Geological Quarterly, 2025, 69, 14
DOI: https://doi.org/10.7306/gq.1787

Assessing the potential of secondary raw materials from hard coal and iron ore mining waste disposal sites using machine learning

Joanna FAJFER[1, *], Michał ROLKA[2] and Paulina KOSTRZ-SIKORA[1]

[1] Polish Geological Institute – National Research Institute, Rakowiecka 4, 00-975 Warszawa, Poland; ORCID: 0000-0002-7382-5383 [J.F.], 0000-0001-7633-9930 [P.K.-S.]

[2] Polish Geological Institute – National Research Institute, Upper Silesian Branch, Królowej Jadwigi 1, 41-200 Sosnowiec, Poland

Fajfer, J., Rolka, M., Kostrz-Sikora, P., 2025. Assessing the potential of secondary raw materials from hard coal and iron ore mining waste disposal sites using machine learning. Geological Quarterly, 69, 14; https://doi.org/10.7306/gq.1787

Associate Editor: Karol Zglinicki

Machine learning is the science of how algorithms and systems improve knowledge and performance through experience. The paper discusses the possibility of applying machine learning algorithms (an AI subfield) to assess the secondary raw material potential of mining waste and tailings accumulated in hard coal and iron ore disposal sites. Applied machine learning algorithms are available for the Python language. Classification algorithms were used only (supervised machine learning). The first stage of studies included defining the category and features of the objects selected. The learning process used training data, whereas test data were applied to check the model efficiency. Analysis of the results obtained indicates that machine learning seems a promising tool to assess the secondary raw materials potential of waste accumulated in dumps and heaps, despite achieving a prediction accuracy at the level of 0.75 for hard coal objects and 0.85 for iron ore objects. It has been assumed that Gaussian NB, GaussianProcessClasifier and MLP Classifier algorithms of supervised learning show the highest prediction results. This suggests that machine learning may be used as a tool supporting the decision process, the result of which will be the economic use of waste accumulated in heaps, dumps and disposal sites.

Key words: machine learning algorithms, mining waste, secondary raw material potential, disposal sites.

INTRODUCTION

Industrial development has resulted in the increased production and exploitation of natural raw materials, and has also produced a large amount of waste both in Poland and Europe. This waste has been mostly managed through accumulation on waste dumps, heaps, tailing ponds and industrial landfills (Garbarino et al., 2020; Fajfer and Kostrz-Sikora, 2022a). Changing legal provisions and increased ecological awareness have steered waste management toward the circular economy model, in which waste, including that deposited in the environment, may serve as a valuable source of secondary raw materials (Nieć, 1999; Kozioł and Kawalec, 2008; Bellenfant et al., 2013; Danthurebandara et al., 2015; Falagán et al., 2017; Careddu et al., 2018; Spoorena et al., 2020; Araya et al., 2021; Dino et al., 2023).

Utilization of waste accumulated in heaps, dumps and industrial landfills has been carried out in Poland since the 1970s (e.g. Paluch and Morawski, 1975; Łukaszyk and Tabor, 1979; Nowok and Skiba, 1979), but it was not until the late 1990s that this direction began to develop rapidly (Kozioł and Uberman, 1996; Makowski and Fajfer, 1997; Fajfer and Kostrz-Sikora, 2022a). The use of material accumulated at these sites has led to the amount of waste in the environment decreasing by ~15% compared to 2000 (from 2.1 billion (metric) tons to 1.6 billion tons). The area occupied by these sites also decreased by ~34% (from 12,000 hectares to just over 8,000 hectares) (GUS, 2001; GUS, 2023; Environment 2023, 2023).

A similar trend has also been observed in the European Union. Analysis of data from 2004-2020 has indicated that the amount of waste deposited in industrial landfills, heaps and dumps decreased by ~35%, from 963.5 billion tons to 634 billion tons (Waste Statistic, 2020).

Waste accumulated in heaps, dumps and disposal sites (in particular from the mining, steel and energy industries) has wide application in substituting for natural resources in various branches of the economy (Fig. 1), for instance, in engineering construction (Rosik-Dulewska and Karwaczyńska, 2008; Birnemanns et al., 2015; Różański, 2019; Fisher and Barron, 2019; Cobîrzan et al., 2022), hydraulic engineering (Potempa and Szlugaj, 2007: Rosik-Dulewska and Karwaczyńska, 2008; Ratajczak et al., 2015; Fisher and Barron, 2019), road construction (Han and Johnson, 1995; Kozioł and Uberman, 1996; Wowkonowicz et al., 2018; Segoui et al., 2023), mining (Palarski

* Corresponding author, e-mail: joanna.fajfer@pgi.gov.pl

Received: December 3, 2024; accepted: April 17, 2025; first published online: July 8, 2025

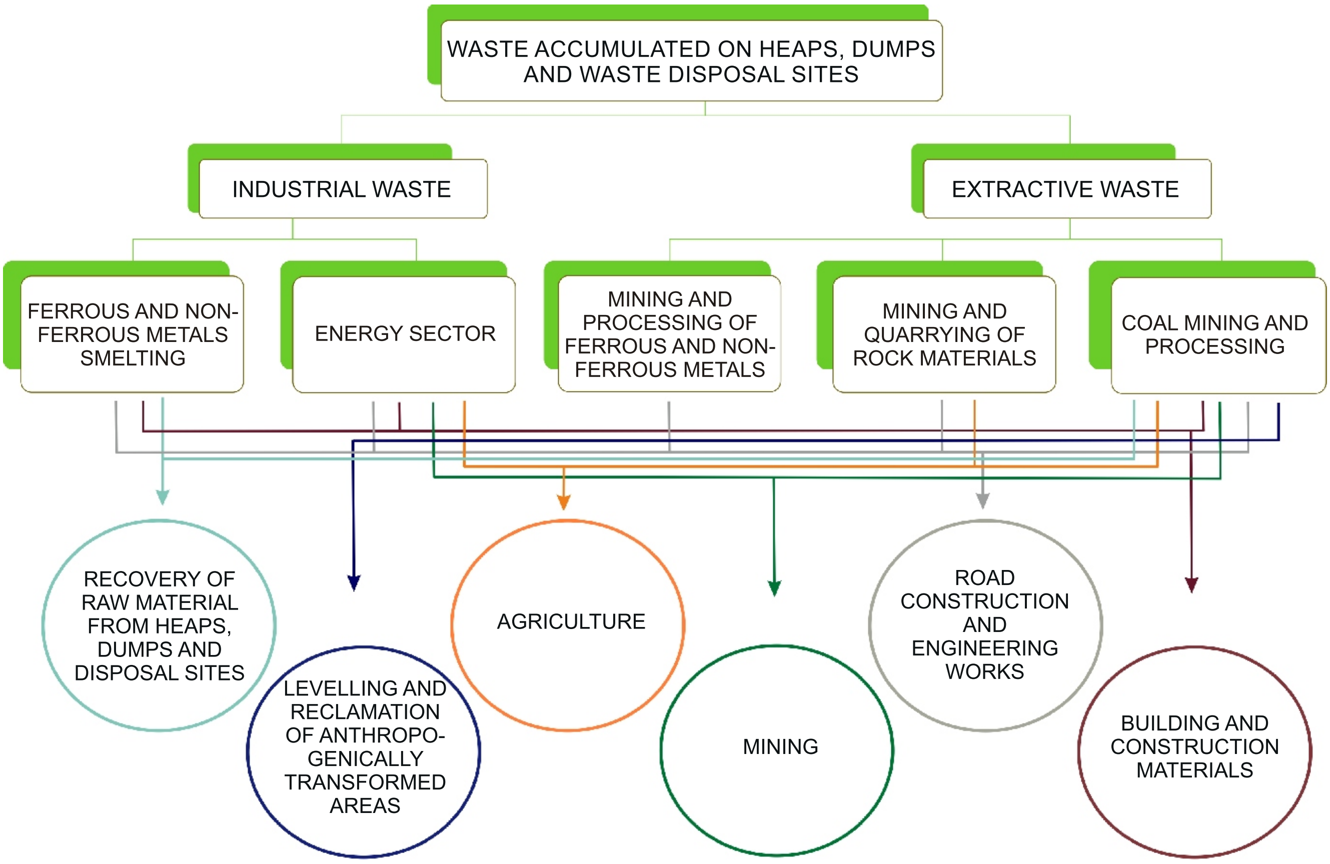


Fig. 1. Fields of application of mining and industrial waste accumulated in heaps, dumps and industrial landfills

et al., 1987; Plewa et al., 2005; Pozzi and Mzyk, 2007; Kinnunen et al., 2018), agriculture (Pietras, 1979; Ratajczak and Korona, 2020), and in levelling and reclamation of anthropogenically transformed areas (Ratajczak et al., 2015; Kozioł et al., 2015; Fajfer et al., 2022). Waste from the mining of non-ferrous minerals, including that from non-ferrous metallurgy, is a potential source of metallic raw materials (Lucarini et al., 2020; Lim and Alorro, 2021; Nwaila et al., 2021; Oliveira et al., 2022; Adrianto et al., 2023; Rosario-Beltré et al., 2023), which are presently highly desirable in industry.

In the past, the usage of material accumulated in heaps, dumps and industrial landfills resulted mainly from the fact that the waste had to fulfill technical requirements indispensable for the technologies available at that time. In effect, some heaps, dumps and other landfills were reclaimed and the accumulated material was used as a product. The sites in which the accumulated waste did not fulfill these expected conditions were reclaimed and integrated into the landscape, to perform various functions for natural, recreational, and touristic purposes. At present, due to better developed recovery methods, they are considered as a potential source of secondary raw materials (Blengini et al., 2019; Fajfer and Kostrz-Sikora, 2022b; Fajfer et al., 2022) with a considerable economic value, e.g. the value of potentially recoverable metallurgical slags in the old Pagida smelter landfill in Romania has been estimated at 1175.7440 million USD (Ilutiu-Varvara and Aciu, 2022); the value of metal mine tailings accumulated in landfills in Finland has been estimated at several hundred million euros (Kinnunen et al., 2022), and the raw material value of rock waste from kaolinite mining accumulated in old dumps in Italy has been assessed at ~1 333.800 million euros (Dino et al., 2021).

Waste extraction from various sites (including reclaimed disposal sites) is a complex issue as regards technological and investment-related, as well as environmental and social, aspects. It requires analysis of archival data (if available), obtaining relevant administrative permits, performing raw data analysis of accumulated waste, selecting appropriate technologies, and meeting the conditions of economic viability of waste extraction.

