



머신러닝을 활용한 세라믹 정밀여과 파일럿 플랜트의 파울링 조기 경보 방법

An early fouling alarm method for a ceramic microfiltration pilot plant using machine learning

탁도현 · 김동건 · 전종민 · 김수한*

Dohyun Tak · Dongkeon Kim · Jongmin Jeon · Suhan Kim*

부경대학교 지속가능공학부 토목공학전공

Major of Civil Engineering, Division of Sustainable Engineering, Pukyong National University

ABSTRACT

Fouling is an inevitable problem in membrane water treatment plant. It can be measured by trans-membrane pressure (TMP) in the constant flux operation, and chemical cleaning is carried out when TMP reaches a critical value. An early fouling alarm is defined as warning the critical TMP value appearance in advance. The alarming method was developed using one of machine learning algorithms, decision tree, and applied to a ceramic microfiltration (MF) pilot plant. First, the decision tree model that classifies the normal/abnormal state of the filtration cycle of the ceramic MF pilot plant was developed and it was then used to make the early fouling alarm method. The accuracy of the classification model was up to 96.2% and the time for the early warning was when abnormal cycles occurred three times in a row. The early fouling alarm can expect reaching a limit TMP in advance (e.g., 15-174 hours). By adopting TMP increasing rate and backwash efficiency as machine learning variables, the model accuracy and the reliability of the early fouling alarm method were increased, respectively.

Key words: Early fouling alarm method, Ceramic microfiltration (MF), Pilot plant, Machine learning, Decision tree

주제어: 파울링 조기경보 방법, 세라믹 정밀여과, 파일럿 플랜트, 머신러닝, 의사결정나무

Received 31 August 2023, revised 10 October 2023, accepted 11 October 2023.

*Corresponding author: Suhan Kim (E-mail: suhankim@pknu.co.kr; Fax: 82-51-629-6063, Tel. 82-51-629-6065)

탁도현 (석사과정) / Dohyun Tak (Master Student)

부산광역시 남구 용소로 45, 48513
45, Yongso-ro, Nam-gu, Busan 48513, Republic of Korea

김동건 (박사과정) / Dongkeon Kim (Ph.D. Student)

부산광역시 남구 용소로 45, 48513
45, Yongso-ro, Nam-gu, Busan 48513, Republic of Korea

전종민 (박사후연구원) / Jongmin Jeon (Postdoctoral Researcher)

부산광역시 남구 용소로 45, 48513
45, Yongso-ro, Nam-gu, Busan 48513, Republic of Korea

김수한 (교수) / Suhan Kim (Professor)

부산광역시 남구 용소로 45, 48513
45, Yongso-ro, Nam-gu, Busan 48513, Republic of Korea

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

pp. 253-259

pp. 261-269

pp. 271-279

pp. 281-288

1. 서론

막여과 공정은 반투과성 막을 이용해 대상 물질을 걸러내는 기법을 말한다. 막여과 정수시설은 기존 정수처리과정 대비 시설 집적화, 운전 자동화와 안정적인 생산수질 등의 장점을 갖고 있다.

막여과 플랜트 운영에서 가장 주의해야 할 점은 막 파울링(Fouling)이다(Kim, 2013). 일반적으로 막여과 정수 플랜트는 생산수 유량과 연동한 펌프 인버터 조절을 통해 정유량 운전을 실시하는데, 파울링이 발생하게 되면 막차압(TMP; Trans-membrane pressure)이 상승하게 된다 (Choi, 2007). 즉, 막차압 상승을 통해 파울링 정도를 계량할 수 있다.

막차압이 한계치 이상으로 상승하면 화학세척을 실시한다. 화학세척을 위해 약품을 준비해야 하고, 안정적인 수량 공급을 위해 막여과 플랜트 계열 별 운전 스케줄 조절을 해야 하기 때문에 막차압 모델 등을 통해 화학세척 시기를 미리 예측할 수 있으면 편리하다 (Kim, 2013; Kim et al., 2014; Kim et al., 2022). 그런데, 예기치 못한 급격한 차압 상승이 발생하는 경우를 예측하는 연구 사례는 거의 없는 실정이다. 본 연구에서는 머신러닝을 활용하여 급격한 파울링 발생이 일어나기 전에 경고를 할 수 있는 방법을 개발하고자 한다.

머신러닝 기술은 막여과 공정에서 파울링, 막차압 상승 예측에 활용되고 있다. Shetty and Chellam (2003)은 인공신경망(Artificial neural network; ANN) 알고리즘을 사용해 나노여과(Nanofiltration; NF) 운전 과정에서 파울링으로 인한 막차압의 변화를 93%의 정확도 및 5% 미만의 절대 상대오차로 예측하였다. Kim et al.(2021)은 ANN 알고리즘을 사용하여 정삼투 막모듈 실험 데이터의 플럭스를 예측, 분석하였다. Dalmau et al. (2013)에서는 Decision tree 알고리즘을 사용해 정밀여과(Microfiltration; MF) 파일럿 실험 데이터의 막차압 변화를 시계열 예측하여 0.96의 결정계수를 얻었다. Yoon et al.(2021)은 Long-short term memory(LSTM) 알고리즘을 사용해 막여과 운전 중 차압, 생산수질 등을 예측하였다. Kim et al.(2020)의 연구에서는 역삼투여과 실험에서 펌프의 진동 데이터를 분석 및 전처리하여 막의 파울링을 파악하였다. Schmitt et al. (2018)은 Anoxic-aerobic membrane bioreactor의 차압을 예측하기 위하여 ANN 알고리즘을 사용하였다.

