



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2021년01월25일  
(11) 등록번호 10-2206737  
(24) 등록일자 2021년01월19일

(51) 국제특허분류(Int. Cl.)  
F03B 11/00 (2006.01) G06N 20/00 (2019.01)  
(52) CPC특허분류  
F03B 11/008 (2013.01)  
G06N 20/00 (2019.01)  
(21) 출원번호 10-2020-0004456  
(22) 출원일자 2020년01월13일  
심사청구일자 2020년01월13일  
(56) 선행기술조사문헌  
JP06066241 A\*  
KR102051226 B1  
KR101991296 B1  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
한국과학기술원  
대전광역시 유성구 대학로 291(구성동)  
한국수력원자력 주식회사  
경상북도 경주시 양북면 불국로 1655  
(72) 발명자  
이필승  
대전광역시 유성구 대학로 291 (구성동)  
신소진  
대전광역시 유성구 대학로 291 (구성동)  
현철규  
대전광역시 유성구 대학로 291 (구성동)  
(74) 대리인  
장수현

전체 청구항 수 : 총 16 항

심사관 : 최진환

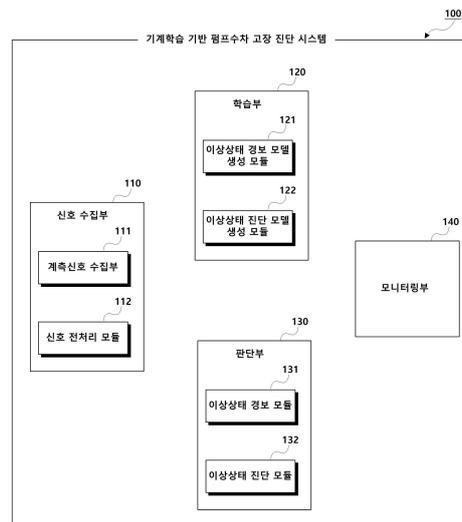
(54) 발명의 명칭 기계학습 기반 펌프수차 고장 진단 방법 및 시스템

(57) 요약

본 발명은 기계학습 기반 펌프수차 고장 진단 방법 및 시스템에 관한 것으로서, 더욱 상세하게는 사전 및 실시간 학습한 기계학습 모델을 이용하여 양수발전소 펌프수차의 고장 유무 및 종류를 진단하기 위한 방법 및, 그 방법을 수행하는 시스템에 관한 것이다.

본 발명에 의하면, 사전 및 실시간 학습한 기계학습 모델을 이용해 양수발전소 펌프수차의 고장 유무 및 종류를 진단하도록 하여 과도한 예방정비를 최소화하며, 이로써 유지보수 비용 및 돌발고장으로 인한 손실비용을 절감하고, 전문가에 의한 운전 데이터 분석을 최소화해 작업량을 감소시킬 수 있도록 한다.

대표도 - 도1



(52) CPC특허분류  
*F05B 2260/80* (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	No. 2017-Tech-11
부처명	기타(정부수탁)
과제관리(전문)기관명	한국수력원자력(주)
연구사업명	연구사업
연구과제명 개발(2019)	IoT 기반 상태감시 및 지능형 자가진단을 통한 펌프수차 이상 징후 예측기술
기 여 율	1/1
과제수행기관명	한국과학기술원
연구기간	2019.02.01 ~ 2020.01.31

---

**명세서**

**청구범위**

**청구항 1**

기계학습 기반 펌프수차 고장 진단 시스템으로서,

복수의 센서에서 계측된 신호 데이터(이하 '계측 데이터'라 한다)를 수집하고, 이에 대한 전처리를 수행하는 신호 수집부;

정상상태 또는 이상상태 여부를 판단하는 이상상태 경보 모델을 생성하는 이상상태 경보 모델 생성 모듈을 구비하는 학습부; 및,

상기 학습부에서 생성된 이상상태 경보 모델을 이용하여 정상상태 또는 이상상태 여부를 판단하는 이상상태 경보 모듈을 구비하는 판단부

를 포함하고,

상기 이상상태 경보 모델 생성 모듈은,

다변량 계측 데이터에 대한 평균, 공분산 및, 상기 다변량 계측 데이터의 확률밀도함수 값에서 이상상태와 정상상태를 구분하는 경계값인 신뢰한계값을 산출함으로써 이상상태 경보 모델을 생성하는,

기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 2**

청구항 1에 있어서,

상기 학습부는,

정상/이상 여부 및, 이상일 경우 이상상태 종류를 분류 산출하는 이상상태 진단 모델을 생성하는 이상상태 진단 모델 생성모듈

을 더 구비하고,

상기 판단부는,

상기 학습부에서 생성한 이상상태 진단 모델을 이용하여 정상/이상 여부 및, 이상일 경우 이상상태 종류를 분류 산출하는 이상상태 진단 모듈

을 더 구비하는 것을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 3**

청구항 1 또는 청구항 2에 있어서,

기계학습 기반 펌프수차 고장 진단 시스템의 운행 중 이상 여부를 모니터링하고, 상기 판단부에서 산출된 결과 및, 상기 모니터링에 의해 확인된 실제 이상 여부 결과를 상기 학습부로 전달하는 모니터링부

를 더 포함하는 것을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 4**

청구항 1에 있어서,

상기 신호 수집부의 계측 데이터 전처리에는,

노이즈 제거, 데이터 분류, 데이터 정규화 기능을 포함하는 것

을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 5**

청구항 4에 있어서,  
 상기 데이터 분류 기능은,  
 양수(펌프) 및 발전(터빈) 기동 데이터로의 분류를 포함하는 것  
 을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 6**

청구항 5에 있어서,  
 상기 데이터 분류 기능은,  
 상기 양수(펌프) 기동 데이터 및 발전(터빈) 기동 데이터 각각에 대하여, 휴동 구간, 변동 구간 및 가동 구간으  
 로의 분류를 더 포함하는 것  
 을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 7**

청구항 4에 있어서,  
 상기 데이터 정규화 기능은,  
 분류된 데이터 별로 정규화를 수행하는 것  
 을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 8**

삭제

**청구항 9**

청구항 1에 있어서,  
 상기 다변량 계측 데이터의 확률밀도함수는,

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

로 결정되고,  
 $\mathbf{x}$ 는 계측 데이터,  $\boldsymbol{\mu}$ 는 계측 데이터의 평균,  $\Sigma$ 는 계측 데이터의 공분산,  $d$ 는 데이터의 차원을 의미하는 것  
 을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 10**

청구항 1에 있어서,  
 다변량 계측 데이터에 대한 평균, 공분산은,  
 정상상태의 계측 데이터를 이용하여 산출되는 것  
 을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 11**

