



Contents lists available at ScienceDirect

Journal of Rock Mechanics and Geotechnical Engineering

journal homepage: www.jrmge.cn

Full Length Article

Multiclass stand-alone and ensemble machine learning algorithms utilised to classify soils based on their physico-chemical characteristics

Eyo Eyo ^{a,b,*}, Samuel Abbey ^a^a Faculty of Environment and Technology, Department of Geography and Environmental Management, Civil Engineering Cluster, University of the West of England, Bristol, BS16 1QY, UK^b Fugro GB Marine Geotechnical Services Limited, Wallingford, OX10 9RB, UK

ARTICLE INFO

Article history:

Received 14 April 2021

Received in revised form

21 July 2021

Accepted 20 August 2021

Available online 14 November 2021

Keywords:

Soil classification

Physico-chemistry

Soil plasticity

Machine learning

Logistic regression (LR)

Machine learning ensembles

Artificial neural network (ANN)

ABSTRACT

This study has provided an approach to classify soil using machine learning. Multiclass elements of stand-alone machine learning algorithms (i.e. logistic regression (LR) and artificial neural network (ANN)), decision tree ensembles (i.e. decision forest (DF) and decision jungle (DJ)), and meta-ensemble models (i.e. stacking ensemble (SE) and voting ensemble (VE)) were used to classify soils based on their intrinsic physico-chemical properties. Also, the multiclass prediction was carried out across multiple cross-validation (CV) methods, i.e. train validation split (TVS), *k*-fold cross-validation (KFCV), and Monte Carlo cross-validation (MCCV). Results indicated that the soils' clay fraction (CF) had the most influence on the multiclass prediction of natural soils' plasticity while specific surface and carbonate content (CC) possessed the least within the nature of the dataset used in this study. Stand-alone machine learning models (LR and ANN) produced relatively less accurate predictive performance (accuracy of 0.45, average precision of 0.5, and average recall of 0.44) compared to tree-based models (accuracy of 0.68, average precision of 0.71, and recall rate of 0.68), while the meta-ensembles (SE and VE) outperformed (accuracy of 0.75, average precision of 0.74, and average recall rate of 0.72) all the models utilised for multiclass classification. Sensitivity analysis of the meta-ensembles proved their capacities to discriminate between soil classes across the methods of CV considered. Machine learning training and validation using MCCV and KFCV methods enabled better prediction while also ensuring that the dataset was not overfitted by the machine learning models. Further confirmation of this phenomenon was depicted by the continuous rise of the cumulative lift curve (LC) of the best performing models when using the MCCV technique. Overall, this study demonstrated that soil's physico-chemical properties do have a direct influence on plastic behaviour and, therefore, can be relied upon to classify soils.

© 2022 Institute of Rock and Soil Mechanics, Chinese Academy of Sciences. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Soil classification is considered as one of the fundamental steps in the planning, design, development, and implementation of various infrastructural projects and related undertakings that have to do with static and dynamic interactions with the ground (Kaliakin, 2017). This is mainly because the results obtained from classifying soils could serve as valuable indices in determining mechanical properties, such as compressibility, permeability, shear

strength, and swelling. Depending on the aim or type of construction project envisaged, the class or category of soils required may vary as a result of either their plasticity properties or particle sizes.

Two major groups of soil classification systems are usually adopted in the practice of soil science and engineering. The first category often relies on textural classification based mainly on soil particle size distribution (PSD) originally proposed and standardised by the United States Department of Agriculture (USDA). In the second category, the engineering behaviour of soils is mainly considered; hence, classification is based on both PSD and plasticity properties. The methods and procedures of soil groupings in this case are all captured and standardised in the American Association of State Highway and Transportation Officials (AASHTO) as well as the unified soil classification system (USCS). Moreover, it is also imperative to add that several researchers are continuously

* Corresponding author. Faculty of Environment and Technology, Department of Geography and Environmental Management, Civil Engineering Cluster, University of the West of England, Bristol, BS16 1QY, UK.

E-mail address: eyo.eyo@uwe.ac.uk (E. Eyo).

Peer review under responsibility of Institute of Rock and Soil Mechanics, Chinese Academy of Sciences.

devising other means of classifying soils (Moreno-Maroto et al., 2021).

Although the classification of soils through their textural characteristics is important, particle sizes do not completely represent the general properties of soils because soils having similar PSD can possess different physical characteristics (Casagrande, 1948; Moreno-Maroto et al., 2021). Therefore, the focus of this article is the classification of soils based on their plasticity behaviour. Plasticity is simply defined as the capacity of soil to be moulded into any shape without rupturing or cracking. Soil plasticity is mostly established by carrying out the Atterberg limits test and, more specifically, by determining the liquid limit (LL) and plastic limit (PL), representing the upper and lower moisture content values for which the soil can exhibit plasticity.