본 연구에서는 앞서 소개된 머신러닝 알고리즘 중 이진 분류에 특화된 Decision tree를 사용하였고, 운전 데이터는 부산의 B-정수장 원수를 이용한 세라믹 정밀여과 파일럿 플랜트 데이터를 활용하였다(Dong-Ah Geological Engineering, 2012). 세라믹 정밀여과는 초기 설치비용은 높으나, 막의 내구성이 뛰어나 높은 압력과 투과유속으로 여과를 진행할 수 있다. 또한, 내화학성이 강해 세척 시 고농도의 화학약품을 사용할 수 있어 파울링으로부터의 회복률이 높은 특징을 갖고 있다(You et al., 2014; Park et al., 2015).

2. 연구 방법

2.1 세라믹 정밀여과 파일럿 플랜트 운영 및 데이터 수집

부산시 B-정수장 원수를 사용한 세라믹 정밀여과 공정 파일럿 플랜트에 사용된 막모듈 사양은 Table 1과 같다(Dong-Ah Geological Engineering, 2012). 파일럿 테스트는 2012년 2월 20일부터 2012년 7월 30일까지 약 5개월 진행되었으며, 180분의 정유량 여과(여과 플럭스 120 L/m²h)와 2.5분의 역세를 번갈아가며 운전되었다. 운전 기간 동안의 여과 cycle은 약 1,100회였고, 원수 수온은 3.9-25.2 °C, pH는 5.5 - 8.2, 탁도는 4.6 - 24.0 NTU였다.

정수장 원수는 A, B 2개 계열로 분리되어 각각 응집-혼화를 거친 후 여과된다. A, B 계열의 혼화기는 각각 back-mixer, inline mixer였다. 그리고, 두 계열 모두 유지 세정(Chemical enhanced backwash; CEB)을 1.5일 간격으로 시행하였다. 유지세정 약품으로는 황산

Table 1. Ceramic MF module specification

Membrane module specification	
Type	Ceramic Microfiltration
Manufacture	Metawater, Japan
Nominal pore size	0.1 μm
Length	0.1 m
Diameter	0.03 m
Channels per module	55
Module area	0.04 m ²
Configuration	Dead-end, inside out

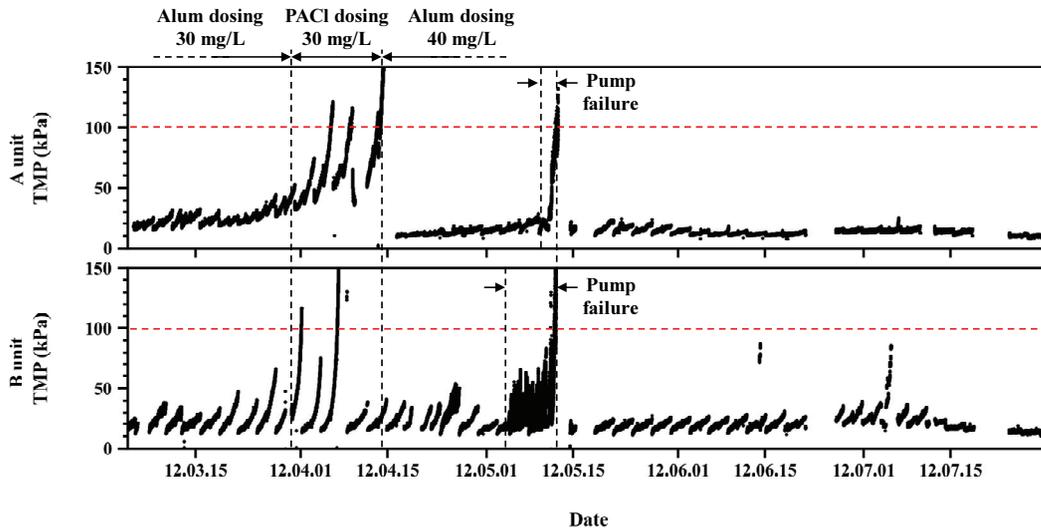


Fig. 1. Trans-membrane pressure during the pilot test period.

(H₂SO₄: 500 mg/L)과 차아염소산(NaOCl: 200 mg/L)을 번갈아 사용했고, 여과 공정을 멈춘 상태에서 막 모듈 안에 각 화학약품을 15분간 침지시킨 후 역세척하였다. 화학 세척(Clean in place; CIP)은 차압이 한계값(100 kPa)을 돌파했거나 예상치 못한 급격한 차압 상승이 일어났을 때 시행했는데, 황산 9,800 mg/L, 차아염소산 3,000 mg/L에 각 2시간 막모듈을 침지시킨 후 순환 세척 하였다.