Application of machine learning algorithms may be used as support for the decision-making process with regard to predicting the raw material potential of sites in which mining waste and tailings were accumulated. Machine learning is the science of algorithms and systems improving their knowledge and performance with experience (Flach, 2012). Its essence is to create an automatic system (algorithm), which can learn and improve based on the accumulated data (Sarker, 2021a). The task of algorithms in the learning process is finding hidden relations among a large amount of data, which will directly and/or indirectly have impact on the solution of a given problem. Depending on the available data and expected results, regression, classification or clustering are used in the machine learning process (Fig. 2). In this paper, classification algorithms were used only. The aim of the regression and classification process is to predict the values of the decision attribute based on the values of the remaining attributes. In case of classification, the decision attribute has a discrete value (binary, nominal) (Szeliga, 2019; Król-Nowak and Kotarba, 2022).

Algorithms of machine learning are now utilized in almost every branch of life and science, and the possibilities of their practical application are continuously increasing. In geology and related disciplines, their training is performed, e.g. in order

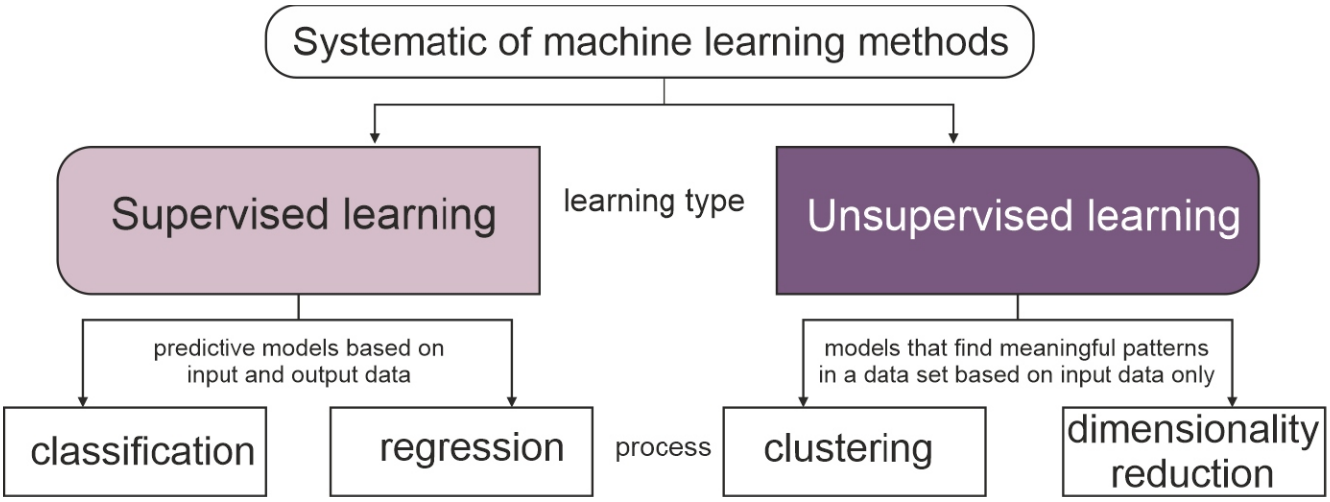


Fig. 2. Systematics of machine learning methods and processes

to recognize and characterize geological structures related with specific types of ore deposits (e.g. Male and Duncan, 2020; Topór, 2021), but also to predict geophysical phenomena causing natural disasters such as earthquakes (e.g. Ridzwan and Yusoff, 2023) or volcanic eruptions (e.g. Malfante et al., 2018). Other examples involving the usage of machine learning include, for example, predicting terrain subsidence in areas transformed by mining activities (e.g. Ambrožič and Turk, 2003), improvement of processes aimed at the recovery of some natural resources, e.g. copper through leaching (Flores and Leiva, 2021), and assessment of the impact posed by mining waste on the environment (e.g. due to the release of toxic elements) (e.g. Trifi et al., 2022; Chongchong et al., 2023). Artificial Intelligence tools – beside solving numerous complex problems that appear in the extraction and processing of ores – may also find application in the optimization of waste management processes and supply reliable data on their quality parameters. With regard to this issue, solutions were proposed by Herrera et al. (2023), although, after a wide literature search, these authors emphasized that the problem requires further analyses.

Despite the wide range of publications noted above, devoted to the utility of machine learning in solving problems related to geology and mining, these did not include the use of machine learning techniques aimed at recognition of the secondary raw material potential of waste derived from mining activities accumulated in heaps, dumps and disposal sites. Presented research therefore assesses the secondary raw material potential of these sites using machine learning. This method has numerous advantages, of which the minimization of fieldwork costs is the most significant one. It may indicate sites with secondary raw material potential, which should help the decision-making process in selecting a particular site for further economic and environmental analyses. The process is not time-consuming or restricted by meteorological conditions by comparison with traditional assessment methods, and does not require significant amounts of funding. The machine learning algorithms applied (including defined categories and ascribed features) may be used to assess the potential of other groups of mining waste disposal facilities.

STUDY AREA

Mining waste disposal sites are characterized by variable surface areas and large variations in the quality and quantity of accumulated waste resulting, for instance, from the technological processes applied during raw material mining and processing. The range of data describing the particular facilities is also variable.

In the frame of activities performed by PIG-PIB in 2012, 368 mining waste facilities (active and historical) were inventoried in which potential secondary raw materials were accumulated. At 252 of these facilities, the accumulated waste was derived from mining and processing of metallic raw materials, at 104 the waste was from mining and processing of energy-based raw materials, and at 12 of them the waste was from mining and processing of chemical raw materials (Kostrz-Sikora et al., 2013). Earlier data had documented 350 industrial landfills, in which the accumulated waste was generally from the energy, metallurgy, metal and machinery industries (KPGO, 2002). In turn, the most recent data indicate the presence of 148 mining waste facilities (KPGO, 2023).

The selection of sites for study was made based on earlier research performed by PIG-PIB (Kostrz-Sikora et al., 2013; Fajfer and Kostrz-Sikora, 2022b; Baza Hałdy, 2024), available published data (Ratajczak, 1998; Ratajczak and Korona, 2000), archival material, and expert knowledge. Of particular interest were objects with historical exploitation of raw materials, because the technologies used in the past were less precise/advanced and produced more waste. In effect, the waste accumulated on dumps and landfills in the past may contain secondary raw materials that may be substitutes for natural resources. Inactive, reclaimed facilities, as well as active facilities in which waste is accumulated at present, were analysed during the study.

Industrial waste landfills were removed from consideration due to the wide range of waste types, restricted methods of application resulting from the specific chemical composition of the waste created during industrial processes, and the limited available archival data.

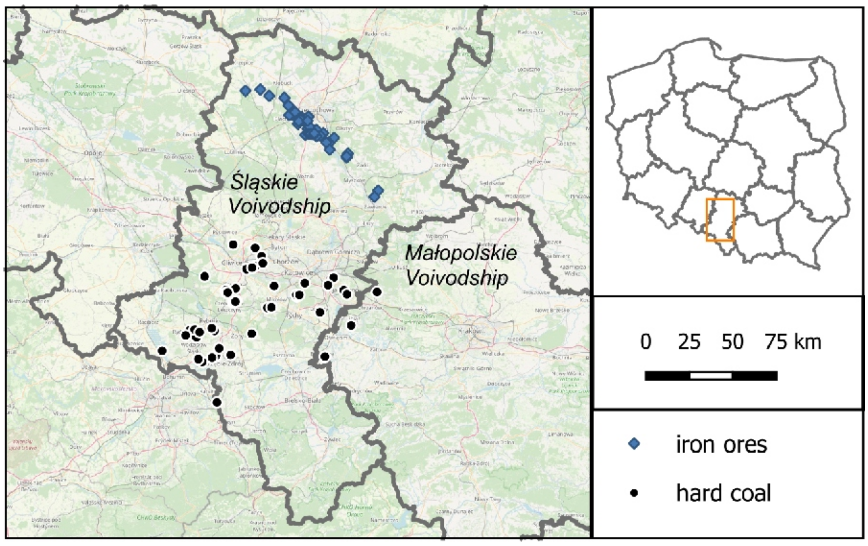


Fig. 3. Location of the sites analysed to test their secondary raw material potential

In effect, two groups of sites were selected for detailed studies (due to the large amount of available data): sites with waste from mining and processing of hard coal mining (43) and sites with waste from mining and processing of iron ore (40). The sites include those that are active, closed-down, and reclaimed, and are located in the Silesian and Małopolska voivodships (Fig. 3). Approximately 654.5 million tons of waste accumulated in the sites selected for the study.

METHODS

Assessement of the secondary raw material potential of sites containing waste from hard coal and iron ore mining, and processing by the application of machine learning, was performed according to the following procedure (Fig. 4):

DATA AVAILABILITY AND COLLECTION

The first stage of studies included the definition of criteria and characteristics describing the facilities selected. Due to the large variability of the sites (e.g. different sources of waste, variable waste chemistry, different technologies of extraction and processing of raw materials), it was not possible to assign a fixed and invariant set of features describing the criteria for each site. In the end, the facilities were described by seven criteria:

1. Source of waste (described by category-related features, e.g. underground mining, surface mining, type of waste: mining waste, tailings, waste from raw material processing).
2. Type of facility (described by category-related features, e.g.: complete – incomplete, historical – modern, closed – operated, non-reclaimed – reclaimed).

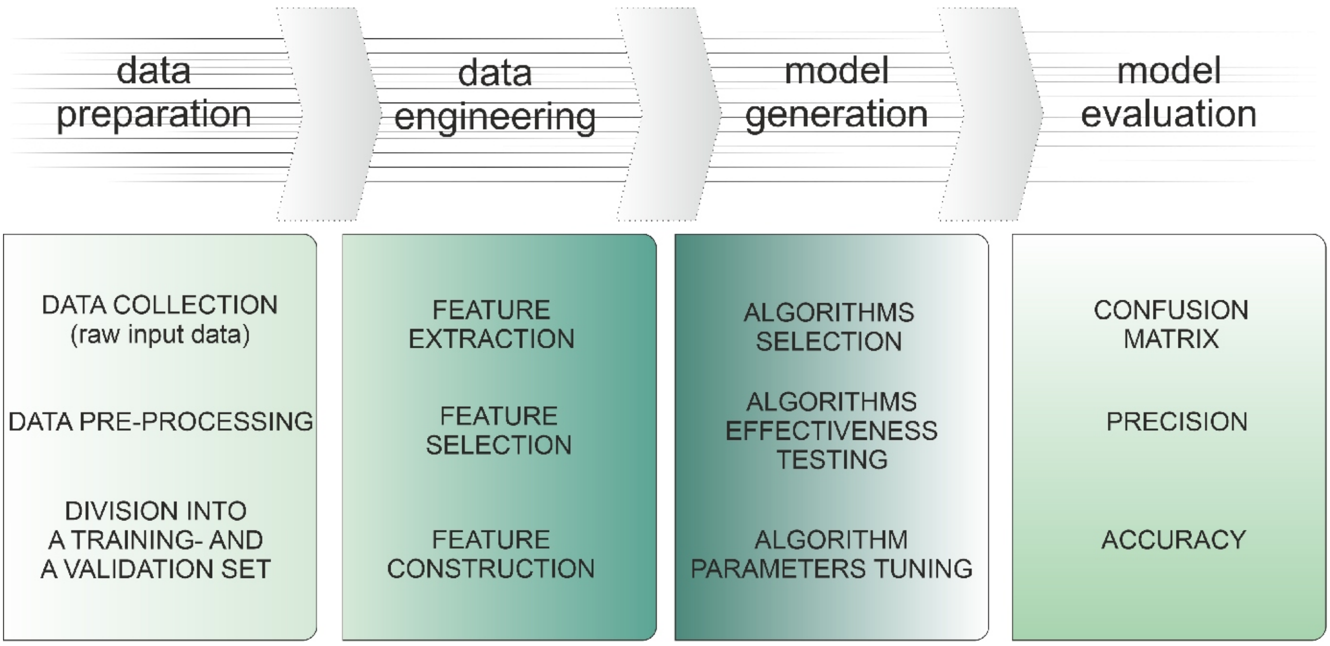


Fig. 4. The machine learning procedure used to assess the secondary raw material potential of sites with mining and processing waste

3. Characteristics of the site (described by numerical features, e.g. area, amount of accumulated waste, time of waste deposition, spontaneous combustion).
4. Waste characteristics (described by category-related/descriptive features, e.g.: chemical composition).
5. Development of the deposit (described by category-related features, e.g. management, type of mining resource, form of deposit).
6. Utilization of waste accumulated at the facilities (described by category-related features, e.g.: current utilization, past utilization, directions of utilization, performed suitability studies in each direction).
7. Existing functions of the site relative to the environment and society (described by category-related features, e.g. cultural heritage, tourism potential, recreation and sports potential, protected areas).