Most previous systems of soil classification are based on an indirect estimation and correlation of LL , PL , and plasticity index (PI) on charts, which in some situations are subject to a few setbacks most especially in borderline cases (Casagrande, 1948; Saito and Miki, 1975; Polidori, 2007, 2009; Moreno-Maroto et al., 2017; Moreno-Maroto and Alonso-Azcárate, 2018). These techniques rely mostly on the activities of either first remoulding or working the soil and then testing or laborious physical measurements in order to indirectly infer soil classifications (Cai et al., 2011; Bol, 2013; Shahri et al., 2015; Eyo et al., 2019). Pham et al. (2021) attempted a classification of soils by correlating properties such as specific gravity, moisture content, void ratio, clay content, and Atterberg limit with their plasticity categories. Even though the accuracy of prediction had some improvements, their model indicated that about 12.5% of the soil samples were neither captured nor properly identified. Hence, in addition to the challenges mentioned above, some of the problems encountered in their method of soil classification could have been as a result of not considering the core intrinsic properties of the soils.

Soil plasticity is affected mainly by mineralogy and chemical compositions (Polidori, 2015; Spagnoli et al., 2017; Okeke et al., 2021). According to Moreno-Maroto et al. (2021), these factors can provide even more information on the plasticity behaviour of the soils than those stated previously. Besides, studies on the relationships between these factors and engineering behaviours, such as shear strength, swelling, compressibility, and compaction, have been reported (e.g. Christidis, 1998; Al-Rawas, 1999; Yilmaz, 2004; Abbey et al., 2019, 2020, 2021; Eyo et al., 2019, 2020, 2021). In view of this and the aforementioned shortcomings, this article aims to provide a novel approach to soil classification based on a consideration of their physico-chemical characteristics (i.e. cation exchange capacity (CEC), carbonate content (CC), specific surface area (SSA), and clay fraction (CF)) using machine learning.

Machine Learning is artificial intelligence (AI) paradigm that is gradually gaining attraction and popularity within the civil engineering discipline. However, the focus of machine learning technique application in geotechnical engineering at present tends to be mostly on regression-based problems related to soil engineering macro-behaviours, such as shear strength, unconfined compressive strength, resilient modulus of elasticity, compressibility, swelling, compaction, and stabilisation (e.g. Kayadelen et al., 2009; Ikizler et al., 2010; Liao et al., 2011; Tekin and Akbas, 2011; Yilmaz and Kaynar, 2011; Bekhor and Livneh, 2014; Tinoco et al., 2014; Mozumder and Laskar, 2015; Zhang and Goh, 2016; Goh et al., 2017; Mozumder et al., 2017; Soleimani et al., 2018; Gajurel et al., 2019; Ermias and Vishal, 2020; Hanandeh et al., 2020; Eyo and Abbey, 2021; Zhang et al., 2020, 2021).

In summary, the major objective of this study is the application of stand-alone machine learning algorithms (logistic regression (LR) and artificial neural network (ANN)), decision tree ensembles

(decision forest (DF) and decision jungle (DJ)), and meta-ensemble models (stacking ensemble (SE) and voting ensemble (VE)) to the soil classification. Unlike regression, the task of machine learning classification is to employ a decision-making technique to the class membership of an unknown item of data based on the entirety of the dataset (Dreiseitl and Ohno-Machado, 2002; Pham et al., 2021). For the nature of soil classification envisaged, with the predictor variable containing potentially more than two categorical features (due to wide soil plasticity ranges), multiclass elements of machine learning models adopted herein shall be used in the soil classification. Also, prediction shall be carried out across multiple cross-validation (CV) methods for the first time.

2. Methodology

2.1. Database generation and integration

Both numerical and categorical high-quality data collected from careful experiments on soils (both clayey and silty soils) with wide-ranging plasticity properties were used in this study. The soil classes represented by the USCS are based on the Casagrande chart of PI versus LL (see Fig. 1). The dividing line (or 'A-Line') expressed as $PI = 0.73 (LL - 20)$ is the line separating inorganic clays, located above the 'A-Line', and the rest of the soil materials (i.e. organic soils, silts, and fine sandy soils) located below. A series of symbols that allow the soils to be categorised or classified into different groups are further explained in Table 1. It should be noted that the soils used in this study are those of inorganic clayey and silty nature. Highly standardised measurement techniques and testing procedures used for determining physico-chemical properties of the soils were adopted (e.g. Smith et al., 1985; Dexter, 1990; Abduljawwad, 1994; Cerato, 2001; Kalkan and Akbulut, 2004; Sridharan and Gurtug, 2004; Senol et al., 2006; Arnepalli et al., 2008; Venkat et al., 2008; Yazdandoust and Yasrobi, 2010; Gaidzinski et al., 2011; Ngun et al., 2011; Erzin and Gunes, 2013; Bayat et al., 2015; Mahmoudi et al., 2016; Mehta and Sachan, 2017; Akgün et al., 2018; Spagnoli and Shimobe, 2019). The physico-chemical properties of the soils (i.e. CEC, CC, SSA, and CF ($< 2 \mu\text{m}$)) derived from these measurements shall serve as inputs or independent variables for machine learning prediction. To achieve the goal (i.e. determining soil classifications in terms of their plasticity), Atterberg limits of the soils were appropriately used to depict the machine learning predictor categorical variables (soil classes). Table 2 shows some of the important components of descriptive statistics obtained from the raw and independent soil data features. Also, Table 2 confirms the wide plasticity ranges of the soils used in this study. The values of LL range between approximately 18% and 155%, while that of PI range between about 1% and 94%, indicating that soils of low and high plasticity are represented according to the USCS. Frequency distributions of the raw data collected from the literature are depicted in Fig. 2. Although the data are quite sparse in some bin counts, especially for SSA and CEC, CF seems to possess a more uniform distribution. Among these features, both CC and CF seem to have less deviation from the mean of the values.

