



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0113421
(43) 공개일자 2019년10월08일

(51) 국제특허분류(Int. Cl.)
B01D 65/10 (2006.01) B01D 61/02 (2006.01)
B01D 65/02 (2006.01)
(52) CPC특허분류
B01D 65/10 (2013.01)
B01D 61/025 (2013.01)
(21) 출원번호 10-2018-0036081
(22) 출원일자 2018년03월28일
심사청구일자 2018년03월28일

(71) 출원인
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)
(72) 발명자
김준하
광주광역시 북구 첨단과기로 123 (오룡동) 광주과
학기술원 지구환경공학부
임승지
광주광역시 북구 첨단과기로 123 (오룡동) 광주과
학기술원 지구환경공학부
기서진
광주광역시 북구 경양로119번길 10-10(신안동)
(74) 대리인
한상수

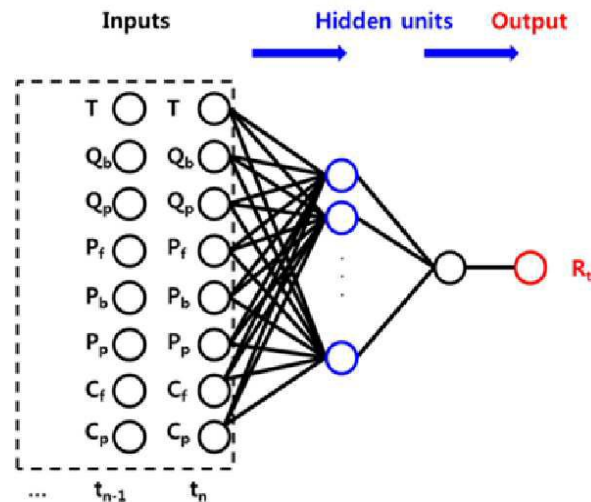
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법 및 이를 이용한 장치

(57) 요약

본 발명의 일 실시예는 역삼투 공정의 센서로부터 수집된 모니터링 데이터의 데이터베이스를 구축하는 단계; 상기 데이터베이스를 이용하여 역삼투막의 오염 예측 모델을 구성하는 단계; 상기 예측 모델을 이용하여 미래 예측 시점 및 입력 변수의 종류에 따라 입력 데이터의 양을 변화시켜 예측 정확도를 계산하는 단계; 및 상기 예측 정확도 중 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정하는 단계;를 포함하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법을 제공하여 정확한 막 오염 정도를 예측함으로써 역삼투막 세정 및 교체 시점을 결정하는데 기여할 수 있다.

대표도 - 도2



(52) CPC특허분류

B01D 65/02 (2013.01)

C02F 1/441 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1615009164(GM10740)

부처명 국토교통부

연구관리전문기관 국토교통과학기술진흥원

연구사업명 플랜트연구사업

연구과제명 지능형 SWRO 공정 운영 알고리즘 개발

기 여 율 1/1

주관기관 성균관대학교 산학협력단

연구기간 2017.03.01 ~ 2017.12.03

명세서

청구범위

청구항 1

역삼투 공정의 센서로부터 수집된 모니터링 데이터의 데이터베이스를 구축하는 단계;

상기 데이터베이스를 이용하여 역삼투막의 오염 예측 모델을 구성하는 단계;

상기 예측 모델을 이용하여 미래 예측 시점 및 입력 변수의 종류에 따라 입력 데이터의 양을 변화시켜 예측 정확도를 계산하는 단계; 및

상기 예측 정확도 중 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정하는 단계;를 포함하는 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

청구항 2

제1항에 있어서,

상기 역삼투 공정의 센서로부터 수집된 모니터링 데이터는 유입수 온도, 유입수 TDS 농도, 운전 압력, 유입수 유량, 생산수 유량, 생산수 TDS 농도, 생산수 압력 및 농축수 압력 중 어느 하나 이상인 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

청구항 3

제2항에 있어서,

상기 역삼투 공정의 센서로부터 수집된 모니터링 데이터 중 하나 이상의 결측이 있는 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

청구항 4

제1항에 있어서,

상기 데이터베이스를 이용하여 역삼투막의 오염 예측 모델을 구성하는 단계는 기계학습을 통하여 이루어지는 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

청구항 5

제4항에 있어서,

상기 기계학습은 시간지연 신경망을 통하여 이루어지는 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

청구항 6

제5항에 있어서,

상기 시간지연 신경망은 숨김층이 1층 내지 3층인 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

청구항 7

제1항에 있어서,

상기 미래 예측 시점이 현재에서 멀어질수록 입력 데이터의 양이 증가하는 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

청구항 8

제1항에 있어서,

상기 입력 데이터는 결측 데이터를 포함하고, 상기 결측 데이터가 많아질수록 입력 데이터의 양이 증가하는 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

청구항 9

제1항에 있어서,

상기 예측 정확도는 하기 수식 (1)에 따라 계산되는 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

$$NSE(\text{Nash} - \text{Sutcliffe Efficiency}) = 1 - \left[\frac{\sum_{t=1}^T (Q_m^t - Q_o^t)^2}{\sum_{t=1}^T (Q_o^t - Q_o^{mean})^2} \right] \quad (1)$$

(여기서, Q_m^t 는 t에서 모의값, Q_o^t 는 t에서 관측값, Q_o^{mean} 은 관측값의 평균값임)

청구항 10

제1항에 있어서,

상기 예측 정확도 중 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정하는 단계에서 상기 데이터의 양은 일(日) 단위로 결정되는 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

청구항 11

제1항에 있어서,

상기 결정된 데이터의 양을 역삼투막 오염 예측에 실시간으로 제공하는 단계;를 더 포함하는 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

청구항 12

제1항에 있어서,

상기 역삼투 공정의 센서로부터 수집된 모니터링 데이터의 데이터베이스를 구축하는 단계;와 상기 데이터베이스를 이용하여 역삼투막의 오염 예측 모델을 구성하는 단계; 사이에 입력값을 정규화(normalizing)하는 단계;를 더 포함하는 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법.