76 or 105 features were attributed to particular categories respectively. A larger number of features (105) was defined for disposal sites in which waste from hard coal mining and processing was accumulated. This waste was generated and stored in the largest amounts. The waste is characterized by parameters which allow for their wide application in various branches of industry while also being significant barriers as regards these purposes. In turn, a lower number of features (76) was indicated for facilities accumulating waste from the mining and processing of iron ore. These sites have been clearly recognized and characterized in the past. Mining of iron ore terminated in the 1980s, and so archival data were used in the definition of their categories and features (Ratajczak, 1998).

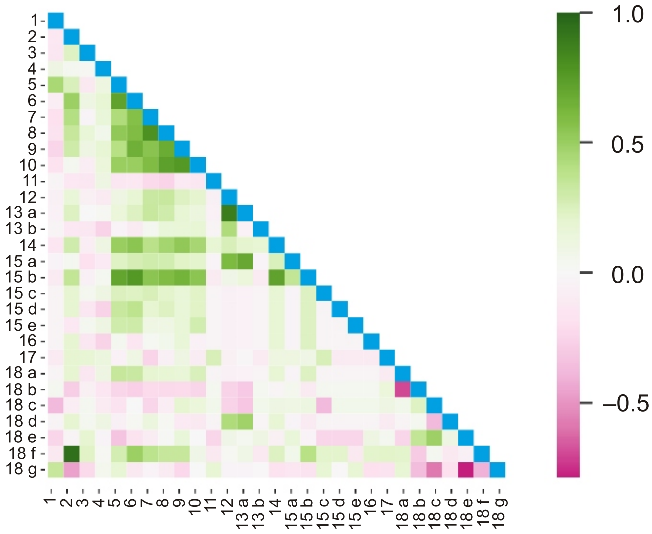
The datasets were balanced (for iron ore sites class frequencies are 22/18 and for hard coal sites, 25/18). Based on possibility of waste recovery as regards profitability (e.g. sites complete, a large amount of accumulated waste, chemical analyses available, no fire hazards) and their other functions in society (e.g. cultural heritage, UNESCO sites) the class labels were determined.

DATA PRE-PROCESSING

At this stage, the raw data were processed to a form suitable for machine learning algorithms (Grus, 2018). The input raw datasets contained data in the form of numerical data (single numbers and their ranges), text categories (yes/no/missing data, etc.), descriptive categories (waste lithology, mineral type, age). Numerical data were left unchanged. The number ranges were separated into two positions (from-to). Columns containing text descriptions of the facilities (e.g. object name) were removed. These data were retained in the source table to enable subsequent addition of descriptions to the output data. Numerical values as integers were ascribed to the text features. Moreover, features corresponding to 'missing data' were replaced by a fixed numerical value (-1). Multi-value descriptive categories were separated using OneHotEncoding, i.e., additional categories were created for each value occurring in a given feature with values 1 or 0 representing the occurrence or non-occurrence of a given category value.

The next step in initial data processing was analysis of variance in each column performed using the VarianceThreshold function from the Python scikit-learn package. At this stage, columns with low variance (<0.05) of the data were also removed from the analysis.

Because the occurrence of mutually correlated data columns can reduce the accuracy of learning, the dataset prepared was subject to correlation analysis, using tools available in the pandas library for Python (Raschka and Mirjalili, 2021; Pandas, 2024).



1 – method of mining operation; 2 – type of waste; 3 – rehabilitation; 4 – method of using the facility; 5 – the area; 6 – amount of waste; 7 – start of waste disposal; 8 – end of waste disposal; 9 – depth of the deposit; 10 – depth of deposit exploitation; 11 – Ni content in waste; 12 – past use of waste from the facilities; 13 – past direction of use of waste: a – cement production, b – construction ceramic; 14 – waste use in the present; 15 – current direction of use of waste: a – cement production, b – construction ceramic, c – construction, d – aggregate production, e – others; 16 – waste analysis in the past; 17 – the cultural and landscape potential; 18 – litho logy: a – clays, b – waste from the sorting process, c– sands, d – sandstones, e – clay sandstones, f – siderites, g – others

Fig. 5. Correlation heat map of the input criteria for facilities with waste from iron ore mining and processing (p-value<0.05)

This process restricts the dimension of the input data and the demand for computing power. A correlation matrix was created for each pair of data columns using Pearson's linear correlation. For each pair of columns positively or negatively correlated at the level of 100%, only one was retained. The correlation between data for facilities with iron ore mining and processing waste is visualized as a 'thermal map' (Fig. 5).

Additionally, an initial assessment of utilizing the accumulated waste, based on expert knowledge, was made for facilities where hard coal and iron ore mining and processing waste was stored. Initially, the input data for tests of supervised learning were subdivided into dissociated datasets: learning (75% input data) and test (25% input data). However, in order to avoid problems linked with the relatively small size of the input dataset during the tests, in further proceedings a cross-validation with the stratified k-fold method was used, with k=4, 6 and 8; this separates the input dataset into a preset number of parts and uses all combinations of these parts as the learning and test datasets, so all records from the input dataset are used at least once in the test set. Moreover, using the stratified k-fold method assures equal representation of each target class in test set.

Table 1

An overview of the classification algorithms used

Group of ML algorithms	ML algorithm	Description
Tree-Based Ensembles	AdaBoostClassifier, GradientBoostingClassifier, RandomForestClassifier, ExtraTreesClassifier	These algorithms combine multiple decision trees to create strong predictive models, reducing overfitting and improving accuracy through ensemble techniques.
	DecisionTreeClassifier, ExtraTreeClassifier	These algorithms create tree-like structures to make decisions based on feature values, with ExtraTreeClassifier adding randomness for further diversification.
Support Vector Machines (SVMs)	LinearSVC, NuSVC, SVC	These algorithms find optimal hyperplanes to separate data into classes, with variations in kernel functions and parameterization for linear and non-linear boundaries.
Generalized Linear Models	RidgeClassifier, SGDClassifier, Perceptron	These algorithms use linear models with regularization or iterative optimization to classify data, suitable for linearly separable problems or large datasets.
Neural Networks	MLPClassifier	This algorithm uses multi-layer perceptrons, a type of neural network, to learn complex non-linear relationships between features and target variables.
Gaussian Processes and Naive Bayes	GaussianProcessClassifier	This algorithm uses Gaussian processes to model the probability distribution of the target variable, providing probabilistic predictions.
	GaussianNB, ComplementNB, BernoulliNB	These algorithms apply Bayes' theorem with different assumptions about feature distributions (Gaussian, multinomial, or Bernoulli) for probabilistic classification.

APPLICATION OF THE MACHINE LEARNING APPROACH

Algorithms of supervised machine learning from the Python scikit-learn package were used to predict the parameters influencing the secondary raw material potential of untested post-mining sites. Due to the fact that the problem analysed can be brought down to object classification such as: yes/no, fulfils/does not fulfil, 0/1, only classifier algorithms were used for supervised learning. Supervised learning algorithms (Table 1) used in the first stage represent different models of machine learning, from simple linear models (RidgeClassifier, SGDClassifier, Perceptron) through decision trees (DecisionTreeClassifier, ExtraTreeClassifier) operating on the principle of a series of conditions, Naive Bayesian classifiers (GaussianNB, ComplementNB, BernoulliNB) that assume mutual independence of variables, the support vector method – SVM (LinearSVC, NuSVC) that separates objects belonging to different classes in multidimensional space, the Nearest Neighbors method (KneighborsClassifier, RadiusNeighborsClassifier), and the multilayer perceptron network (MLPClassifier) classified as a neural network. Algorithms that enhance the performance of a single algorithm (ExtraTreesClassifier, AdaBoostClassifier, RandomForestClassifier, GradientBoostingClassifier) were also applied, as well as algorithms that compose results from several component algorithms, such as the VotingClassifier algorithm and the StackingClassifier algorithm that treat component results as input data for the final estimator (Géron, 2018; Scikit-learn, 2024).

TRAINING AND EVALUATION

A training dataset was used to construct the machine learning model, and later, in order to test the model's efficiency, tests were performed with the test dataset. The process required several iterations until acceptable results were obtained (e.g. cross-validation). In order to remove the potential negative impact of data variability, tests were performed on both raw data

Table 2

An overview of the scaling algorithms used

Scaling Algorithm	Description
Normalizer	Scales each sample (row) to the unit norm, useful when the direction of the data is important, not its magnitude.
StandardScaler	Standardizes features by removing the mean and scaling to unit variance, commonly used in algorithms that assume a normal distribution of data.
MinMaxScaler	Scales features to a specific range (default 0–1), useful when we need to constrain data to a specific interval.
MaxAbsScaler	Scales each feature so that its maximum absolute value is 1, useful for sparse data or when we want to preserve zeros.
RobustScaler	Scales features using median and quartiles, robust to outliers in the data.
QuantileTransformer	Transforms features to have a quantile distribution, which can be useful when data has non-linear distributions.