To enable comprehensive analysis and evaluation of the soils' plasticity categories generated from multiple sources, thorough integration and alignment of the data were imperative. All the datasets were appropriately appended by putting together the elements with similar attributes and features. An appraisal of the importance of each independent feature by assessing their relative importance in the machine learning prediction before and after the process of data training and validation was conducted and reported in subsequent sections of this study.

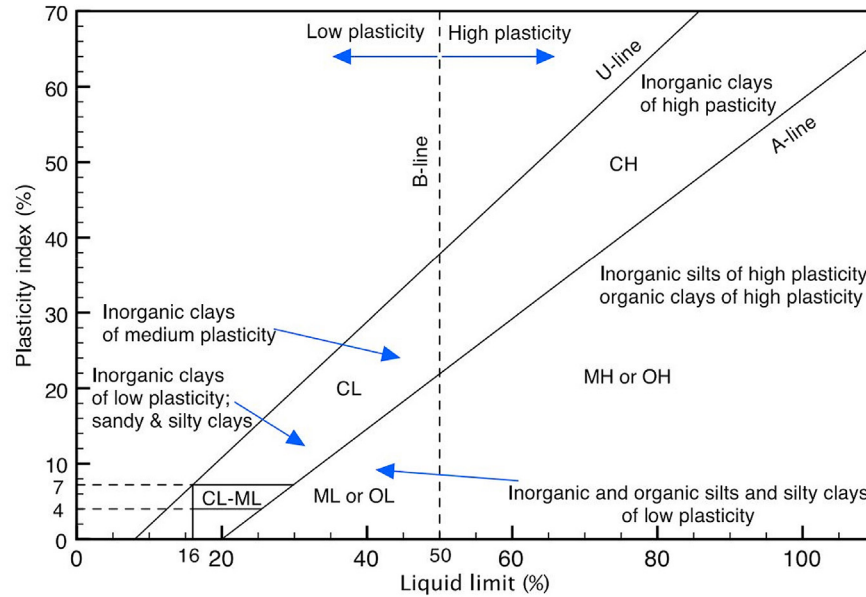


Fig. 1. Casagrande–USCS plasticity chart (Reprinted from Soil Mechanics: Calculations, Principles, and Methods, Victor N. Kaliakin, Example problems related to soil identification and classification, Copyright (2017), with permission from Elsevier).

2.2. Data wrangling and CV

The data used in the supervised machine learning classification in this study do not contain any missing value. Although it is sometimes considered essential (depending on the variability of the dataset) to transform the dependent features before training and validation, it is unnecessary given these data features are categorical in nature. However, it is important to employ the technique of CV to ensure an improvement of the training dataset as well as a drastic reduction of the possibility that some coincidental variables receive more importance in the multiclass prediction (Joshi, 2020). Also, by using this method, overfitting of the machine learning models could be avoided, and the predictive performance is enhanced (DeRousseau et al., 2019). The techniques of CV adopted in this study are train-validation split (TVS), k -fold cross-validation (KFCV), and Monte Carlo cross-validation (MCCV).

TVS ensures that the dataset is radically and randomly split with each of the dataset used for training, testing, and validation. In this study, 80% of the dataset from the source were utilised to effectively train the machine learning multiclass algorithms, whereas the remaining 20% were used to test and validate the predictive performance of the models. This method is based on the recommendations from most machine learning studies that the validation set should be 10%–30% of the total dataset (Han et al., 2020; Joshi, 2020).