청구항 1에 있어서,

상기 신뢰한계값은,

계측 데이터 중, 검증 데이터를 이용해 정밀도(precision), 재현율(recall), F1(재현율과 정밀도의 조화평균) 등을 이용해 성능이 가장 좋은 지점의 확률밀도함수값으로 결정되는 것을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 12**

청구항 2에 있어서,

상기 이상상태 진단 모듈에서 사용하는 이상상태 진단 모델은 인공신경망이 이용되고, 정상/이상 여부 및, 이상 일 경우 이상 유형 분류를 산출하며,

상기 이상상태 진단 모델은 상기 이상상태 진단 모델 생성 모듈에서 기계학습에 의해 생성되는 것을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 13**

청구항 12에 있어서,

상기 이상상태 진단 모델은,

정상/이상 여부 및 산출된 이상 유형 분류 각각에 대한 확률값을 산출하며,

확률값이 가장 높은 정상/이상 여부 또는 이상 유형 분류를 판별해 모니터링부로 전달하는 것을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 14**

청구항 13에 있어서,

상기 이상상태 진단 모델이 상기 모니터링부로 전달하는 데이터에는,

분류된 각 정상/이상 여부의 확률값 또는 이상 유형 분류의 확률값을 더 포함하는 것

을 특징으로 하는 기계학습 기반 펌프수차 고장 진단 시스템.

**청구항 15**

청구항 1의 기계학습 기반 펌프수차 고장 진단 시스템이 수행하는 펌프수차 고장 진단 방법으로서,

(a) 복수의 센서에서 계측된 신호 데이터(이하 '계측 데이터'라 한다)를 수집하고, 이에 대한 전처리를 수행하는 단계;

(b) 정상상태 또는 이상상태 여부를 판단하는 이상상태 경보 모델을 생성하는 단계; 및,

(c) 상기 이상상태 경보 모델을 이용하여 정상상태 또는 이상상태 여부를 판단하는 단계

를 포함하고,

상기 단계(b)에서는,

다변량 계측 데이터에 대한 평균, 공분산 및, 상기 다변량 계측 데이터의 확률밀도함수 값에서 이상상태와 정상상태를 구분하는 경계값인 신뢰한계값을 산출함으로써 이상상태 경보 모델을 생성하는,

기계학습 기반 펌프수차 고장 진단 방법.

**청구항 16**

기계학습 기반 펌프수차 고장 진단 시스템이 펌프수차 고장 진단을 수행하기 위한, 비일시적 저장매체에 저장된 컴퓨터 프로그램으로서,

비일시적 저장매체에 저장되며, 프로세서에 의하여,

- (a) 복수의 센서에서 계측된 신호 데이터(이하 '계측 데이터'라 한다)를 수집하고, 이에 대한 전처리를 수행하는 단계;
  - (b) 정상상태 또는 이상상태 여부를 판단하는 이상상태 경보 모델을 생성하는 단계; 및,
  - (c) 상기 이상상태 경보 모델을 이용하여 정상상태 또는 이상상태 여부를 판단하는 단계
- 가 실행되도록 하는 명령을 포함하고,
- 상기 단계(b)에서는,
- 다변량 계측 데이터에 대한 평균, 공분산 및, 상기 다변량 계측 데이터의 확률밀도함수 값에서 이상상태와 정상상태를 구분하는 경계값인 신뢰한계값을 산출함으로써 이상상태 경보 모델을 생성하는,
- 기계학습 기반 펌프수차 고장 진단 시스템이 펌프수차 고장 진단을 수행하기 위한, 비일시적 저장매체에 저장된 컴퓨터 프로그램.

**청구항 17**

- 기계학습에 기반하여 펌프수차 고장 진단을 수행하는 시스템으로서,
- 적어도 하나의 프로세서; 및
- 컴퓨터로 실행가능한 명령을 저장하는 적어도 하나의 메모리를 포함하되,
- 상기 적어도 하나의 메모리에 저장된 상기 컴퓨터로 실행가능한 명령은, 상기 적어도 하나의 프로세서에 의하여,
- (a) 복수의 센서에서 계측된 신호 데이터(이하 '계측 데이터'라 한다)를 수집하고, 이에 대한 전처리를 수행하는 단계;
  - (b) 정상상태 또는 이상상태 여부를 판단하는 이상상태 경보 모델을 생성하는 단계; 및,
  - (c) 상기 이상상태 경보 모델을 이용하여 정상상태 또는 이상상태 여부를 판단하는 단계
- 가 실행되도록 하고,
- 상기 단계(b)에서는,
- 다변량 계측 데이터에 대한 평균, 공분산 및, 상기 다변량 계측 데이터의 확률밀도함수 값에서 이상상태와 정상상태를 구분하는 경계값인 신뢰한계값을 산출함으로써 이상상태 경보 모델을 생성하는,
- 기계학습 기반 펌프수차 고장 진단 수행 시스템.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 기계학습 기반 펌프수차 고장 진단 방법 및 시스템에 관한 것으로서, 더욱 상세하게는 사전 및 실시간 학습한 기계학습 모델을 이용하여 양수발전소 펌프수차의 고장 유무 및 종류를 진단하기 위한 방법 및, 그 방법을 수행하는 시스템에 관한 것이다.

**배경 기술**

[0002] 청평 양수발전소(1980), 삼랑진 양수발전소(1985) 등 양수발전소 노후화에 따라 발전설비의 실시간 상태진단 필요성이 높아지고 있다. 전력상황에 따른 부하 변동 및 빈번한 기동 정지가 발생하는 양수발전에는 임계값 기준으로 기동 정지(경보)를 판단하는 기법들은 적용하기 어려운 상황이다. 즉, 실제 부하에 맞는 최적의 임계값 설정에 어려움이 존재하는 것이다. 예를 들어, 사고 예방을 위해 임계값을 보수적으로 설정하면 불필요한 기동정지가 발생할 수 있다. 또한, 국내 양수발전소 이상상태 경보시스템의 경우 각각의 센서로부터 측정된 값을 단일 형태로 분석하여 사용하기 때문에 이상상태 분석이 어렵다. 나아가, 정상가동을 벗어난 설비의 이상만을 판단하는 경보시스템이므로 상태진단이 어렵다. 따라서 양수발전소의 고장으로 인한 비계획적 비용 손실을 줄일 수 있는 이상상태 감지 및 종류 분석이 가능한 시스템이 필요한 실정이다.

선행기술문헌

특허문헌

[0003] (특허문헌 0001) KR 10-2027389 B1

발명의 내용

해결하려는 과제

[0004] 본 발명은 이와 같은 문제점을 해결하기 위해 창안된 것으로서, 사전 및 실시간 학습한 기계학습 모델을 이용해 양수발전소 펌프수차의 고장 유무 및 종류를 진단하도록 하여 과도한 예방정비를 최소화하며, 이로써 유지보수 비용 및 돌발고장으로 인한 손실비용을 절감하고, 전문가에 의한 운전 데이터 분석을 최소화해 작업량을 감소시킬 수 있도록 하는데 그 목적이 있다.

