

광업 데이터의 시계열 분석을 통해 실리카 농도를 예측하기 위한 머신러닝 모델

이승훈* · 윤연아** · 정진형** · 심현수** · 장태우* · 김용수*†

*경기대학교 산업경영공학과

**경기대학교 일반대학원 산업경영공학과

A Machine Learning Model for Predicting Silica Concentrations through Time Series Analysis of Mining Data

Lee, Seung Hoon* · Yoon, Yeon Ah** · Jung, Jin Hyeong** · Sim, Hyun su** · Chang, Tai-Woo*
· Kim, Yong Soo*†

*Department of Industrial and Management Engineering, Kyonggi University

**Department of Industrial and Management Engineering, Kyonggi University Graduate School

ABSTRACT

Purpose: The purpose of this study was to devise an accurate machine learning model for predicting silica concentrations following the addition of impurities, through time series analysis of mining data.

Methods: The mining data were preprocessed and subjected to time series analysis using the machine learning model. Through correlation analysis, valid variables were selected and meaningless variables were excluded. To reflect changes over time, dependent variables at baseline were treated as independent variables at later time points. The relationship between independent variables and the dependent variable after n point was subjected to Pearson correlation analysis.

Results: The correlation (R^2) was strongest after 3 hours, which was adopted as a dependent variable. According to root mean square error (RMSE) data, the proposed method was superior to the other machine learning methods. The XGboost algorithm showed the best predictive performance.

Conclusion: This study is important given the current lack of machine learning studies pertaining to the domestic mining industry. In addition, using time series analysis in mining data will show further improvement. Before establishing a predictive model for the proposed method, predictions should be made using data with time series characteristics. After doing this work, it should also improve prediction accuracy in other domains.

Key Words: Mining Data, Froth Flotation, Time Series Forecasting, Data Reconstruction

● Received 13 september 2020, 1st revised 16 september 2020, accepted 21 september 2020

† Corresponding Author(kimys@kgu.ac.kr)

© 2020, Korean Society for Quality Management

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.

※ 본 논문은 2020년도 경기대학교 대학원 연구원장학생 장학금 지원에 의하여 수행되었음.

1. 연구 배경 및 목적

최근 4차 산업혁명으로 인해 인공지능, IoT(Internet of things), 전기자동차 등과 관련된 핵심부품 제조에 사용되는 광물 수요가 급증하고 있으며(Jang, 2019), 광물의 품질관리에 대한 중요성이 대두되고 있다. 또한, 데이터 기술 처리 기술이 고도화됨에 따라 데이터의 활용성이 증가하여 빅데이터에 관한 관심이 점차 높아지고 있다(Lee et al. 2016). 그러나 빅데이터 분석 방법 중 하나인 머신러닝을 국내 광업 분야에서 도입한 사례를 찾아보기 어렵고, 기초 연구도 부족한 상황이다. 따라서 머신러닝 기술을 국내 광업 프로세스에 도입하기 위해서는 기존의 활용사례들을 조사하고 분석할 필요가 있다(Lee et al., 2019).

광업이란 지하 또는 지표상의 원석, 광맥, 광맥 층으로부터 가치 있는 광물을 가공하고 만드는 모든 산업 활동을 말하며, 광업 프로세스는 원하는 광물을 얻기 위한 일련의 공정이다. 특히 선탄광은 광물을 유용하게 이용하기 위해 주로 물리적, 화학적으로 목적 광물을 다른 광물과 분리해 유용 광물의 품위를 높이며 유해성분을 미리 제거하는 작업으로서 채광과 제련의 중간 공정으로 종류는 비중선광, 자력선광, 부유선광 등이 있다. 부유선광은 광물 표면이 가지는 특유한 성질 차이를 활용하여 비중과는 상관없이 특정 광물들을 물 위에 띄우고 나머지 광물들은 물속에 남겨서로 분리하는 방법이다. 즉, 공기를 광액 속에 주입하여 소수성 광물 알갱이가 거품에 붙어 수면에 떠오르게 하고, 친수성 광물 알갱이는 물속에 가라앉게 하는 것이다.