The KFCV method divides the entire dataset (N_{KFCV}) into k subsets with equal size. One of the subsets is used for training, while the remaining ones are for validation (n_v). The process is then repeated for k times while excluding the k -subset in each iterative cycle (see Eq. (1)). Ten-fold CV (i.e. $k = 10$) is used in this research for the training and validation of the classification data.

$$N_{KFCV} = kn_v \quad (1)$$

MCCV technique combines the concepts of TVS and KFCV methods. MCCV divides the entire dataset (N_{MCCV}) into two sets through sampling without replacing one of the datasets. The training is then carried out on the non-replaced subset (n_t), while validation is performed on the remaining ones (n_v). The MCCV may preclude the running of iterations, unlike the KFCV (see Eq. (2)). In this study, the proportion of validation data is 20%, in addition to ten-fold CV.

$$N_{MCCV} = n_t + n_v \quad (2)$$

3. Machine learning algorithms

3.1. LR

LR is based on the least square function of linear regression, where several correlations between one or more independent (or explanatory) variables and predictor (or dependent) variables are established by

$$Y_n = \beta_0 + \beta_1 X_{1n} + \beta_2 X_{2n} + \dots + \beta_m X_{mn} + \epsilon_n \quad (3)$$

where Y_n is the predictor variable; $X_{1n}, X_{2n}, \dots, X_{mn}$ are the independent variables; β_0 is a constant; and $\beta_1, \beta_2, \dots, \beta_m$ are the coefficients of regression; and ϵ_n is the error term.

On the other hand, LR tends to determine the conditional probability that the predictor variable $Y = 1$ given the independent variable X . The probability distribution for LR is generally given by the hypothesis function:

$$P(Y = 1|X = x) = p(x) \quad (4)$$

LR then tends to model the logarithm odds of an event linearly as follows:

$$\log_{10} \left[\frac{p(x)}{1 - p(x)} \right] = \beta_0 + \beta_1 X_{1n} + \beta_2 X_{2n} + \dots + \beta_m X_{mn} \quad (5)$$

where the solution for the even occurring $p(x)$ is given as

$$p(x) = \frac{1}{1 + \exp(\beta_0 + \beta_1 X_{1n} + \beta_2 X_{2n} + \dots + \beta_m X_{mn})} \quad (6)$$

For every single data point, there exists some feature vectors x_i as well as a class of observation y_i whose probability is either p (if $y_i = 1$) or $1 - p$ (if $y_i = 0$) given the likelihood of the event as

$$l(\beta) = \sum_{i=1}^n [y_i x_i \beta - \log_{10}(1 + e^{x_i \beta})] \quad (7)$$

Table 1
Casagrande–USCS description of soil plasticity.

Symbol	Plasticity class
CL	Inorganic clays of low to medium plasticity
CH	Inorganic clays of high plasticity
ML	Inorganic silts of low plasticity
OL	Organic silty clays of low plasticity
MH	Inorganic silts of high plasticity
OH	Organic clays of high plasticity

Note: The first letter 'C' in the symbol column represents clays, 'M' represents silts, and 'O' represents organic; the second letter 'L' represents low to medium plasticity, and 'H' represents high plasticity.

3.2. ANN

ANN is a type of machine learning algorithm that is modelled after the human brain. Hence, the functioning and problem-solving processes of an ANN mimic that of the human fundamental neural network unit called the neuron (Fig. 3a). The neuron is the brain's simple information processing unit that is created to receive and process input signals generated from other surrounding neurons (connected at a junction called synapsis) through an input path referred to as the dendrite to an output path called the axon. It is worth mentioning that although the ANN structure is based on the human nervous system, which is said to possess billions of neurons, the ANN does utilise only a few hundreds of its neurons in practical geotechnical engineering problems (Das, 2013).

The neurons are described as processing nodes or elements in the ANN's mathematical model. Hence, a network having an input vector consisting of a single element x_l ($l = 1, \dots, N_i$) would be transmitted through some form of connections that are multiplied by a set of weights to produce the hidden unit z_j as

$$z_j = \sum_{l=1}^{N_i} w_{jl}x_l + b_{j0} \quad (j = 1, \dots, N_h) \quad (8)$$

where N_i is the number of input units; N_h is the number of hidden units consisting of weighted inputs $\left(\sum_{l=1}^{N_i} w_{jl}x_l\right)$ and bias (b_{j0}); w_{jl} is weight; x_l is input.

In order to allow for nonlinearity in the network system, the inputs would have to pass through a layer of some transfer (activation) function, f which then generates:

$$s_j = f\left(\sum_{l=1}^{N_i} w_{jl}x_l + b_{j0}\right) \quad (9)$$

Fig. 3b depicts a typical three-layer ANN architecture with the input, hidden, and output layers.