과제의 해결 수단

[0005] 이와 같은 목적을 달성하기 위하여 본 발명에 따른 기계학습 기반 펌프수차 고장 진단 시스템은, 복수의 센서에서 측정된 신호 데이터(이하 '측정 데이터'라 한다)를 수집하고, 이에 대한 전처리를 수행하는 신호 수집부; 정상상태 또는 이상상태 여부를 판단하는 이상상태 경보 모델을 생성하는 이상상태 경보 모델 생성 모듈을 구비하는 학습부; 및, 상기 학습부에서 생성된 이상상태 경보 모델을 이용하여 정상상태 또는 이상상태 여부를 판단하는 이상상태 경보 모듈을 구비하는 판단부를 포함하고, 상기 이상상태 경보 모델 생성 모듈은, 다변량 측정 데이터에 대한 평균, 공분산 및, 상기 다변량 측정 데이터의 확률밀도함수 값에서 이상상태와 정상상태를 구분하는 경계값인 신뢰한계값을 산출함으로써 이상상태 경보 모델을 생성한다.

[0006] 상기 학습부는, 정상/이상 여부 및, 이상일 경우 이상상태 종류를 분류 산출하는 이상상태 진단 모듈을 생성하는 이상상태 진단 모듈 생성모듈을 더 구비하고, 상기 판단부는, 상기 학습부에서 생성한 이상상태 진단 모듈을 이용하여 정상/이상 여부 및, 이상일 경우 이상상태 종류를 분류 산출하는 이상상태 진단 모듈을 더 구비할 수 있다.

[0007] 상기 기계학습 기반 펌프수차 고장 진단 시스템은, 기계학습 기반 펌프수차 고장 진단 시스템의 운행 중 이상 여부를 모니터링하고, 상기 판단부에서 산출된 결과 및, 상기 모니터링에 의해 확인된 실제 이상 여부 결과를 상기 학습부로 전달하는 모니터링부를 더 포함할 수 있다.

[0008] 상기 신호 수집부의 측정 데이터 전처리에는, 노이즈 제거, 데이터 분류, 데이터 정규화 기능을 포함할 수 있다.

[0009] 상기 데이터 분류 기능은, 양수(펌프) 및 발전(터빈) 기동 데이터로의 분류를 포함할 수 있다.

[0010] 상기 데이터 분류 기능은, 상기 양수(펌프) 기동 데이터 및 발전(터빈) 기동 데이터 각각에 대하여, 휴동 구간, 변동 구간 및 가동 구간으로의 분류를 더 포함할 수 있다.

[0011] 상기 데이터 정규화 기능은, 분류된 데이터 별로 정규화를 수행할 수 있다.

[0013] 상기 다변량 측정 데이터의 확률밀도함수는,

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)$$

로 결정될 수 있고,  $\mathbf{x}$ 는 측정 데이터,  $\boldsymbol{\mu}$ 는 측정 데이터의 평균,  $\Sigma$ 는 측정 데이터의 공분산,  $d$ 는 데이터의 차원을 의미한다.

[0014] 다변량 측정 데이터에 대한 평균, 공분산은, 정상상태의 측정 데이터를 이용하여 산출될 수 있다.

[0015] 상기 신뢰한계값은, 측정 데이터 중, 검증 데이터를 이용해 정밀도(precision), 재현율(recall), F1(재현율과 정밀도의 조화평균) 등을 이용해 성능이 가장 좋은 지점의 확률밀도함수값으로 결정될 수 있다.

[0016] 상기 이상상태 진단 모듈에서 사용하는 이상상태 진단 모듈은 인공지능망이 이용되고, 정상/이상 여부 및, 이상

일 경우 이상 유형 분류를 산출할 수 있으며, 상기 이상상태 진단 모델은 상기 이상상태 진단 모델 생성 모듈에서 기계학습에 의해 생성될 수 있다.

- [0017] 상기 이상상태 진단 모델은, 정상/이상 여부 및 산출된 이상 유형 분류 각각에 대한 확률값을 산출할 수 있으며, 확률값이 가장 높은 정상/이상 여부 또는 이상 유형 분류를 판별해 모니터링부로 전달할 수 있다.
- [0018] 상기 이상상태 진단 모델이 상기 모니터링부로 전달하는 데이터에는, 분류된 각 정상/이상 여부의 확률값 또는 이상 유형 분류의 확률값을 더 포함할 수 있다.
- [0019] 본 발명의 다른 측면에 따르면, 기계학습 기반 펌프수차 고장 진단 시스템이 수행하는 펌프수차 고장 진단 방법은, (a) 복수의 센서에서 계측된 신호 데이터(이하 '계측 데이터'라 한다)를 수집하고, 이에 대한 전처리를 수행하는 단계; (b) 정상상태 또는 이상상태 여부를 판단하는 이상상태 경보 모델을 생성하는 단계; 및, (c) 상기 이상상태 경보 모델을 이용하여 정상상태 또는 이상상태 여부를 판단하는 단계를 포함하고, 상기 단계(b)에서는, 다변량 계측 데이터에 대한 평균, 공분산 및, 상기 다변량 계측 데이터의 확률밀도함수 값에서 이상상태와 정상상태를 구분하는 경계값인 신뢰한계값을 산출함으로써 이상상태 경보 모델을 생성한다.
- [0020] 본 발명의 또 다른 측면에 따르면, 기계학습 기반 펌프수차 고장 진단 시스템이 펌프수차 고장 진단을 수행하기 위한, 비일시적 저장매체에 저장된 컴퓨터 프로그램은, 비일시적 저장매체에 저장되며, 프로세서에 의하여, (a) 복수의 센서에서 계측된 신호 데이터(이하 '계측 데이터'라 한다)를 수집하고, 이에 대한 전처리를 수행하는 단계; (b) 정상상태 또는 이상상태 여부를 판단하는 이상상태 경보 모델을 생성하는 단계; 및, (c) 상기 이상상태 경보 모델을 이용하여 정상상태 또는 이상상태 여부를 판단하는 단계가 실행되도록 하는 명령을 포함하고, 상기 단계(b)에서는, 다변량 계측 데이터에 대한 평균, 공분산 및, 상기 다변량 계측 데이터의 확률밀도함수 값에서 이상상태와 정상상태를 구분하는 경계값인 신뢰한계값을 산출함으로써 이상상태 경보 모델을 생성한다.
- [0021] 본 발명의 또 다른 측면에 따르면, 기계학습에 기반하여 펌프수차 고장 진단을 수행하는 시스템은, 적어도 하나의 프로세서; 및 컴퓨터로 실행가능한 명령을 저장하는 적어도 하나의 메모리를 포함하되, 상기 적어도 하나의 메모리에 저장된 상기 컴퓨터로 실행가능한 명령은, 상기 적어도 하나의 프로세서에 의하여, (a) 복수의 센서에서 계측된 신호 데이터(이하 '계측 데이터'라 한다)를 수집하고, 이에 대한 전처리를 수행하는 단계; (b) 정상상태 또는 이상상태 여부를 판단하는 이상상태 경보 모델을 생성하는 단계; 및, (c) 상기 이상상태 경보 모델을 이용하여 정상상태 또는 이상상태 여부를 판단하는 단계가 실행되도록 하고, 상기 단계(b)에서는, 다변량 계측 데이터에 대한 평균, 공분산 및, 상기 다변량 계측 데이터의 확률밀도함수 값에서 이상상태와 정상상태를 구분하는 경계값인 신뢰한계값을 산출함으로써 이상상태 경보 모델을 생성한다.