청구항 13

역삼투 공정에서 유입수 인자 및 생산수 인자를 수집하기 위한 데이터 수집부;

상기 유입수 인자 및 생산수 인자를 입력 받아 역삼투막 오염 예측 모델을 구동하는 구동부;

상기 구동부에 미래 예측 시점 및 입력 인자의 종류에 따라 입력 데이터의 양을 변화시켜 예측 정확도를 계산하고 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정하는 연산부; 및

상기 연산부에서 결정된 데이터의 양을 실시간으로 구동부에 전달하는 피드백부;를 포함하는 것을 특징으로 하는 역삼투막 오염 예측 장치.

청구항 14

제13항에 있어서,

상기 역삼투막 오염 예측 모델은 시간지연 신경망을 이용하는 것을 특징으로 하는 역삼투막 오염 예측 장치.

청구항 15

제13항에 있어서,

상기 연산부가 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정할 때, 상기 데이터의 양은 일(日) 단위로 결정되는 것을 특징으로 하는 역삼투막 오염 예측 장치.

발명의 설명

기술 분야

- [0001] 본 발명은 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법 및 이를 이용한 장치에 관한 것으로, 보다 상세하게는 역삼투 공정의 유지관리에 필요한 막 오염 예측을 위해 기계학습 기반의 역삼투막 오염 예측 모델을 이용하고, 역삼투막 오염 예측 모델의 정확도 향상을 위하여 표본 입력 데이터의 양을 결정하는 방법과 이를 이용한 역삼투막 오염 예측 장치에 관한 것이다.

배경 기술

- [0002] 해수담수화 기술은 최근 대두되기 시작한 물 부족 문제를 해결할 수 있는 기술로서, 계절이나 기상조건에 상관없이 다량의 수자원의 확보가 가능한 해수를 담수화하는 기술을 말한다. 해수담수화 기술은 댐 건설에 비해 공사기간이 짧고 초기 투자비가 작으므로 지속적인 수자원 확보를 위한 대체 기술로 자리 잡고 있으며 저에너지, 대형화, 안정성, 친환경 등을 추구하는 기술이다.
- [0003] 해수담수화 플랜트란 바닷물을 증발시키거나 막을 통과하여 염분뿐만 아니라 다수의 무기 염류를 제거하여 공업용수, 식수 등 인류가 사용할 수 있는 담수를 생산하는 공정과 설비를 총칭한다.
- [0004] 담수화 방식은 크게 열에너지와 물의 증발현상을 이용한 증류/증발법과 막의 차별성과 선택적 투과 능력을 이용한 역삼투(Reverse Osmosis; RO)법으로 나눌 수 있다. 해수담수화 플랜트의 주요 공법 중 역삼투(RO) 방식, 즉 해수 역삼투(Seawater Reverse Osmosis; SWRO) 방식은 기존의 다단 플래쉬 증류법(Multi-Stage Flash Distillation; MSF)이나 다중효용 증발법(Multiple-Effect Evaporation; ME)보다 에너지 효율이 높아 최근 상업실적이 증가하고 있다. 역삼투법은 압력을 이용하여 물은 통과시키고 용질은 투과시키지 않는 역삼투막에 해수를 가압하여 담수를 분리하는 공법이다.
- [0005] 역삼투막을 이용한 해수담수화 공정에서는 역삼투가 진행될수록 막의 오염이 발생한다. 막 오염은 유입수에 존재하는 물질들이 막에 부착하여 분리막의 성능을 저하시키는 것을 말하며, 특히 해수담수화에서는 고압조건(50 내지 70bar)으로 인하여 유입수에 존재하는 물질들이 막에 강하게 축적된다. 막이 오염되면 전체 공정의 효율이 급격하게 저하되므로 막 오염을 제어하는 것은 역삼투 공정에 있어서 매우 중요하다. 막 오염이 진행되면 주기적으로 막을 세정하거나, 세정으로 해결할 수 없을 정도가 되면 막을 교체해야 한다.
- [0006] 문제는 막 오염이 진행되는 정도를 파악하여 막 세정시기나 교체시기를 결정해야 하는데, 막 오염 정도를 실시간으로 모니터링하거나 예측하는데 한계가 있다는 것이다. 유입수나 생산수의 흐름이 감소하거나 압력이 감소하는 정도를 측정하여 간접적으로 막 오염의 진행정도를 파악할 수 있으나, 막 자체의 오염정도를 판단하는 것은 아니고, 이렇게 간접적으로 막 오염을 판단하여 막 세정이나 교체시기를 결정하는 경우 부적절한 시기의 세정이나 교체가 발생하여 불필요한 유지관리 비용을 초래하게 된다. 이를 해결하기 위하여 막 오염 정도를 파악하는데 있어서 알고리즘을 이용한 접근 방법이 모색되어 왔다.
- [0007] 종래기술에서는 역삼투 공정으로 유입되는 유입원수의 수질, 공정의 운전 조건, 그리고 생산수 수질을 측정하여 데이터 베이스를 구축하고, 구축된 데이터 베이스를 기반으로 막 오염 정도에 대한 지표인 막간 차압, 수투과량 변화 그리고 막 저항 등의 변화를 예측하고 이를 통해서 공정의 성능 변화를 예측하는 기술을 제공한다.
- [0008] 그러나 미래 예측 시점 및 입력 데이터의 종류에 상관없이 고정된 입력 데이터의 양을 이용하기 때문에 최고의 정확도가 유지되지 않을뿐더러, 특히 입력 데이터에서 예측되는 변수 발생 등 예측 과정에서 발생하는 데이터의 실시간 변화를 정확하게 반영하지 못하기 때문에 예측 정확도가 낮아지는 문제가 있다.

