KSMER Journal of The Korean Society of Mineral and Energy Resources Engineers

Vol. 59, No. 3 (2022) pp. 265-275, https://doi.org/10.32390/ksmer.2022.59.3.265

ISSN 2288-0291(print)
ISSN 2288-2790(online)

연구논문

머신러닝 기법을 이용한 광산 지반침하 예측 성능의 비교분석

한호상[1] · 최교익[1] · 서장원[2]*

[1]강원대학교 에너지자원융합공학과 석사과정, [2]강원대학교 에너지공학부 (에너지자원융합공학전공) 조교수

Comparative Analysis of Prediction Performance of Mine Subsidence Using Machine Learning Techniques

Hosang Han[1], Kyoik Choi[1] and Jangwon Suh[2]*

[1]M.S. course, Department of Energy and Mineral Resources Engineering, Kangwon National University, Samcheok, Korea
[2]Assistant Professor, Department of Energy and Mineral Resources Engineering, Kangwon National University, Samcheok, Korea

***Corresponding Author.** Jangwon Suh, jangwonsuh@kangwon.ac.kr, Dept. of Energy and Mineral Resources Engineering, Kangwon National University, Samcheok, Korea

Received
21 April 2022

Final version Received
10 June 2022

Accepted
27 June 2022

Abstract

In this study, the prediction of mining-induced subsidence is analyzed and compared using various machine learning models. Factors affecting the occurrence of subsidence are identified from eight and 1,730 sets of subsidence data. Five machine learning models are selected, i.e., Adaboost, artificial neural networks, the k-nearest neighbor, random forest, and the support vector machine, which are frequently used in studies related to geohazard prediction. In addition, the stacking technique is applied to five algorithms based on 10 combinations, and the predictive performance of each ensemble method is evaluated and compared. To evaluate the classification performance of the machine learning technique applied in this study, recall is used as an evaluation index, which describes the ratio of the predicted ground subsidence instead of the area under curve used previously. Based on the values of recall, the random forest demonstrates the best performance (with a recall of 0.955). The recall is expected to be a more reliable evaluation index for predicting subsidence occurrences compared with other indices.

Key words : Mine subsidence, Machine learning, Random forest, Recall, Area Under the Curve

요 약

본 연구에서는 폐광 지역의 지반침하 발생 예측에 활용되는 다양한 머신러닝 기법의 예측 성능을 비교·분석하였다. 광산 지반침하 발생 여부에 영향을 미치는 인자는 총 8개로 구성하였고, 1,730개의 지반침하 데이터를 획득하였다. 머신러닝 기법은 지반재해 예측에 대한 문헌에서 주로 활용되는 Adaboost, ANN(Artificial Neural Network), kNN(k-Nearest Neighbor), RF(Random Forest), SVM(Support Vector Machine) 등 5개를 선정하였다. 또한, 5개 알고리즘에 Stacking 기법을 적용하여 10개 조합으로 나누어 각각의 앙상블 기법들의 예측 성능을 비교하였다. 본 연구에서 적용된 머신러닝 기법의 분류 성능 평가 지표로는 기존에 사용되던 AUC(area under the curve)가 아닌 실제로 발생된 지반침하를 맞게 예측하는 비율을 의미하는 Recall에 중점을 두었다. Recall에 기반한 결과, RF 기법이 0.955로 가장 높은 성능을 나타내었다. 이는 지반침하 발생 여부를 예측하는데 기존 기법보다 신뢰성 있는 평가 기준이 될 수 있을 것으로 기대된다.

주요어 : 광산 지반침하, 머신러닝, 랜덤포레스트, 재현율, 곡선하면적비율

Copyright ⓒ 2022, The Korean Society of Mineral and Energy Resources Engineers

서 론

폐광산 지역에서 발생되는 지반침하는 지상에 위치한 철도 및 도로와 같은 국가기간시설의 안정성뿐만 아니라 인명과 부동산, 동산에도 피해를 미칠 수 있으며, 해당 지역의 개발계획에도 장애요인이 된다(Kratzsch, 1983; Suh *et al*., 2010; Waltham *et al*., 2011). 2019년 광해통계연보에 따르면 국내의 경우 4,677개의 폐광산이 존재하며, 702건의 지반침하가 발생된 것으로 보고되는 등 이는 폐광지역 안전문제의 주요 이슈로 꼽힌다(Mine Reclamation Corporation, 2019). 따라서 광산 지반침하로 인한 피해를 최소화하고 이를 효과적으로 관리하기 위해서는 과학적 기법들을 활용하여 광산 지반침하를 사전에 평가 및 예측할 필요가 있다.

광산 지반침하 평가 및 예측을 위해 다양한 분석 기법들이 국내·외의 다수의 연구자들에 의해 제안되었다. 이러한 연구들은 대부분 지리정보시스템(geographic information systems, GIS) 환경에서 지반침하 발생과 영향인자간의 상관성 분석에 확률·통계 기법, 전문가 시스템, 머신러닝 기법들이 적용된 사례에 해당한다. 확률·통계 기법으로는 Frequency ratio 또는 Weight of Evidence를 이용한 연구(Choi *et al*., 2007; Oh and Lee, 2010; Oh *et al*., 2011; Suh *et al*., 2013; Son *et al*., 2015; Suh *et al*., 2016)가 주를 이루며 Logistic regression 기법이 적용된 사례(Kim *et al*., 2006)도 보고되었다. 전문가 시스템 적용 사례로는 Fuzzy theory를 이용한 연구(Choi *et al*., 2010; Park *et al*., 2012; Park *et al*., 2014)가 있고, 여러 개의 기법을 결합하여 지반침하를 예측한 연구(Oh and Lee, 2011; Bui *et al*., 2018; Oh *et al*., 2019)도 발표되었다.

최근에는 지반침하 예측에 머신러닝 기법을 접목한 사례들을 다수 확인할 수 있었다. NN(neural network)을 이용한 연구(Kim *et al*., 2009; Lee *et al*., 2012)가 있었으며, Decision Tree를 이용한 연구(Lee and Park, 2013), SVM (support vector machine) 기반 연구(Bui *et al*., 2018), 그리고 Bayes net이나 naive Bayes를 활용한 연구(Oh *et al*., 2019)도 진행된 바 있다. 이러한 연구들은 예측 성능의 지표로 실제 지반침하 발생지와 미발생지 전체에 대한 예측 정확도를 평가하는 AUC(area under the curve)를 사용하였다는 점에서 개선의 여지가 있다. 일반적으로 연구지역에서 지반침하 발생지에 비해 미발생지가 차지하는 면적 비율이 높기 때문에 AUC 결과는 실제 지반침하 미발생지를 잘 예측하는 것에 더 영향을 많이 받을 수 있다. 반면에 머신러닝 기법의 분류 예측에 대한 성능 평가 지표 중 하나인 Recall의 경우, 적용된 기법이 실제 발생된 지반침하를 얼마나 올바르게 맞추었는지에 중점을 둔다.