Table 3

Results of machine learning tests using supervised algorithms for facilities with waste from hard coal mining and processing

Group of ML algorithms	ML algorithm	Scaling algorithms							
		No scaling	Normalizer	Standardscaler	Maxabsscaler	Minmaxscaler	Powertransformer	Quantile-transformer	Robustscaler
Decision Trees	DecisionTreeClassifier	0.49	0.57	0.54	0.49	0.49	0.51	0.54	0.49
	ExtraTreeClassifier	0.51	0.59	0.60	0.65	0.65	0.54	0.54	0.58
Ensemble	AdaBoostClassifier	**0.70**	0.49	0.65	0.65	0.68	0.65	**0.70**	0.65
	ExtraTreesClassifier	**0.70**	0.52	0.65	**0.70**	**0.70**	0.68	**0.72**	**0.70**
	GradientBoostingcClassifier	0.68	0.57	0.68	0.68	0.68	0.68	0.68	0.65
	RandomForestClassifier	**0.70**	0.54	0.68	0.68	0.65	**0.72**	0.68	**0.70**
GaussianProcess	GaussianProcessClassifier	0.63	0.51	0.59	0.63	0.68	0.44	**0.75**	0.40
Linear	Perceptron	0.50	0.40	0.58	0.67	0.65	0.44	0.68	0.58
	RidgeClassifier	0.46	0.49	0.46	0.54	0.56	0.49	0.54	0.44
	SGDClassifier	0.45	0.42	0.54	0.60	0.58	0.51	0.58	0.49
Naive Bayes	BernoulliNB	0.50	0.40	0.58	0.67	0.65	0.44	0.68	0.58
	ComplementNB	0.46	0.49	0.46	0.54	0.56	0.49	0.54	0.44
	GaussianNB	0.45	0.42	0.54	0.60	0.58	0.51	0.58	0.49
Nearest Neighbour	KNeighborsClassifier	0.54	0.47	0.50	0.60	0.60	0.54	0.56	0.35
Neural network	MLPClassifier	0.58	0.58	0.65	0.56	0.68	0.56	**0.70**	0.56
SVC	LinearSVC	0.62	0.49	0.44	0.54	0.51	0.39	0.54	0.47
	NuSVC	0.44	0.34	0.58	0.49	0.67	0.56	**0.70**	0.58
	SVC	0.58	0.49	0.58	0.51	0.65	0.58	**0.70**	0.49

and data subjected to scaling and normalization algorithms, i.e. StandardScaler, RobustScaler, MinMaxScaler, MaxAbsScaler, PowerTransformer, QuantileTransformer and Normalizer from the scikit-learn library (Table 2). A Python script was written to test the algorithms, combining scaling and normalization algorithms with machine learning algorithms (pipeline). Pairs of algorithms were subject to cross-validation (stratified k-fold) by calculating the average value of prediction accuracy for 4, 6 and 8 splits of the dataset.

RESULTS AND DISCUSSION

Tests performed for facilities with waste and tailings from hard coal mining (Table 3) have indicated that the most accurate prediction matches after the learning process were obtained using the AdaBoostClassifier, ExtraTreesClassifier, RandomForestClassifier, GaussianNB, GaussianProcessClassifier, MLPClassifier, NuSVC and SVC algorithms. In turn, tests performed for facilities with iron ore mining waste and tailings have shown that the most accurate prediction matches were obtained after application of the GaussianProcessClassifier, SGDClassifier, GaussianNB, MLPClassifier, BernoulliNB, ComplementNB, KNeighborsClassifier, ExtraTreeClassifier and SVC algorithms (Table 4). The results in the following tables are the means of accuracy values obtained from cross-validation tests performed on full sets.

The analyses performed have indicated that the highest matching value for facilities with hard coal mining waste and tailings was 0.75. The observed low quality of prediction may suggest that the dataset lacks criteria relevant for the selection of prospective objects. In the case of sites with iron ore mining waste and tailings the best fit was 0.85. The prediction quality obtained may indicate that the dataset contains criteria significant for selecting a heap or dump as a prospective facility. Due to the diversity of the dataset (criteria and attributed features) for the two groups of sites, it is not possible to clearly identify a single algorithm which would generate the best results for the selection of prospective sites. The closest results for both groups of sites are generated by the GaussianProcessClassifier algorithm (from the GaussianProcess group of algorithms). Individual analysis of each group of facilities showed that for sites with hard coal mining waste and tailings, the best prediction results were obtained using the Gaussian NB algo-

Table 4

Results of machine learning tests using supervised algorithms for facilities with waste from iron ore mining and processing

Group of ML algorithms	ML algorithm	Scaling algorithms						
		No scaling	Normalizer	Standardscaler	Maxabsscaler	Minmaxscaler	Quantile-transformer	Robustscaler
Decision Trees	DecisionTreeClassifier	0.60	0.58	0.58	0.58	0.63	0.60	0.60
	ExtraTreeClassifier	0.75	0.67	0.70	**0.80**	0.68	0.70	0.73
Ensemble	AdaBoostClassifier	0.75	0.67	0.73	0.78	0.75	0.75	0.75
	ExtraTreesClassifier	0.78	0.70	0.75	0.73	0.78	0.73	0.78
	GradientBoostingcClassifier	0.55	0.63	0.55	0.58	0.58	0.58	0.53
	RandomForestClassifier	0.78	0.67	0.75	0.78	0.70	0.73	0.75
Gaussian Process	GaussianProcessClassifier	0.43	0.77	0.75	**0.83**	**0.80**	**0.80**	**0.85**
Linear	Perceptron	0.70	0.73	0.68	0.65	0.77	0.73	0.68
	RidgeClassifier	0.65	0.77	0.65	0.73	0.68	0.67	0.65
	SGDClassifier	0.73	0.63	0.75	0.65	**0.85**	0.68	0.70
Naive Bayes	BernoulliNB	0.68	0.68	0.77	0.68	0.75	0.75	**0.80**
	ComplementNB					**0.80**		
	GaussianNB	0.78	0.78	0.78	0.78	0.78	**0.83**	0.78
Nearest Neighbour	KNeighborsClassifier	0.75	0.75	0.75	0.73	0.75	0.78	**0.80**
Neural network	MLPClassifier	0.58	0.77	0.78	**0.80**	0.78	0.78	**0.83**
SVC	LinearSVC	0.73	0.77	0.70	0.70	0.75	0.70	0.75
	NuSVC	0.67	0.70	0.73	0.75	0.75	0.75	0.73
	SVC	0.75	0.77	0.75	**0.80**	0.75	0.75	**0.80**

rithm, whereas for sites with iron ore mining waste and tailings, the best prediction results were achieved using the GaussianProcessClassifier algorithm with the application of input scaling algorithms. Analysis of the results shows that the application of the scaling algorithms MaxAbsScaler, MinMaxScaler, RobustScaler and QuantileTransformer had a positive influence on increasing the prediction precision.

MULTI-LAYER PERCEPTRON ALGORITHM

The next stage of the analysis was testing whether the Multi-Layer Perceptron algorithm would be applicable in the determination of prospective sites in both groups analysed. Analysis of the effectiveness of the Multi-Layer Perceptron algorithm has shown that for the facilities with hard coal mining waste and tailings, the results obtained reached a maximum efficiency of 0.70 prediction at proportions of the test to learning datasets of 1:3, and 0.79 at a ratio of 1:7 in the parameter combinations tested (Table 5). For sites with iron ore mining waste and tailings, for some combinations the results exceeded the 0.80 threshold, with a maximum 0.85 for a 1:7 set ratio (Table 6). The best results were obtained when working on data rescaled using the MinMaxScaler algorithm for hard coal facilities, whereas for the iron ore sites the effect of scaling on the variation of the results was smaller, and similar results were obtained with the RobustScaler, StandardScaler and MaxAbsScaler algorithms. The effect of a small number of input data on the variability of the results depending on the subdivision of the dataset into the learning and test parts is clear. The test results show that the large size of the hidden layer is not required, as the results in the range of the upper accuracy limit are achieved even for a layer composed of 10 neurons, and increase in their number does not result in an increase in prediction accuracy. Similar results were obtained for each of the activation functions: tanh, relu, identity and logistic. The most effective algorithm for optimizing the weights solver is sgd (stochastic gradient descent); only slightly worse results for both site types were obtained in the case of the lbfgs algorithm.

Table 5

Selected results of the Multi-Layer Perceptron algorithm for facilities with waste from hard coal mining and processing

Scaler	Number of neurons in the hidden layer	Activation function	Solver algorithm	cv=4	cv=8
MM	40	tanh	sgd	0.70	0.77
MM	10	relu	sgd	0.70	0.72
MM	10	tanh	sgd	0.70	0.72
MM	20	identity	sgd	0.70	0.72
MM	160	relu	sgd	0.67	0.77
MM	160	tanh	sgd	0.67	0.75
MM	10	logistic	lbfgs	0.63	0.79
MM	20	tanh	sgd	0.63	0.77

MM – MinMaxScaler; cv – number of subdivisions of the input dataset into training and test data using cross-validation. The dimension of the training dataset for cv=N is (N-1)/N of the input dataset; the dimension of the testing dataset is 1/N

Table 6

Selected results of the Multi-Layer Perceptron algorithm for facilities with waste from iron ore mining and processing

Scaler	Number of neurons in the hidden layer	Activation function	Solver algorithm	cv=4	cv=8
R	30	logistic	sgd	0.83	0.85
MA	25	identity	sgd	0.80	0.83
SS	10	logistic	sgd	0.78	0.85
SS	15	logistic	sgd	0.78	0.85
MA	10	identity	sgd	0.78	0.83
MA	10	relu	sgd	0.78	0.83
MA	20	tanh	sgd	0.78	0.83
MA	25	relu	sgd	0.78	0.83
MA	35	identity	sgd	0.78	0.80

R – RobustScaler, MA – MaxAbsScaler, SS – StandardScaler; cv – number of subdivisions of the input dataset into training and test data using cross-validation. The dimension of the training dataset for cv=N is (N-1)/N of the input dataset; the dimension of the testing dataset is 1/N

The networks constructed, having a large number of neurons in the hidden layers, may be overfitted. Therefore, it is important to disregard them as a valid model.

ENSEMBLE ALGORITHMS

For facilities with hard coal and iron ore mining waste and tailings, learning tests were made with the application of algorithms that compose results from multiple component algorithms.

Four input basic algorithms (MLPClassifier with a MinMaxScaler scaling algorithm, ExtraTreesClassifier with a RobustScaler scaling algorithm, GaussianNB with a MinMaxScaler scaling algorithm, and DecisionTreeClassifier with a StandardScaler scaling algorithm) were selected for the VotingClassifier and StackingClassifier component algorithms for the first group of sites. For the second group of sites, the basic algorithms were: MLPClassifier with a MaxAbsScaler scaling algorithm, BernoulliNB with a RobustScaler scaling algorithm, GaussianNB without a scaling algorithm, and GaussianProcessClassifier with a Normalizer scaling algorithm. For the compositional algorithms, cross-validation tests were performed with splitting of the input dataset into 4 and 8 parts. The aim of these tests was assessing whether application of these algorithms allows one to obtain prediction results better than for individual component algorithms. It was observed that for hard coal disposal sites there was a slight improvement in the prediction results (0.74 for the 4-part split and 0.75 for the 8-part split) in comparison to the results shown in Table 3. For iron ore waste facilities there was no significant improvement compared to the results from Table 4. The values achieved were 0.78 for the 4-part split and 0.83 for the 8-part split. The suitability of these algorithms for the application chosen was not found satisfactory.