실무에서 관리, 야금 및 제어하는 운영자들은 거품 부유선광 시스템에서 최적의 광석 회수 품질을 달성을 목적으로 올바른 조치를 위해 실험실에서 테스트를 수행한다. 연구자들은 측정 가능한 입력 값이 주어질 때, 맥석 입자를 해방하는 상대적으로 효율적인 방법을 우회하는 지속적인 해결책을 찾는 방법에 대해서 수많은 다른 기술들을 연구했지만, 최적의 광석 회수 품질을 달성하기 위한 철광석이나 실리카 농축액의 비율을 실시간으로 추정하는 노력은 거의 수행되지 않았다. 따라서 실시간으로 의사결정을 위한 신뢰할 수 있는 부유선광 시스템을 구축하기 위해 모든 확률적 특성을 통합하는 효율적인 예측 모델을 도출해야 한다. 그럴 뿐만 아니라 공정 시간별로 예측하기 위해서 시계열 특성을 이용한다. 시계열 데이터란 시간에 따른 순서를 갖는 데이터로, 시간의 흐름에 따라 값이 변한다. 특히 생산 분야의 경우 정확한 생산 예측이 결과적으로 이익과 비용에 영향을 주기 때문에 정확한 예측을 위한 방법론이 연구되었으며 실제로 적용되고 있다. 이를 바탕으로 시간별로 나누어진 광업 프로세스 데이터를 시간의 흐름에 따른 분석을 통해 예측하고자 한다.

본 논문에서는 제 2장에서 동일한 데이터를 활용하여 예측한 방법론 및 통계적 분석과 빅데이터 분석 기법, 광업 프로세스에서 부유선광과 연관된 분석 및 예측에 관한 연구 그리고 시계열 데이터를 기반으로 한 예측모델 구축 후 비교 분석한 문헌들을 소개하고자 한다. 제 3장에서는 방법론과 머신러닝 기법을 소개한다. 제 4장을 통해 실험 결과 및 분석을 보이고 마지막 5장에서는 결론을 제시한다.

2. 관련 문헌 연구

관련 문헌 연구는 크게 세 가지로 나누어 수행하였다. 첫 번째로, 본 연구에서 사용한 같은 데이터를 활용하여 예측한 방법론에 관한 문헌 연구, 통계적 분석과 빅데이터 분석 기법을 활용한 문헌연구를 수행하였다. Kwame(2019)는 머신러닝을 사용하여 거품 처리 공장의 실리카 농축액 비율을 실시간으로 예측하는 경우의 실현 가능성을 평가하였고 철광석 추출과 관련된 변수가 통계적으로 유의한지 알아내어 거품 발생 시 실리카 농축액의 비율을 예측하는

모델을 제안하였다. Yoon(2019)은 딥러닝 기반의 군집화 모델인 DEC(Deep embedding clustering: DEC)와 JDEC(Jointly deep embedding clustering: JDEC)를 기반으로 다중센서 시계열 데이터를 군집화 하는 방법을 제안하였다. Jeong(2018)은 빅데이터 특성에 따른 상관성 측도의 조건을 제시하고 이 조건에 부합하는 상호정보의 사용을 제안하였다. Jang et al.(2020)은 여러 통계적 가정을 요구하지 않으며 데이터에 대한 유연한 적합이 가능한 비모수적 방법을 고려하여 작품 가격 평가를 위해 기계 학습 분야에서 개발된 예측 모델을 방법론으로 제안하였다. Tak et al.(2018)은 항공 무기체계의 축적된 빅데이터를 활용하여 레이다 시스템의 SRU(Shop replacement unit) 단위에 대한 고장률 예측방안을 제시하였다.