본 연구에서는 광산 지반침하 발생 예측을 위해 다양한 머신러닝 기법을 적용하고, 예측 기법의 성능 평가 지표로 ROC 기반의 AUC가 아닌 Recall 관점에서의 평가를 제안한다. 이를 위하여 5종류의 알고리즘을 선정하고 파라미터 최적화를 진행해 각 기법의 성능을 비교분석하고, 앙상블 기법의 일종인 Stacking 기법을 이용하여 10종류의 알고리즘 조합에 대한 Recall 수치를 비교·분석한다. 추가로 타 연구 논문과의 비교를 통하여 본 연구에서 개발한 머신러닝 기법의 성능을 비교하고자 한다.

연구지역

연구지역은 강원도 정선군 고한읍 고한 16리 두문동 일대이며, 해당 지역에는 강원도 고한읍과 태백시를 잇는 38번 국도가 위치한다. 이 지역의 지리좌표는 북위 37°12'~37°13', 동경 128°53'10"~128°54'10"이고, 1967년과 1989

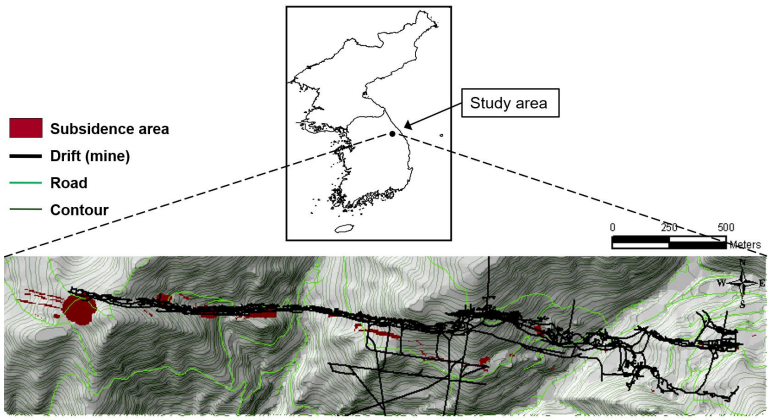

Fig. 1. Location of mine drifts and subsidence occurrences in study area (Suh *et al*., 2010).

년 사이 약 20여 년간 채탄이 진행되었던 세원탄광이 위치한 곳이다. 이 탄광은 심도 약 660 m까지 개발하여 약 756,000톤의 석탄을 채탄하였으며, 국도의 직하부에는 채굴적이 분포하는 것으로 조사되었다. Fig. 1은 본 연구지역의 지형과 도로, 지반침하의 발생 위치, 갱도의 분포를 보여준다. 해당 지역에서 발생한 지표침하는 상당수 산악에서 발생하였으며, 국도 주변에서의 발생 사례는 도로 근방 산악지에서 석탄층의 주향 방향으로 다수의 함몰형 침하가 발생된 것을 관찰할 수 있었다. 또한, 주변 지역의 일부 가옥에서도 침하 또는 균열 등이 발생된 것으로 확인되었다 (Coal Industry Promotion Board, 2005).

연구방법

Fig. 2는 광산 지반침하 발생 여부 예측을 위한 머신러닝 기법들의 성능 평가 비교 연구에 대한 절차를 보여준다. 이는 데이터 취득 및 분류, 교차검증, 머신러닝 기법의 선정과 최적화, 예측 성능 평가 및 비교·분석 과정으로 구성된다.

데이터 취득 및 분류

우선 지반침하의 발생 여부를 예측하기 위하여 연구지역의 데이터를 구축하고, 이를 훈련 데이터와 테스트 데이터로 분류한다. 전체 데이터는 8개의 입력 인자와 1개의 출력 인자로 구성된다. 입력 인자는 지반침하 발생에 영향을 미치는 것으로 알려진 변수 8개(갱도 심도와 갱도 밀도, 갱도로부터의 거리, 도로로부터의 거리, RMR(rock mass rating), 지하수위, 경사도, 강우누적흐름양)이고, 출력 인자는 지반침하 발생 여부를 나타낸다. 지반침하의 발생 여부에 따른 발생을 Positive, 미발생을 Negative로 구분하였으며, 8개 영향인자에 대한 자세한 설명은 Suh *et al*.(2010)을 참고할 수 있다. Table 1은 연구에 사용된 입력 인자 데이터의 기초 통계치를 보여준다.

또한 본 연구에서 수집된 데이터는 연구지역을 가로 5 m, 세로 5 m 격자로 구성된 래스터 형식이며, 각 격자에는 8개의 영향인자 값과 1개의 지반침하 발생 여부를 포함하고 있

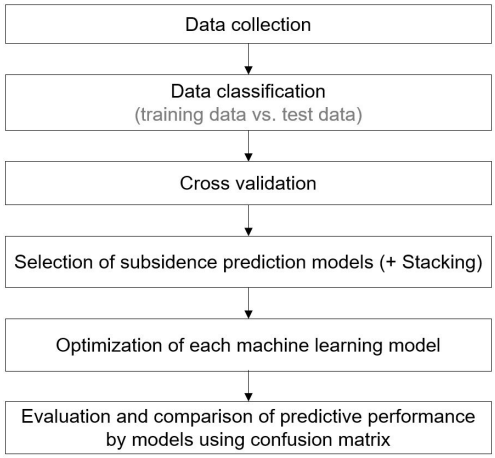

Fig. 2. Flowchart for comparing performance of subsidence prediction models in this study.

Table 2. Amount of data for predicting subsidence in this study

Data type		Data size
Data set	Original data	93,800
	Training data (80%)	75,040
	Test data (20%)	18,760
Class	Subsidence occurrence	1,730
	Non-subsidence occurrence	92,070

Table 1. Basic statistics of factors used to develop machine learning techniques

Factor	Statistics				
	Min	Median	Mean	Max	STDEV.*
Drift depth (m)	0.00	0.00	7.96	508.00	38.40
Drift density	0.00	0.00	0.05	0.94	0.11
Distance from drift (m)	0.00	84.98	121.49	607.40	119.86
Distance from road (m)	0.34	48.35	63.14	289.51	52.71
Rock mass rating	32.98	37.98	39.86	52.03	5.38
Groundwater level (m)	13.55	41.03	35.65	53.59	12.73
Slope (°)	0.00	23.96	23.58	68.34	9.42
Flow accumulation	0.00	48.00	635.57	217644.70	6998.09

*STDEV: standard deviation

다. 이에 따라 연구지역에서 총 9개의 인자로 구성된 1개의 격자가 총 93,800개 수집되었으며, 전체 데이터 93,800개에 계층적 샘플링을 적용하여 훈련 데이터와 테스트 데이터 80%:20%의 비율로 설정 및 분리하였다. 따라서 연구에 사용된 데이터 개수는 Table 2와 같다.