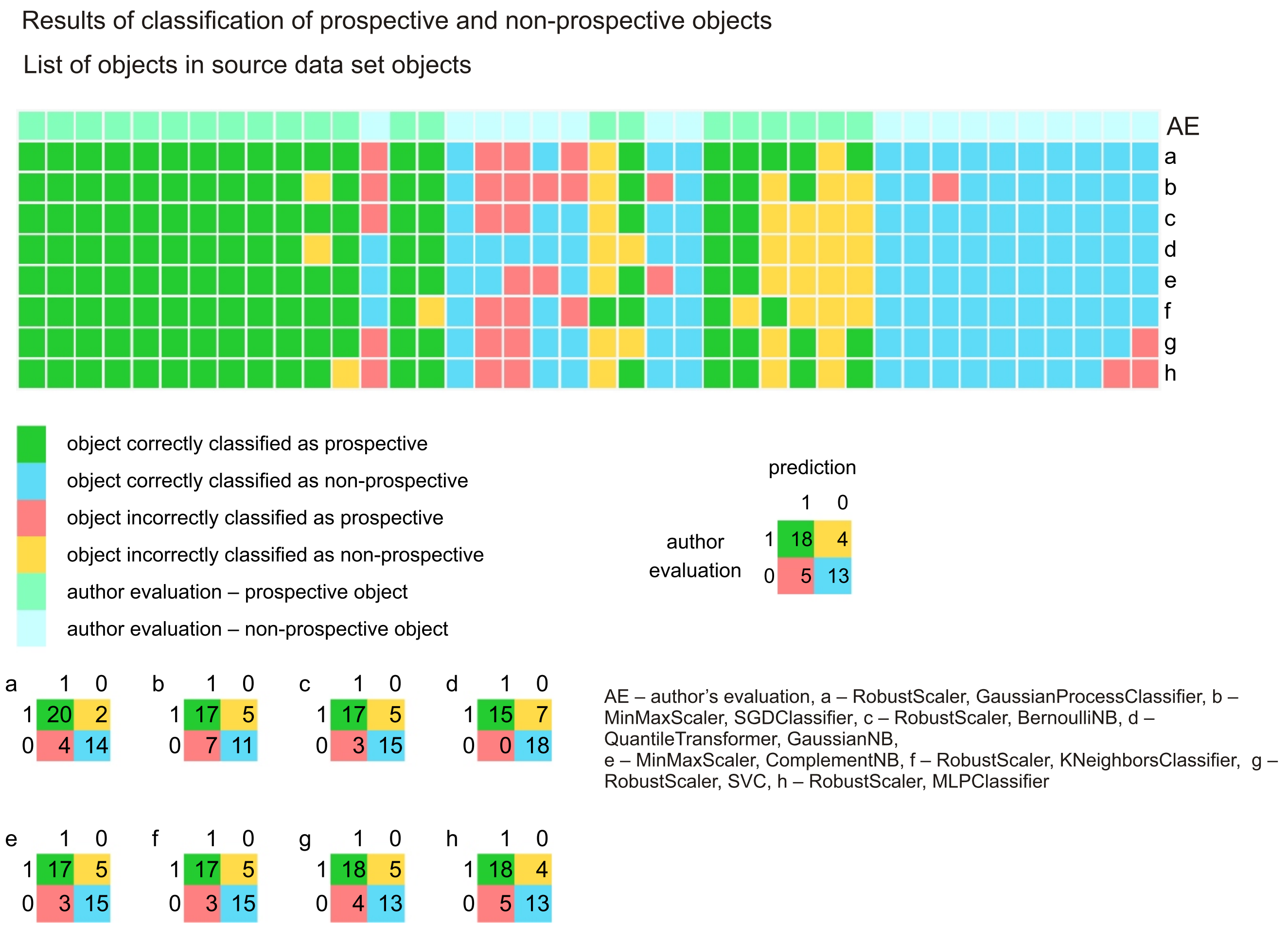


Fig. 6. Confusion matrix for facilities with waste from iron ore mining and processing

VALIDATION OF PREDICTION ACCURACY

The entire dataset of facilities with waste from hard coal and iron ore mining and processing was subject to cross-validation. The main goal of this procedure was assessment of the effectiveness of the learning process with application of the earlier-selected learning algorithms returning the best results. The effectiveness of prediction is shown in Figures 6 and 7. The results obtained were compared with the expert assessment of the secondary raw material potential of individual facilities. Objects that were incorrectly classified (as prospective or non-prospective) by the machine learning algorithms in relation to the expert assessment may be subject to analysis in order to determine the features influencing such classification; this refers in particular to objects classified incorrectly by most algorithms. Information on the number and types of performed errors was obtained by using the confusion matrix to display the results.

The ROC curve (Receiver Operating Characteristic curve) is a graphical representation used to evaluate the performance of a binary classification model. It plots the True Positive Rate (TPR) (sensitivity) against the False Positive Rate (FPR) at various threshold levels, showing the trade-off between correctly identifying positive cases and incorrectly classifying negative cases as positive. A perfect classifier would have a curve passing through the top-left corner, while a random classifier would follow the diagonal line from (0.0) to (1.1). The AUC (Area Under the Curve) is a single metric summarizing the ROC curve. It represents the probability that a randomly chosen positive instance is ranked higher than a randomly chosen negative instance. An AUC of 1.0 indicates a perfect model, 0.5 corresponds to random guessing, and values below 0.5 suggest worse-than-random performance. Achieved AUC values are larger than 0.8. The models are effective in distinguishing between positive and negative classes (Fig. 8).

THE RFE PROCEDURE AND PERMUTATION IMPORTANCE

Using the RFE algorithm (recursive feature elimination) from the scikit-learn library, the previously tested machine learning algorithms returning the best results (prediction accuracy at the level of 0.70 for hard coal sites and 0.80 for iron ore sites) were analysed in order to find features with the greatest impact on prediction accuracy. The RFE algorithm subsequently removes the features from the input dataset and checks the resulting changes in the prediction results. The algorithm creates a list of features with the highest impact on prediction accuracy. Using the RFE method resulted in the formation of most significant features lists for 4 algorithms: ExtraTreeClassifier, LinearSVC, SGDClassifier, and GaussianProcessClassifier (not all algorithms can be tested using the RFE method). For hard coal disposal sites the most common features were: cause fire hazards on the site by waste in the past, usage of the object, and area occupied by the object. In turn, for the iron ore disposal sites, these features included: amount of waste, past usage of waste from the facilities, and current direction of waste usage: construction ceramics.

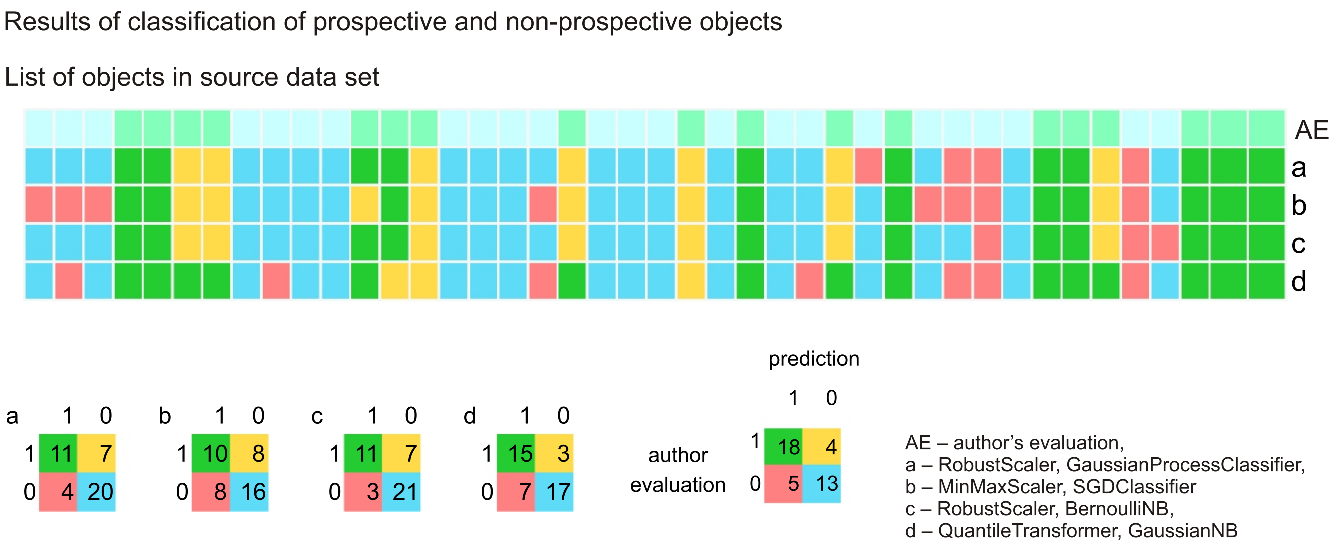


Fig. 7. Confusion matrix for facilities with waste from hard coal mining and processing