선행기술문헌

특허문헌

- [0009] (특허문헌 0001) 대한민국 등록특허 제 10-1187416 호

발명의 내용

해결하려는 과제

- [0010] 본 발명이 이루고자 하는 기술적 과제는 상기한 문제점을 해결하기 위한 것으로, 데이터의 변화 경향성을 반영하여 예측 정확도를 향상시키고, 예측되는 변수가 있을지라도 예측 정확도를 일정 수준 이상으로 유지할 수 있는 역삼투막 오염 예측 모델의 표본 입력 데이터 양을 결정하는 방법을 제공하는 것이다.
- [0011] 본 발명이 이루고자 하는 기술적 과제는 이상에서 언급한 기술적 과제로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0012] 상기 기술적 과제를 달성하기 위하여, 본 발명의 일 실시예는 역삼투 공정의 센서로부터 수집된 모니터링 데이터의 데이터베이스를 구축하는 단계; 상기 데이터베이스를 이용하여 역삼투막의 오염 예측 모델을 구성하는 단계; 상기 예측 모델을 이용하여 미래 예측 시점 및 입력 변수의 종류에 따라 입력 데이터의 양을 변화시켜 예측 정확도를 계산하는 단계; 및 상기 예측 정확도 중 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정하는 단계;를 포함하는 것을 특징으로 하는 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법을 제공한다.
- [0013] 본 발명의 일 실시예에 있어서, 상기 역삼투 공정의 센서로부터 수집된 모니터링 데이터는 유입수 온도, 유입수 TDS 농도, 운전 압력, 유입수 유량, 생산수 유량, 생산수 TDS 농도, 생산수 압력 및 농축수 압력 중 어느 하나 이상일 수 있다.
- [0014] 본 발명의 일 실시예에 있어서, 상기 역삼투 공정의 센서로부터 수집된 모니터링 데이터 중 하나 이상의 결측이 있을 수 있다.
- [0015] 본 발명의 일 실시예에 있어서, 상기 데이터베이스를 이용하여 역삼투막의 오염 예측 모델을 구성하는 단계는 기계학습을 통하여 이루어질 수 있다.
- [0016] 본 발명의 일 실시예에 있어서, 상기 기계학습은 시간지연 신경망을 통하여 이루어질 수 있다.
- [0017] 본 발명의 일 실시예에 있어서, 상기 시간지연 신경망은 숨김층이 1층 내지 3층일 수 있다.
- [0018] 본 발명의 일 실시예에 있어서, 상기 미래 예측 시점이 현재에서 멀어질수록 입력 데이터의 양이 증가할 수 있다.
- [0019] 본 발명의 일 실시예에 있어서, 상기 입력 데이터는 결측 데이터를 포함하고, 상기 결측 데이터가 많아질수록 입력 데이터의 양이 증가할 수 있다.
- [0020] 본 발명의 일 실시예에 있어서, 상기 예측 정확도는 하기 수학적 식 (1)에 따라 계산되는 것일 수 있다.

$$\text{NSE(Nash - Sutcliffe Efficiency)} = 1 - \left[\frac{\sum_{t=1}^T (Q_m^t - Q_o^t)^2}{\sum_{t=1}^T (Q_o^t - Q_o^{\text{mean}})^2} \right] \quad (1)$$

[0022] (여기서, Q_m^t 는 t에서 모의값, Q_o^t 는 t에서 관측값, Q_o^{mean} 은 관측값의 평균값임)

- [0023] 본 발명의 일 실시예에 있어서, 상기 예측 정확도 중 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정하는 단계에서 상기 데이터의 양은 일(日) 단위로 결정될 수 있다.
- [0024] 본 발명의 일 실시예에 있어서, 상기 결정된 데이터의 양을 역삼투막 오염 예측에 실시간으로 제공하는 단계;를 더 포함할 수 있다.
- [0025] 본 발명의 일 실시예에 있어서, 상기 역삼투 공정의 센서로부터 수집된 모니터링 데이터의 데이터베이스를 구축하는 단계;와 상기 데이터베이스를 이용하여 역삼투막의 오염 예측 모델을 구성하는 단계; 사이에 입력값을 정규화(normalizing)하는 단계;를 더 포함할 수 있다.
- [0026] 상기 기술적 과제를 달성하기 위하여, 본 발명의 다른 실시예는 역삼투 공정에서 유입수 인자 및 생산수 인자를 수집하기 위한 데이터 수집부; 상기 유입수 인자 및 생산수 인자를 입력 받아 역삼투막 오염 예측 모델을 구동

하는 구동부; 상기 구동부에 미래 예측 시점 및 입력 인자의 종류에 따라 입력 데이터의 양을 변화시켜 예측 정확도를 계산하고 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정하는 연산부; 및 상기 연산부에서 결정된 데이터의 양을 실시간으로 구동부에 전달하는 피드백부;를 포함하는 것을 특징으로 하는 역삼투막 오염 예측 장치를 제공한다.

[0027] 본 발명의 다른 실시예에 있어서, 상기 역삼투막 오염 예측 모델은 시간지연 신경망을 이용하는 것일 수 있다.

[0028] 본 발명의 다른 실시예에 있어서, 상기 연산부가 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정할 때, 상기 데이터의 양은 일(日) 단위로 결정될 수 있다.