데이터 훈련 및 검증을 위한 교차검증

머신러닝 기법을 이용한 데이터 훈련 과정에서 일부 또는 특정 자료만 활용할 경우 샘플링 과정에서 편향(bias)이 생기거나 전체 훈련 데이터의 특성을 반영하지 못하는 단점이 있다. 이러한 문제를 극복하기 위하여 본 연구에서는 전체 훈련 데이터를 모두 검증 데이터로도 사용할 수 있는 계층적 샘플링의 일종인 K-폴드 교차검증(K-fold cross validation) 방법을 사용하였다. 이 방법은 훈련 데이터 집합으로부터 K개의 상호 배타적인 같은 크기의 부분 집합(fold)을 만든 후에 K-1개의 폴드를 훈련용으로, 나머지 1개의 폴드를 검증용으로 활용하는 방식을 반복적으로 수행함으로써 편향을 최소화하고 훈련 과정에 대한 신뢰도를 확보할 수 있는 장점이 있다(Jung and Choi, 2021). 본 연구에서는 K값을 5로 설정한 5겹 교차검증을 수행하였다.

머신러닝 기법 선정

본 연구에서는 Python 기반의 오픈소스 머신러닝 및 데이터 가시화 툴킷 중 하나인 Orange Data Mining을 활용하여 머신러닝 기법의 지반침하 분류 예측 성능을 평가하였다. 광산 지반침하 발생 여부를 예측하기 위하여 동일 주제를 다룬 기존 문헌에서 많이 활용된 기법을 검토하였고, 본 연구에서는 AdaBoost(adaptive boost)와 ANN(artificial neural network), kNN(k-nearest neighbor), RF(random forest), SVM(support vector machine) 등 총 5종류의 기법을 선정하였다.

- AdaBoost: 선형적으로 결합한 약한 분류기(weak classifier)를 반복적으로 이용하여 높은 성능을 갖춘 강한 분류기(strong classifier)로 구성하는 기법으로 가장 대중적인 알고리즘 중 하나이다(An and Kim, 2010). 생성된 강한 분류기는 잘못 분류된 데이터가 잘 분류된 데이터보다 많으므로 잘못 분류된 데이터에 초점을 두고, 각 반복에서 갱신된 가중치를 덧셈 또는 뺄셈 연산하여 최종 가중치를 계산한다. 이후 최종 알고리즘은 각 가중치를 고려한 총 가중치로 나누어 계산한다(Hakim *et al*., 2020).
- ANN: 인간의 두뇌가 정보를 처리하는 방식에서 영감을 얻어, 훈련 데이터로부터 반복적인 학습 과정을 거쳐 해당 데이터의 패턴을 도출하고 이를 일반화하는 알고리즘이다(Ahatonovic-Kustrin and Beresford, 2000). 이 기법은 여러 뉴런이 일렬로 나열된 레이어(층)를 통하여 데이터의 값을 출력하는데 이 레이어는 입력층과 은닉층, 출력층으로 구성된다. 또한, 여러 층의 구조를 가진 인공신경망을 MLP(multi-layer perceptron)이라고 하며, 해당 기법은 복잡한 물리적인 이론식을 계산하기 어려운 문제에 응용할 경우 전통적인 통계 방법보다 효율적이고 정확하게 해결할 수 있는 장점이 있다(Koo *et al*., 2019). Orange Data Mining에서의 ANN 기법은 sklearn의 MLP 알고리즘을 기반으로 한다.
- kNN: 알고리즘의 구현이 단순하고 성능이 뛰어나 데이터 마이닝 분야와 머신러닝 분야에 널리 사용되는 알고리즘이다(Zhang *et al*., 2017). 이때 k값은 사용자가 정의하는 상수이며, 해당 값에 따라 기법의 예측값이 변화한다. kNN은 데이터를 특정 값으로 분류하기 위해 훈련 데이터 내에서 해당 데이터에 가장 근접한 k개의 훈련 데이터 중 최빈값의 레이블을 할당하여 해당 데이터를 분류한다.
- RF: 주어진 데이터의 특징 중 가장 영향력 있는 특징을 찾기 위하여 정보 엔트로피나 지니 계수를 이용해 특정 영향력의 대소를 비교하는 의사결정나무를 다수 생성하는 알고리즘이다. 다수 생성된 의사결정나무는 최고의 영향을 가진 결과를 선정하는데 해당 과정을 RF라고 칭한다(Breiman, 2001). 단일 의사결정나무와 달리 해당 기법의 경우 각 노드에 표본을 뽑은 데이터 이외의 데이터를 제외하여 의사결정나무를 분기하며 이 과정에서 편향이 한 번 더 증가하여 과대적합의 가능성을 감소시킨다.
- SVM: 단순한 알고리즘에 비하여 우수한 성능을 보여주는 기법으로, 분류와 회귀에 다량 사용되는 지도학습 방법이다(Jakkula, 2006). 해당 기법은 벡터 공간에서 데이터를 두 개의 범주로 구분 짓는 결정 경계(decision boundary)를 찾아내어 이와 가장 가까운 데이터에서 맞닿는 서포트 벡터(support vector)의 거리를 계산한다. 이때의 거리를 마진(margin)이라 칭하며, 이를 최대로 갖는 결정 경계를 찾는 것을 목적으로 한다. 선형적으로 분류할 수 없는 공간적 데이터의 경우 비선형 분류를 사용하여야 하므로 저차원 공간에서 고차원 공간으로 이동시켜주는 매핑 함수(mapping function)와 커널 함수(kernel function)를 사용하여 계산한다(Nam *et al*., 2016).