**발명의 효과**

- [0022] 본 발명에 의하면, 사전 및 실시간 학습한 기계학습 모델을 이용해 양수발전소 펌프수차의 고장 유무 및 종류를 진단하도록 하여 과도한 예방정비를 최소화하며, 이로써 유지보수 비용 및 돌발고장으로 인한 손실비용을 절감하고, 전문가에 의한 운전 데이터 분석을 최소화해 작업량을 감소시킬 수 있도록 하는 효과가 있다.

**도면의 간단한 설명**

- [0023] 도 1은 본 발명의 기계학습 기반 펌프수차 고장 진단 시스템의 구성을 도시한 도면.
- 도 2는 본 발명의 본 발명의 기계학습 기반 펌프수차 고장 진단 방법을 수행하는 순서도.
- 도 3은 펌프 및 터빈의 기동에 대한 모니터링 데이터를 나타낸 그래프.
- 도 4는 각 모니터링 데이터에 대한 시간별 변화값을 정규화하여 나타낸 그래프.
- 도 5는 2개의 모니터링 데이터를 각각 정규화하여 각각 x축 및 y축으로 나타낸 그래프.
- 도 6은 2개의 모니터링 데이터에 대하여 정규분포로 나타내어지는 확률밀도함수를 나타낸 도면.
- 도 7은 이상상태 경보 모델이 이상상태를 판단하는 단계를 나타내는 순서도.
- 도 8은 이상상태 진단 모델이 이상상태의 종류를 판단하는 단계를 나타내는 순서도.
- 도 9는 이상상태 진단 모델의 내부 레이어에 대한 일 실시예로서의 도면.

**발명을 실시하기 위한 구체적인 내용**

- [0024] 이하 첨부된 도면을 참조로 본 발명의 바람직한 실시예를 상세히 설명하기로 한다. 이에 앞서, 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이거나 사전적인 의미로 한정해서 해석되어서는 아니되며, 발명자는 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념을 적절하게 정의할 수 있다는 원칙에 입각하여 본 발명의 기술적 사상에 부합하는 의미와 개념으로 해석되어야만 한다. 따라서, 본 명세서에 기재된 실시예와 도면에 도시된 구성은 본 발명의 가장 바람직한 일 실시예에 불과할 뿐이고 본 발명의 기술적 사상을 모두 대변하는 것은 아니므로, 본 출원시점에 있어서 이들을 대체할 수 있는 다양한 균등물과 변형예들이 있을 수 있음을 이해하여야 한다.
- [0026] 도 1은 본 발명의 기계학습 기반 펌프수차 고장 진단 시스템(100)의 구성을 도시한 도면이고, 도 2는 본 발명의 본 발명의 기계학습 기반 펌프수차 고장 진단 방법을 수행하는 순서도이다.
- [0027] 이하, 도 1 및 도 2를 참조하여서는 기계학습 기반 펌프수차 고장 진단 시스템(100)의 구성 및 수행하는 방법에 대하여 전체적 및 개략적으로 설명하도록 하며, 각 구성 및 방법의 세부 단계에 대하여 필요한 부분에 대하여는 도 3 내지 도 9를 참조하여 더욱 상세히 설명하기로 한다.
- [0028] 기계학습 기반 펌프수차 고장 진단 시스템(100)은, 신호 수집부(110), 학습부(120) 및 판단부(130)를 포함하고, 모니터링부(140)를 더 포함할 수 있다.
- [0029] 신호 수집부(110)는 하나 또는 복수의 센서를 구비하는 계측신호 수집부(111)와 신호 전처리부(112)를 포함한다. 계측신호 수집부(111)에서 수집(S201)하는 계측신호 데이터, 즉, 계측 데이터의 종류는 온도, 변위, 속도, 전류 등 다양한 신호를 포함한다. 수집된 계측 데이터들은 신호 전처리 모듈(112)로 전달되어 데이터 전처리가 수행된다(S202). 데이터 전처리에는 노이즈 제거, 데이터 분류, 데이터 정규화가 포함된다.
- [0030] 학습부(120)는 신호 수집부(110)에서 전처리된 데이터를 이용해 학습을 수행하며, 이상상태 경보 모델 생성 모듈(121) 및 이상상태 진단 모델 생성 모듈(122)을 포함한다. 이상상태 경보 모델은 이상상태의 유무를 판별하는 모델이고 이상상태 진단 모델은 발생한 이상상태의 종류를 판단하는 모델이다. 학습부(120)에서 기계학습에 의하여 생성된 이상상태 경보 모델(S203) 및 이상상태 진단 모델(S204)은 판단부(130)로 전달되어 이상상태 판단 및 진단에 사용된다.
- [0031] 판단부(130)는 이상상태 경보 모듈(131) 및 이상상태 진단 모듈(132)을 포함하며, 각각 학습부(120)에서 생성되고 전달받은 이상상태 경보 모델 및 이상상태 진단 모델을 이용해 신호 수집부(110)에서 전달받은 계측 데이터들의 이상상태 유무 및 종류를 판별한다. 이상상태 경보 모듈은 이상상태 경보 모델에 의해 이상상태 유무를 판별하며(S205), 이상상태 진단 모듈은 이상상태 진단 모델에 의해 정상/이상 여부 및, 이상일 경우 이상상태 종류를 분류한다(S206).
- [0032] 기계학습 기반 펌프수차 고장 진단 시스템(100)이 모니터링부(140)를 포함하는 경우, 이상상태 경보 모듈(131)에서 판단된 정상/이상 여부 정보 및, 이상상태 진단 모듈(132)에서 판단된 이상상태 종류에 대한 정보는 모니터링부(140)로 전달되며(S207), 모니터링부(140)는 실제 발생한 정상/이상 여부 정보 및 이상상태 종류를 모니터링에 의해 파악할 수 있다(S208). 모니터링부(140)는, 이상상태 경보 모듈(131) 및 이상상태 진단 모듈(132)로부터 전달받은 정상/이상 여부 정보 및 이상상태 종류에 대한 정보와, 실제 확인된 정상/이상 여부 정보 및 이상상태 종류에 대한 정보를 다시 학습부(120)로 전달할 수 있으며(S209), 이로써 학습부(120)는 이상상태 경보 모델 또는 이상상태 진단 모델의 학습을 더욱 수행하여 각 모델을 업데이트함으로써 실시간 학습의 신뢰도를 높일 수 있다.
- [0034] 도 3은 펌프 및 터빈의 기동에 대한 모니터링 데이터를 나타낸 그래프이다.
- [0035] 신호 수집부(110)의 신호 전처리 모듈(112)은, 도 2를 참조하여 전술한 바와 같이 노이즈 제거, 데이터 분류, 데이터 정규화를 수행한다.
- [0036] 신호 전처리 모듈(112)은 수집된 계측 데이터에 대하여 노이즈 제거 후, 데이터 구간 분리를 수행한다. 양수 발전소 운행시 양수(펌프) 및 발전(터빈) 기동 데이터의 경향이 상이하므로, 효율적인 학습 및 진단을 위해, 도 3(a) 및 도 3(b)와 같이 모니터링 데이터를 우선 양수(펌프) 및 발전(터빈) 기동 데이터로 분류하여 학습 및 진단에 사용한다.
- [0037] 또한 펌프 데이터(도 3(a)) 및 터빈 데이터(도 3(b)) 각각에서, 회전속도를 기준으로 휴동 구간(회전속도 0 RPM), 변동 구간(회전속도 증가 또는 감소하는 구간), 가동 구간(회전속도가 일정한 구간) 등의 세 구간으로 분류해서 학습 및 진단에 사용한다.