발명의 효과

[0029] 본 발명의 실시예에 따르면, 역삼투 공정의 운영관리 업무 중 기계학습을 통해서 의사결정권자가 원하는 공정의 막 오염 정도를 예측할 때, 막 오염 정도 예측 및 예측 정확도 향상을 위해 필요한 표본 데이터의 양을 제안하여, 정확한 막 오염 정도를 예측함으로써 역삼투막 세정 및 교체 시점을 결정하는데 기여할 수 있다.

[0030] 본 발명의 효과는 상기한 효과로 한정되는 것은 아니며, 본 발명의 상세한 설명 또는 특허청구범위에 기재된 발명의 구성으로부터 추론 가능한 모든 효과를 포함하는 것으로 이해되어야 한다.

도면의 간단한 설명

[0031] 도 1은 종래기술에 따른 인공지능망의 구조를 간략히 나타낸 흐름도이다.

도 2는 본 발명에 따른 시간지연 신경망의 구조를 간략히 나타낸 흐름도이다.

도 3은 본 발명의 일 실시예에 따른 예측 정확도를 나타낸 그래프이다.

도 4는 본 발명의 비교예에 따른 예측 정확도를 나타낸 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0032] 이하에서는 첨부한 도면을 참조하여 본 발명을 설명하기로 한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 따라서 여기에서 설명하는 실시예로 한정되는 것은 아니다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0033] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결(접속, 접촉, 결합)"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 부재를 사이에 두고 "간접적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 구비할 수 있다는 것을 의미한다.

[0034] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0035] 이하 첨부된 도면을 참고하여 본 발명의 실시예를 상세히 설명하기로 한다.

[0037] 이하, 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법에 대하여 설명한다.

[0038] 본 발명에 따른 역삼투막 오염 예측 모델의 표본 입력 데이터 양 결정 방법은 역삼투 공정의 센서로부터 수집된 모니터링 데이터의 데이터베이스를 구축하는 단계; 상기 데이터베이스를 이용하여 역삼투막의 오염 예측 모델을 구성하는 단계; 상기 예측 모델을 이용하여 미래 예측 시점 및 입력 변수의 종류에 따라 입력 데이터의 양을 변화시켜 예측 정확도를 계산하는 단계; 및 상기 예측 정확도 중 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정하는 단계;를 포함한다.

[0039] 상기 역삼투 공정의 센서로부터 수집된 모니터링 데이터는 유입수 온도, 유입수 TDS 농도, 운전 압력, 유입수

유량, 생산수 유량, 생산수 TDS 농도, 생산수 압력 및 농축수 압력 중 어느 하나 이상일 수 있다.