머신러닝 기법 최적화

본 연구에서는 전술한 5종류의 머신러닝 기법과 각 알고리즘 2개씩에 앙상블(ensemble) 기법을 적용한 조합(10종)을 개발하여 지반침하 예측 성능을 비교하였다. 각 기법의

예측 분류 성능을 비교하기 전에 각각의 기법에 대하여 최소의 오차율을 갖는 설정을 찾기 위한 과정인 최적화를 진행하였다. 최적화는 머신러닝 기법에 영향을 미치는 매개변수를 조정해가며 적절한 특징을 도출하는 시기까지 수행하는데 만약 데이터에서 충분한 특징을 찾아내지 못하고 기법을 적용하면 과소적합(underfitting)이 발생할 가능성이 증가한다. 반면에 과도한 양의 특징을 찾아내어 기법을 적용하면 결괏값이 과대적합(overfitting)될 가능성이 높아진다. 따라서 각 기법별로 최적화를 수행하기 위해 옵션과 변수 설정을 변화시켜가며 훈련 데이터의 학습률을 높이는 작업을 수행한다. Table 3은 각 머신러닝 기법의 최적화 과정에서 변화시킨 옵션과 파라미터를 정리한 것이다.

Adaboost 기법의 경우 파라미터인 Fixed seed for random generator와 Number of estimators와 Learning rate, Classification Algorithm, Regression loss function 등의 수치를 구간을 두어 증가시키더라도 성능 평가 수치가 변화하지 않아 특정 파라미터에 의해 영향을 미치지 않는 것으로 판단하였다. ANN 기법의 경우 파라미터인 Neurons in hidden layers를 경곗값인 100에서 1,000까지 5구간으로 나누어 다른 파라미터인 Activation과 Optimizer, Maximal number of iterations를 변화시키며 최적화를 수행하였다. 이후 다른 변수들의 변화에 따른 평가 수치를 나열하고, 각 구간의 평가 수치 상위 2개를 선택하여 동일 구간에서 Maximal number of iterations를 100에서 1,000까지 100단위로 증가시키며 평가 기준값을 관찰하였다. Neurons in hidden layers의 경곗값인 100 이하의 수치는 최적화된 수치보다 낮았음을 확인하였고, 포함되지 않은 Regularization의 경우 값의 변화가 생기더라도 결과 수치에 영향을 미치지 않아 제외하였다.

kNN 기법의 경우 파라미터인 Number of neighbors를 1에서 5까지 진행한 후 10에서 100까지 3구간으로 나누어 최적화를 수행하였다. 이외에도 Metric과 Weight도 값을 변경하며 기법의 예측 성능에 대한 민감도를 분석하였다. RF 기법의 경우 구간을 두어 수치를 변경하여도 결과에 영향을 미치지 않는 파라미터인 Do not split subsets smaller than을 제외한 나머지 5가지의 파라미터를 변화시키며 최적화를 수행하였다. Number of trees는 1에서 100까지 4구간으로 나누었으며, Replicable training 적용 여부, Balance class distribution 적용 여부, Limit depth of individual trees의 수치를 한계 구간인 1에서 10까지 변화시키며 최적화하였다. SVM 기법의 경우 기준값에 영향을 미친다고 판단된 파라미터인 커널 함수의 방식과 g, c, d, Optimization parameters의 iteration limit를 변화시키며 최적화를 수행하였다. 각각 g의 경우 1에서 10까지, c의 경우 2에서 10까

Table 3. Parameter setting for each machine learning model used in this study

Algorithm	Parameter	Min	Max	Step
Adaboost	Fixed seed for random generator	0	1,000	100
	Number of estimators	1	100	10
ANN	Activation	Identity / Logistic / tanh / ReLu		
	Optimizer	L-BFGS-B / SGD / Adam		
	Neurons in hidden layers	100	1,000	100
	Maximal number of iterations	100	1,000	100
kNN	Metric	Euclidean / Manhattan / Chebyshev /Mahalanobis		
	Weight	Uniform / Distance		
	Number of neighbors	1	100	1
RF	Replicable training	Checked / Unchecked		
	Balance class distribution	Checked / Unchecked		
	Number of trees	1	100	25
	Limit depth of individual trees	1	10	1
SVM	Kernel	Linear / Polynomial / RBF / Sigmoid		
	g	1	10	1
	c	2	10	1
	d	3	10	1
	Iteration limit	10	100	10

지, d의 경우 3에서 10까지, iteration limit의 경우 10에서 100까지 증가시키며 기준값이 달라지는 정도를 측정하였다. 이외의 파라미터인 Cost와 epsilon, Tolerance의 경우 변화하는 수치에 따라 성능 평가 값에 영향을 미치지 않아 제외하였다.

머신러닝 기법 성능평가

머신러닝 기법의 성능평가 방법은 최종 예측값의 형태가 연속형인지 이산형인지에 따라 달라진다. 본 연구에서는 지반침하 발생 여부를 예측하는 분류 문제에 해당하기 때문에 이에 적합한 성능평가 지표를 사용하였다. 기존의 지반침하 평가 및 예측 연구에는 주로 AUC가 보편적으로 활용되었다. 이 지표는 ROC(receiver operating characteristic) 곡선이 이진 존재-부재 변수로 변환될 때 모든 임곗값에 대한 전체 기법의 성능을 요약한다(Lobo *et al*., 2008). ROC 곡선은 이진 분류 문제에서 임곗값이 변화할 때 민감도(Recall)와 1-특이도(Specificity)를 각각 축으로 놓고 해당 조합이 어떠한 형식으로 변하는가를 보여주는 그래프이다. 또한 여러 기법의 ROC 곡선은 비교할 때는 곡선 아랫부분의 면적인 AUC(area under curve)을 이용하며, 해당 면적이 클수록 성능이 우수한 기법이라고 판단한다.

지반침하 발생 여부에 대한 분류 문제와 같은 기법의 성능을 평가하기 위해서는 주로 혼동행렬(confusion matrix)을 이용한다. 혼동행렬의 평가 요소로는 각각 TP(true positive)와 TN(true negative), FP(false positive), FN(false negative)이 존재하며, 해당 네 가지 요소들을 이용하여 적용된 기법을 성능을 평가한다(Fig. 3). TP와 TN은 실제로 발생된 사건을 올바르게 예측한 것이며, FP와 FN의 경우 실제 발생 여부를 다르게 예측한 것을 의미한다. 상기 혼동행렬의 4가지 요소들로부터 계산될 수 있는 네 가지 개념이 있는데 이는 각각 Classification accuracy(CA; 정확도)와 Precision(정밀도), Recall(재현율), F1 Score이다. 우선 CA(Fig. 3(a))는 가장 일반적인 머신러닝 기법의 성능 평가 지표로써 전체 데이터 중 실제로 긍정(positive)인 것을 긍정(positive)으로 예측하고 부정(negative)인 것을 부정(negative)으로 정확하게 예측하였는지를 보여주는 값이다. Precision(Fig. 3(b))의 경우 기법의 예측값이 얼마나 정확한가를 보여주는 지표로, 기법이 Positive라 예측 분류한 값 중 실측값이 Positive인 비율을 의미한다. Recall(Fig. 3(c))은 실제 실측값 중 기법이 예측한 실측값을 나타내는 지표이며, 실측값이 Positive인 값 중 기법이 Positive인 비율을 표시한다. F1 Score(Fig. 3(d))는 Precision과 Recall 모두를 고려한 수치이며, 이는 두 수치를 조화평균 내어 계산하는 지표이다.