3.3. Tree-ensemble models

Decision tree models are based on a concept that involves repeated or continuous splitting (branching) of input data by adhering to a set of formal rules or criteria that aims to maximise the separation of these data, hence, resulting in a typical tree-like structure (Fig. 4). Each splitting of the input data leads to a decrease in the system's entropy and a maximisation of the splits. This formal rule (or information gain criterion) is common in most decision tree networks. The estimate of the probability distribution $P(m|n)$ in this case would be the ratio of m class elements to all the elements of the leaf nodes containing data item n (Dreiseitl and Ohno-Machado, 2002). However, depending on the kind of problem or application, there are some differences in how the decision

Table 2
Relevant statistics of machine learning independent features and soils' consistency limits.

Statistical parameter	CC (%)	CEC (meq/100 g)	SSA (m ² /mg)	CF (%)	LL (%)	PL (%)	PI (%)
Mean	17.58	28.49	122.01	40.99	43.53	24.49	19
Std. dev.	19.86	22.63	79.08	19.08	19.75	7.68	14.07
Kurtosis	1.77	6.55	2.89	−0.44	9.09	3.8	7.52
Skewness	1.44	2.3	1.41	0.16	2.14	1.38	2.08
Range	88.1	120.6	438	91	137.3	49	93.3
Minimum	0.1	5.5	11	4	17.7	12	0.7
Maximum	88.2	126.1	449	95	155	61	94

Note: Std. dev. means the standard deviation.

tree machine learning models are constructed or used. For the sake of this research, the multiclass elements of both DF and DJ are considered.

3.3.1. DF

The DF as a tree ensemble is typically created in order to minimise the fluctuations or instability which would have possibly existed if a solo tree structure has been utilised in machine learning prediction. The multiclass classification DF relies on the concept of bagging or 'bootstrapping aggregation' to perform its function efficiently. Bagging, in this case, is defined simply as a technique of training a dataset by bringing together multiple tree algorithms in a 'bag' (Kang et al., 2021). One of the main setbacks of using a DF is its susceptibility to overfitting due to its various small biases and wide variance.

3.3.2. DJ

DJ is regarded sometimes as some form of an extension of the random forest. DJ comprises an aggregation or ensemble of deep-rooted decision directed acyclic graphs (DAGs), which is a technique that ensures that compact and precise machine learning classifiers are obtained (Shotton et al., 2013). Because the merging of several trees is performed, the decision from the DAG typically tends to possess a low memory footprint, thus making them rather phenomenal in their general performance. The multiclass element of DJ has the advantage of being non-parametric and thus can effectively represent nonlinear decision boundaries. Moreover, DJ can select some integrated features and perform classification problems while also be robust in resisting noisy features during data training.

3.4. Meta-ensembles

Models or meta-ensembles are machine learning paradigms that combine various models, some of which are mentioned above in machine learning prediction. The basic idea is to utilise the strength of each model to estimate different patterns in the data. This can be carried out to improve prediction accuracy through the technique of either majority or weighted voting or stacking generalisation depending on the classification problem envisaged.

3.4.1. VE

This method of machine learning may not necessarily require any prerequisite from the aggregated classifiers or models. Hence, in most cases, the method of voting neither relies on any assumption of prior knowledge of how individual model behaves nor does it require training on very large quantities of results of representative recognitions from the classifiers (Kim and Upneja, 2021). Weighted voting combining the best weights from the above-mentioned models is proposed for the multiclass problem in this study (the structure is shown in Fig. 5a).