[0039] 도 4는 각 모니터링 데이터에 대한 시간별 변화값을 정규화하여 나타낸 그래프이고, 도 5는 2개의 모니터링 데이터를 각각 정규화하여 각각 x축 및 y축으로 나타낸 그래프이다.

[0040] 신호 전처리부(112)는 도 3을 참조하여 전술한 바와 같은 데이터 분류, 즉, 펌프, 터빈 별 기동 데이터의 분류 및 그와 같은 각 기동 데이터에서 휴동 구간, 변동 구간, 가동 구간 등의 구간 분류를 수행한 후, 각 분류된 데이터에 대한 데이터 정규화를 수행한다. 데이터 정규화를 수행하는 이유는, 계측 데이터와 관련한 여러 요인이 학습에 영향을 미치기 때문이다. 즉, 온도, 진동, 전력 등 모니터링 데이터의 단위가 서로 다른 요인과, 외부 환경 변화와 같은 요인 등이 학습에 영향을 미치게 되는 이유이다. 이러한 문제를 해결하기 위해 일정한 간격으로 센서별 모니터링된 계측 데이터의 정규화를 실시하는 것이다.

[0041] 도 4(a)는 2016년 7월부터 2016년 12월까지의 측정된 온도 데이터를 시간축(sec)에 나타낸 그래프이고, 도 4(b)는 도 4(a)의 데이터를 2개월 간격으로 정규화한 온도 데이터이다. 도 4에 표시한 수식에서 x는 계측된 온도 데이터이며,  $x_{min}$ 은 각 2개월에 대한 계측 온도 데이터 중 최소값,  $x_{max}$ 은 해당 2개월에 대한 계측 온도 데이터 중 최대값을 나타내고, z는 해당 x값이 정규화된 값을 의미한다.

[0042] 도 5(a)는 측정된 데이터 세트(set)를 하나의 그래프로 나타낸 것이다. 즉, 도 5(a)의 각 점들은 한 시점에 측정된 온도(temperature)와 진동(vibration) 데이터로서, (온도, 진동) 좌표를 나타낸다.

[0043] 도 5(b)의 각 점들은, 도 5(a)의 온도 및 진동 데이터를 각각 정규화한 데이터로서 (온도, 진동) 좌표를 나타낸다.

[0045] 도 6은 2개의 모니터링 데이터에 대하여 정규분포로 나타내어지는 확률밀도함수를 나타낸 도면이고, 도 7은 이상상태 정보 모델이 이상상태를 판단하는 단계를 나타내는 순서도이다.

[0046] 일반적으로 고장 데이터 갯수는 정상 데이터 갯수 보다 매우 적기 때문에, 많은 수의 정상 데이터를 효과적으로 이용할 수 있는 다수의 변수에 대한 정규분포 형태의 확률밀도함수, 즉, 다변량 정규분포(다변량 가우시안 분포로도 불림)를 이용한다. 도 6(a)는 그와 같은 다변량 정규분포의 일 예를 도시한 것이다. 편의상 확률밀도함수 그래프를 전면과 측면상에 붉은색과 파란색으로 일부만 도시하였으나, 실제로는 수평면상의 각 (variable 1, variable 2) 좌표점에 대하여 확률밀도함수 값이 존재하여 모두 도시하면, 3차원 그래프가 된다.

[0047] 또한 도 6에서는 도면으로 나타낼 수 있도록, 편의상 2개의 변수(variable 1, variable 2)에 대한 확률밀도함수로서의 정규분포를 도시하였으나, 일반적으로 N개의 변수(variable 1, variable 2, ..., variable N)에 대한 확률밀도함수로서의 정규분포를 사용할 수 있다. 여기서 '변수'란 온도, 진동, 회전속도 등, 다양한 종류의 센서에 의한 측정값들을 의미한다.

[0048] 도 6(b)는 도 6(a)의 각 다변량 변수 좌표점 (variable 1, variable 2) 들을 나타낸 것이며, 초록색 범위 내부에 있는 다변량 데이터 좌표점들은, 확률밀도함수 값이 일정 값 이상인 점들로서, 정상상태에 있다고 판단하며, 초록색 범위 외부에 있는 다변량 데이터 좌표점들은, 확률밀도함수 값이 해당 값 미만인 점들로서, 이상상태에 있다고 판단한다. 이와 같은 확률밀도함수의 '일정 값'을 '신뢰한계값'이라 한다.

[0050] 다변량 정규분포 확률밀도함수는 [수학식 1]과 같다.

[0052] [수학식 1]

$$P(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

[0053] 여기서  $\mathbf{x}$ 는 계측 데이터,  $\boldsymbol{\mu}$ 는 계측 데이터의 평균,  $\Sigma$ 는 계측 데이터의 공분산, d는 데이터의 차원을 의미한다.  $\mathbf{x}$ 는 다변량 데이터로서 N개의 센서에 의한 계측 데이터일 경우, (variable 1, variable 2, ..., variable N)과 같이 표현될 수 있으며, 이 경우  $\boldsymbol{\mu}$ 도 마찬가지로 ( $\mu_1, \mu_2, \dots, \mu_N$ )과 같이 표현될 수 있다. 이때  $d=N$ 이된다. 공분산(covariance)  $\Sigma$ 는 다변량 확률변수의 상관정도를 나타내는 값이다.

