화를 측정하였다. SNR의 정의는 다음과 같다 [44].

여기서 P_{signal} P_{noise} 는 각각 신호와 노이즈의 힘을 의미한다. 백색 가우시안 노이즈의 추가 범위는 -4dB 에서 0dB 까지 조정하였다. 여기서 0dB 는 노이즈가 들어가지 않은 상태를 의미한다. 두 번째 실험의 의도는 인공지능 알고리즘들이 노이즈에 얼마나 취약한지를 보여주는 의도가 있었다. 두 번째 실험은 [표 4-2]과 [표 4-3]을 통해 확인할 수 있다. [표 4-2]에 나타난 실험 결과는 시각화 데이터인 [그림 4-10]에 나타나 있다. 5개의 비교 그룹에서 dB 에 의한 Accuracy와 F1-score를 비교하였다. 예상대로 성능이 매우 떨어지는 결과를 확인할 수 있었다. 특히 SSA가 없는 실험에서 대부분의 인공지능 모델은 -4dB 에서 70%미만의 성능까지 나타남을 확인할 수 있었다.

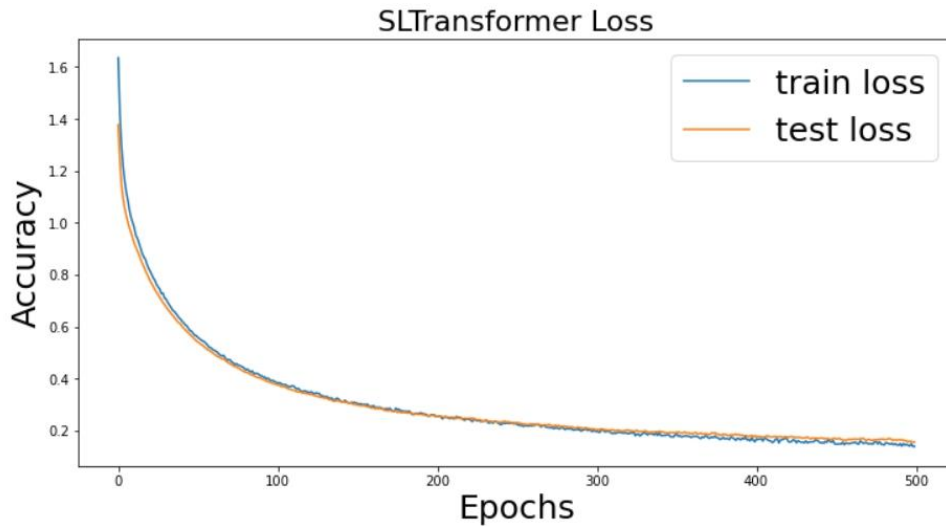
세 번째 실험은 SSA 알고리즘을 적용하는 실험이다. 여기서 SSA가 노이즈에 얼마나 강한지를 확인하는 것이 주 목적이다. 이 실험에서는 흥미로운 사실을 관찰할 수 있었다. 소음이 없는 실험의 최고성능 95%보다 더 높은 수치의 96%를 확인할 수 있었다. 이 결과는 SSA 알고리즘이 데이터를 분해하고 재구성하는 과정에서 시계열 데이터가 인공지능 신경망에 더 적합하게 특성인자를 만들었다는 또 하나의 가설을 세울 수 있었다. 결론적으로 본 실험에서는 SL 트랜스포머가 비교모델에 비해 전반적으로 높은 Accuracy와 F1-score를 기록함을 볼 수 있으며, SSA 알고리즘이 노이즈에 강건하다는 것을 확인할 수 있었다.

[표 4-1] SL 트랜스포머와 비교 모델의 Accuracy와 F1-score 표.