- [0040] 상기 역삼투 공정의 센서로부터 수집된 모니터링 데이터 중 하나 이상의 결측이 있을 수 있다.
- [0041] 역삼투를 이용한 해수담수화 플랜트에 있어서 유입원수는 해수가 되고, 생산수는 담수가 되는데, 유입수의 온도, 압력 또는 유량을 조절함으로써 생산수의 유량 또는 수질을 향상시킬 수 있다. 수질은 통상적으로 총 용존성 고형물질(Total Dissolved solids; TDS) 단위로 판단한다.
- [0042] 유입수의 온도, 압력 또는 유량의 제어에 따른 생산수의 유량, 수질 데이터는 수일 내지는 수년에 걸쳐 축적될 수 있고, 이를 데이터베이스로 구축하여 본 발명에 이용한다. 모니터링 데이터의 종류는 상술한 것에 한정되지 않고, 막 오염 예측에 필요한 정도로 확장 또는 부가될 수 있다.
- [0043] 상기 데이터베이스를 이용하여 역삼투막의 오염 예측 모델을 구성하는 단계는 기계학습을 통하여 이루어질 수 있고, 기계학습은 시간지연 신경망을 통하여 이루어질 수 있다.
- [0044] 기계학습은 인공지능의 한 분야로, 컴퓨터가 학습할 수 있도록 하는 알고리즘과 기술을 개발하는 분야를 말한다. 컴퓨터에 명시적인 프로그램 없이 학습하는 능력을 부여하는 것이며, 인공지능망 모델은 기계학습의 한 분야이다.
- [0045] 도 1은 종래기술에 사용되는 인공신경망의 구조를 개략적으로 나타낸 흐름도이다. 인공신경망은 인간의 뇌가 패턴을 인식하는 방식을 모사한 알고리즘이다. 도 1에서 보는 원은 인간의 신경망을 구성하는 뉴런을 모사한 것이고, 노드라고 불린다. 노드들은 시냅스의 결합으로 연결되어 네트워크를 형성한다. 노드들이 여러 개 배열되어 층을 형성하며, 층은 입력층과 숨김층, 출력층으로 형성된다.
- [0046] 노드에서는 실제로 연산이 일어나는데, 이 연산 과정은 인간의 신경망을 구성하는 뉴런에서 일어나는 과정을 모사하도록 설계되어 있다. 노드는 일정 크기 이상의 자극을 받으면 반응을 하는데, 그 반응의 크기는 입력 값과 노드의 계수(또는 가중치)를 곱한 값과 대략 비례한다. 일반적으로 노드는 여러 개의 입력을 받으며 입력의 개수만큼 계수를 가지고 있다. 따라서 이 계수를 조절함으로써 여러 입력에 다른 가중치를 부여할 수 있다. 최종적으로 곱한 값들은 전부 더해지고 그 합은 활성화 함수의 입력으로 들어가게 된다. 활성화 함수의 결과가 노드의 출력에 해당하며 이 출력값이 궁극적으로 분류나 회귀 분석에 쓰이게 된다.
- [0047] 인공신경망 학습의 목적은 출력의 오류를 최소화 하는 것이다. 학습이 시작되기 전에 네트워크 상의 모든 계수를 초기화한다. 그리고 반복적으로 데이터를 보내주어서 학습을 한다. 만일 학습이 잘 되었다면 계수는 적절한 값으로 업데이트 되어 있을 것이고, 그 인공신경망으로 각종 분류와 예측이 가능하다. 학습과정 내부에서는 같은 원리의 계수 업데이트가 반복적으로 일어난다. 계수 업데이트의 원리는 우선 계수를 추정하고 그 계수를 사용했을 때 발생하는 에러를 측정한 뒤 그 에러에 기반해서 계수를 약간씩 업데이트 하는 것이다. 먼저 입력이 신경망에 들어오면, 현재 상태의 계수를 이용하여 결과를 출력한다. 그리고 이렇게 추정한 값을 실제 정답과 비교한다. 정답과 추정값의 차이가 오차이며, 신경망은 오차를 측정하고 이 오차를 반영해서 계수를 보정하는 것이다. 이 과정을 계속 반복하는 것이 학습 과정이다.
- [0048] 종래기술에서는 인공신경망을 학습시켜 모델을 구축하고, 이 모델을 이용하여 막 오염 정도에 대한 지표인 막간차압, 수투과량 변화 그리고 막 저항 등의 변화를 예측하고 이를 통해서 공정의 성능 변화를 예측하였다.
- [0049] 그러나 인공신경망은 표본 데이터의 크기가 고정된 상태에서 결과를 예측하였으므로, 표본 데이터의 크기에 변화가 있는 경우 예측의 정확도가 감소하는 문제가 있다. 인공신경망에 표본 입력 데이터의 크기를 고려하는 기능이 추가된 자율 학습을 수행하는 것이 시간지연 신경망이며, 표본 데이터의 크기에 따라서 과거의 데이터를 고려한 예측을 진행한다.
- [0050] 도 2는 본 발명의 시간지연 신경망의 구조를 개략적으로 나타낸 흐름도이다. 인공신경망과 비교하면, 시간지연 신경망은 입력층에 차이가 있다. 인공신경망에서는 입력층에 한 줄의 노드가 배치되어 있고, 각 노드는 동일 시간대에 하나의 데이터만을 숨김층의 노드들로 전달한다. 그러나 시간지연 신경망은 입력 데이터 결측이 있는 경우를 고려하여 t_n 시간의 데이터와 t_{n-1} , t_{n-2} , ...의 데이터를 하나로 묶어 입력한다. 인공신경망의 경우 입력 데이터에 결측이 있는 경우 해당 노드는 다음 층으로 전달할 신호가 없게 되나, 시간지연 신경망의 경우 기준 시점으로부터 임의의 과거 시점까지의 데이터를 하나로 묶어 입력하므로 입력 데이터 일부에 결측이 발생하더라도 입력층의 노드가 다음 층으로 신호를 전달할 수 없게 되는 확률이 줄어들게 된다.
- [0051] 시간지연 신경망에서는 기준 시점으로부터 임의의 과거 시점까지의 데이터를 하나로 묶어 입력하므로, 시간의

흐름에 따라 연속적으로 입력되는 데이터를 처리하는데 한계가 있다. 따라서 시간의 흐름에 따라 연속적으로 입력되는 데이터를 처리하려면 미리 하나로 묶을 시간 간격을 정하고 이를 묶어주는 작업이 필요하다. 다만 연속적인 데이터의 연속성을 유지하기 위하여 시간 간격에 따라 나눈 데이터를 단순히 묶는 것이 아니라 일정한 시간 간격의 데이터를 일정 부분씩 겹치도록 전처리 한 뒤 사용하는 것이 보통이다. 예를 들어 $t_n, t_{n-1}, t_{n-2}, \dots$ 의 데이터가 있고, t_n 부터 t_{n-3} 까지, t_{n-4} 부터 t_{n-7} 까지 시간 간격으로 나누어 데이터를 묶는다고 한다면 첫 번째 데이터 묶음은 t_n 부터 t_{n-3} 까지, 두 번째 데이터 묶음은 t_{n-2} 부터 t_{n-5} 까지, 세 번째 데이터 묶음은 t_{n-4} 부터 t_{n-7} 까지의 데이터를 포함하게 된다. 즉, t_n 부터 t_{n-3} 까지, t_{n-4} 부터 t_{n-7} 까지 두 번의 시간 간격 동안 세 묶음의 데이터를 전송해야 하고, 시간지연이 발생하는 것이다.

[0052] 본 발명에 있어서 시간지연 신경망은 숨김층이 1층 내지 3층인 것일 수 있다. 신경망의 숨김층은 여러 층을 가질 수 있고, 각각의 층마다 다른 특징을 학습할 수 있다. 이러한 특징들을 조합하여 더욱 복잡한 데이터를 분류 및 군집화 할 수 있는 것이다. 본 발명에서는 역삼투막 성능 예측이 목적이고, 막간 차압, 수투과량 변화 그리고 막 저항 등의 요소를 고려하여 역삼투막의 성능을 예측하게 되므로, 숨김층에 여러 개의 층을 두어 각각의 특징을 학습시킬 수 있고, 하나의 층으로 학습을 시키는 것도 가능하다. 다만 적어도 하나의 숨김층은 반드시 필요하고, 3개층을 초과하여 숨김층이 존재하면 학습에 너무 긴 시간이 소요되고 시스템 자원을 많이 필요로 하게 되므로 바람직하지 않다.