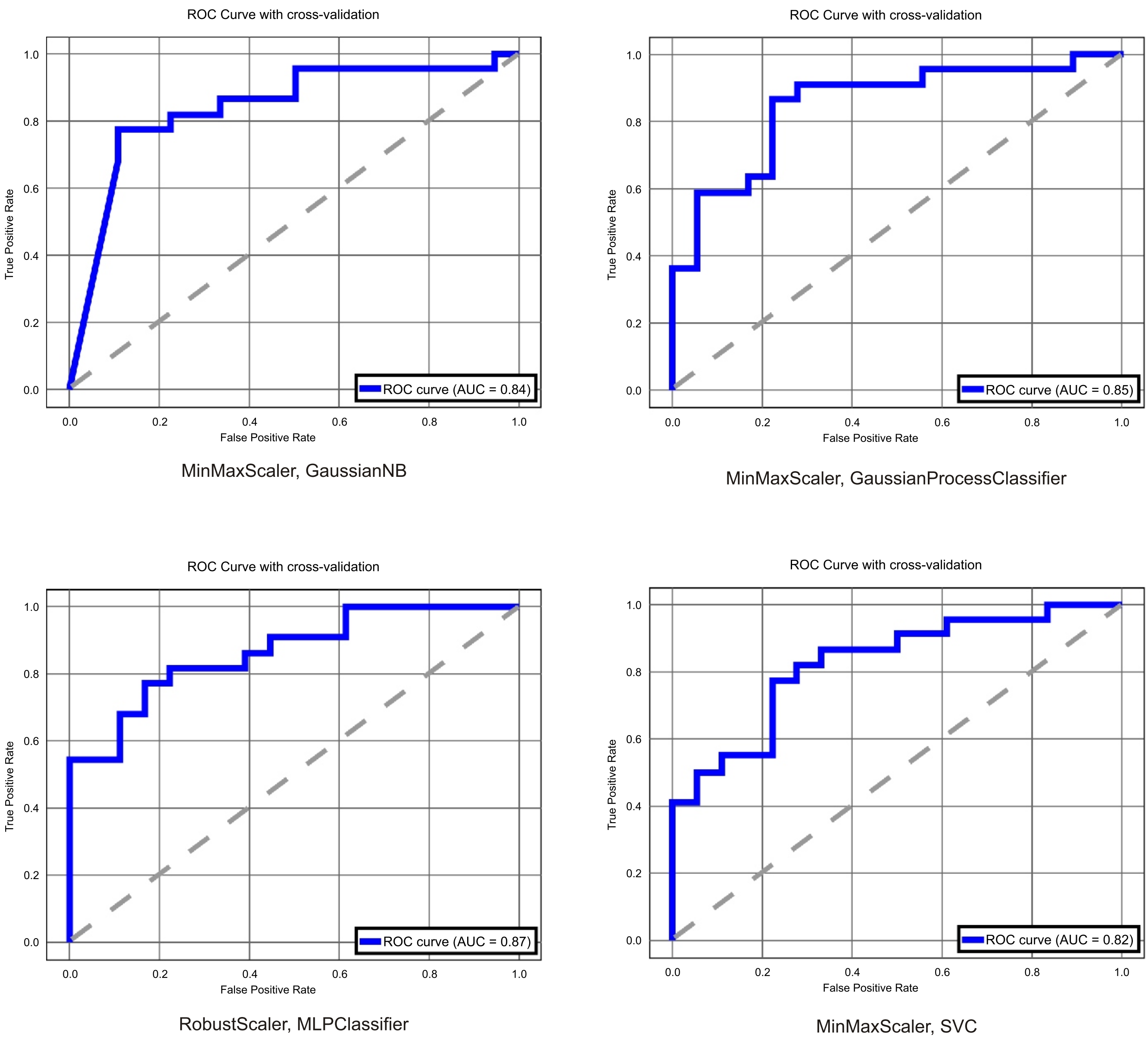
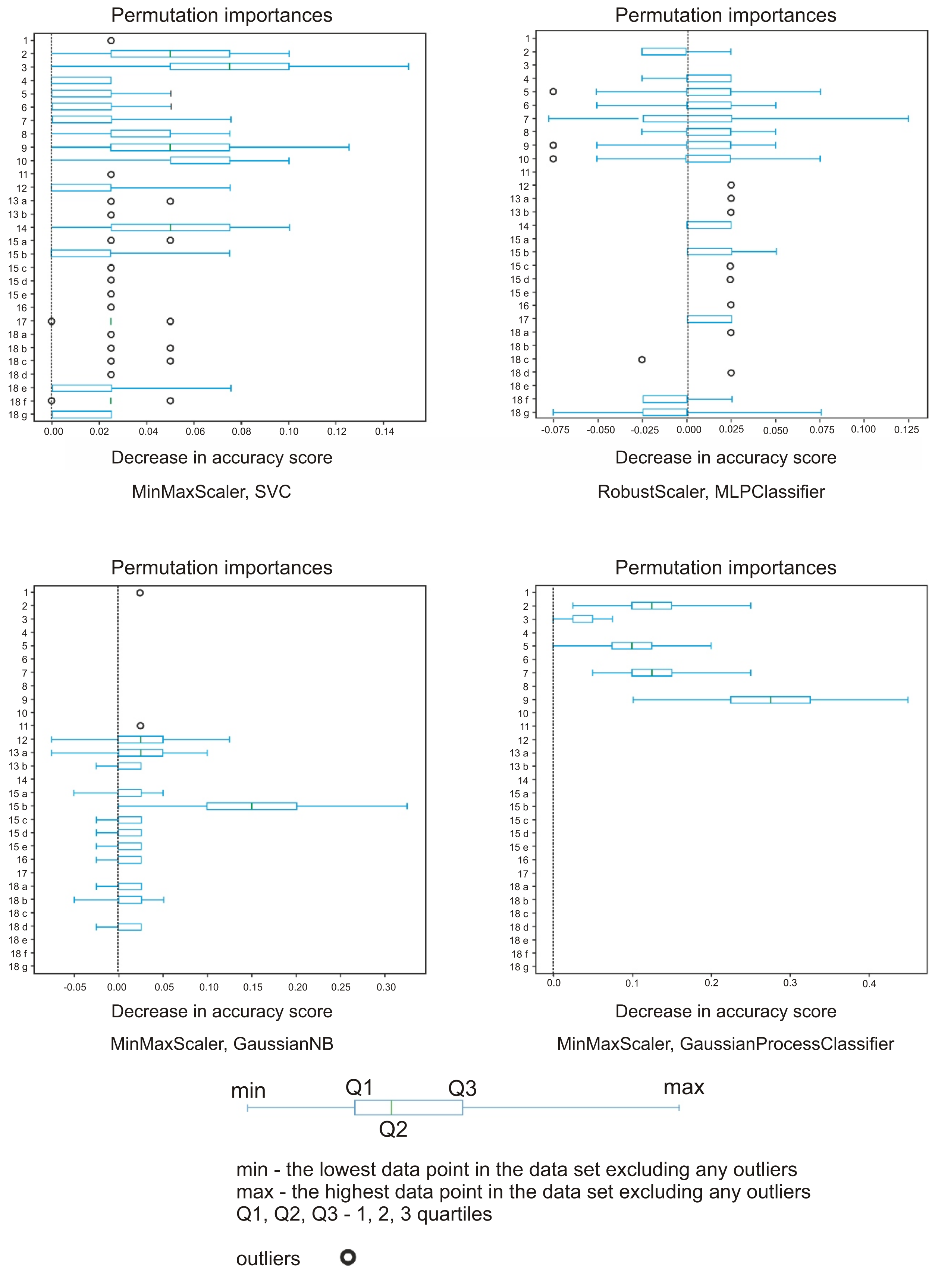


Fig. 8. The ROC curves for selected algorithms



For an explanation of the numbers on the Y-axis, see the Legend for Fig. 5

Fig. 9. Examples of feature importance obtained by the permutation_importance function

Detailed tests assessing feature validity for the accuracy were performed both for the hard coal and iron ore sites for the supervised learning algorithms returning the highest results.

The permutation_importances function was used for the earlier tested algorithms (MLPClassifier, GaussianNB and SVC), which cannot be analysed using the RFE algorithm due to the lack of a ranking list of relevant features. For comparison, this procedure was also applied for the GaussianProcessClassifier algorithm. This function changes the values of individual features of the input data and tests the effect of these changes on the resulting prediction score compared to the results obtained using unmodified data (Scikit-learn, 2024).

In the case of iron ore facilities, the performed assessment of the importance of features for the prediction accuracy for the algorithms returning the highest results has indicated that among the significant features occur those linked with the characteristics of the raw material, waste disposal site characteristics (age, amount of waste, condition, purpose of the facility), and the predicted usage of waste. For the GaussianProcessClassifier algorithm, the most significant features were: usage of the facility, area of the facility, deposit depth, deposit exploitation depth, amount of waste, time of waste deposition. For the SGDClassifier algorithm, the relevant features included: usage of the facility, direction of current waste usage – construction ceramics, deposit depth, deposit exploitation depth, end of waste disposal.

Analysis of the input criteria using the RFE algorithm and the permutation_importance function is aimed at determining the list of features with a significant impact on the prediction results, thus reducing the workload spent on preparing the input data.

Examples of results obtained for the iron ore facilities using the permutation_importance function for the MLPClassifier algorithm (with RobustScaler scaling algorithm), for algorithm SVC (with MinMaxScaler scaling algorithm), algorithm GaussianProcessClassifier with MinMaxScaler scaling algorithm, and GaussianNB algorithm with MinMaxScaler scaling algorithm are shown below (Fig. 9).

The values obtained using the permutation_importance function show the effect of modifying the value of a given feature on the decrease in prediction accuracy. The negative values indicate the increase in the prediction accuracy after data permutation. This points to the low validity or lack of validity of a given feature for the model accuracy. This phenomenon occurs in small datasets, where the learning process is unable to detect the low importance of a given feature, and in reality has a random effect on the accuracy of the model. The small dataset volume also has influence on the differences existing in the validity of features in different algorithms, and thus the overrepresentation of the validity of some features with regard to the underestimation of the validity of other features.

CONCLUSIONS

The economic value of mining waste and tailings accumulated in heaps, dumps and disposal sites (active, reclaimed and historical) is significant. However, the assessment of these objects with regard to the recovery of raw materials is a process requiring large financial outlays. The development of machine learning technologies and the results achieved in a PIG-PIB project indicate that this is a promising tool for predicting the secondary raw material potential of the sites analysed. Moreover, successful usage of machine learning methods requires access to an appropriately large dataset and appropriate data processing. Moreover, assessing the secondary raw material potential of these objects is a process requiring large financial outlays and is time-consuming.

The results obtained, due to the failure to achieve a 100% prediction certainty, indicate that machine learning cannot be the only tool used to undertake decisions on the reuse of the facilities analysed. It may, however, support the decision-making process, supplying information on the potentially useful and potentially non-useful facilities for further usage. Application of machine learning methods may contribute to reducing the financial outlays indispensable for performing in-situ studies on the facilities in the initial phase of the investment. Additionally, the process is not time-consuming and restricted e.g. by other conditions having impact on its realization compared to other traditional methods that are presently used. The machine learning algorithms applied (including the defined criteria and their features) may be used to assess the potential of other groups of industrial landfills. Nevertheless, there are restrictions in the application of this method resulting from the need to obtain good quality data, feature selection, or potential errors resulting from inappropriately selected algorithms for the calculations.

The results of classification facilities from iron ore mining and processing showed that the outcomes were greatly influenced by features such as the area of the disposal sites and the amount of accumulated waste. Consideration should be given to identifying and introducing an additional feature in future studies, indicating the prospects of the sites. In the case of hard coal mining and processing facilities, the analysis was difficult due to the larger number of features, as well as the lower quality of data than in the previous case.

Experience gained during these studies may be useful for other, similar applications, e.g. the environmental classification of facilities or areas which can be described by a set of criteria. Because the investigations are novel, as shown by the lack of published reports on machine learning applications for the recognition of the secondary raw material potential of waste accumulated in heaps, dumps and disposal sites, their continuation seems appropriate. Future studies should concentrate on the identification of the most effective input data for machine learning algorithms, the availability and quality of the necessary data, as well as creating uniform standards for datasets describing particular groups of accumulated waste.

The study was financed from funds allocated to the statutory activities of PGI-NRI (Project No. 61.3794.2300.00.0).