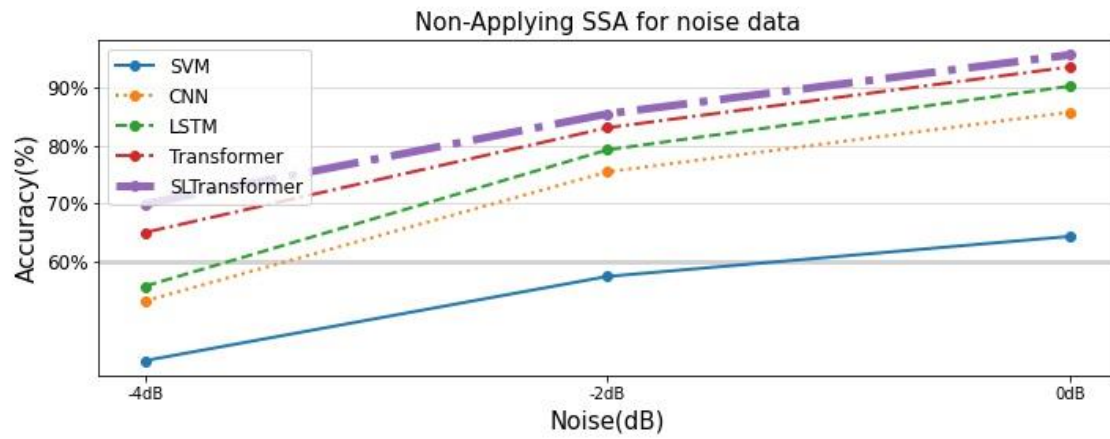
Model	SVM	CNN	LSTM	Transformer	SLTransformer
Accuracy	64.32	85.69	90.15	93.47	95.54
F1-score	57.44	84.44	89.72	92.31	94.47



[그림 4-8] SL 트랜스포머의 Accuracy 그래프.



[그림 4-9] SL 트랜스포머의 Loss 그래프.



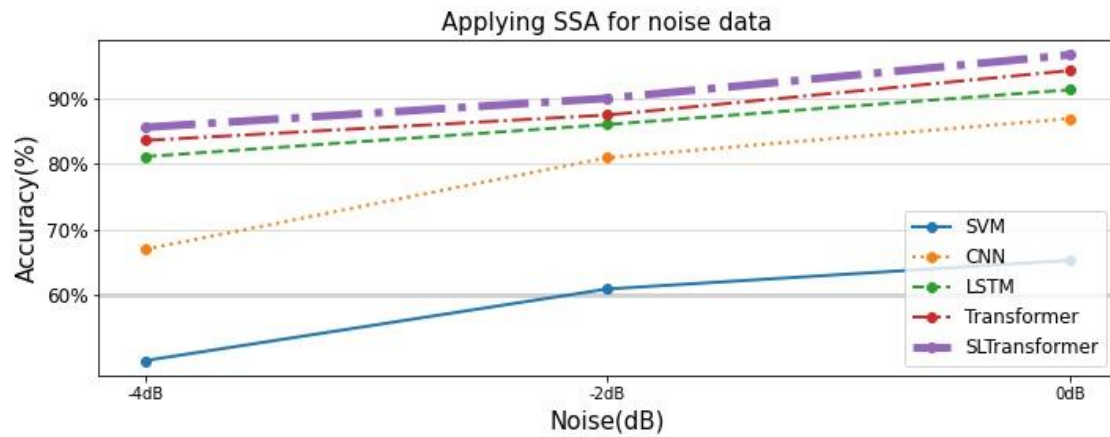
[그림 4-10] SSA 미적용 Accuracy 그래프.

[표 4-2] SSA 알고리즘 미적용 Accuracy 표.

Model	-4dB	-2dB	0dB
SVM	42.95	57.43	64.32
CNN	53.21	75.42	85.69
LSTM	55.76	79.21	90.15
Transformer	65.00	83.01	93.47
SLTransformer	69.89	85.40	95.54

[표 4-3] SSA 알고리즘 미적용 F1-score 표.

Model	-4dB	-2dB	0dB
SVM	54.79	55.44	57.44
CNN	59.44	61.15	84.44
LSTM	61.55	64.31	89.72
Transformer	64.16	68.51	92.31
SLTransformer	68.13	74.10	94.47



[그림 4-11] SSA 적용 Accuracy 그래프.

[표 4-4] SSA 알고리즘 적용한 Accuracy 표.

Model	-4dB	-2dB	0dB
SVM	50.10	60.98	65.33
CNN	67.00	80.95	86.88
LSTM	81.10	85.97	91.25
Transformer	83.55	87.42	94.17
SLTransformer	85.55	89.95	96.55

[표 4-5] SSA 알고리즘 적용한 F1-score 표

Model	-4dB	-2dB	0dB
SVM	49.98	59.15	64.17
CNN	65.43	79.95	85.79
LSTM	77.42	82.44	90.01
Transformer	82.11	85.17	93.15
SLTransformer	84.00	86.37	95.33

제5장 결 론

1. 연구결과

본 논문은 베어링 고장 진단 과제를 해결하기 위한 SSA-SL 트랜스포머 프레임워크를 제안한다. 베어링 고장 진단을 진행하면서, 기존 연구의 한계를 발견하고 다음과 같은 가설을 제시할 수 있었다. 기존의 많은 연구들은 공장의 실제 환경을 가정하지 않았다. 실제 공장에서는 다양한 소음이 발생한다. 일반적으로 깨끗한 데이터만 수집하기 어려운 실정이다. 이에 따라 베어링 데이터에 노이즈가 삽입되는 상황으로 실제 공장 환경을 가정하고 구현하였다. 잡음이 섞인 데이터를 극복해내는 과정이 본 논문의 주요 내용이다.