두 번째로, 광업 프로세스에서 부유선광과 연관된 분석 및 예측에 대한 문헌 연구를 수행하였다. Sawyerr et al.(1998)은 기계적 플로팅 셀에서 발생하는 기본적인 가스 분산 메커니즘을 조사하였고 기계적 플로팅 셀의 버블 크기 분포 예측모델을 도출하였다. Bushell et al.(2012)는 자동화된 광물학 시스템을 통해서 스캐닝 전자현미경, 백스캐터 전자 이미지, 이미지 분석 및 에너지 분산분석의 조합을 사용해 광석의 광물학적 변화를 정확히 파악하였다. Gorain et al.(1995)은 기계적 플로팅 셀에서 첫 번째 순서 플로팅 속도 상수와 버블 표면 영역 플럭스 사이의 선형 관계를 관찰했다.

세 번째로, 시계열 데이터를 기반으로 비교 분석에 관한 문헌연구를 수행하였다. Hur et al.(2009)은 항공수요 예측 분야에서 ARIMA(Auto regressive integrated moving average) 모형, 일변량 그리고 다변량 계절형 시계열 모형을 구축하여 분석하였다. Seo et al.(2003)은 ARIMA 모형을 이용하여 구한 예측 값을 바탕으로 X11-PROC를 사용하여 계절조정 및 성분을 분해하여 최종적으로 계절 요인, 추세요인, 추세 순환 요인, 불규칙요인, 그리고 계절변동 조정 시계열의 값을 구하는 과정과 방법을 제시하였다. Park et al.(2017)은 MODWT(Maximum overlap discrete wavelet transform) 시계열 데이터를 적용하고 심층 신뢰망(Deep belief network, DBNs)를 이용하여 예측하였다. Heimes et al.(2008)은 RNN(Recurrent neural network) 기법을 이용하여 시계열 데이터를 사용해 항공기 엔진의 잔여 수명을 예측하였다.

3. 연구 프로세스

3.1 방법론

3.1.1 데이터 소개

본 논문에서는 부유선광에서 실리카 농축액의 비율을 실시간으로 추정하기 위해 만든 데이터를 사용하였다. 데이터는 2017년 3월10일 1시부터 2017년 3월 16일 5시, 2017년 3월 29일 12시부터 2017년 9월 9일 23시까지로 구성되어있고 총 737,453개로 이루어져있다. 본 논문에서 3월 16일 5시 이전의 데이터가 2017년 9월 9일 23시 이후의 데이터에 영향을 충분히 줄 수 있다고 생각하여 중단된 일자는 없다고 가정하였다. 데이터는 날짜변수(date), 공정에 투입하는 입력변수(% Silica Feed, % Iron Feed) 2개, 프로세스 흐름으로 생기는 변수(Starch Flow~%Iron Concentrate) 20개로 독립변수는 총 23개, 실리카 농축액의 비율을 나타낸 종속변수(%Silica Concentrate) 1개로 이루어져 있다. 입력변수는 각각 X1, X2로 나타내며 프로세스 변수는 X3부터 X22까지로 나타낸다. 데이터는 1시간 간격으로 입력변수, 출력변수데이터는 180개의 값이 동일하지만 process(총 21개)안에서는 서로 다른 180개로 이루어져 있으며 Figure 1.과 같다.

date	Input		process				output
	X1	X2	X3	...	X21	X22	Y
2017-03-10 1:00	55.2	16.98	3019.53		446.37	66.91	1.31
2017-03-10 1:00	55.2	16.98	3024.41		489.382	66.91	1.31
2017-03-10 1:00				:			
2017-03-10 1:00	55.2	16.98	3118.46		448.911	66.91	1.31
2017-03-10 1:00	55.2	16.98	3147.27		481.052	66.91	1.31

Figure 1. Data Configuration

3.1.2 연구 프로세스

실리카 농축액의 비율을 실시간으로 추정하기 위해 제안하는 연구 프로세스는 Figure 2.와 같다. 변수들 간의 관련성 여부와 정도를 알아보기 위해 독립변수와 종속변수간의 상관관계 분석을 수행하였다. 또한, 기존 데이터에서의 종속변수를 결정계수(R^2)가 가장 높게 나온 3시간 후의 데이터를 사용하고 기존의 종속변수, 1시간 후, 2시간 후의 데이터를 추가하여 데이터를 재구성한다.