[0053] 본 발명에 있어서 예측 모델을 이용하여 미래 예측 시점 및 입력 변수의 종류에 따라 입력 데이터의 양을 변화시켜 예측 정확도를 계산하는 단계는 현재 시점으로부터 미래 예측 시점까지 얼마나 멀리 떨어져 있는지 또는 입력 변수의 종류가 얼마나 되는지에 따라 입력 데이터의 양을 달리하여 구축된 역삼투막 오염 예측 모델에 입력하여 처리해가면서 예측 정확도를 계산하는 것이다. 상기 미래 예측 시점이 현재에서 멀어질수록 입력 데이터의 양이 증가하는 것을 특징으로 한다.

[0054] 시간지연 신경망을 통하여 역삼투막 오염 예측 모델을 구성하는 단계나 예측 정확도를 계산하는 단계 모두 데이터를 입력하고 출력값을 얻는 것은 동일하지만, 역삼투막 오염 예측 모델을 구성하는 단계는 상술한 바와 같이 각 노드의 계수를 초기화하고 데이터 입출력을 반복하면서 각 노드의 계수를 업데이트 하는 것에 목적이 있다면, 예측 정확도를 계산하는 단계는 후술할 최고의 정확도를 나타낼 입력 데이터의 양을 결정하는데 목적이 있다.

[0055] 현재 시점으로부터 미래 예측 시점이 떨어져 있는 시간차만큼 현재 시점으로부터 과거 시점까지의 입력 데이터를 역삼투막 오염 예측 모델에 입력해 가면서 입력 데이터의 크기에 따른 예측 정확도를 계산하게 되고, 이 때 입력 데이터의 크기는 현재 시점으로부터 미래 예측 시점 또는 그에 대응하는 과거 시점까지의 시간차, 즉 시간의 길이에 대응하여 정해질 수 있다. 다만 상술한대로 시간의 길이가 2배가 된다고 하더라도 실제 묶어서 입력해야 하는 데이터는 일정부분씩 겹치게 하여 전처리 되므로, 실제 입력되는 데이터의 크기는 2배 이상이 될 것이고, 시간의 길이에 정비례하는 것은 아니다.

[0056] 입력 데이터에는 결측 데이터가 포함되어 있을 수 있고, 결측 데이터가 많아질수록 입력 데이터의 양이 증가하게 된다. 입력 데이터가 증가하는 경우 상술한 바와 같이 과거 데이터를 참조하여 결측값을 보완할 수 있게 된다.

[0057] 예측 정확도의 계산은 하기 수학적 식 (1)에 따라 계산될 수 있다.

[0058]
$$NSE(\text{Nash} - \text{Sutcliffe Efficiency}) = 1 - \left[\frac{\sum_{t=1}^T (Q_m^t - Q_o^t)^2}{\sum_{t=1}^T (Q_o^t - Q_o^{mean})^2} \right] \quad (1)$$

[0059] (여기서, Q_m^t 는 t에서 모의값, Q_o^t 는 t에서 관측값, Q_o^{mean} 은 관측값의 평균값임)

[0060] 상기 식에서 모의값은 역삼투막 오염 예측 모델에 의해 산출된 값이고, 관측값은 모니터링 데이터에 의한 실제 값이다. NSE값이 1에 가까울수록 예측 정확도가 높은 것이며 1에서 멀수록 예측 정확도가 낮은 것이다.

[0061] 상기 예측 정확도 중 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 최종적으로 결정할 수 있고, 이후 역삼투막 오염 예측 모델의 구동에 상기 결정된 양을 단위로 하여 데이터를 입력할 수 있다. 상기 데이터의 양은 일(日) 단위로 결정되는 것이 바람직하다. 시간 단위로 결정되는 경우 예측 정확도의 계산 및 데이터 양을 결정하는데 너무 많은 시간이 소요될 수 있고, 일(日)을 넘어 주 단위로 결정되는 경우 막 오염 예측에 따른 세정 및 교체 주기 결정에 실효성이 떨어지게 되므로 바람직하지 않다.