[0056] 학습부(120)의 이상상태 정보 모델 생성 모듈(121)은, 이상상태 정보 모듈(131)에서 사용되는 이상상태 정보 모델을 생성하는데, 이와 같은 이상상태 정보 모델에서는 도 7의 순서도(S701 내지 S705)에 나타난 바와 같이 특정 계측 데이터에 대해, [수학식 1]을 이용하여 산출한 확률밀도함수값이 신뢰한계값 미만인 경우 이상상태, 이상인 경우 정상상태로 판단하므로, 결국 이상상태 정보 모델을 구성하기 위해서는 평균, 공분산, 신뢰한계값이

필요하다. 즉, 이상상태 경보 모델 생성 모듈(121)은 계측 데이터로부터 평균, 공분산, 신뢰한계값을 산출함으로써 이상상태 경보 모델을 생성하는 것이다.

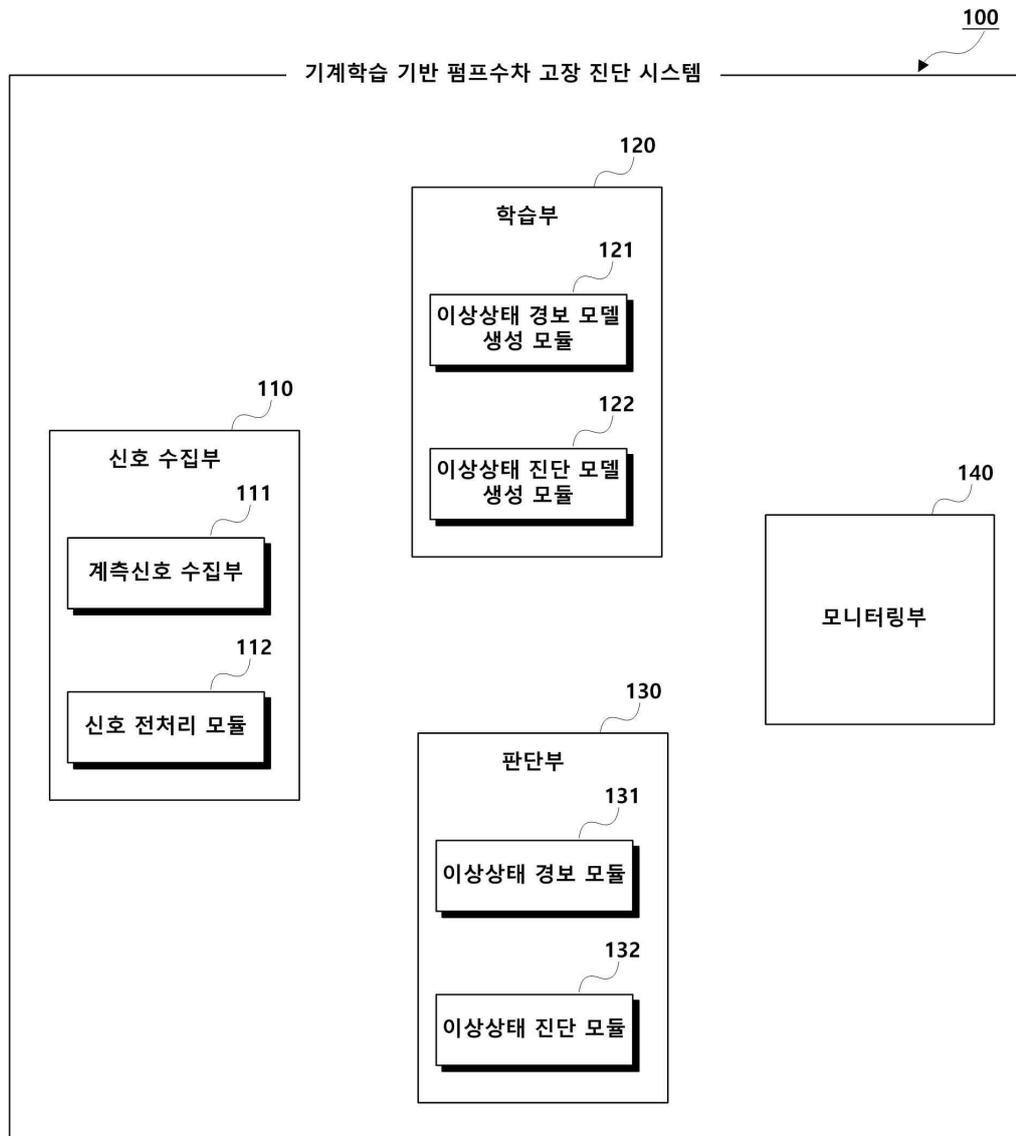
- [0057] 계측 데이터는 학습 데이터, 검증(validation) 데이터 및 평가(test) 데이터로 나누는데, 학습 데이터는 다변량 정규 분포를 구성하기 위한 정상상태 데이터로서, 검증 데이터 및 평가 데이터를 제외한 데이터를 의미한다. 이러한 정상상태 데이터인 학습 데이터를 이용하여 평균( $\mu$ )과 공분산( $\Sigma$ )을 산출한다.
- [0058] 검증 데이터는 신뢰한계값을 결정하기 위한 정상상태 및 이상상태 데이터이며, 평가 데이터는 생성된 모델의 성능을 평가하기 위한 정상상태 및 이상상태 데이터이다.
- [0059] 검증 데이터를 이용한 신뢰한계값 결정 방법은 다음과 같다.
- [0060] 즉, 검증 데이터를 이용해 정밀도(precision), 재현율(recall), F1(재현율과 정밀도의 조화평균) 등을 이용해 성능이 가장 좋은 지점의 확률밀도함수값으로 신뢰한계값을 결정하는 것이다. 고장을 양성이라 정의하면, 정밀도란, '그 중 실제 고장인 경우의 수/고장이라고 판단된 경우의 수'를 의미하고, 재현율이란 '그 중 고장이라고 판단되었던 경우의 수/실제 고장인 경우의 수'를 의미한다.
- [0061] 예를 들어 신뢰 한계값을  $10^{200}$ 에서  $10^{-200}$ 까지 변화시켜가며 F1 값을 계산하고, F1이 가장 큰 신뢰한계값을 선택할 수 있다.
- [0064] 도 8은 이상상태 진단 모델이 이상상태의 종류를 판단하는 단계를 나타내는 순서도이고, 도 9는 이상상태 진단 모델의 내부 레이어에 대한 일 실시예로서의 도면이다.
- [0065] 도 8을 참조하면, 이상상태 경보 모듈(131)을 거친 계측 데이터는 이상상태 진단 모듈(132)로 전달된다. 그 계측 데이터는 이상상태 진단 모듈(132)에 의하여 정상/이상 여부 및, 이상상태의 종류를 판단한다. 이상상태 진단 모듈(132)에서 사용하는 이상상태 진단 모델은 인공지능망이 이용될 수 있는데, 이러한 이상상태 진단 모델은 이상상태 진단 모델 생성 모듈(122)에서 기계학습에 의해 생성된다.
- [0066] 이러한 이상상태 진단 모델을 구성하는 인공지능망의 예가 도 9에 예시되어 있는데, 이상상태 진단 예시 모델은 입력층에서 전처리된 데이터를 받아 여러개의 은닉층(hidden layer), 출력층(output layer) 및 softmax 함수를 통과해 확률값이 가장 높은 클래스를 판별해 모니터링부(140)로 전달한다. 이때 분류된 정상/이상 여부의 확률값 또는 각 고장(이상) 유형 분류(클래스)의 확률값도 함께 출력하여 전달함으로써, 이상상태 경보 모듈 및 이상상태 진단 모듈의 결과가 상이할 경우 고장 진단시스템(100)의 운전원의 선택에 도움을 준다.