재구성한 데이터를 바탕으로 머신러닝 기법을 사용하기 위해서는 훈련셋(training set)과 테스트셋(test set)을 나누어야 한다. 경험적 분석에 따르면 데이터의 20~30%를 테스트에 할당하고 나머지 70~80%를 훈련에 사용하면 최상의 결과를 얻을 수 있다(Afshin et al., 2018). 또한, 반복된 k-겹(k-fold) 교차검증이 편향을 작게 유지하면서 추정의 정확도를 증가시킨다고 밝혔다(Molinaro et al., 2005). 여러 종류의 데이터 세트에 10-겹 교차검증은 편향과 분산 사이에 가장 좋은 절충점을 주었다(Kohavi, 1995). 따라서 본 논문에서 데이터를 훈련 70%, 테스트 30%로 나누었으며 그리고 10-겹 교차검증을 채택하였다. 또한 재구성한 데이터를 예측 모델링하기 위해 머신러닝 기법인 다중선형회귀분석, 랜덤포레스트, XG부스트를 채택하였고 예측 정확성 비교는 RMSE를 통해 비교하였다.

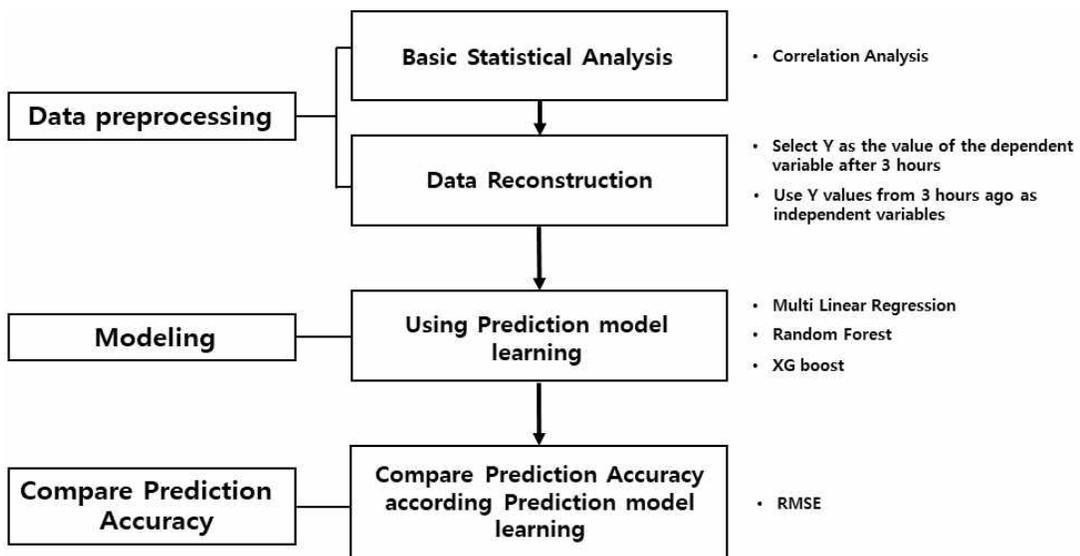


Figure 2. Research Process

3.1.3 데이터 재구성

데이터는 1시간마다 180개의 Input, Process, Output이 타점 되어있고 line-by-line 형식으로 1시에 측정된 데이터는 1시에 결과가 나오는 형태이다. 데이터 전처리 이후 시간의 흐름적 특성을 이용하여 기존 데이터에서의 종속 변수, 기존 데이터의 종속변수를 제외하고 1시간이 지난 후부터의 데이터, 기존 종속변수와 1시간이 지난 후를 제외한 2시간 후 데이터, 앞에 기존 종속변수와 1~2시간이 지난 후부터의 데이터를 제외한 3시간 후 데이터 등 반복적으로 n번 반복하였다. 회귀분석을 통해서 R²와 RMSE로 비교해본 결과는 **Table 1.**과 같으며, 가장 큰 값을 보인 3시간 후를 종속변수로 채택하는 전처리 과정을 한 후 예측 모델링을 수행하였다.