REFERENCES

Adrianto, L.R., Ciacci, L., Pfister, S., Hellweg, S., 2023. Toward sustainable reprocessing and valorization of sulfidic copper tailings: Scenarios and prospective LCA. Science of the Total Environment, **871**, 162038; https://doi.org/10.1016/j.scitotenv.2023.162038

Ambrožič, T., Turk, G., 2003. Prediction of subsidence due to underground mining by artificial neural networks. Computers & Geosciences, **29**: 627–637; https://doi.org/10.1016/S0098-3004(03)00044-X

Araya, N., Ramirez, Y., Kraslawski, A., Cisternas, L.A., 2021. Feasibility of re-processing mine tailings to obtain critical raw materials using real options analysis. Journal of Environmental Management, **284**, 112060; https://doi.org/10.1016/j.jenvman.2021.112060

Baza Hałdy, 2024. https://geologia.pgi.gov.pl/haldy/

Bellenfant, G., Guezennec, A.G., Bodenan, F., D'Hugues, P., Cassard, D., 2013. Reprocessing of Mining waste:combining environmental management and metal recovery? In: Mine Closure 2013: Proceedings of the Eighth International Seminar on Mine Closure, Australian Centre for Geomechanics, Cornwall (eds. M. Tibbett, A.B. Fourie and C. Digby): 571–582; https://doi.org/10.36487/ACG_rep/1352_48_Bellenfant

Binnemans, K., Jones, P.T., Blanpain, B., Van Gerven, T., Pontikes, Y., 2015. Towards zero-waste valorisation of rare-earth-containing industrial process residues: a critical review. Journal of Cleaner Production, **99:** 17–38; https://doi.org/10.1016/j.jclepro.2015.02.089

Blengini, G.A., Mathieux, F., Mancini, L., Nyberg, M., Viegas, H.M. (eds.), 2019. Recovery of critical and other raw materials from mining waste and landfills: state of play on existing practices. European Commission, Joint Research Centre, Publications Office, Luxembourg. ISBN 978-92-76-08568-3; https://data.europa.eu/doi/10.2760/600775

Careddu, N., Dino, G.A., Danielsen, D.W., Přikryl, R., 2018. Raw materials associated with extractive industry: an overview. Resource Policy, **59:** 1–6; https://doi.org/10.1016/j.resourpol.2018.09.014

Cobîrzan, N., Muntean, R., Thalmaier, G., Felseghi, R.A., 2022. Recycling of Mining Waste in the Production of Masonry Units. Materials, **15**, 594; https://doi.org/10.3390/ma15020594

Danthurebandara, M., Van Passel, S., Vanderreydt, I., Van Acker, K., 2015. Assessment of environmental and economic feasibility of Enhanced Landfill Mining. Waste Management, **45**: 434–447; https://doi.org/10.1016/j.wasman.2015.01.041

Dino, G., Cavallo, A., Faraudello, A., Piercarlo, R., Mancini, S., 2021. Raw materials supply: kaolin and quartz from ore deposits and recycling activities. The example of the Monte Bracco area (Piedmont, Northern Italy). Resources Policy, **74**, 102413; https://doi.org/10.1016/j.resourpol.2021.102413

Dino, G., Mancini, S., Casale, M., Lasagna, M., 2023. Use and recovery of mining waste and tailings for sustainable management of works related to extractive activities. EGU General Assembly 2023, Vienna, Austria, 23–28, EGU23-6013; https://doi.org/10.5194/egusphere-egu23-6013

Environment 2023, 2023. Statistics Poland, Warsaw; https://stat.gov.pl/obszary-tematyczne/srodowisko-energia/srodowisko/ochrona-srodowiska-2023,1,24.html

Eurostat, 2020. Treatment of waste by waste category, hazardousness and waste management operations;

https://ec.europa.eu/eurostat/databrowser/view/env_wastrt/default/table?lang=en/ (accessed 6.11.2023).

Fajfer, J., Kostrz-Sikora, P., 2022a. Initial cost effectiveness of extracting waste accumulated on the rehabilitation of waste facilities and the concept of circular economy (in Polish with English summary). Przegląd Geologiczny, **70:** 190–201; https://www.pgi.gov.pl/dokumenty-pig-pib-all/publikacje-2/przeglad-geologiczny/2022/3-marzec-2/9120-wstepna-ocena-oplacalnosci-wydobycia-odpadow-nagromadzonych-w-zrekultywowanych-obiektach-a-koncepcja-gospodarki-o-obiegu-zamknietym/file.html

Fajfer, J., Kostrz-Sikora, P., 2022b. Economic considerations of a circular economy based on waste accumulated on the rehabilitation of waste facilities.In: Mineral Deposits Safeguarding as a Basis of Mineral Raw Materials Safety. Book of abstracts. 10–11 May 2022, Cracow, Poland, Publishing House MEERI PAS: 21–22. ISBN 978-83-964171-5-2; https://min-pan.krakow.pl/wydawnictwo/wp-content/uploads/sites/4/2022/04/2022-International-Conference.-Mineral-deposits-safeguarding-as-a-basis-of-mineral-raw-materials-safety-Book-of-abstracts.pdf

Fajfer, J., Kostrz-Sikora, P., Radwanek-Bąk, B., Krasuska, J., 2022. Prospects and challenges in use of mining waste accumulated in the closed and abandoned heaps and dumps. In: XXXI Conference in the series Aktualia i perspektywy gospodarki surowcami mineralnymi, Book of abstracts. 16–18 November 2022, Rytro, Poland, Publishing House MEERI PAS: 13–15. ISBN 978-83-964234-2-9; https://konferencja-aktualia.min-pan.krakow.pl/pliki/XXXI_Konferencja-Aktualia_2022_zeszyt_streszczen.pdf

Falagán, C., Grail, B.M., Johnson, D.B., 2017. New approaches for extracting and recovering metals from mine tailings. Minerals Engineering, **106**: 71–78; https://doi.org/10.1016/j.mineng.2016.10.008

Fisher, L.V., Barron, A.R., 2019. The recycling and reuse of steelmaking slags - A review. Resources, Conservation and Recycling, **146**: 244–255; https://doi.org/10.1016/j.resconrec.2019.03.010

Flach, P., 2012. Machine Learning: The Art and Science of Algorithms that Make Sense of Data. Cambridge University Press. 1st Edition. ISBN 978-1-107-09639-4.

Flores, V., Leiva, C., 2021. A comparative study on supervised machine learning algorithms for copper recovery quality prediction in a leaching process. Sensors, **21**, 2119; https://doi.org/10.3390/s21062119

Garbarino, E., Orveillon, G., Saveyn, H.G.M., 2020. Management of waste from extractive industries: the new European reference document on the best available techniques. Resources Policy, **69**, 101782; https://doi.org/10.1016/j.resourpol.2020.101782

Géron, A., 2018. Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow: pojęcia, techniki i narzędzia służące do tworzenia inteligentnych systemów (in Polish). Helion S.A. Publishing House, Gliwice. ISBN: 978-83-8322-423-7.

Grus, J., 2018. Data science od podstaw. Analiza danych w Pythonie (in Polish). Helion S.A. Publishing House, Gliwice. ISBN: 978-83-8322-131-1.

GUS, 2001. Statistical Yearbook of the Republic of Poland 2000, 2001. Statistical Publications House, Warsaw.

GUS, 2023. Statistical Yearbook of the Republic of Poland 2022, 2023. Statistics Poland, Warsaw. ISSN 0867-3217.

Han, Ch., Johnson, A.M., 1995. Waste Products in Highway Construction; https://onlinepubs.trb.org/Onlinepubs/conf/1995/cp6/v2/cp6v2-008.pdf

Herrera, N., Gonzales, M.S., Okkonen, J., Mollehuara, R., 2023. Soft computing application in mining, mineral processing and metallurgy with an approach to using it in mineral waste disposal. Minerals, **13**, 1450; https://doi.org/10.3390/min13111450

Ilutiu-Varvara, D.A., Aciu, C., 2022. Metallurgical wastes as resources for sustainability of the steel industry. Sustainability, **14**, 5488; https://doi.org/10.3390/su14095488

Kinnunen, P., Ismailov, A., Solismaa, S., Sreenivasan, H., Räisänen, M.-L., Levänen, E., Illikainen, M., 2018. Recycling mine tailings in chemically bonded ceramics – a review. Journal of Cleaner Production, **174**: 634–649; https://doi.org/10.1016/j.jclepro.2017.10.280

Kinnunen, P., Karhu, M., Yli-Rantala, E., Kivikytö-Reponen, P., Mäkinen, J., 2022. A review of circular economy strategies for mine tailings. Cleaner Engineering and Technology, **8**, 100499; https://doi.org/10.1016/j.clet.2022.100499

Kostrz-Sikora, P., Bliźniuk, A., Fajfer, J., Rolka, M., 2013. Inwentaryzacja zamkniętych i opuszczonych obiektów unieszkodliwiania odpadów wydobywczych (in Polish). In: XXIII Conference in the series Aktualia i perspektywy gospodarki surowcami mineralnymi, The Bulletin of The Mineral and Energy Economy Research Institute of the Polish Academy of Sciences, **85:** 199-205, 6–8 November 2013, Zakopane, Poland, Publishing House MEERI PAS; https://min-pan.krakow.pl/wp-content/uploads/sites/4/2017/12/15-12-kostrz-blizniuk-fajfer-.pdf

Kozioł, W., Kawalec, B., 2008. Kruszywa alternatywne w budownictwie (in Polish). Nowoczesne Budownictwo Inżynieryjne, **4**: 34–37; https://nbi.com.pl/content/uploads/assets/NBI-pdf/2008/4_19_2008/pdf/11_kruszywa%20_alternatywne_w_budownictwie.pdf

Kozioł, W., Uberman, R., 1996. Możliwości i warunki zagospodarowania odpadów z górnictwa i energetyki w drogownictwie, zwłaszcza do budowy autostrad i dróg ekspresowych (in Polish). Przegląd Geologiczny, **44**: 701–709; https://geojournals.pgi.gov.pl/pg/article/view/18571

Kozioł, W., Cieplińska, A., Machniak, Ł., Borcz, A., 2015. Kruszywa w budownictwie. Cz. 2. Kruszywa alternatywne (in Polish). Nowoczesne Budownictwo Inżynieryjne, **5**: 35–40; https://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-aa40467f-5fec-428c-810c-3a09b9f3816c

KPGO, 2002. National Waste Management Plan (in Polish). Załącznik do Uchwały Nr 219 Rady Ministrów z dnia 29 października 2002 r. w sprawie krajowego planu gospodarki odpadami. Monitor Polski, 2002, Nr 11, poz. 159; https://isap.sejm.gov.pl/isap.nsf/download.xsp/WMP20030110159/O/M20030159.pdf

KPGO, 2023. National Waste Management Plan 2028 (in Polish). Załącznik do Uchwały Nr 96 Rady Ministrów z dnia 12 czerwca 2023 r. w sprawie Krajowego planu gospodarki odpadami 2028. Monitor Polski, 2023, poz. 702; https://isap.sejm.gov.pl/isap.nsf/download.xsp/WMP20230000702/O/M20230702.pdf

Krzywicki, T., 2022. Systemy uczące się: sztuczne sieci neuronowe i uczenie głębokie (in Polish). The University of Warmia and Mazury Publishing House, Olsztyn, Poland. ISBN: 978-83-8100-354-4

Lim, B., Alorro, R.D., 2021. Technospheric Mining of Mine Wastes: A Review of Applications and Challenges. Sustainable Chemistry, **2:** 686–706; https://doi.org/10.3390/suschem2040038

Lucarini, M., Carta, R., Fumanti, F., Martarelli, L., Serra, M., 2020. The Italian database GeMMA: from monitoring production to cataloguing mining wastes, a starting point for recovering critical raw materials from abandoned mines? European Geologist, **49:** 27–31; https://doi.org/10.5281/zenodo.3938887

Łukaszyk, W., Tabor, A., 1979. Gospodarka odpadami w górnictwie węgla kamiennego (in Polish). In: Materiały Sympozjum Naukowo-Technicznego: Stan i perspektywy gospodarki odpadami w środowisku, **6**: 39–52. NOT Publishing House, Warszawa.