Fig. 1은 운전 기간 동안의 온도보정차압 데이터를 나타내는데, 수온에 의한 차압 변화를 보정하여 파울링의 영향만을 보기 위해서이다 (Kim et al., 2022). 운전 초기에는 응집제로 Alum 30 mg/L를 사용했고, 이를 Poly Aluminum Chloride(PACl) 30 mg/L로 바꾸면서 급격한 차압 상승 등 운전이 어려움을 겪었다. 수차례의 화학세척(CIP) 후 응집제를 Alum 40 mg/L로 바꾸면서 운전이 안정을 찾았는데, 펌프 고장이 발생하면서 차압이 급격하게 상승하는 구간이 나타났다.

2.2 파울링 조기 경보를 위한 여과 cycle 특성변수

2.1절에 언급된 응집제 종류 및 농도 변경, 펌프 고장 등 특별한 이벤트 발생으로 차압이 급격히 상승하는 경우는 일반적인 막차압 예측 방법으로 파악하기 어렵다. 본 연구에서는 여과 cycle의 특성을 파악하여 정상 cycle과 비정상 cycle로 분류하는 방법을 통해 파울링 조기 경보를 하고자 한다. 비정상 cycle이 연속

적으로 발생하면 그 이후에 차압이 급격히 상승하여 한계 차압(본 연구에서는 100 kPa로 정의)에 도달할 가능성이 높다는 가설을 세우고, 2.1절에 설명된 세라믹 정밀여과 플랜트 데이터를 활용해 이 가설을 검증해보고자 한다. Table 2는 여과 cycle 특성변수를 정리한 것이다.

여과 cycle 특성 변수는 크게 기본 변수와 파생 변수로 나눌 수 있다. 기본 변수는 Table 2의 탁도(Turbidity), pH, 여과 지속시간(t), 초기 차압(P_i), 최종 차압(P_f), 초기 생산유량(Q_i), 최종 생산유량(Q_f)을 말한다. 각 기본 변수들은 계측기로부터 1분에 1개씩 측정되는데, cycle별로 1개의 대푯값을 추출한다. 탁도와 pH는 여과 지속 시간 동안 계측된 데이터들의 평균값

Table 2. Characteristic parameters for a membrane filtration cycle

Index	Description
Turbidity	Average turbidity (NTU)
pH	Average pH
Δt	Cycle duration (min)
P_i	Initial TMP (kPa)
P_f	Final TMP (kPa)
Q_i	Initial flow rate (L/min)
Q_f	Final flow rate (L/min)
BE	Backwash efficiency
$\Delta P/\Delta t$	TMP increasing rate (kPa/min)

을 대표값으로 사용하였다. 유량과 차압의 경우 여과 시작 시점과 종료 시점의 값을 사용하는 것이 원칙이지만 펌프가 켜지거나 꺼지는 도중에 계측되는 경우의 값은 신뢰할 수 없기 때문에 cycle 내 전체 데이터 중 앞 뒤 2개씩의 데이터를 제외한 시작 및 종료 시점의 값을 사용하였다.

파생 변수는 Table 2의 막차압 상승속도($\Delta P/\Delta t$), 역세척 효율(Backwash efficiency; BE)인데, 이들은 장기적인 막차압 상승과 연관성이 높은 변수이다 (Kim, 2013). BE는 식 (1)을 통해 구해지며, $n-1$ 번째 시작 시점의 막차압($P_{n-1,i}$)과 n 번째 시작 시점의 막차압($P_{n,i}$) 간의 비율을 의미한다.

$$BE = \frac{P_{n,i}}{P_{n-1,i}} \quad (1)$$

막차압 상승속도는 식 (2)을 통해 구해지며 P_i 는 각 cycle 여과 시작 시점의 막차압, P_f 는 각 cycle 여과 종료 시점의 막차압, t 는 cycle의 여과 지속 시간을 의미한다.

$$\frac{\Delta P}{\Delta t} = \frac{P_f - P_i}{\Delta t} \quad (2)$$

2.3 머신러닝 적용 방법

머신러닝 모델은 학습변수에 따라 성능이 크게 달라질 수 있어 적절한 학습변수를 선택하는 것이 중요하다. 일반적으로 예측 대상 변수와 후보 변수들의 관계를 분석하고 이를 기반으로 학습변수를 선택하게 된다 (Pi et al., 2019). 이러한 변수 선택 과정은 모델의 일반화 능력을 향상시키고, 과적합을 방지하는 데 중요한 역할을 한다. 2.2절에 설명된 파생 변수인 BE와 막차압 상승속도는 기본 변수인 막차압, 여과 지속 시간을 사용해 계산되는 일종의 종속 변수이다. 종속 변수는 해당 변수와 상관성이 높은 독립 변수를 학습할 시 머신러닝의 학습변수로 사용되지 않더라도 반영될 가능성이 있으며, 모델의 복잡성을 줄이기 위해 제외하는 것이 타당할 수도 있다 (Blum and Langlay, 1997; Hall and smith, 1998).