- [0062] 본 발명은 상기 결정된 데이터의 양을 역삼투막 오염 예측에 실시간으로 제공하는 단계를 더 포함할 수 있다. 역삼투 공정에서 모니터링 데이터를 실시간으로 수집하고, 이를 이용하여 역삼투막 오염 예측 모델의 계수를 실시간으로 보정해 가면서 최고의 정확도를 나타내는 데이터의 양을 실시간으로 오염 예측 모델에 제공한다면 막 오염 예측의 정확도를 지속적으로 향상시킬 수 있다.
- [0063] 또한 본 발명은 상기 역삼투 공정의 센서로부터 수집된 모니터링 데이터의 데이터베이스를 구축하는 단계;와 상기 데이터베이스를 이용하여 역삼투막의 오염 예측 모델을 구성하는 단계; 사이에 입력값을 정규화(normalizing)하는 단계;를 더 포함할 수 있다.
- [0064] 정규화(normalizing)란 데이터를 일정한 규칙에 따라 변형하여 이용하기 쉽게 하는 절차로, 데이터의 규격을 통일하는 절차라고 할 수 있다. 역삼투 공정의 센서로부터 수집된 모니터링 데이터는 종류에 따라 제각각의 범위를 가지고 있고, 이를 그대로 시간지연 신경망에 입력하여 기계학습을 하는 경우 너무 넓은 데이터 범위로 인하여 데이터간 상관관계를 파악하고 학습하는데 오랜 시간이 걸릴 수 있다. 또한 중복데이터의 입력으로 신경망에 불필요한 로드가 걸릴 수도 있다. 이를 해결하기 위하여 모든 데이터를 일정 범위 내로 들어오도록 변환하거나, 센서 오류 등으로 평균값에서 크게 벗어나는 값들을 제외시킬 수도 있다. 이러한 과정을 거쳐 산출된 값들을 기계학습에 이용하면 역삼투막 오염 예측 모델의 구성 효율을 높일 수 있다.
- [0066] 이하, 역삼투막 오염 예측 장치에 대하여 설명한다.
- [0067] 본 발명은 또한 상기 방법을 이용한 역삼투막 오염 예측 장치를 제공한다. 본 발명에 따른 역삼투막 오염 예측 장치는 역삼투 공정에서 유입수 인자 및 생산수 인자를 수집하기 위한 데이터 수집부; 상기 유입수 인자 및 생산수 인자를 입력 받아 역삼투막 오염 예측 모델을 구동하는 구동부; 상기 구동부에 미래 예측 시점 및 입력 인자의 종류에 따라 입력 데이터의 양을 변화시켜 예측 정확도를 계산하고 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정하는 연산부; 및 상기 연산부에서 결정된 데이터의 양을 실시간으로 구동부에 전달하는 피드백부;를 포함하는 것을 특징으로 한다.
- [0068] 상기 역삼투막 오염 예측 모델은 시간지연 신경망을 이용하는 것이며, 상기 연산부가 최고의 예측 정확도를 가질 때의 입력 데이터의 양을 결정할 때, 상기 데이터의 양은 일(日) 단위로 결정할 수 있다.
- [0069] 수집부에서 수집된 데이터는 일 단위 내지 년 단위로 축적되어 데이터베이스로 구축될 수 있다. 구축된 데이터베이스를 구동부에 연결하여 역삼투막 오염 예측 모델을 구동하여 계수를 업데이트 하고, 연산부에서 데이터의 양을 변화시켜가며 예측 정확도를 계산하여 최고의 정확도를 가질 때의 데이터 양을 결정할 수 있다. 결정된 데이터의 양은 구동부에 피드백 되어 결정된 데이터 양만큼 역삼투막 오염 예측 모델에 입력되도록 구성할 수도 있다.
- [0071] 이하, 실시예를 참조하여 본 발명을 설명한다.
- [0073] **실시예 1**
- [0074] 역삼투막 해수담수화 공정에서 유입수 온도, 유입수 TDS 농도, 운전 압력, 유입수 유량, 생산수 유량 및 생산수 TDS 농도 데이터를 모니터링하여 2년간의 데이터를 데이터베이스화 하였다. 시간지연 신경망 모델을 구동하여 상기 데이터베이스의 값들을 임의의 일정한 시간 간격으로 입력하여 계수를 보정하여 역삼투막 오염 예측 모델을 구축하였다. 구축된 모델에 데이터의 양을 2일, 4일, 6일 간격으로 입력하여 1일 이후의 결과 예측 정확도를 계산하여 하기 표 1에 나타내었다. 가장 높은 예측 정확도를 가지는 시간 간격을 정하여 데이터의 양을 결정하였다.
- [0076] **실시예 2**
- [0077] 생산수 압력 및 농축수 압력 데이터를 추가한 것을 제외하고는 상기 실시예 1과 동일하게 1일 이후의 결과 예측 정확도를 계산하여 하기 표 1에 나타내었다. 가장 높은 예측 정확도를 가지는 시간 간격을 정하여 데이터의 양을 결정하였다. 결정된 데이터의 양에 대한 시간에 따른 예측 정확도를 도 3에 나타내었다.

비교예 1

역삼투막 해수담수화 공정에서 유입수 온도, 유입수 TDS 농도, 운전 압력, 유입수 유량, 생산수 유량 및 생산수 TDS 농도 데이터를 모니터링하여 2년간의 데이터를 데이터베이스화 하였다. 인공신경망 모델을 구동하여 상기 데이터베이스의 값들을 입력하여 계수를 보정하여 역삼투막 오염 예측 모델을 구축하였다. 구축된 모델에 데이터를 입력하여 1일 이후의 결과 예측 정확도를 계산하여 하기 표 1에 나타내었다.

비교예 2

생산수 압력 및 농축수 압력 데이터를 추가한 것을 제외하고는 상기 비교예 1과 동일하게 1일 이후의 결과 예측 정확도를 계산하여 하기 표 1에 나타내었다. 시간에 따른 예측 정확도를 도 4에 나타내었다.

표 1

구 분	Input 변수의 개수	시간 지연 (Time Delay)	Train (NSE)	Test (NSE)
비교예 1	6	0	0.93	0.68
실시예 1	6	2	0.97	0.75
	6	4	0.98	0.85
	6	6	0.97	0.63
비교예 2	8	0	0.98	0.53
실시예 2	8	2	0.99	0.66
	8	4	0.99	0.86
	8	6	0.99	0.51

비교예 1 및 비교예 2에서는 인공신경망을 사용하여 시간지연이 일어나지 않으므로 시간 지연 값이 0이다.

6개의 데이터를 가지고 예측 시 시간 지연이 4일 때 가장 높은 정확도 0.85를 나타내었다.

8개의 데이터를 가지고 예측 시 시간 지연이 4일 때 가장 높은 정확도 0.86을 나타내었다.

시간 지연이 없는 비교예 1 및 비교예 2보다 실시예들의 정확도가 높은 것을 알 수 있다.

실시예 3

데이터의 양을 4일, 6일, 8일 간격으로 입력하여 7일 이후의 결과 예측 정확도를 계산한 것을 제외하고는 상기 실시예 1과 동일하게 데이터의 양을 결정하였다.