Makowski, A., Fajfer, J., 1997. Zastosowanie odpadów do budowy autostrad i towarzyszącej im infrastruktury (in Polish). In: Szkoła Gospodarki Odpadami '97. Centrum PPGSMiE PAN Publishing House. ISBN: 83-86286-78-4.

Male, F., Duncan, I.J., 2020. Lessons for machine learning from the analysis of porosity-permeability transforms for carbonate reservoirs. Journal of Petroleum Science and Engineering, **187**, 106825; https://doi.org/10.1016/j.petrol.2019.106825

Malfante, M., Dalla, Mura, M., Metaxian, J.-P., Mars, J.I., Macedo, O., Inza, A., 2018. Machine Learning for Volcano-Seismic Signals: Challenges and Perspectives. IEEE Signal Processing Magazine, **35**: 20–30; https://doi.org/10.1109/MSP.2017.2779166

Nieć, M., 1999. Anthropogenic mineral deposits (in Polish with English summary). Przegląd Geologiczny, **47:** 93–98; https://geojournals.pgi.gov.pl/pg/article/download/15660/13232

Nowok, S., Skiba, E., 1979. Stan i perspektywy gospodarki odpadami górnictwa i hutnictwa metali nieżelaznych (in Polish). In: Materiały Sympozjum Naukowo-Technicznego: Stan i perspektywy gospodarki odpadami w środowisku, **6**: 53–73. NOT Publishing House, Warszawa.

Nwaila, G.T., Ghorbani, Y., Zhang, S.E., Frimmel, H.D., Leon, C., Tolmay, K., Rose, D.H., Nwaila, P.C., Bourdeau, J.E., 2021. Valorisation of mine waste - Part I: Characteristics of, and sampling methodology for, consolidated mineralised tailings by using Witwatersrand gold mines (South Africa) as an example. Journal of Environmental Management, **295**, 113013; https://doi.org/10.1016/j.jenvman.2021.113013

Oliveira, M., Escobar, A.G., Relvas, J., Pinto, A., 2022. Perspectives on tailings (re)mining at the Neves Corvo deposit, Portugal: valorisation of the Cerro do Lobo Tailings Storage Facility, challenges and opportunities. In: (Re)mining Extractive Waste, a New Business?, 17–18 May 2022, Mechelen, Belgium, Book of Abstracts Symposium, 10–11. ISBN: 9789464594966; https://re-mine.eu/wp-content/uploads/2022/05/ReMining.pdf

Palarski, J., Musioł, S., Kaczmarczyk, A., 1987. Możliwości utylizacji odpadów drobnofrakcyjnych poprzez deponowanie ich w górotworze i wyrobiskach podziemnych (in Polish). Zeszyty Naukowe Politechniki Śląskiej, Seria: Górnictwo, **158**: 39–49; https://delibra.bg.polsl.pl/Content/42061/BCPS_45999_1987_Mozliwosci-utylizacj.pdf

Paluch, J., Morawski, S., 1975. Rekultywacja terenów poprzemysłowych w Polsce ze szczególnym uwzględnieniem Śląskiej Niecki Węglowej (in Polish). Zeszyty Problemowe Postępów Nauk Rolniczych, **166**: 83–97.

Pandas, 2024. Python Data Analysis Library; https://pandas.pydata.org/ (accessed 1.02.2024)

Pietras, Z., 1979. Gospodarka odpadami przemysłu energetycznego (in Polish). In: Materiały Sympozjum Naukowo-Technicznego: stan i perspektywy gospodarki odpadami w środowisku, **6**: 89–125. NOT Publishing House, Warszawa.

Plewa, F., Popczyk, M., Jendruś, R., 2005. Waste management according to formal and legal conitions in underground mines (in Polish with English summary). Zeszyty Naukowe Politechniki Śląskiej, Seria: Górnictwo, **270:** 463–476; http://delibra.bg.polsl.pl/Content/74536/BCPS-83796_2005_Zagospodarowanie-odp.pdf

Pozzi, M., Mzyk, T., 2007. Ocena wpływu na środowisko odpadów ulokowanych w wyrobiskach podziemnych kopalń w świetle modelowania hydrogeochemicznego (in Polish). Górnictwo i Geoinżynieria, **31:** 461–474; http://journals.bg.agh.edu.pl/GORNICTWO/2007-03-1/GG_2007_3-1_37.pdf

Potempa, M., Szlugaj, J., 2007. The direction of utilization of gangue obtained during production of the hard coal in the Piast coal mine (in Polish with English summary). Prace Naukowe Instytutu Politechniki Wrocławskiej, **119:** 159–171; https://www.researchgate.net/publication/351345736_KIERUNKI_WYKORZYSTANIA_SKAL_PLONNYCH_W_KWK_PIAST

Qi, C., Wu, M., Liu, H., Linag, Y., Liu, X., Lin, Z., 2023. Machine learning exploration of the mobility and environmental assessment of toxic elements in mining-associated solid wastes. Journal of Cleaner Production, **401**, 136771; https://doi.org/10.1016/j.jclepro.2023.136771

Raschka, S., Mirjalili, V., 2021. Python. Machine learning i deep learning: Biblioteki Scikit-learn i TensorFlow2. Helion SA Publishing House, 3rd edition, Gliwice. ISBN: 978-283-7001-2.

Ratajczak, T., Korona, W., 2000. Mineralogical-chemical and technological characteristics of mineral matter from dumps after mining of iron ores the Częstochowa region (Central Poland) (in Polish with English summary). Przegląd Geologiczny, **48**: 607–616; https://geojournals.pgi.gov.pl/pg/article/download/15304/12972

Ratajczak, T., Hycnar, E., Bożęcki, P., 2015. Mineralogical criterion as a part of suitability assessment of some Polish clay deposits for construction of waterproofing apertures (in Polish with English summary). IGSMiE PAN Publishing House, Cracow. ISBN: 978-83-62922-53-6; https://min-pan.krakow.pl/wp-content/uploads/sites/4/2017/12/SRM-194.pdf

Ridzwan, N.S.M., Yusoff, S.H.M., 2023. Machine learning for earthquake prediction: a review (2017–2021). Earth Science Informatics, **16**: 1133–1149; https://doi.org/10.1007/s12145-023-00991-z

Rosario-Beltré, A.J., Sánchez-España, J., Rodríguez-Gómez, V., Fernández-Naranjo, F.J., Bellido-Martín, E., Adánez-Sanjuán, P., Arranz-González, J.C., 2023. Critical Raw Materials recovery potential from Spanish mine wastes: a national-scale preliminary assessment. Journal of Cleaner Production, **407**, 137163; https://doi.org/10.1016/j.jclepro.2023.137163

Rosik-Dulewska, C., Karwaczyńska, U., 2008. Methods of Leaching Contaminants from Mineral Waste in the Aspect of its Potential Utilization in Hydrotechnical Construction (in Polish with English summary). Annual Set The Environment Protection, **10:** 205–219; https://ros.edu.pl/index.php/en/?view=article&id=745

Różański, Z., 2019. Recovery of the raw materials as an element of coal waste management and reusing of landfill sites. Mineral Resources Management, **35**: 147–162; https://doi.org/10.24425/gsm.2019.128542

Sarker, I.H., 2021a. Deep learning: a comprehensive overwiew on techniques, taxonomy, applications and research directions. SN Computer Science, **2**, 420; https://doi.org/10.1007/s42979-021-00815-1

Sarker, I.H., 2021b. Machine learning: algorithms, real-world applications and research directions. SN Computer Science, **2**, 160; https://doi.org/10.1007/s42979-021-00592-x

Scikit-learn, 2024. Scikit-learn: machine learning in Python – scikit-learn 1.4.1 documentation; https://scikit-learn.org/stable/ (accessed 1.02.2024)

Segoui, P., Safhi, A.e.M., Amrani, M., Benzaazoua, M., 2023. Mining wastes as road construction material: a review. Minerals, **13**, 90; https://doi.org/10.3390/min13010090

Spoorena, J., Binnemansb, K., Björkmalm, J., Breemersch, K., Dams, Y., Folens, K., González-Moya, M., Horckmans, L., Komnitsas, K., Kurylak, W., Lopez, M., Mäkinen, J., Onisei, S., Oorts, K., Peys, A., Pietek, G., Pontikes, Y., Snellings, R., Tripiana, M., Varia, J., Willquist, K., Yurramendi, L., Kinnunen, P., 2020. Near-zero-waste processing of low-grade, complex primary ores and secondary raw materials in Europe: technology development trends. Resources, Conservation and Recycling, **160**, 104919; https://doi.org/10.1016/j.resconrec.2020.104919

Szeliga, M., 2019. Praktyczne uczenie maszynowe (in Polish). PWN Scientific Publishing House, Warsaw. ISBN: 978-83-01-19232-7.

Topór, T., 2021. Application of machine learning algorithms to predict permeability in tight sandstone formations. Nafta-Gaz, 5**:** 283–292; https://doi.org/10.18668/NG.2021.05.01

Trifi, M., Gasmi, A., Carbone, C., Majzlan, J., Nasri, N., Dermach, M., Charef, A., Elfil, H., 2022. Machine learning-based prediction of toxic metals concentration in an acid mine drainage environment, Northern Tunisia. Environmental Science and Pollution Research, **29:** 87490–87508; https://doi.org/10.1007/s11356-022-21890-8

Waste Statistics, 2020. https://ec.europa.eu/eurostat/statistics-explained/SEPDF/cache/1183.pdf (accessed 6.11.2023).

Wowkonowicz, P., Bojanowicz-Bablok, A., Gworek, B., 2018. Use of thermal metallurgy and mining waste in road construction (in Polish with English summary). Annual Set The Environment Protection, **20**: 1335–1349; https://ros.edu.pl/index.php/en/?view=article&id=987