이를 확인하기 위해 학습변수를 기본 변수로만 사용한 Model 1과 막차압 상승에 직접적인 영향을 주는

파생 변수와 기본 변수를 모두 학습변수로 사용한 Model 2를 구분하여 지도학습을 실시하였다. 지도학습을 위해서는 여과 cycle을 정상/비정상으로 미리 구분해야 한다. 전체 운전 기간의 막차압 데이터 및 사이클 특성 데이터를 분석하여, 주변 cycle 대비 막차압 상승속도가 높거나, 역세척이 잘 안되었거나, 또는, 한계 차압까지 도달한 시점과 가까운 시기의 여과 cycle들을 비정상적으로 분류하여 지도학습의 정답지로 활용하였다. 이에 대한 구체적인 예시로 PACl 30 mg/L를 응집제로 사용한 기간, 펌프 고장 기간의 cycle 데이터가 있다 (Fig. 1). 분류 결과 1100개의 cycle 데이터 중 A 계열은 71개, B 계열은 150개의 데이터가 비정상적으로 분류되었다.

Decision Tree 알고리즘은 여러 입력변수의 동질성에 기반하여 분류 혹은 예측 모델을 개발하는 데 널리 사용되는 데이터 마이닝 기법이다. Decision Tree 알고리즘은 시작 지점인 뿌리(Root) 노드, 변수를 상호 배타적인 두 개의 하위 집합으로 분할하는 내부(Internal)노드, 분할의 최종 결과를 표현하는 잎(Leaf)노드, 각 노드에서 일어난 분할의 결과를 나타내는 가지(Branch)로 구성된 나무 모양의 구조(Tree)를 형성한다 (Fig. 2). 그리고 최종적으로 구축된 트리의 뿌리노드에서 가장 하위의 잎 노드까지의 길이를 깊이(Depth)라고 한다.

모델이 구축될 때, 분할은 변수의 순도에 근거하여 사전 결정된 순도를 만족하는 하위 노드를 형성할 때까지 계속된다. 대부분의 경우, 모든 잠재적 입력변수가 모델을 구축할 때 사용되는 것은 아니며 특정 변수가 여러 번 사용될 수도 있다 (Song and Lu, 2015).

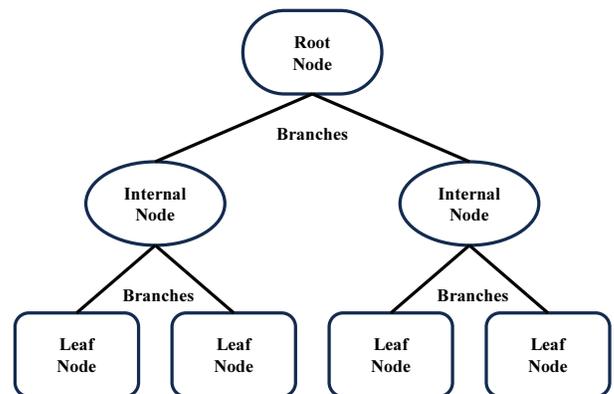


Fig. 2. Schematic of the decision tree algorithm.



본 연구에 사용한 CART(Classification And Regression Trees) 알고리즘은 데이터 셋의 패턴을 파악하고 예측을 수행하기 위한 노드 분할 기준으로 Gini 불순도를 사용하며 식 (3)을 통해 구해진다 (Jang and Park, 2021).

$$G_i = 1 - \sum_{k=1}^n p_{i,k}^2 \quad (3)$$

$p_{i,k}$ 는 노드 i에 배치된 데이터 중 클래스 k의 확률분포이다. Gini 불순도는 노드에 배치된 데이터의 클래스가 균등하게 분포된 정도를 의미하는데 0일 경우 노드가 하나의 클래스만 할당되었음을, 1일 경우 모든 클래스가 동일한 분포로 할당되었음을 의미한다 (Ko, 2011). 모델은 학습 과정에서 최종적으로 Gini 불순도가 0 혹은 그에 근접한 잎 노드를 생성할 때까지 뿌리 노드로부터 분할해 나간다.

2.4 파울링 조기 경보 시뮬레이션

본 연구에서 개발된 파울링 조기 경보 방법이 적용되기 위해서는 2.2, 2.3절에서 설명된 학습 데이터가 반드시 필요하다. 그러나 실제 플랜트 운전 시에는 초기에는 학습 데이터가 없거나 부족하다가, 운전이 계속될수록 늘어나게 된다. 이러한 상황을 모사하기 위해, 파울렛 플랜트 전체 여과 cycle 1,100여개를 100개 단위로 나누어 단계적으로 학습하여 모델의 정확도 향상 여부를 확인하였다. 즉, 첫 100개 cycle 데이터가 수집되는 동안은 학습만 하고, 그 다음 100개 cycle 데이터에 대해서는 이전 100개 cycle 데이터로 만들어진 모델의 정확도를 검증하면서 누적 200개의 cycle 데이터를 이용해 모델을 개선하였고, 이 과정을 여과 cycle 수가 100개 단위로 늘어남에 따라 반복적으로 수행하였다.