상기 문단에서 ROC 곡선 기반의 AUC와 서술한 4개 지표는 머신러닝 기법의 성능을 평가하는 데 주로 사용되고 있다. 그러나 본 연구에서는 실제 발생한 현상을 기법이 얼마나 맞추었는가에 중점을 두고자 하여, ROC 곡선 기반의 AUC가 아닌 Recall에 초점을 맞추는 것이 알맞다고 판단하였다.

또한, 단일 머신러닝 기법들에 대하여 Stacking 기법을 적용하였는데 이 기법은 여러 기본적인 기법들을 조합하여 개선된 성능을 나타내는 메타 기법을 생성하는 앙상블 기법의 일종이다. 앙상블 기법은 보편적으로 다양한 알고리즘을 사용하여 각각의 조합을 생성해 서로의 장점을 강화

(a)

		Predicted	
		Positive	*Negative*
Actual	*Positive*	TP	FN
	Negative	FP	TN

$$CA = \frac{TP + TN}{TP + FN + FP + TN}$$

(b)

		Predicted	
		Positive	*Negative*
Actual	*Positive*	TP	FN
	Negative	FP	TN

$$Precision = \frac{TP}{TP + FP}$$

(c)

		Predicted	
		Positive	*Negative*
Actual	*Positive*	TP	FN
	Negative	FP	TN

$$Recall = \frac{TP}{TP + FN}$$

(d)

		Predicted	
		Positive	*Negative*
Actual	*Positive*	TP	FN
	Negative	FP	TN

$$F1\ score = \frac{2 * Precision * Recall}{Precision + Recall}$$

Fig. 3. Formulas used to evaluate prediction performance of machine learning model.

하고 약점을 보완하기 위하고자 할 때 이용하며, 이는 보다 높은 분류 정확도를 획득하기 위하여 이용된다. Stacking 기법은 기반으로 사용된 기법의 결과를 훈련 데이터의 특징으로 메타 기법에 학습시킴으로써 앙상블 기법의 가중치를 추정한다(Yeon, 2020). 따라서 본 연구에서는 최적화된 5가지의 기법을 기반으로 하여 각 기법을 2개씩 조합한 10개의 모델에 대하여 Stacking 기법을 적용하였다. Fig. 4는 Stacking 기법의 적용 방법에 대한 개념도를 보여준다.

연구결과 및 해석

Table 4는 최적화를 수행한 각 머신러닝 기법별 분류 예측 성능 지표 중 테스트 데이터에 대한 Recall과 ROC 곡선 기반의 AUC 값을 보여준다. Recall 지표 기준에서는 RF 모델이 가장 높은 예측 성능(0.955)을 보였는데, 이는 테스트 데이터 중 실제 발생된 지반침하 격자(146개)의 95.5%(140개)를 올바르게 예측한 것으로 해석할 수 있다. 이후 SVM(0.869), kNN(0.829), Adaboost(0.790)가 뒤를 이었다. 마지막으로 ANN의 경우 0.653을 기록하여 타 기법들에 비해 다소 낮은 수치를 나타내었다. 한편 Adaboost와 ANN, kNN의 경우 AUC 값이 Recall 값보다 다소 높았지만 RF와 SVM의 경우에는 Recall 값이 AUC 값보다 높게 나타났다. 특히 SVM의 경우 Recall 값이 AUC 값에 비하여 약 1.9배 높게 나타났다.

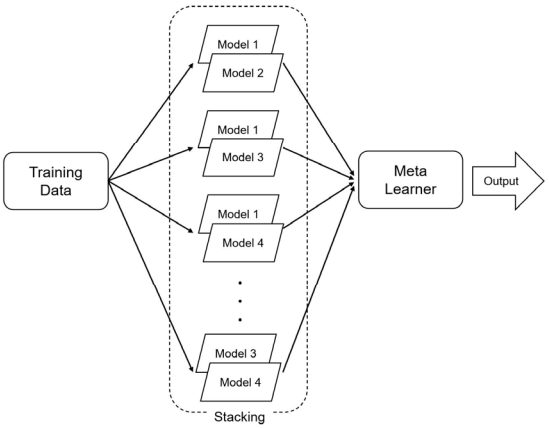

Fig. 4. Stacking example based on underlying machine learning models.

Table 4. Recall and AUC values based on optimized models

Model	Recall	AUC
Adaboost	0.790	0.893
ANN	0.653	0.990
kNN	0.829	0.911
RF	0.955	0.927
SVM	0.865	0.462

각 기법에 최적화를 수행한 결과, Adaboost의 경우 Number of estimators는 50이고, Learning rate가 1.00000, Fixed seed for random generator는 700, Classification algorithm은 SAMME.R, Regression loss function은 Linear일 때 Recall 수치가 가장 높았다. ANN의 경우 Neurons in hidden layers가 100이고, Activation이 Logistic, Solver가 L-BFGS-B, Regularization(α)이 0.0001, Maximal number of iterations가 600일 때 가장 높은 Recall 수치를 보였다. kNN의 경우 Number of neighbors가 2이고, Metric은 Mahalanobis, Weight는 Uniform일 때 가장 높은 수치를 나타냈다. RF의 경우 Number of trees는 10이고, Number of attributes considered at each split는 8, Limit depth of individual trees는 3, Do not split subsets smaller than은 5일 때의 성능이 가장 좋았다. SVM의 경우 Cost는 1, epsilon은 0.1, Kernel 함수는 Polynomial, g는 8, c는 1, d는 3이며, Numerical tolerance는 0.001, iteration limit는 100일 때 가장 높은 수치를 보여주었다.