**부호의 설명**

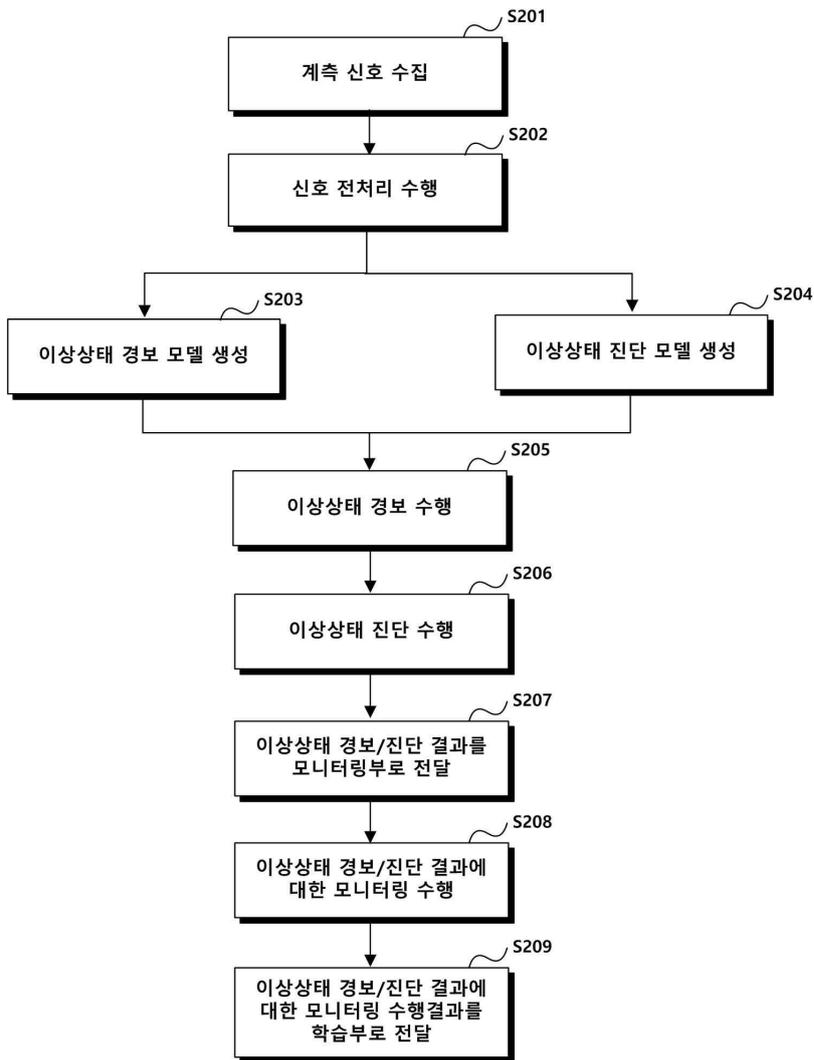
- [0067] 100: 기계학습 기반 펌프수차 고장 진단 시스템

도면

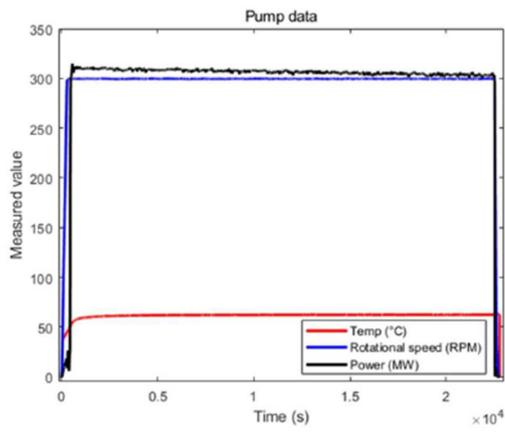
도면1



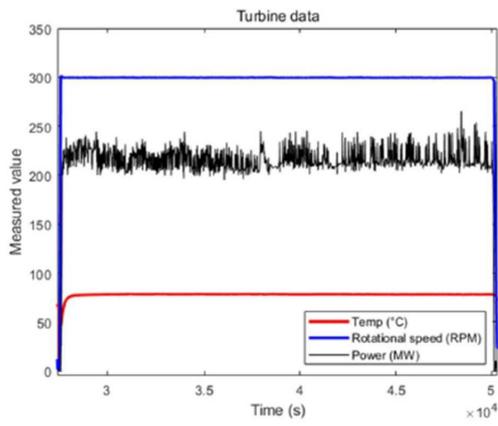
도면2



도면3

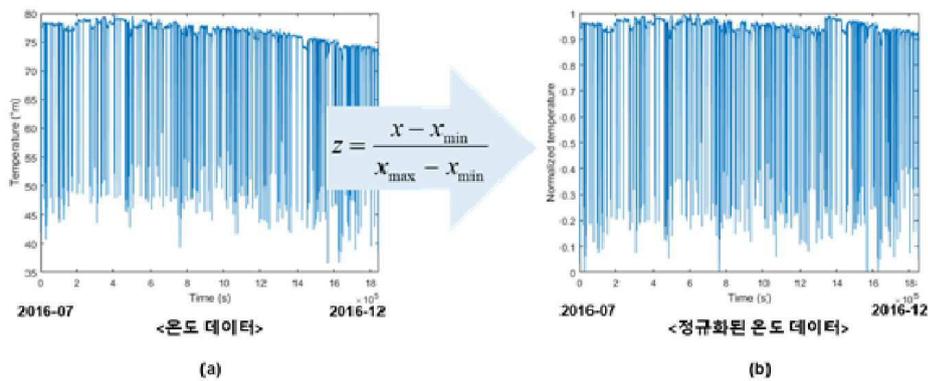


(a)