모델의 정상/비정상 분류 성능을 분석하기 위한 지표로 오차 행렬(Error Matrix)을 사용하였다. 실제 양성 값을 양성으로 예측한 경우 True Positive(TP), 실제 음성 값을 음성으로 예측한 경우는 True Negative(TN), 실제 양성 값을 음성으로 예측한 경우 False Negative(FN), 그리고 실제 음성 값을 양성으로 잘못 예측한 경우 False Positive(FP), 네 가지 경우의 수로 분류한다. 오차 행렬을 통해 계산할 수 있는 지표는

정확도(Accuracy) 재현율(Recall), 특이도(Sensitivity) 등이 있다. 정확도는 식 (4)을 통해 구해지며 예측의 정확도를 의미한다.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

재현율은 식 (5)를 통해 구해지며 실제 positive를 positive로 예측한 비율을 의미한다.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

특이도는 식 (6)를 통해 구해지며 실제 negative를 negative로 예측한 비율을 의미한다.

$$\text{Sensitivity} = \frac{TN}{TN + FP} \quad (6)$$

본 연구에서는 정상 데이터를 negative, 비정상 데이터를 positive로 분류하였다. 실제 막여과 시설 운영에서 비정상을 정상으로 예측하는 경우, 즉 False negative의 경우는 급격한 차압 상승을 미리 발견하여 화학세척 시기를 놓칠 수 있는 등 막여과 플랜트 운전에서 심각한 차질을 불러일으킬 수 있다. 따라서 재현율을 높이는 방향(즉, False negative를 최소화)으로 모델을 개발하였다.

3. 결과 및 고찰

3.1 모델 파라미터 최적화

과적합은 학습 데이터에 대한 예측 정확도에 비해 학습되지 않은 데이터, 즉 테스트 데이터에 대한 예측 정확도가 떨어지는 경우를 말한다. 모델이 지나치게 복잡하거나 단순한 경우, 적절한 학습변수를 사용하지 않은 경우, 학습 데이터의 개수가 적은 경우 등의 상황에 발생하게 되는데 과적합이 일어난 모델은 일반적인 데이터에 대한 예측을 수행하기 힘들다 (Song and Lu, 2015). Decision tree 모델의 복잡도는 모델의 깊이가 깊어질수록 커지는데 학습 과정에서 최대 깊이(Max depth) 파라미터를 통하여 조절할 수 있다. 모든 학습변수를 사용한 Model 2의 최대 깊이를 1부터 15까지 변경하며 전체 데이터셋의 70%(770개)를 모델

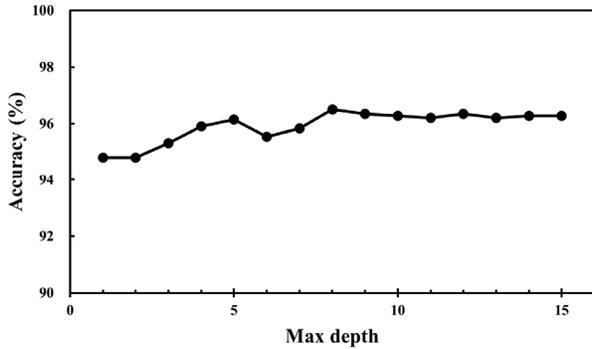


Fig. 3. Accuracy of model prediction by max depth.

에 학습시킨 후, 나머지 30%(330개)를 예측하게 하여 정확도를 구하였다 (Fig. 3).

예측 결과, 최대 깊이가 8일 때 예측 정확도는 96.5%로 가장 높았으나 그 이상 깊이가 증가할수록 예측 정확도가 떨어지는 경향을 볼 수 있다. 이는 모델이 지나치게 복잡하게 되어 과적합이 일어났음을 의미한다. 또한 최대 깊이 1~5 범위에서는 예측 정확도가 증가하다가(최대 깊이 5일 때 96.2%) 최대 깊이 6에서는 떨어진 것으로 보아 이미 최대 깊이 6에서 과적합이 일어났을 가능성이 높다. 최대 깊이 8에서도 변곡점이 관찰되나, 최대 깊이 5일 경우에 비해 과적합의 위험이 크다. 따라서, 본 연구의 decision tree 분류 모델의 최대 깊이를 5로 설정하였다.

3.2 여과 cycle 정상/비정상 예측 정확도 분석

Fig. 4는 Model 1, Model 2의 cycle 구간별 정상/비정상 예측 정확도를 나타낸다. Model 1의 초기 정확도는 A 계열 36%, B 계열 14%였으며, Model 2의 경우 A 계열 87%, B 계열 69%로 비교적 Model 1보다 정확도가 높았다. 학습 데이터가 누적될수록 두 모델의 정확도는 향상되어 최종적으로 100%의 정확도를 달성했다.

Table 3는 정답지의 정상 건수(Exp_0)와 비정상 건수(Exp_1), 예측 결과의 정상 건수(Pre_0)와 비정상 건수(Pre_1)를 사용한 오차행렬이다. 오차 행렬 확인 결과, cycle No. 700 이후부터 비정상 cycle의 개수는 8개로 비정상 cycle 개수가 220개인 cycle No. 700 이전 대비 비정상 cycle의 개수가 적었다. 데이터 개수가 너무 적으면 데이터 불균형 현상 때문에 정확도가 높아질 수 있는데, 모델의 정확도가 비정상적으로 높은 이