Fig. 5는 각 머신러닝 기법의 Recall 결괏값에 가장 높은 영향을 미친다고 판단한 파라미터의 변화(최적화)에 따른 Recall 값의 변화를 보여준다. 즉, Fig. 5에서 각 기법별로 최적화를 수행한 때의 Recall 값이 Table 3에 기법별 Recall 값으로 정리된 것이다. Adaboost 기법의 경우 파라미터 중 조정 가능한 옵션 수치의 변화가 Recall 값에 큰 영향을 미치지 않는 것으로 판단하였으며, 이에 따라 상기한 수치에서의 Recall 값이 0.790으로 가장 높게 나타났다(Fig. 5(a)). ANN 기법의 경우 파라미터의 조합에 따른 Recall 값을 기반으로 간단한 선형회귀식을 도출하고, 이를 기반으로 Maximal number of iterations의 변화에 따른 Recall 값을 도시하였다(Fig. 5(b)). kNN 기법의 경우 최적화 수행 과정에서 가장 높은 Recall 값을 보이는 Mahalanobis와 Uniform일 때를 기준으로 Number of neighbors를 증가시킨 결과를 표시하였다(Fig. 5(c)). RF 기법 또한 최상위 평균값을 보이는 변수들을 고정한 후 의사결정나무의 복잡도를 결정하는 Limit depth of individual trees를 증가시킨 모습을 나타내었으며(Fig. 5(d)), SVM 기법의 경우 영향을 미치지 않는다고 판단된 변수들을 배제하고 나머지 파라미터의 변화에 따른 결과를 종합 분석하여 가장 높은 영향력을 미치는 것으로 판단된 c 값의 변화에 따른 Recall 수치를 도시하였다(Fig. 5(e)). 결과적으로 파라미터 변화를 통한 최적화 과정을 통해 에 따른 Recall 수치를 비교해본 결과 각 기법에 따라 최대 0.4에서 최소 0.015까지 상승된 것으

로 분석되었다.

각 머신러닝 기법을 2개씩 조합으로 짝지어 Stacking을 적용하였을 때의 Recall 결과는 Fig. 6과 같다. 기본적으로 Adaboost와 앙상블을 구성한 조합 사례들이 타 조합들에 비하여 높은 성능을 보였으나, 단일 기법을 적용하였을 때(Table 3)와 달리 전체적으로 Recall 값이 하락한 것을 확인할 수 있다. Adaboost와 SVM 기법이 조합된 경우의 Recall 값이 최대치(0.790)로 나타났지만, 이 경우에도 Adaboost 기법 하나만을 적용했을 때와 Recall 값이 같은 것으로 나타났다. 즉, 앙상블 과정을 통해 각 기법 간의 단점이 서로 보완되어 Recall 값이 증가할 것으로 기대한 것과는 상반된 결과라고 볼 수 있다. 이는 앙상블 기법에 사용되는 이론을 이중으로 적용함에 따라 성능이 하락한 것으로 사료된다. 앙상블 기법의 대표적 이론 두 가지는 배깅(bagging)과 부스팅(boosting)이며, 본 연구에 사용된 기법 중 RF의 경우 여러 개의 의사결정나무를 배깅하여 예측을 실행하는 기법

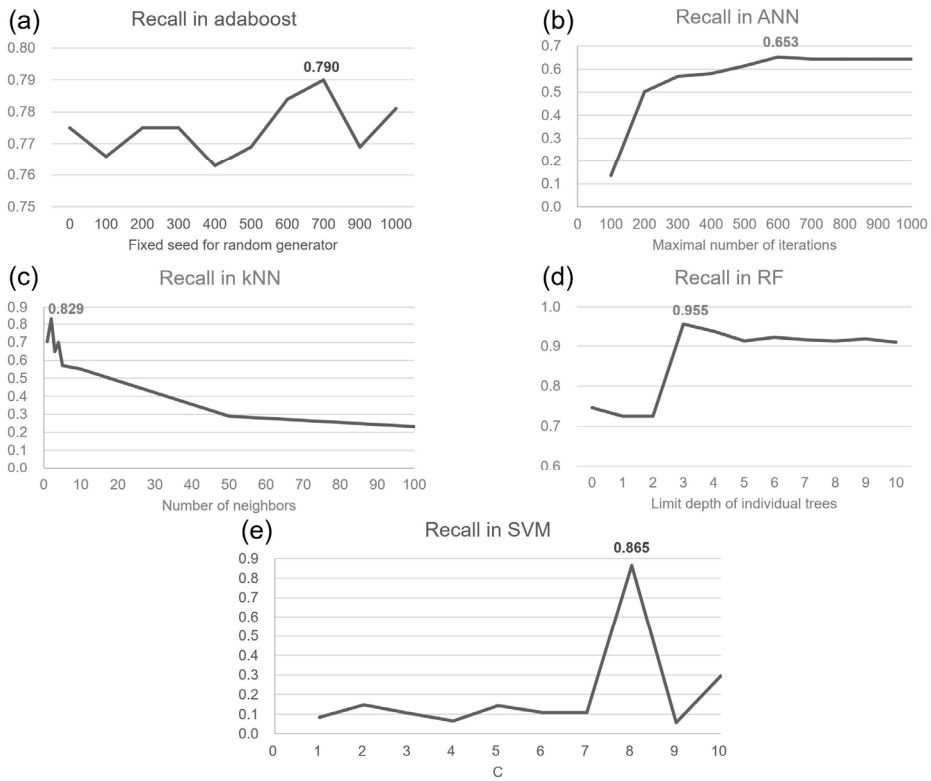

Fig. 5. Variations in recall value based on parameter changes (optimization). (a) Adaboost; (b) ANN; (c) kNN; (d) RF; (e) SVM.

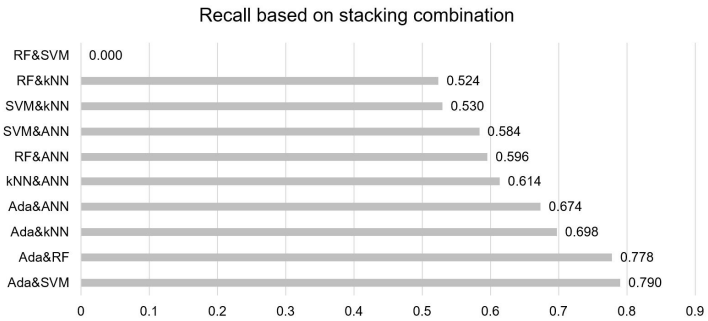

Fig. 6. Recall based on stacking combination of machine learning models.

이다. 또한, Adaboost의 경우 약한 학습기의 오류 데이터에 가중치를 부여하면서 부스팅을 수행하는 기법이며, 이때 해당 기법의 약한 학습기로는 의사결정나무를 사용한다. 이는 결국 앙상블 기법을 진행한 기법들은 중복된 이론이 적용됨에 따라 더 큰 편차를 갖게 되어 성능이 하락한 것으로 사료된다.