본 논문에서 제안된 방법은 전처리 기술의 재 조명과 새로운 인공지능 모델의 제안이다. 재조명된 전처리 기술인 SSA 전처리 방법을 사용하면 노이즈를 분해하고 재구성하여 노이즈 제거를 위한 강력한 프레임워크를 구축할 수 있다. 그리고 기존 인공지능 모델인 트랜스포머 모델에 대해 새롭게 접근하였다. 트랜스포머 모델 내부에 회귀속성을 지닌 LSTM 모델을 적용하였고 LSTM 내부 활성화 함수로 시계열 데이터에 적합하다 판단한 Swish 활성화 함수를 사용하였다. 이 인공지능 모델은 이전에 접근하던 방식과는 다른 접근 방식이었다. 우리는 다양한 메트릭을 통해 SL 트랜스포머의 성능을 입증하였다. 그 가설을 증명하기 위해 세 가지 실험이 진행되었다. 첫 실험은 CWRU 데이터 셋에 SL 트랜스포머를 사용하여 메트릭 점수를 도출했다. 이 실험에서는 바닐라 트랜스포머보다 성능이 우수하다는 것을 확인할 수 있었다. 두번째 실험은 CWRU 데이터 셋에 노이즈를 첨가하였다. 노이즈 첨가 정도에 따라 dB별로 나누어 확인했다. 두번째 실험은 인공지능 모델이 가설대로 소음에 취약하다는 것을 확인할 수 있었다. 세번째 실험은 두번째 실험에 SSA 전처리 공정을 추가했다. SSA를 통해 노이즈를 분해하고 재구성한 결과 인공지능 신경망의 성능이 크게 향상되었으며, 특히 SSA-SL 트랜스포머는 96%의 성능을 보였다.

본 논문을 통해 후속 연구자들과 엔지니어들은 다음과 같은 것들을 얻기를 바란다. 인공지능 프레임워크에 데이터 노이즈를 처리할 때 SSA 방법론을 사용하면 좋은 성능을 기대할 수 있다. 그리고 우리는 SL 트랜스포머 방법론을 제시한다. 우리는 이 모델을 더욱 더 발전시키기를 기대한다. 실험 조건과 매개변수는 3장과 4장에서 확인할 수 있다. 엔지니어는 본 논문의 방법을 통해 센서로 데이터를 수집할 때 발생하는 노이즈를 해결할 수 있을 것으로 기대할 수 있다.

2. 향후 연구 및 계획

본 연구에서는 CWRU 공공 베어링 데이터 셋을 활용하여 실험을 진행하였다. 향후 연구 및 계획으로는 공공 베어링 데이터 셋보다, 실제 공장의 환경에서도 본 연구의 프레임워크가 적용되는지를 살펴보는 것이 주요 목표다. 또한 SL 트랜스포머에서 셀프어텐션 내부 리니어 계층을 LSTM 블록으로 대체하는 것이 목표기도 하다.

참 고 문 헌

* 본 장의 내용은, 2022년 Open Access Journal MDPI Electronics의 Volume 11, issue 9, 1504에 게재되었음을 명시합니다.