(b)

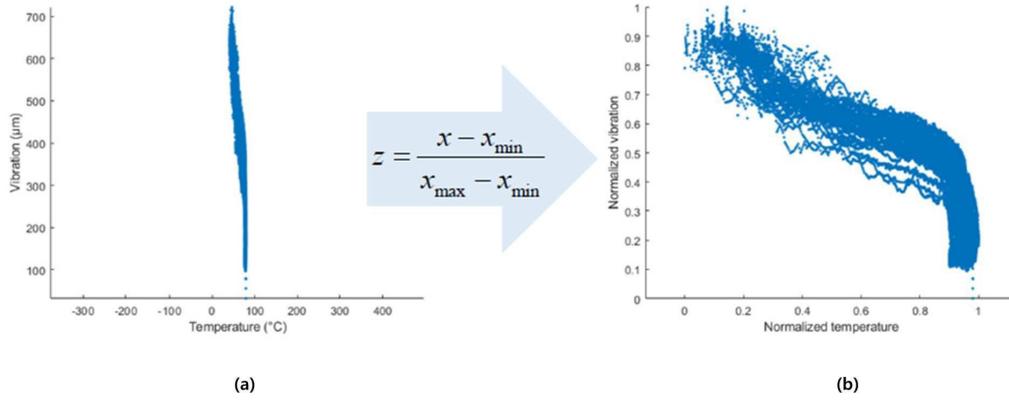
도면4



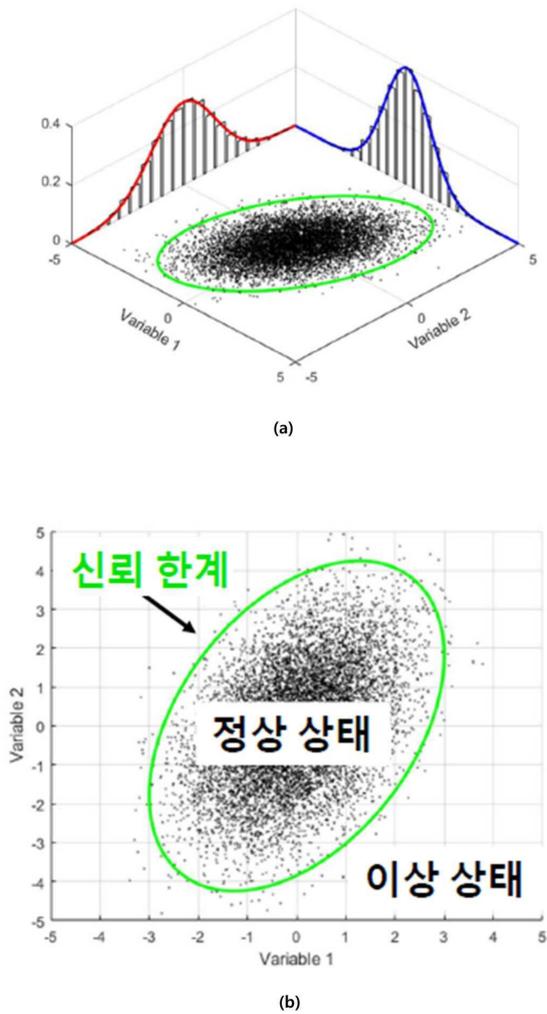
(a)

(b)

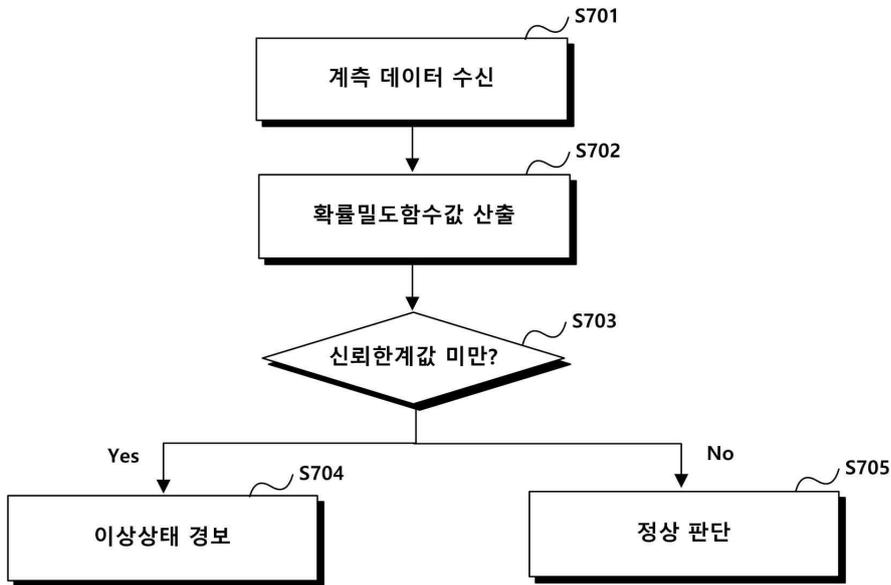
도면5



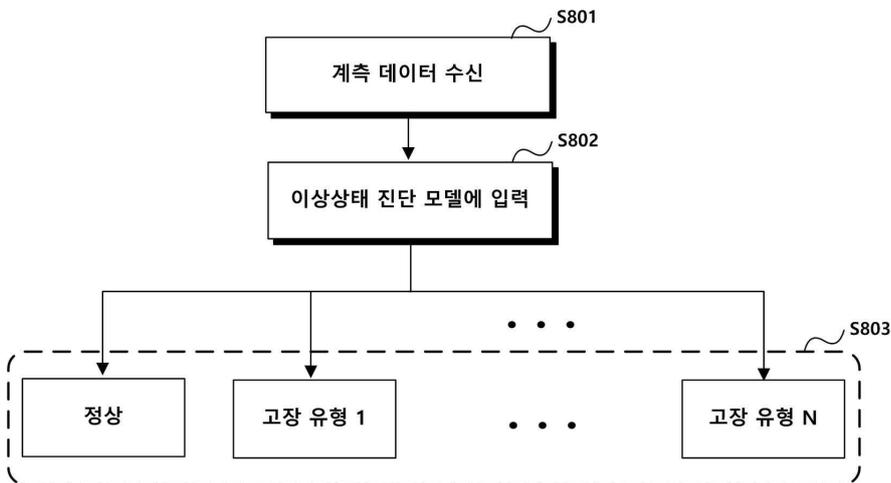
도면6



도면7



도면8



도면9