토 의

본 연구에서는 기존에 지반침하 발생 예측 성능을 평가하는 지표로 보편적으로 사용되던 ROC 곡선 기반의 AUC를 대신하여 Recall 값에 주안을 두고 머신러닝 기법들의 지반침하 예측 성능을 비교 분석하였다. 분류 예측 성능 지표 중 하나인 ROC 곡선 AUC는 보편적으로 사용되며, 이는 예측 기법이 광산 지반침하가 발생한 지역과 발생하지 않은 지역 전체를 맞게 분류하였는지를 평가한다. 예를 들어, Table 3에서 ANN 기법의 Recall 값은 상대적으로 낮고, AUC 값은 상대적으로 높은 것을 확인할 수 있다. 이 결과는 ANN 기법이 지반침하가 발생된 지역을 정확하게 예측하기보다는 데이터 개수가 많은 지반침하 미발생 지역에 대해 맞게 예측했기 때문이라고 볼 수 있다. 반면 예측 성능 평가 지표 중 본 연구에서 중점을 둔 Recall은 실제 지반침하가 발생한 지역만을 대상으로 이를 얼마나 정확하게 예측하는지를 의미한다. 즉, 지반침하가 발생하지 않은 지역은 예측의 정답 여부를 성능 지표에 반영하지 않는다. 따라서 지반침하가 발생하지 않은 지역까지 포함하는 AUC가 아닌 실제 지반침하 발생 지역을 정확하게 예측하는 것이 중요하다고 판단하였다. 그러한 이유로 본 연구에서는 기 사용되던 ROC 곡선 AUC가 아닌 Recall에 초점을 두어 머신러닝 기법의 예측 성능을 평가하였다.

또한, 혼동행렬 중 기법의 예측 성능을 평가하는 다른 요소인 Precision의 경우 예측 기법이 지반침하가 발생하였다고 판단한 데이터 중 실제 지반침하가 발생한 비율을 나타낸다. 이는 기법에 주안점을 두어 기법이 지반침하가 발생한다고 예측한 데이터 중 실제로 발생하였는지를 평가하는 것이지만, Recall의 경우 실제 지반침하 발생에 초점을 맞추어 기법을 평가한다. 따라서 혼동행렬의 평가 요소인 Precision과 Recall을 비교하였을 때, 기법의 예측 성능을 평가하는 본 연구의 목적에 따라 기법의 재현율을 나타내는 Recall에 초점을 두어 기법의 성능을 평가하는 것이 적합하다고 판단하였다.

본 연구에서 제안한 머신러닝 기법의 예측 성능의 정도를 평가하기 위하여 기존 연구 문헌들과의 비교를 수행하였다. 그러나 다른 연구 논문들은 기법의 성능 평가 기준으로 Recall에 기반한 사례가 희박하여 부득이하게 본 연구에서 Recall 값이 가장 높았던 RF의 AUC를 타 연구들과 비교하였다(Table 5). 연구지역과 사용 데이터의 양, 고려한 지반침하 영향인자 등의 차이는 존재하나 본 연구의 예측 성능(0.927)은 타 연구 논문들의 기법 성능과 유사한 수준을 나타냄을 확인하였다.

결 론

본 연구에서는 광산 지반침하 발생 여부를 평가하기 위해 5개의 머신러닝 모델과 10개의 앙상블 기법을 적용하고, 예측 성능을 비교 분석하였다. 각 머신러닝 기법의 옵션과 파라미터를 최적화하고 분류 예측 성능 지표로써 Recall 수치에 중점을 두고 기법을 평가하였다. 분석 결과 전체 15개 예측모델 중 RF 기법이 가장 높은 성능(0.955)을 보였으며 이는 실제 지반침하 발생 지역 중 95.5%의 발생지를 올바르게 예측한 것으로 해석할 수 있다.

다만 본 연구결과는 연구지역과 사용된 데이터의 종류에 따라 달라질 수 있으므로 모든 광산에서 지반침하 예측 기법으로 RF 모델이 가장 적합하다고 볼 수는 없다. 따라서 타 광산지역에서의 머신러닝 기법의 적용성을 평가하기 위해서는 추가적인 연구가 필요할 것으로 사료되며 본 연구와의 비교를 통하여 범용적인 지반침하 발생 예측 모델을 개발하는 것이 필요하다고 판단된다. 또한, 본 연구에 사용된 지반침하 영향인자들에 더하여 해당 지역의 지질학적, 지형학적, 암석학적 특징들을 더욱 상세히 반영하는 자료들이 확보될 경우 예측 모델의 성능을 개선하기 위한 다양한 연구와 노력이 필요할 것이다.

Table 5. Comparison of prediction performance between existing studies and this study

	Prediction algorithm	AUC
Mohammady *et al*. (2019)		0.770
Ebrahimy *et al*. (2020)		0.798
Pourghasemi and Saravi (2019)	RF	0.930
Sekkeravani *et al*. (2022)		0.953
This study		0.927

사 사

이 논문은 2021년도 정부(산업통상자원부)의 재원으로 해외자원개발협회의 지원을 받아 수행된 연구임(데이터사이언스 기반 석유·가스 탐사 컨소시엄).

References

Ahatonovic-Kustrin, S. and Beresford, R., 2000. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research, *Journal of Pharmaceutical and Biomedical Analysis*, 22(5), p.717-727.

An, T.K. and Kim, M.H., 2010. A New Diverse AdaBoost Classifier, *2010 International Conference on Artificial Intelligence and Computational Intelligence*, IEEE, Sanya, China, p.359-363.

Breiman, L., 2001. Random forests, *Machine Learning*, 45(1), p.5-32.

Bui, D.T., Shahabi, H., Shirzadi, A., Chapi, K., Pradhan, B., Chen, W., Khosravi, K., Panahi, M., Ahmad, B.B., and Lee, S., 2018. Land subsidence susceptibility mapping in south korea using machine learning algorithms, *Sensors*, 18(8), 2464p.

Choi, J.K., Kim, K.D., Lee, S., and Won, J.S., 2010. Application of a fuzzy operator to susceptibility estimations of coal mine subsidence in Taebaek City, Korea, *Environmental Earth Sciences*, 59(5), p.1009-1022.

Choi, J.K., Kim, K.D., Lee, S., Kim, I.S., and Won, J.S., 2007. Prediction of ground subsidence hazard area using GIS and probability model near abandoned underground coal mine, *Economic and Environmental Geology*, 40(3), p.295-306.

Coal Industry Promotion Board, 2005. *Detailed Investigation Report on the Ground Stability of Jeongseon Jeongam Area*, p.2-9.

Ebrahimy, H., Feizizadeh, B., Salmani, S., and Azadi H., 2020. A comparative study of land subsidence susceptibility mapping of Tasuj plane, Iran, using boosted regression tree, random forest and classification and regression tree methods, *Environmental Earth Sciences*, 79(10), p.1-12.

