

기계학습을 이용한 국내산 플라이애시의 반응성 예측

The Prediction of Reactivity of Domestic Fly Ashes
Using Machine Learning

박 우 영* 문 주 혁**
 Park, Wooyoung Moon, Juhyuk

ABSTRACT

In this study, a machine learning model was built from domestic fly ashes' chemical compositions (wt%) and sum of amorphous aluminosilicate(wt%) calculated by X-ray fluorescence(XRF) analysis and quantitative X-ray diffraction(QXRD) analysis. It is the machine learning(ML) model that predicts fly ashes' sum of amorphous aluminosilicate from their chemical compositions. An ensemble technique is applied to the ML model that is re-trained by a meta model. The meta model is the combined model of Boosting, Random Forest and multi-layer perceptron(MLP). The metric of performance of the ML model is a coefficient of determination, the coefficient of determination of train set is 0.747 and the coefficient of determination of test set is 0.340.

요 약

본 연구에서는 X-선 형광 분석법(XRF)과 정량 X-선 회절 분석법(QXRD)을 통해 계산한 국내산 플라이애시의 화학 조성값(wt%)과 비정질 알루미늄실리케이트 함량(wt%)에 대한 기계학습 모델을 구축하였다. 플라이애시의 화학 조성값으로 그것의 비정질 알루미늄실리케이트 함량을 예측하는 모델이다. 해당 기계학습 모델에서는 부스팅, 랜덤 포레스트, 다층 퍼셉트론(MLP) 모델을 합친 메타모델로 재학습시키는 앙상블 기법을 사용하였다. 모델의 정확도는 결정계수로 나타냈으며, 학습 데이터의 결정계수는 0.747, 테스트 데이터의 결정계수는 0.340이다.

1. 서 론

플라이애시는 비정질 알루미늄실리케이트 함량이 높을수록 반응성이 높기 때문에 콘크리트 품질을 예측하기 위해서는 플라이애시의 비정질 알루미늄실리케이트 함량을 아는 것이 중요하다. 하지만, 국내 화력발전소에서 생산된 플라이애시의 정량적 x-선 회절 분석에 관한 연구가 부족한 실정이다.

2. 사용 재료 및 실험 방법

본 연구의 기계학습 모델에서 사용된 데이터는 62종의 국내산 플라이애시로, 17개 국내의 논문의 데이터를 활용하였다. 플라이애시의 화학조성값 (1)Al₂O₃, (2)CaO, (3)Fe₂O₃, (4)SiO₂, (5)MgO, (6)Na₂O+0.658K₂O 을 입력값으로 받아, 기계학습 모델의 출력값인 비정질 알루미늄실리케이트 함량을 예측하였다. 62개의 데이터는 계층적 샘플링을 통해 훈련 데이터 49개, 테스트 데이터 13개로 나누었다.

* 정회원, 서울대학교, 멀티스케일구조재료연구실, 석사과정, uyoung@snu.ac.kr

** 정회원, 서울대학교, 멀티스케일구조재료연구실, 부교수, juhyukmoon@snu.ac.kr

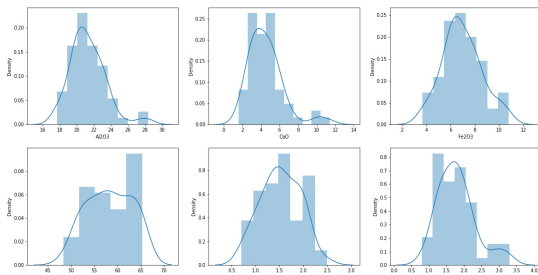


그림 1. 플라이애시 62종의 데이터 분포

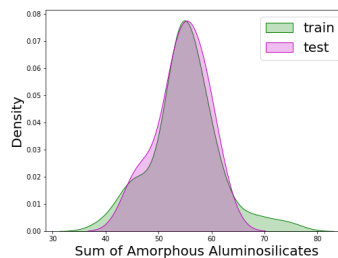


그림 2. 출력값의 분포에 따른 데이터 샘플링

3. 실험결과 및 고찰

그림 3은 기반모델인 부스팅, 랜덤 포레스트, 다층 퍼셉트론으로 각 모델에서 예측한 비정질 알루미늄노실리케이트 값과 실제값간의 분포를 나타낸다. 그림 4는 3개의 기반모델을 합친 메타모델에 의해 예측한 값과 실제값간의 분포를 나타낸다. 실제값-예측값이 $y=x$ 선 위에 모두 분포할 때 완벽한 모델이 구축되었을 의미하므로, 해당 기계학습 모델은 정확한 비정질 알루미늄노실리케이트 함량 값을 예측하지 못함을 뜻한다.

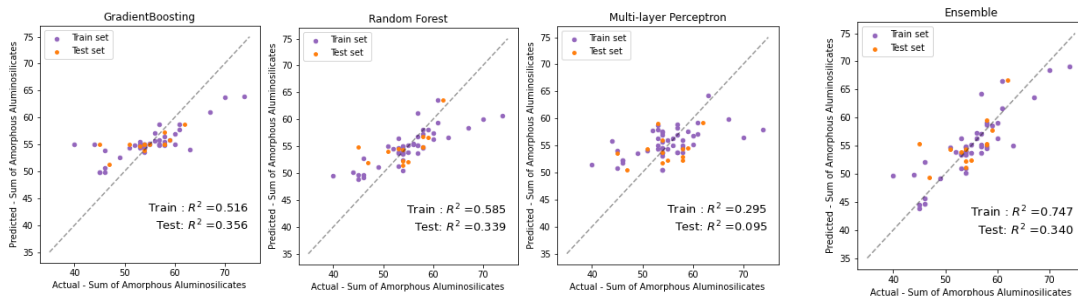


그림 3. 각 3개의 기반모델에서 예측한 비정질 알루미늄노실리케이트 그림 4. 메타모델에서 예측한 비정질 알루미늄노실리케이트

4. 결 론

양상불 기법을 이용해 플라이애시의 화학 조성에 따른 반응성을 예측하고자 하였다. 플라이애시의 반응성을 예측하기 위해 비정질 알루미늄노실리케이트 함량을 예측하는 기계학습 모델을 구축하였다. 다중 기계학습 모델을 결합한 메타 모델로 학습하였음에도 불구하고 높은 성능을 보이지 못했다. 정량 X-선 회절 분석법이 사용된 국내산 플라이애시의 데이터 수가 충분하지 않아, 훈련 데이터에 과적합(over-fitting) 되어 테스트 데이터에서 일반화된 성능을 보이지 못한 것으로 보인다.

감사의 글

본 연구는 한국연구재단의 신진연구지원 사업(NRF-2021 R1A2C400194412)의 지원을 받아 수행되었음.

참고문헌

1. Song, Y., Yang, K., Chen, J., Wang, Xaizin., Sant, G., & Bauchy, M. (2021). Machine Learning Enables Rapid Screening of Reactive Fly Ashes Based on Their Network Topology. *ACS Sustainable Chemistry & Engineering*, 2021 9 (7), 2639-2650