Table 1. R² square and RMSE of Linear Regression with Significant Variable

	Y _{t+3}	Y _{t+2}	Y _{t+1}	y
R ²	0.2862	0.2803	0.2785	0.1662
RMSE	0.9735	0.9779	0.9802	1.0243

모델링 수행 시 가장 영향이 적었던 독립변수인 date, % Iron Concentrate를 제거하고 시계열 특성을 가진 기존 시간 종속변수, 1시간 그리고 2시간 후의 종속변수 결과를 추가하여 총 24개의 변수를 사용하였다. 모델링에 사용한 머신러닝 기법은 총 3가지로 다중회귀분석(Multiple linear regression), 랜덤 포레스트(Random forest), XG 부스트(XG boost)이다. 랜덤포레스트와 XG 부스트 학습 시 Dataiku data science studio(DSS)을 사용하였으며 다중회귀분석 학습 시에는 Python 3.7을 사용하여 최적화를 수행하였다.

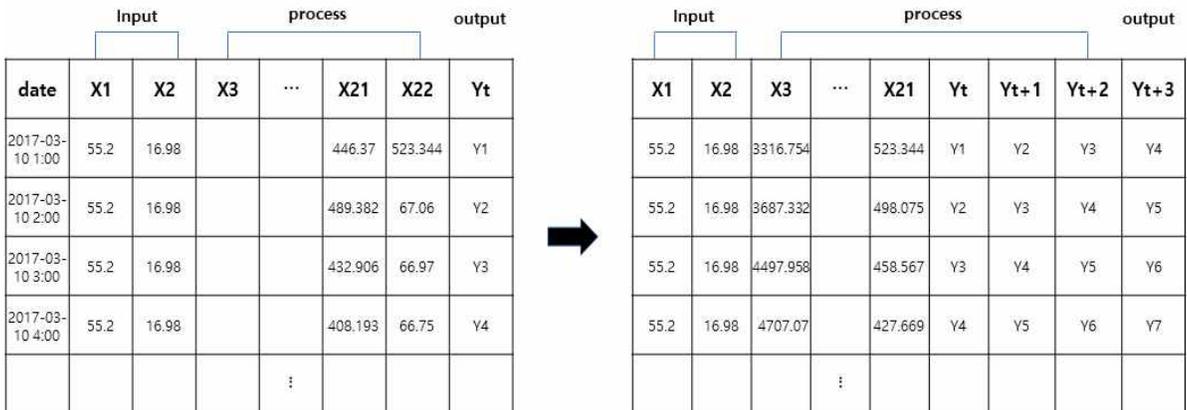


Figure 3. Proposed Method

3.2 머신러닝 기법

3.2.1 다중회귀분석

다중 회귀분석은 변수간의 함수관계를 통계적 방법을 이용하여 종속변수를 설명하기 위해 두 개 이상의 독립변수를 사용하는 선형회귀모형이다. 일반적으로 자연적 현상을 설명하는 데 있어서 종속변수는 두 개 이상의 독립변수

에 의해 좌우되는 경우가 많으며, 설명력 있는 독립변수를 여러 개 선택하여 이들의 함수로 종속변수를 나타냈을 때 회귀식의 정확도가 향상된다.