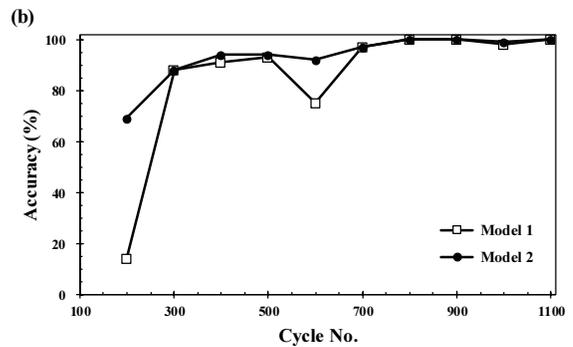
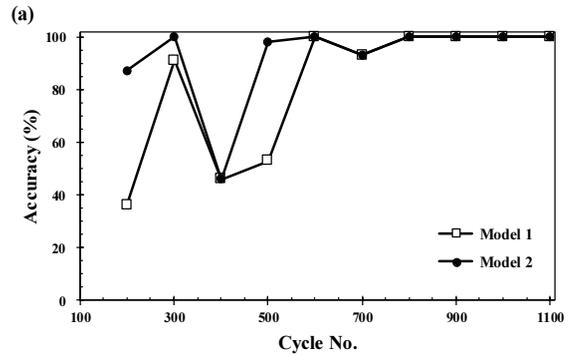


Fig. 4. Prediction accuracy of Models 1 and 2 for (a) A unit and (b) B unit.

유가 데이터 불균형 때문인 것으로 사료된다.

전체 데이터를 다 사용한 Model 1의 재현율과 특이도는 각각 72.1%, 84.6%로 계산되었으며, Model 2의 재현율과 특이도는 각각 85.0%, 93.5%로 계산되었다. 즉, Model 2가 Model 1보다 false negative와 false positive의 수치가 낮음을 의미한다.

기본 변수만 사용한 Model 1에 비하여 파생 변수도 함께 학습한 Model 2의 초기 정확도, 데이터 누적에 따른 정확도 향상, 재현율과 특이도 모두 우수하였다. 특히 두 모델의 초기 예측 정확도의 차이가 큰 편이었는데, 파생 변수가 있는 경우 데이터 개수가 동일하더라도 더 높은 성능의 모델을 만들 수 있다는 것을 의미한다.

3.3 파울링 조기경보 성능 평가

모델의 예측(정상/비정상 분류) 정확도는 전체 데이터 대상 최대 96.2%였지만 여과 cycle의 정상/비정상 분류만으로는 파울링 조기 경보는 불가하였다. 예를 들어, 모델에 의해 여과 cycle이 비정상적으로 판별된 경우가 3회 미만 연속으로 발생한 경우에는, 그 이후에



Table 3. Error matrix of model prediction results

Cycle No.		Model 1				Model 2			
		A unit		B unit		A unit		B unit	
		Pre_0	Pre_1	Pre_0	Pre_1	Pre_0	Pre_1	Pre_0	Pre_1
100-200	Exp_0	36	64	13	86	87	13	68	31
	Exp_1	0	0	0	1	0	0	0	1
200-300	Exp_0	91	9	80	9	100	0	88	1
	Exp_1	0	0	3	8	0	0	11	0
300-400	Exp_0	26	53	74	4	26	53	76	2
	Exp_1	1	20	5	17	1	20	4	18
400-500	Exp_0	43	47	82	2	88	2	79	5
	Exp_1	0	10	5	11	0	10	1	15
500-600	Exp_0	100	0	60	1	100	0	55	6
	Exp_1	0	0	24	15	0	0	2	37
600-700	Exp_0	88	5	85	3	88	5	85	3
	Exp_1	2	5	0	12	2	5	0	12
700-800	Exp_0	100	0	100	0	100	0	100	0
	Exp_1	0	0	0	0	0	0	0	0
800-900	Exp_0	100	0	99	0	100	0	99	0
	Exp_1	0	0	0	1	0	0	0	1
900-1000	Exp_0	100	0	93	0	100	0	93	0
	Exp_1	0	0	2	5	0	0	1	6
1000-1100	Exp_0	100	0	100	0	100	0	100	0
	Exp_1	0	0	0	0	0	0	0	0

여과 cycle이 정상으로 돌아와 급격한 차압 상승 이후 한계 차압(100 kPa)을 돌파하는 경우가 드물었다. 즉, 여과 cycle이 비정상인 경우가 한번이나 두 번 연속 나왔을 때 경보를 올린다면 이는 오보가 될 것이다.

모델을 통해 여과 사이클의 정상/비정상 분류 패턴을 검토한 결과, 비정상 cycle이 3회 연속으로 발생하는 조기 경고 규칙을 설정한다면 보다 신뢰성 높은 파울링 조기 경보가 될 수 있음을 발견하였다. Fig. 5는 A, B 계열의 온도보정차압과 파울링 조기 경보를 함께 나타낸 것이며 오보는 붉은색, 경보는 푸른색으로 표시하였다. 3.2절에 논의된 바와 같이 Model 1은 분류 정확도도 떨어지지만, 이를 파울링 조기 경보에 적용했을 때 잦은 오보가 발생하므로 파울링 조기 경고 신뢰도가 낮다.