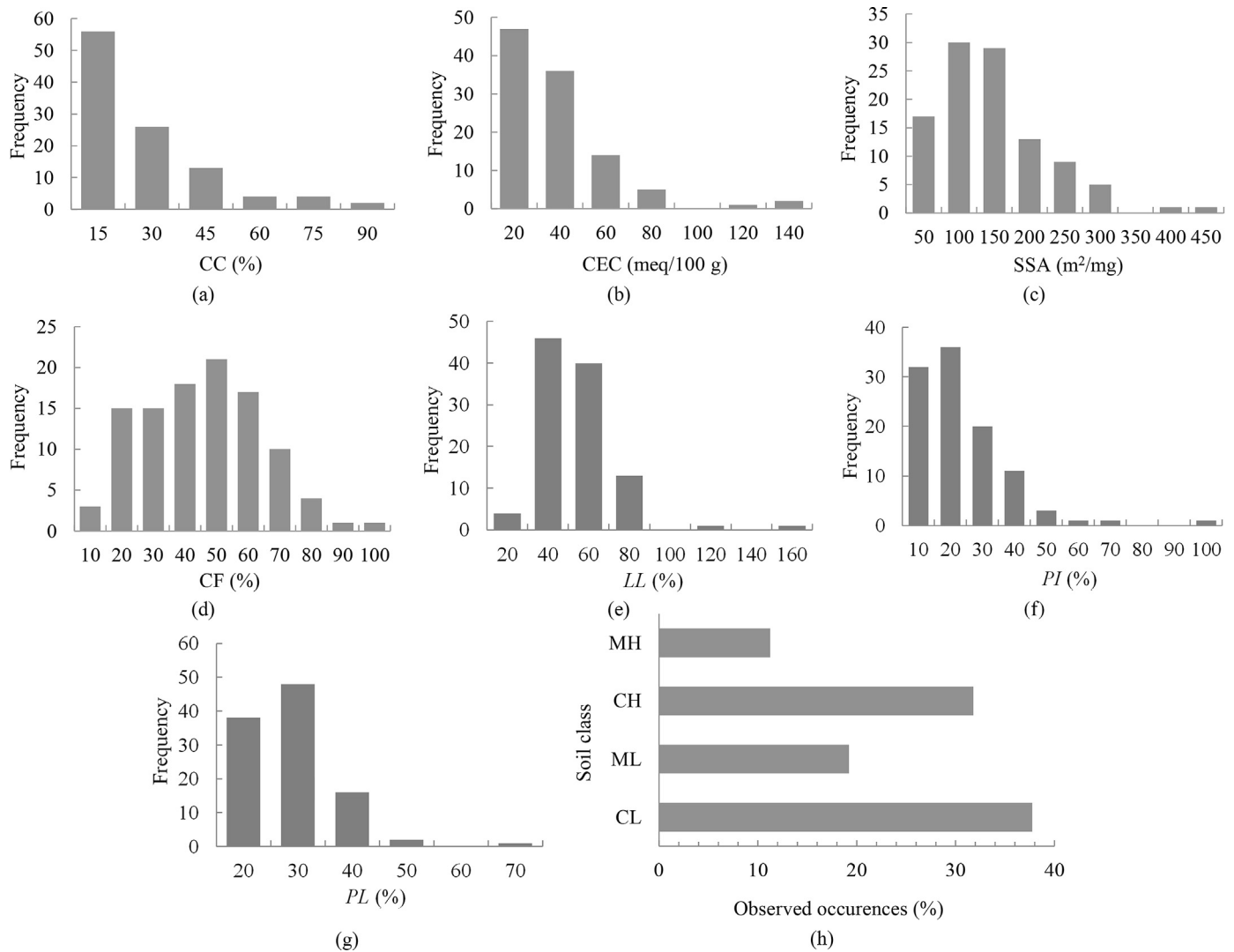


Fig. 2. Frequency distributions of machine learning features and soils' consistency limits: (a) CC, (b) CEC, (c) SSA, (d) CF, (e) LL, (f) PI, (g) PL, and (h) Distribution of soil class labels.

3.4.2. SE

SE uses two-level models or classifiers and includes both base learners and meta-classifiers for machine learning predictions. SE tends to combine the outputs produced from the base learners by using meta-classifiers to learn the patterns or relationships between the generated outputs. Fig. 5b presents the structure of the SE as proposed, which was used in this study for multiclass classification of the soils based on their physico-chemical characteristics.

3.5. Machine learning model development and implementation

All the processes of machine learning training and validation were implemented through a designer platform that supports programming with python with its libraries (pandas, NumPy, Matlib plots, and scikit-learn) and pipeline development. The dataset records utilised in this study for multiclass classification of the natural soils, the fundamental specifications of some of the important features, and the parameters adopted for both the soils and machine learning models are given in Tables 3 and 4. A flow-chart depicting the methodology in this study for the development

of machine learning pipelines, subsequent evaluation, and the deployment of the models is shown in Fig. 6.

3.6. Performance evaluation criteria

This study utilises the 'discrimination' criterion to assess the quality of the classification algorithms. Discrimination simply measures how well multiple categories or class labels in a dataset are separated. The following discrimination-based performance evaluation metrics are used in this study:

- (1) Accuracy, which is regarded as the overall proportion or correct predictions of an machine learning algorithm and is expressed as

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (10)$$

where true positive (TP) is the total number of collated instances from a positive class given that the true class label is equal to the predicted class label, false positive (FP) is the total number of collated instances from a negative class where the machine learning algorithm has been known to misclassify

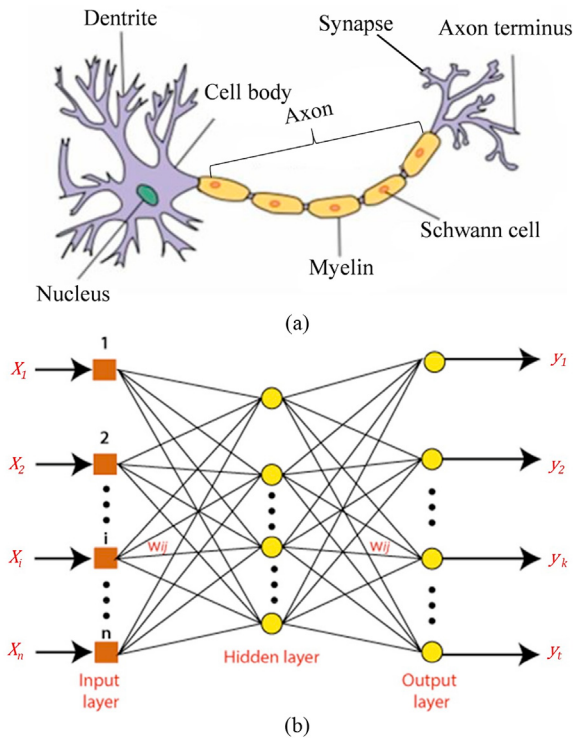


Fig. 3. (a) Human neuron and (b) typical structure of ANN.