하나의 종속변수와 여러 개의 독립변수 사이에는 와 같은 관계식이 성립한다.

$$y_j = \beta_0 + \beta_1 x_{1j} + \dots + \beta_k x_{kj} + \epsilon_j, j = 1, 2, \dots, n \quad \text{Equation(1)}$$

다음 식에서, $\beta_0, \beta_1, \dots, \beta_k$ 는 모집단의 회귀계수로 모수이고 ϵ 는 종속변수를 측정할 때 발생하는 잔차이다. 또한, x_{kj} 는 k 번째 독립변수가 j 번째 관찰점이 되었을 때 가지게 되는 독립변수 값이다(Yoon et al., 2009).

3.2.2 랜덤 포레스트

랜덤 포레스트는 의사 결정 나무의 앙상블 기반 모형으로 Breiman에 의해 제안되었으며, 배깅의 원리와 임의적 특성을 더한 형태이며 빠른 학습속도, 많은 양의 데이터 처리 능력 그리고 이상치에 크게 영향을 받지 않는다는 장점이 있다(L Breiman, 2001). 앙상블 기법은 단일 분류기를 여러 분류기로 결합해 그 분류 모형들로부터 나온 결과들을 다수결 투표를 이용하거나 가중 평균을 내어 분류를 하는 것이다(Kim et al., 2017).

3.2.3 XG 부스트

XG 부스트는 CART(Classification and regression tree)를 기반으로 만들어진 알고리즘으로 의사 결정 나무의 앙상블 기반모형으로 의사결정나무 모형에 부스팅 기법을 이용하여 병렬처리를 통해 학습하는 속도를 개선한 모델이다. CART는 이러한 여러 가지 의사결정나무를 통한 방법론으로 CART의 원리는 Additive learning로 정의되며 다음과 같은 수식으로 표현이 가능하다(Tianqi chen et al., 2016).

$$Y' = a * treeA + b * treeB + c * treeC + \dots \quad \text{Equation(2)}$$

Y' 은 Y 에 대한 예측 값을 a, b, c, \dots 는 각 트리 A, B, C, ...에서 나온 가중치를 뜻한다. 다음과 같은 개념을 XG 부스트와 Gradient boosting tree에 적용시키면 아래 식과 같이 표현할 수 있다.

$$y'_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad \text{Equation(3)}$$

$$obj = \sum_{i=1}^n l(y_i, y'_i) + \sum_{k=1}^K \Omega(f_k) \quad \text{Equation(4)}$$

첫 번째 식에서 y'_i 는 XG부스트에 적용한 예측 값을 의미하며 f_k 는 k 번째 결정나무를 의미한다. 두 번째 식에서 l 은 손실함수를 의미하며 Ω 은 정규화하는 기간을 의미한다. 따라서 여러 개의 의사 결정 나무 모델들을 학습시켜 예측 값을 더한 것으로 결정한다는 개념으로써 더해진 예측 점수들을 이용하여 결론을 내려 과적합이 일어나는 경우나 기존 모델이 잘 설명하지 못하는 부분에 대해서도 보완 할 수 있다(Choi et al., 2016).

5. 결 론

국내 광업 분야에서 머신러닝을 도입한 사례는 많지 않으며 현재까지도 기초적인 연구가 매우 부족한 상황이다. 또한 일반적으로 광업 프로세스 데이터를 분석에 용이하게 전처리하여 사용하는 데 있어서 어려움이 존재한다.