한편, Model 2를 이용한 파울링 조기 경고에서도 오보가 적지 않게 발생했지만, 이는 주로 학습 데이터가 부족했던 운전 초기에 집중적으로 발생했다. A계

열의 경우 첫 번째 한계 차압 돌파(Fig. 5(a)에서 ①) 이전에 4회 오보가 있었지만, 그 이후에는 경보를 올릴 때마다 한계 차압 돌파(Fig. 5(a)에서 ②, ③, ④)가 발생하여 파울링 조기 경고 신뢰도가 높았다. B 계열의 경우 A 계열의 경우보다 적었지만 초기에 2회 오보가 있었다. B 계열에서 첫 번째와 두 번째 한계 차압 돌파 사이(Fig. 5(b)에서 ①과 ② 사이)에 경고 이후 한계 차압까지 상승하지 않아 오보인 것처럼 볼 수도 있지만, 한계 차압까지 상승하지 않은 이유가 화학세척을 시행했기 때문이므로 오보라 볼 수 없다. 그리고, 2012년 4월 25일경 두 차례 경보가 있었는데, 경고 발생 이후 유지 세정으로 차압이 잘 회복되었기 때문에, 이는 오보로 판명하는 것이 타당하다. 결과적으로 차압상승률과 역세회복률 등의 파생 변수를 적용한 모델을 활용한 파울링 조기 경고 방법의 신뢰도가 높다.

pp. 253-259

pp. 261-269

pp. 271-279

pp. 281-288

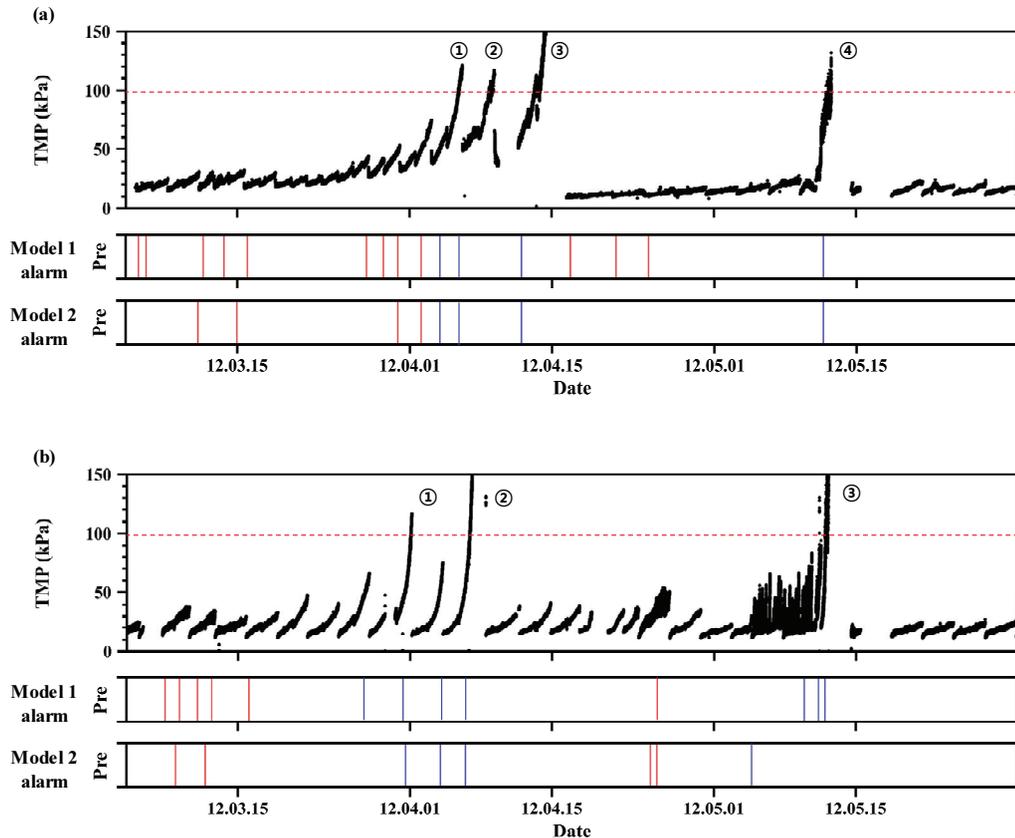


Fig. 5. Early fouling alarm using Models 1 and 2 for (a) A unit and (b) B unit.

4. 결 론

본 연구에서는 세라믹 정밀여과 파일럿 플랜트 운전 데이터를 활용하여 여과 cycle의 정상/비정상 상태를 분류하는 decision tree 모델 및 이를 활용한 파울링 자동경보 방법을 개발하였다. 분류 모델의 정확도는 전체 데이터 대상 최대 96.2%였고, 비정상 cycle이 3회 연속으로 발생하는 조기 경보 규칙을 설정하여 한계 차압(100 kPa) 도달 전 15-174시간 전에 한계 차압 상승 경보(즉, 조기 파울링 경보)를 울리는 데 성공하였다. 또한, 실제 플랜트 운전 초기에 학습 데이터가 충분하지 못한 상황을 모사하여, 100개의 학습 데이터가 추가될 때마다 모델을 업데이트하여 예측 성능 및 파울링 조기 경보 신뢰도가 향상되는 것도 확인하였다. 따라서, 본 연구에서 개발된 사이클 분류 모델 및 파울링 조기 경보 방법이 실제 플랜트에 성공적으로 적용될 수 있음을 기대할 수 있다.