these instances by falsely predicting them as positive, true negative (TN) is the total number of collated instances from a negative class whereby the true class label is known to be equal to the predicted class label, false negative (FN) is the total number of collated instances from a positive class whereby the machine learning algorithm has been known to misclassify these instances by predicting them as negative instead.

- (2) Precision, which is the positive predictive value defined as the total number of predictions that are regarded as being actually corrected out of all the predictions considered to have been made on positive classifications which is expressed as

$$\text{Precision} = \frac{TP}{TP + FP} \quad (11)$$

- (3) Recall, which otherwise known as sensitivity, tends to measure the machine learning model's ability to identify the

proportion or percentage of some relevant data points. It can also be taken as the number of collated instances of the correctly predicted positive class. Recall is mathematically expressed as

$$\text{Recall} = \frac{TP}{TP + FN} \quad (12)$$

3.7. Receiver operating characteristic (ROC) curve

A typical class prediction sensitivity analysis to determine the level of acceptability of the best machine learning models in the identification of the TPs belonging to different class labels or categories is through the ROC curve. ROC is a useful probability assessment tool that enables an evaluation of multiclass classification models by demonstrating the relationship between true positive rates (TPRs) and false positive rates (FPRs) during the course of any change in the decision-making threshold. ROC depicts to a great deal how much a machine learning model can properly differentiate between categories or classes. Used along with the ROC, (area under the curve (AUC)), measures the amount of separability between each class in machine learning prediction. For the ROC, an accurate or perfect machine learning classification indicates a point having coordinates of (0, 1) at the top left-hand side corner on a plot of TPR versus FPR, which is otherwise regarded as the sensitivity or recall of 100% (i.e. without any false negatives in the prediction) shown in Fig. 7. On the other hand, an act of randomly guessing the class or category of a natural soil's plasticity would generate points on or below the straight line (or line of no possible discrimination) extending from the origin (0, 0) to the upper right-hand side corner regardless of the positive or negative base rates which in turn would represent the worst possible condition. In this case, the ROC curve would be almost equal to 0.5, meaning that the machine learning model has no possible

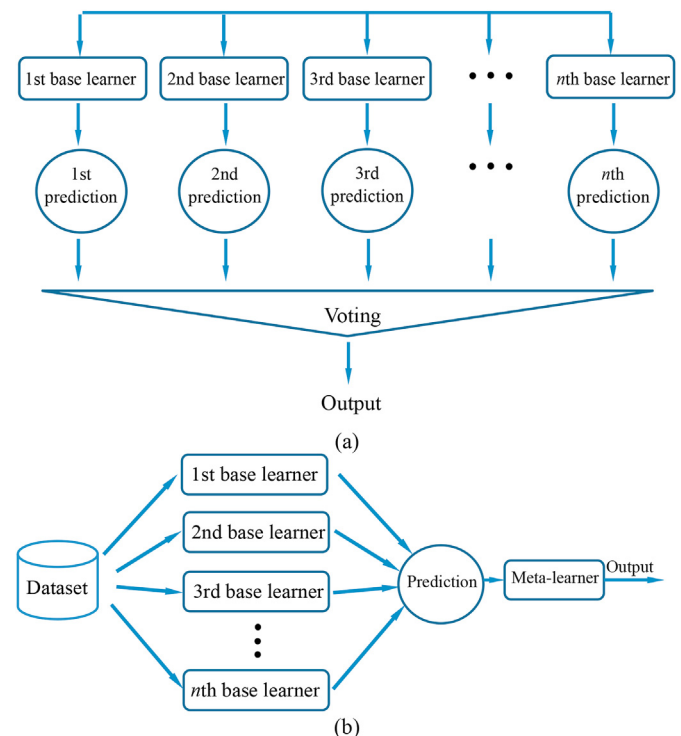


Fig. 5. Typical meta-ensemble architecture for voting and stacking models: (a) VE structure, and (b) SE structure.