- [1] Kalsoom, Tahera, et al. "Advances in sensor technologies in the era of smart factory and industry 4.0." *Sensors* 20.23 (2020): 6783.
- [2] Pech, Martin, Jaroslav Vrchota, and Jiří Bednář. "Predictive maintenance and intelligent sensors in smart factory." *Sensors* 21.4 (2021): 1470.
- [3] Sufian, Amr T., et al. "Six-Gear Roadmap towards the Smart Factory." *Applied Sciences* 11.8 (2021): 3568.
- [4] Sinha, Devarpita, and Rajarshi Roy. "Reviewing cyber-physical system as a part of smart factory in industry 4.0." *IEEE Engineering Management Review* 48.2 (2020): 103–117.
- [5] Büchi, Giacomo, Monica Cugno, and Rebecca Castagnoli. "Smart factory performance and Industry 4.0." *Technological Forecasting and Social Change* 150 (2020): 119790.
- [6] Toma, Rafia Nishat, Alexander E. Prosvirin, and Jong-Myon Kim. "Bearing fault diagnosis of induction motors using a genetic algorithm and machine learning classifiers." *Sensors* 20.7 (2020): 1884.
- [7] Li, Hua, et al. "A bearing fault diagnosis method based on enhanced singular value decomposition." *IEEE Transactions on Industrial Informatics* 17.5 (2020): 3220–3230.
- [8] Kuncan, Melih. "An intelligent approach for bearing fault diagnosis: combination of 1D-LBP and GRA." *Ieee Access* 8 (2020): 137517–137529.

- [9] Hoang, Duy–Tang, and Hee–Jun Kang. "A survey on deep learning based bearing fault diagnosis." *Neurocomputing* 335 (2019): 327–335.
- [10] Yuan, Laohu, et al. "Rolling bearing fault diagnosis based on convolutional neural network and support vector machine." *IEEE Access* 8 (2020): 137395–137406.
- [11] Han, Tian, et al. "Rolling bearing fault diagnosis with combined convolutional neural networks and support vector machine." *Measurement* 177 (2021): 109022.
- [12] Zhang, Xiaolin, et al. "Research on bearing fault diagnosis of wind turbine gearbox based on 1DCNN–PSO–SVM." *IEEE Access* 8 (2020): 192248–192258.
- [13] Zhang, Rongtao, Binbin Li, and Bin Jiao. "Application of XGboost algorithm in bearing fault diagnosis." *IOP Conference Series: Materials Science and Engineering*. Vol. 490. No. 7. IOP Publishing, 2019.
- [14] Neupane, Dhiraj, and Jongwon Seok. "Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review." *IEEE Access* 8 (2020): 93155–93178.
- [15] Zhu, Junjun, et al. "Application of recurrent neural network to mechanical fault diagnosis: a review." *Journal of Mechanical Science and Technology* (2022): 1–16.
- [16] Liu, Han, et al. "Fault diagnosis of rolling bearings with recurrent neural network–based autoencoders." *ISA transactions* 77 (2018): 167–178.
- [17] Eren, Levent, Turker Ince, and Serkan Kiranyaz. "A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier." *Journal of Signal Processing Systems* 91.2 (2019): 179–189.
- [18] Abdelkader, Rabah, Abdelhafid Kaddour, and Ziane Derouiche.

"Enhancement of rolling bearing fault diagnosis based on improvement of empirical mode decomposition denoising method." *The International Journal of Advanced Manufacturing Technology* 97.5 (2018): 3099–3117.

- [19] Chen, Bingyan, et al. "Blind deconvolution assisted with periodicity detection techniques and its application to bearing fault feature enhancement." *Measurement* 159 (2020): 107804.
- [20] Qiao, Zijian, et al. "A second-order stochastic resonance method enhanced by fractional-order derivative for mechanical fault detection." *Nonlinear Dynamics* 106.1 (2021): 707–723.
- [21] Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).
- [22] Han, Kai, et al. "Transformer in transformer." *Advances in Neural Information Processing Systems* 34 (2021).
- [23] Ha, Sooji, et al. "Topic classification of electric vehicle consumer experiences with transformer-based deep learning." *Patterns* 2.2 (2021): 100195.
- [24] Li, Guoqiang, et al. "Rolling bearing fault diagnosis based on wavelet packet transform and convolutional neural network." *Applied Sciences* 10.3 (2020): 770.
- [25] Neupane, Dhiraj, and Jongwon Seok. "Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review." *IEEE Access* 8 (2020): 93155–93178.
- [26] Zhang, Wei, et al. "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load." *Mechanical Systems and Signal Processing* 100 (2018): 439–453.