본 연구에서 개발된 파울링 조기 경보 방법은 근

미래(예: 15-174시간 후)에 한계 차압을 돌파할 수 있다는 경고를 할 수 있지만, 정확하게 언제 돌파하는지에 대한 정보를 주지 못한다. 그 이유는 5개월간의 세라믹 정밀여과 파일럿 플랜트 운전 동안 한계 차압 돌파 이벤트가 A계열 4번, B계열 3번만 발생했기 때문에 이벤트 발생 시기에 대한 예측 모델을 만들기 위한 충분한 데이터가 없었기 때문이다. 향후 더 많은 데이터를 수집하여 시계열(Time-series) 분석, 교차 검증(Cross validation) 기법 등 예측 모델의 성능 개선을 위한 연구가 필요하다.

사 사

이 논문은 부경대학교 자율창의학술연구비(2023년)에 의하여 연구되었음



References

- Blum, A.L., and Langley, P. (1997). Selection of relevant features and examples in machine learning, *Artif. Intell.*, 97, 245 - 271.
- Choi, B. (2007). Influence of conditions for pre-treatment, aeration intensity and methods of backwash on micro-filter fouling in water treatment, Master's Thesis, Yonsei University.
- Dalmau, M., Rodriguez-Roda, I., Ayesa, E., Odriozola, J., Sancho, L., and Comas, J. (2013). Development of a decision tree for the integrated operation of nutrient removal MBRs based on simulation studies and expert knowledge, *J. Chem. Eng.*, 217, 174-184.
- Dong-Ah Geological Engineering (2012). Assessment of applicability by ceramic membrane filtration process at Beomeosa water treatment plant, Research report.
- Hall, M.A., and Smith, L.A. (1998). Practical feature subset selection for machine learning, *Comput. Sci.*, 98, 181 - 191.
- Han, S.W., Lee, C.W., and Kang, L.S. (2003). Improvement of the effectiveness of drinking water treatment using a mixture of polyamine and PACl, *Korean Chem. Eng. Res.*, 41(3), 319-325.
- Jang, D. and Park, M. (2021). Art price prediction using decision tree-based machine learning methods, *Korean Manag. Rev.*, 50(2), 357-381.
- Kim, D., Kim, N., Jeon, J., Kim, J., Lim, J., and Kim S. (2022). Reconsidering clean-in-place criterion for low pressure membrane filtration systems using a model verified by long-term pilot plant operation data, *J. Water Process. Eng.*, 46, 102506.
- Kim, J.Y., Jeon, J., Kim, N., and Kim, S. (2021). Predicting flux of forward osmosis membrane module using deep learning, *J. Korean Soc. Water Wastewater*, 35(1), 93-100.
- Kim, M., Kim, N., Jeon, J., and Kim, S. (2020). Vibration signals at the early stage of fouling in reverse osmosis system, *Desalination Water Treat.*, 183, 81-87.
- Kim, S. (2013). A process diagnosis method for membrane water treatment plant using a constant flux membrane fouling model, *J. Korean Soc. Water Wastewater*, 27(1), 139-146.
- Kim, S., Lim, J., Park, J.Y., and Kim, J.O. (2014). Effect of flux fluctuation on the fouling in membrane water treatment system for smart water grid, *Desalination Water Treat.*, 52, 1028-1034.
- Ko, Y.S. (2011). The Construction methodology of a rule-based expert system using CART-based decision tree method, *J. Korea Inst. Electron. Commun. Sci.*, 6(6), 849-854.
- Park, S.H., Park, Y.G., Lim, J., and Kim, S. (2015). Evaluation of ceramic membrane applications for water treatment plants with a life cycle cost analysis, *Desalination Water Treat.*, 54, 973 - 979.
- Pi, M.G., Shin, I.H., and Min, O.G. (2019). Visual analytics system to help feature selection of machine learning, *Korea Software Congress 2019, Pyeongchang, South Korea*.
- Schmitt, F., Banu, R., Yeom, I.T., and Do, K.U. (2018). Development of artificial neural networks to predict membrane fouling in an anoxic-aerobic membrane bioreactor treating domestic wastewater, *Biochem. Eng. J.*, 133, 47 - 58.
- Shetty, G.R., and Chellam, S. (2003). Predicting membrane fouling during municipal drinking water nanofiltration using artificial neural networks, *J. Membr. Sci.*, 217(1), 69 - 86.
- Song, Y.Y. and Lu, Y. (2015). Decision tree methods: applications for classification and prediction, *Shanghai Arch. Psychiatry*, 27(2), 130-135.
- Yoon, N., Kim, J., Lim, J.L., Abbas, A., Jeong, K., and Cho, K.H. (2021). Dual-stage attention-based LSTM for simulating performance of brackish water treatment plant, *Desalination*, 512, 115107.
- You, S.J., Ahan, H.W., Park, S.H., Lim, J., Hong, S.C., and Lee, B.I. (2014). The study on optimum operation conditions of ceramic MF membrane process in Y water treatment plant, *Membr.*, 24, 201-212.

pp. 253-259

pp. 261-269

pp. 271-279

pp. 281-288