Hakim, W.L., Achmad, A.R., and Lee, C.W., 2020. Land subsidence susceptibility mapping in Jakarta using functional and meta-ensemble machine learning algorithm based on time-series InSAR data, *Remote Sensing*, 12(21) 3627p.

Jakkula, V. 2006. *Tutorial on support vector machine (svm)*, School of EECS, Washington State University, 37(2.5), 3p.

Jung, D.H. and Choi, Y.S., 2021. Prediction of blast vibration in quarry using machine learning models, *Tunnel and Underground Space*, 31(6), p.508-519.

Kim, K.D., Lee, S., Oh, H.J., Choi, J.K., and Won, J.S., 2006. Assessment of ground subsidence hazard near an abandoned underground coal mine using GIS, *Environmental Geology*, 50(8), p.1183-1191.

Kim, K.D., Lee, S.R., and Oh, H.J., 2009. Prediction of ground subsidence in Samcheok City, Korea using artificial neural networks and GIS, *Environmental Geology*, 58(1), p.61-70.

Koo, Y.H., Kim, S.M., Oh, M., and Park, H.D., 2019. Estimation of solar irradiance at weather stations in Korea using regionally trained artificial neural network models, *Journal of the Korean Society of Mineral and Energy Resources Engineers*, 56(2), p.155-171.

Kratzsch, H., 1983. *Mining Subsidence Engineering*, Springer, Berlin, Germany, p. 543.

Lee, S. and Park, I., 2013. Application of decision tree model for the ground subsidence hazard mapping near abandoned underground coal mines, *Journal of environmental management*, 127(1), p.166-176.

Lee, S., Park, I., and Choi, J.K., 2012. Spatial prediction of ground subsidence susceptibility using an artificial neural network, *Environmental Management*, 49(2), p.347-358.

Lobo, J.M., Jiménez-Valverde, A., and Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models, *Global ecology and Biogeography*, 17(2), p.145-151.

Mine Reclamation Corporation, 2019. National mine statistics (Current status by mining species), 2022.04.13., http://www.komir.or.kr

Mohammady, M., Pourghasemi, H.R., and Amiri, M., 2019. Land subsidence susceptibility assessment using random forest machine learning algorithm, *Environmental Earth Sciences*, 78(16), p.1-12.

Nam, K.H., Lee, M.J., and Jeong, G.C., 2016. Landslide susceptibility analysis: SVM application of spatial databases considering clay mineral index values extracted from an ASTER satellite image, *The Journal of Engineering Geology*, 26(1), p.23-32.

Oh, H.J. and Lee, S., 2010. Assessment of ground subsidence using GIS and the weights-of-evidence model, *Engineering Geology*, 115(1-2), p.36-48.

Oh, H.J. and Lee, S., 2011. Integration of ground subsidence hazard maps of abandoned coal mines in Samcheok, Korea, *International Journal of Coal Geology*, 86(1), p.58-72.

Oh, H.J., Ahn, S.C., Choi, J.K., and Lee, S., 2011. Sensitivity analysis for the GIS-based mapping of the ground subsidence hazard near abandoned underground coal mines, *Environmental Earth Sciences*, 64(2), p.347-358.

Oh, H.J., Syifa, M., Lee, C.W., and Lee, S., 2019. Land subsidence

susceptibility mapping using bayesian, functional, and meta-ensemble machine learning models, *Applied Sciences*, 9(6), 1248p.

Park, I., Choi, J., Lee, M J., and Lee, S., 2012. Application of an adaptive neuro-fuzzy inference system to ground subsidence hazard mapping, *Computers & Geosciences*, 48(1), p.228-238.

Park, I., Lee, J., and Lee, S., 2014. Ensemble of ground subsidence hazard maps using fuzzy logic, *Central European Journal of Geosciences*, 6(2), p.207-218.

Pourghasemi, H.R., and Saravi, M.M., 2019. Land-subsidence spatial modeling using the random forest data-mining technique, *Spatial Modeling in GIS and R for Earth and Environmental Sciences*, 6, p.147-159.

Sekkeravani, M.A., Bazrafshan, O., Pourghasemi, H.R., and Holisaz, A., 2022. Spatial modeling of land subsidence using machine learning models and statistical methods, *Environmental Science and Pollution Research*, 29, p.28866-28883.

Son, J., Suh, J., Yi, H., Park, H.D., and Lee, S., 2015. GIS-based subsidence hazard analysis on abandoned coal mine sites combining the frequency ratio and radius of influence, *Journal of the Korean Society of Mineral and Energy Resources Engineers*, 52(6), p.567-576.

Suh, J., Choi, Y., and Park, H.D., 2016. GIS-based evaluation of mining-induced subsidence susceptibility considering 3D multiple mine drifts and estimated mined panels, *Environmental Earth Sciences*, 75(10), p.1-19.

Suh, J., Choi, Y., Park, H.D., Kwon, H.H., Yoon, S.H., and Go, W.R., 2010. Application of frequency ratio and analytic hierarchy process to subsidence hazard assessment around abandoned coal mine, *Journal of the Korean Society for Geosystem Engineering*, 47(5), p.690-704.

Suh, J., Choi, Y., Park, H.D., Yoon, S.H., and Go, W.R., 2013. Subsidence Hazard Assessment at the Samcheok Coalfield, South Korea: A Case Study Using GIS, *Environmental and Engineering Geoscience*, 19(1), p.69-83.

Waltham, T., Park, H.D., Suh, J. Yu, M.H., Kwon, H.H., and Bang, K.M., 2011. Collapses of old mines in Korea, *Engineering Geology*, 118, p.29-36.

Yeon. K., 2020. Stacking based ensemble with instance selection in random KNN, *Journal of The Korean Data Analysis Society*, 22(4), p.1303-1312.

Zhang, S., Li, X., Zong, M., Zhu, X., and Cheng, D., 2017. Learning *k* for knn classification, *ACM Transactions on Intelligent Systems and Technology (TIST)*, 8(3), p1-19.

한 호 상

2022년 강원대학교 공학대학 에너지자원융합공학전공 공학사

현재 강원대학교 에너지자원융합공학과 석사과정
(E-mail; qkfehr@gmail.com)

최 교 익

2021년 강원대학교 공학대학 에너지자원융합공학전공 공학사

현재 강원대학교 에너지자원융합공학과 석사과정
(E-mail; rydlr3020@naver.com)

서 장 원

현재 강원대학교 에너지공학부(에너지자원융합공학전공) 조교수
(本 學會誌 第58券 第6号 參照)