따라서 본 논문에서는 광업 프로세스 데이터를 통해 해당 공정에서 불순물 첨가 시 도출되는 실리카 농축액의 비율을 예측하기 위해 머신러닝 기법을 활용하였다. 독립변수들의 상관관계를 확인하기 위해 상관분석을 실시하여 유의미한 변수들을 찾아내고 상관성이 적은 변수들은 제거했다. 그 후 기존의 종속변수, 기존 종속변수를 제외한 1시간 후의 데이터, 기존 종속변수와 1시간 후의 종속변수를 제외한 2시간 후의 데이터를 선정하였다. 마지막으로 앞의 3가지를 제외한 3시간 후의 데이터를 각각의 종속변수로 설정하여 독립변수들과의 선형회귀분석을 진행하였다. 다음으로 가장 R^2 가 우수하게 나온 3시간 후의 데이터를 종속변수로 설정하였다. 예측모델링을 수행하기 위해 다중회귀분석, 랜덤포레스트, XG 부스트 모델을 구축하였으며, 각 예측 모델에 대한 RMSE 값을 구하였다. RMSE값을 비교해 본 결과, 다른 방법론에 비해 제안한 방법론을 적용한 모델의 예측성능이 더 우수한 결과를 보였다.

본 연구를 바탕으로 다른 분야의 광업 데이터에서 좋은 결과를 보이지 못하고 시계열 모델을 활용하기에 어려움이 존재하며 예측 결과가 저조한 경우, 제안한 전처리 방법을 활용한다면 예측 성능 향상을 기대할 수 있을 것이다.

본 논문에서는 일부 머신러닝 기법만을 이용하여 광업 프로세스 데이터를 분석하였으나 추후 연구에서는 광업 프로세스 데이터에서 필요로 할 수 있는 ANN(Artificial neural network)과 같은 추가적인 머신러닝 기법과 더 나아가 딥러닝 중 시계열 분석에 있어서 주목받고 있는 RNN(Recurrent neural network)와 LSTM(Long short-term memory models)등을 활용하여 제안하는 방법론과 비교하여 분석한다면 유의미한 결과를 얻을 수 있을 것이라 사료된다.

REFERENCES

- Afshin, G., Vladik, K., and Olga, K. 2018. Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation. Departmental Technical Reports(CS). 1209.
- Breiman, L. 2001. Random forests, *Machine Learning* 45(1):5-32.
- Chen, T. and Guestrin, C. 2016. Xgboost : A scalable tree boosting system. *Proceedings of the 22nd acm SigkddInternational Conference on Knowledge Discovery and Data mining.* pp. 785-794.
- Choi, S.-H. and Hur, J. 2020. Optimized-XGBoost Learner Based Bagging Model for Photovoltaic Power Forecasting. *The transactions of The Korean Institute of Electrical Engineers* 69(7):978-984.
- Gorain, B. K. FRANZIDIS, J. P. and MANLAPIG, E. V. 1995. Effect of Bubble Size, Gas Holdup and Superficial Gas Velocity on Metallurgical Performance in an Industrial Flotation Cell. JKMRC Report.
- Heimes, F. O. 2008. Recurrent neural networks for remaining useful life estimation. In 2008 international conference on prognostics and health management. IEEE. pp. 1-6.
- Hur, N. K., Jung, J. Y. and Kim, S. 2009. A Study on Air Demand Forecasting Using Multivariate Time Series Models. *The Korean Journal of Applied Statistics* 22.5:1007-1017.
- Jang, D. R. and Park, M. J. 2020. A Study on the Art Price Prediction Model Using the Random Forest. *Journal of Applied Reliability* 20(1):4-42.