실시예 4

데이터의 양을 4일, 6일, 8일 간격으로 입력하여 7일 이후의 결과 예측 정확도를 계산한 것을 제외하고는 상기 실시예 2와 동일하게 데이터의 양을 결정하였다.

비교예 3

7일 이후의 결과 예측 정확도를 계산한 것을 제외하고는 상기 비교예 1과 동일하게 실시하였다.

비교예 4

[0102] 7일 이후의 결과 예측 정확도를 계산한 것을 제외하고는 상기 비교예 2와 동일하게 실시하였다.

[0104] 실시예 3, 4 및 비교예 3, 4에서 계산한 예측 정확도를 하기 표 2에 나타내었다.

표 2

구 분	Input 변수의 개수	시간 지연 (Time Delay)	Train (NSE)	Test (NSE)
비교예 3	6	0	0.92	0.40
실시예 3	6	4	0.94	0.48
	6	6	0.87	0.62
	6	8	0.94	0.45
비교예 4	8	0	0.90	0.45
실시예 4	8	4	0.90	0.53
	8	6	0.92	0.32
	8	8	0.98	0.55

[0108] 비교예 3 및 비교예 4에서는 인공신경망을 사용하여 시간지연이 일어나지 않으므로 시간 지연 값이 0이다.

[0109] 6개의 데이터를 가지고 예측 시 시간 지연이 6일 때 가장 높은 정확도 0.62를 나타내었다.

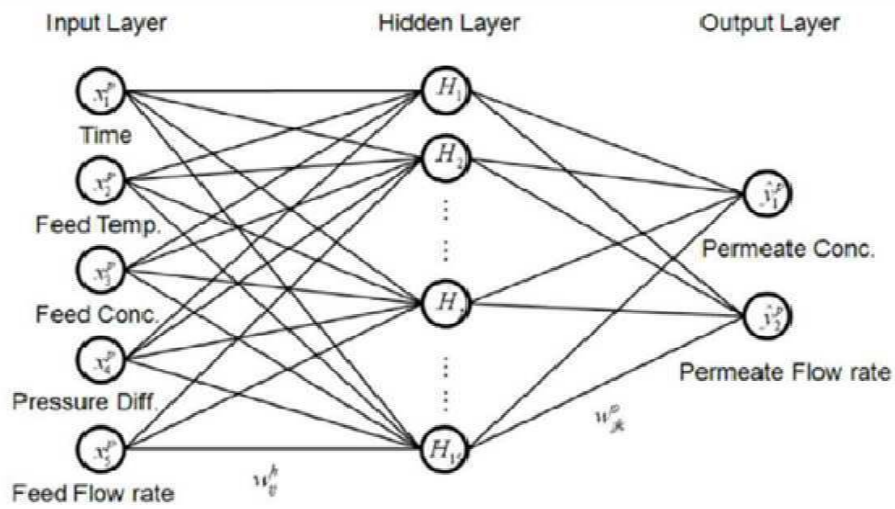
[0110] 8개의 데이터를 가지고 예측 시 시간 지연이 8일 때 가장 높은 정확도 0.55를 나타내었다.

[0112] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

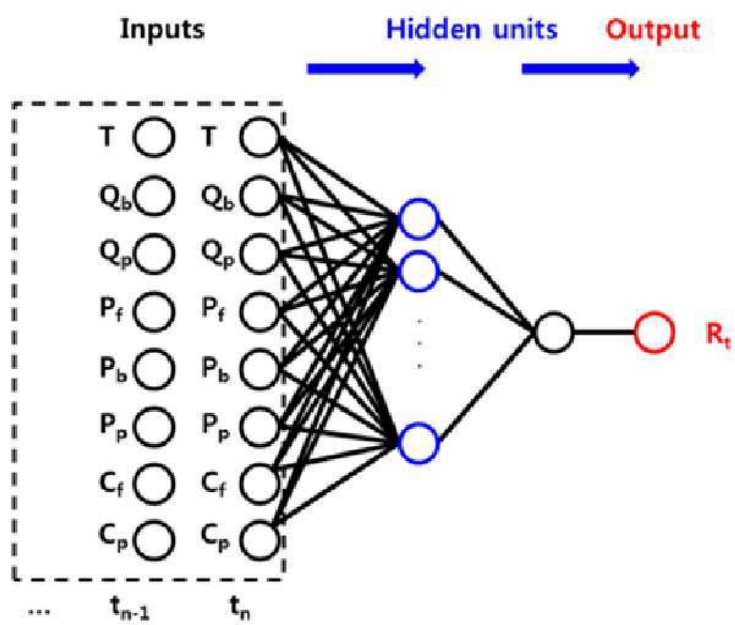
[0113] 본 발명의 범위는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

도면

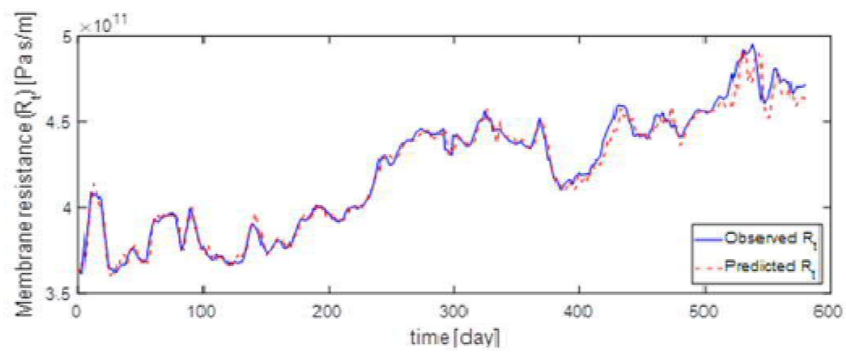
도면1



도면2



도면3



도면4

