Prediction of permeate water flux in forward osmosis desalination system using tree-based ensemble machine learning models

Yinseo Song[a], Jeongwoo Moon[b], Joon Ha Kim[b], Kiho Park[a],*

[a]*School of Chemical Engineering, Chonnam National University, 77 Yongbong-ro, Buk-gu, Gwangju 61186, Republic of Korea, Tel. +82 (0)62 530 1878; email: kiho138@jnu.ac.kr (K. Park)*
[b]*School of Earth Sciences and Environmental Engineering, Gwangju Institute of Science and Technology (GIST), Gwangju 61005, Republic of Korea*

ABSTRACT

This study developed tree-based ensemble machine learning models to predict forward osmosis (FO) permeate flux using extreme gradient boosting (XGBoost) and light-gradient boosting machine (LGBM) methods. The models were trained by approximately 700 data points from the FO experimental data. The results showed that LGBM and XGBoost could predict the FO permeate flux with very high accuracy (>0.95 of R^2) in the test set. Feature analysis using Shapley additive explanations values was performed to identify the influences of input variables on the model output and the correlation between the input variables. The results revealed that water permeability and pressure difference have the most significant variables on the FO permeate flux. The correlation between the operating conditions and water permeability cannot be neglected. In this study, we clarified the applicability of ensemble machine learning models for FO systems and suggested directions for future data collection.

Keywords: Forward osmosis; Machine learning models

* Corresponding author.