- Jang, H. D. 2019. Australian Mining Transformation and Future Prospects in Response to the 4th Industrial Revolution. *Journal of the Korean Society of Mineral and Energy Resources Engineers* 56(5):490-513.
- Jeong, H. S. 2018. Correlation Measure for Big Data. *Journal of Applied Reliability* 18(3):208-212.
- Kim, J. K., Lee, K. B., and Hong, S. G. 2017. ECG-based Biometric Authentication Using Random Forest. *Journal of the Institute of Electronics and Information Engineers* 54(6):100-105.
- Kohavi. and Ron. 1995. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *International Joint Conference on Artificial Intelligence* 14(12):1137-43
- Kwame O, E. 2019. Machine Learning-based Quality Prediction in the Froth Flotation Process of Mining: Master's Degree Thesis in Microdata Analysis.
- Lee, C., Kim, S. M., and Choi, Y. 2019. Case Analysis for Introduction of Machine Learning Technology to the Mining Industry. *Korean Society for Rock Mechanics. TUNNEL AND UNDERGROUND SPACE* 29(1). 1-11.
- Lee, Y. H., Song, M. S., Ha, S. J., Baek, T. H., and Son, S. Y. 2016. Big data Cloud Service for Manufacturing Process Analysis. *The Korean Journal of BigData* 1(1):41-51.
- Molinaro., Annette, M., Richard, S., and Ruth M. P. 2005. Prediction Error Estimation: A Comparison of Resampling Methods. *Bioinformatics (Oxford, England)* 21(15). Oxford University Press: 3301-7.
- Park, K. T., and Baek, J. G. 2017. Time Series Prediction using ARIMA and DBNs with MODWT. *Journal of the Korean Institute of Industrial Engineers* 43.6:474-481.
- Sawyer, C. T. 1998. Prediction of bubble size distribution in mechanical flotation cells. *Journal of the Southern African Institute of Mining and Metallurgy* 98(4):179-185.
- Seo, M. Y. and Rhee, J. T. 2003. A Study on the Seasonal Adjustment of Time Series and Demand Forecasting for Electronic Product Sales. *Journal of Applied Reliability* 3(1):13-39
- Tak, J. H., and Jung, W. 2018. Estimation of Failure Rate of SRU in RADAR System Utilizing Big Data. *Journal of Applied Reliability* 18(4):339-348.
- Yoon, D. H., Kim, S. M., and Kim, D. H. 2019. Clustering of Time Series Data using Deep Learning. *Journal of Applied Reliability* 19(2):167-178.
- Yoon, H. S., Um, M. J., Cho, W. C., and Heo, J. H. 2009. Orographic Precipitation Analysis with Regional Frequency Analysis and Multiple Linear Regression. *Journal of Korea Water Resources Association* 42(6):465-480.

저자소개

- 이승훈** 경기대학교 산업경영공학과에 재학 중이며 주요 관심분야는 신뢰성공학 및 데이터 마이닝 등이다.
- 윤연아** 경기대학교 산업경영공학과를 졸업하고, 동 대학원 데이터공학연구실에서 석사과정에 재학 중이다. 주요 관심 분야는 데이터분석, 신뢰성공학 등이다.
- 정진형** 경기대학교 산업경영공학과에서 학사학위를 취득한 후 동 대학원 석사과정에 재학 중이다. 전공분야는 신뢰성공학 및 데이터 마이닝으로 관련 연구를 수행하고 있다.
- 심현수** 경기대학교 산업경영공학과에서 학사, 석사학위를 취득한 후, 현재 박사과정에 재학중이다. 신뢰성 시험 평가를 주로 연구하며, 특히 가속수명시험 및 가속열화시험 등에 많은 관심을 갖고 있다.
- 장태우** 서울대학교 산업공학과에서 학사를 마쳤으며, 동 대학원에서 산업공학 석사 및 박사 학위를 취득하였다. 한국

전자통신연구원에서 우편물류 정보화/자동화를 연구하였고, 2007년부터 경기대학교 산업경영공학과에 재직 중이다. 주요 연구 관심분야는 스마트공장, 물류/SCM, 시스템 분석 등이 있으며, 현재 경기도 지역협력연구센터(GRRC)인 지능정보융합제조연구센터의 센터장을 맡아서 다양한 스마트공장 프로젝트를 진행 중이다.

김용수 교신저자, KAIST 산업공학과에서 학사, 석사, 박사를 취득한 후 SK텔레콤에서 근무하였다. 현재 경기대학교 산업경영공학과 부교수로 재직 중이며, 품질 및 신뢰성, 기능안전, 통계 및 데이터마이닝 분야를 연구하고 있다.