- [27] Oh, Seokju, Seugmin Han, and Jongpil Jeong. "Multi-scale convolutional recurrent neural network for bearing fault detection in noisy manufacturing environments." *Applied Sciences* 11.9 (2021): 3963.
- [28] Hassani, Hossein. "Singular spectrum analysis: methodology and comparison." (2007): 239–257.
- [29] Golyandina, Nina, Anton Korobeynikov, and Anatoly Zhigljavsky. *Singular spectrum analysis with R*. Springer Berlin Heidelberg, 2018.
- [30] Qiao, Zijian, and Zhengrong Pan. "SVD principle analysis and fault diagnosis for bearings based on the correlation coefficient." *Measurement Science and Technology* 26.8 (2015): 085014.
- [31] Ramachandran, Prajit, Barret Zoph, and Quoc V. Le. "Searching for activation functions." *arXiv preprint arXiv:1710.05941* (2017).
- [32] Greff, Klaus, et al. "LSTM: A search space odyssey." *IEEE transactions on neural networks and learning systems* 28.10 (2016): 2222–2232.
- [33] Yan, Xiaoan, et al. "Reliable fault diagnosis of bearings using an optimized stacked variational denoising auto-encoder." *Entropy* 24.1 (2021): 36.
- [34] Chen, Zhuyun, and Weihua Li. "Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network." *IEEE Transactions on Instrumentation and Measurement* 66.7 (2017): 1693–1702.
- [35] Liu, Hongmei, Lianfeng Li, and Jian Ma. "Rolling bearing fault diagnosis based on STFT-deep learning and sound signals." *Shock and Vibration* 2016 (2016).
- [36] Islam, MM Manjurul, and Jong-Myon Kim. "Automated bearing fault diagnosis scheme using 2D representation of wavelet packet transform and deep convolutional neural network." *Computers in Industry* 106 (2019):

142–153.

- [37] Rodrigues, Paulo Canas, et al. "The decomposition and forecasting of mutual investment funds using singular spectrum analysis." *Entropy* 22.1 (2020): 83.
- [38] Hassani, Hossein, et al. "The effect of data transformation on singular spectrum analysis for forecasting." *Signals* 1.1 (2020): 4–25.
- [39] Karim, Fazle, et al. "LSTM fully convolutional networks for time series classification." *IEEE access* 6 (2017): 1662–1669.
- [40] Karim, Fazle, et al. "Multivariate LSTM-FCNs for time series classification." *Neural Networks* 116 (2019): 237–245.
- [41] Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805* (2018).
- [42] Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *International conference on machine learning*. PMLR, 2015.
- [43] Nguyen, Anh, et al. "An Analysis of State-of-the-art Activation Functions For Supervised Deep Neural Network." *2021 International Conference on System Science and Engineering (ICSSE)*. IEEE, 2021.
- [44] Garbin, Christian, Xingquan Zhu, and Oge Marques. "Dropout vs. batch normalization: an empirical study of their impact to deep learning." *Multimedia Tools and Applications* 79.19 (2020): 12777–12815.
- [45] Chang, Chein-I. "Hyperspectral Target Detection: Hypothesis Testing, Signal-to-Noise Ratio, and Spectral Angle Theories." *IEEE Transactions on Geoscience and Remote Sensing* 60 (2021): 1–23.

ABSTRACT

Research on SSA–SL Transformer Framework for Bearing Fault Diagnosis under Noisy Factory Environments

Seoyeong Lee

Department of Smart Factory Convergence

The Graduate School

Sungkyunkwan University

Among the smart factory studies, we describe defect detection research conducted on bearings, which are elements of mechanical facilities. Bearing research has been consistently conducted in the past however, most of the research has been limited to using existing artificial intelligence models. In addition, previous studies assumed the factory situated in the bearing defect research were insufficient. Therefore, recent research was conducted that applied an artificial intelligence model and the factory environment. The Transformer model was selected as State-of-the-art (SOTA) and was also applied to bearing research. Then, an experiment was conducted with Gaussian noise applied to assume a factory situation. The Swish–LSTM Transformer (SL Transformer) framework was constructed by redesigning the internal structure of the Transformer using the Swish activation function and Long Short–Term Memory (LSTM). Then, the data in noise was removed and reconstructed using

the Singular Spectrum Analysis (SSA) preprocessing method. Based on the SSA–SL Transformer framework, an experiment was performed by adding gaussian noise to the Case Western Reserve University (CWRU) dataset. In the case of no noise, the SL Transformer showed more than 95\% performance, and when noise was inserted, the SSA–SL Transformer showed the better performance than the comparative artificial intelligence models.

Keywords: Bearing Fault Diagnosis, Singular Spectrum Analysis, Transformer, Under Noisy Factory Environments,

碩
士
學
位
請
求
論
文

소
음
환
경
의

공
장
상
황
에
서

베
어
링
결
합
검
사
를
위
한
SS
A
SL
트
랜
스
포
머

프
레
임
위
크
에
관
한
연
구

2
0
2
2

李
書
營