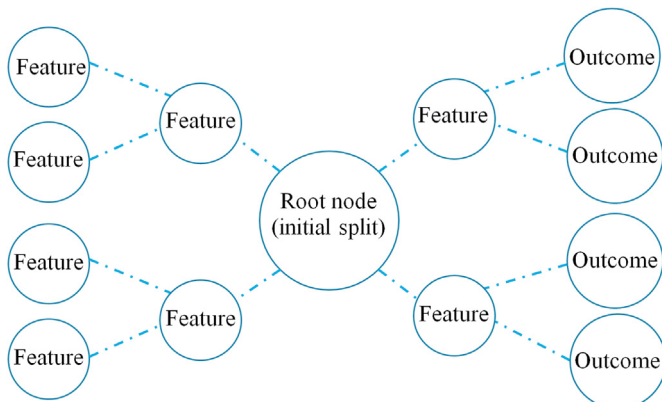


Fig. 4. Structure of a typical decision tree.

Table 3
Base machine learning models' parameter settings.

Multiclass model	Parameter	Option/value
LR	Trainer mode	Single parameter
	L1 regularization weight	1
	L2 regularization weight	1
ANN	Hidden layer specification.	Fully connected
	Number of hidden nodes	100
	Rate of learning	0.005
	Number of learning iterations	100
	Initial learning weight diameter	0.1
	Normalizer	min-max
DF	Resampling method	Bagging
	Trainer mode	Single parameter
	Number of decision trees	8
	Maximum decision tree depth	32
	Number of random splits per node	128
DJ	Minimum sample no. per leaf node	1
	Resampling method	Bagging
	Trainer mode	Single parameter
	Number of DAGs	8
	DAGs maximum depth	32
	DAGs maximum width	128
	Number of optimization steps per DAG layer	2048

discrimination in distinguishing between several positive and negative class labels.

4. Results and discussions

4.1. Assessment of machine learning models' performance

Fig. 8 depicts the performance of the multiclass machine learning algorithms utilised for the prediction of the soils' plasticity categories according to the USCS. A comparison between the numbers of variables selected in the prediction is also depicted in Fig. 8. It is observed that when using all four variables (SSA, CF, CEC, and CC) in the classification problem, both the stand-alone machine learning models (LR and ANN) seem to produce relatively less accurate performance compared to tree-based and meta-ensemble models. Overall, the worst performing model is the LR (accuracy of 0.44, precision of 0.47, and recall of 0.43), with the ANN clearly having the highest accuracy of the two stand-alone models (accuracy of 0.46, average precision of 0.52, and average recall of 0.45). One of the main reasons for LR's worst performance is its intrinsic assumption of linearity in spite of several instances of collinearities between input and response class labels in the multiclass prediction, which leads to their inability to sufficiently learn the features. The inclusion of more 'interaction' product terms to the original independent variables and their covariates can certainly improve the performance of the LR model by causing it to be more nonlinear and robust in the multiclass prediction. However, this approach must be dealt with cautiously because the desire for greater flexibility could be accompanied by a much higher risk of overfitting, reducing further the accuracy of the model (Dreiseitl and Ohno-Machado, 2002). On the other hand, the architectural make-up of the ANN model, as defined previously, means that their perceptron outputs (also referred to as hidden neurons or black-boxes) are inherently nonlinear. Hence, the output of the ANN is a nonlinear function of its inputs which in the context of a multiclass problem would mean that the decision boundary can be nonlinear, thus, making this algorithm relatively more flexible in the multiclass prediction compared to LR.

It is pertinent to add that LR and ANN models are quite similar in many aspects, given that both tend to possess common roots,

especially in most statistical problems. Moreover, these models are regarded as being different from the tree-based ensembles, because as explained previously in Section 2, they can supply both a function f , and a parameter vector α to allow an expression of the probability distribution $P(y|x)$ as

$$P(y|x) = f(x, \alpha) \quad (13)$$

However, what makes the LR different from the ANN is the functional form of expression used. Whereas this is often referred to as the parametric functional method when using the LR, that of the ANN is called non-parametric or semi-parametric. This is an important distinction because the contribution from parameters in an LR can be interpreted. At the same time, this may not often be the case for the parameters or weights of a neural network.

The unique structure of the tree-based models makes them inevitably nonlinear, and according to Fig. 8, they seem to produce slightly higher performance metrics than the stand-alone models. The decision trees are 'white-box' models, so-called because when compared to neural networks, they tend to allow for an interpretation of model parameters within a set of rules. As explained earlier, the tree-based models work by the continuous splitting of the input data according to sets of criteria or rules, which maximises the separation between the data and a corresponding decrease in entropy.

However, Fig. 8 also shows the relative lower performance of the tree-based ensembles (DJ and DF) compared to the meta-ensembles, most especially the VE model (which does outperform the rest of the models). One of the main drawbacks of decision tree algorithms can be traced to their somewhat 'greedy' behaviour at each successive step of construction during the actual training process. A combination of the best variables and an optimal splitting of the same are obtained at each step. However, a single progressive construction or feed-forward step that utilises a combination of the variables may end in a much better or different result while compromising the rest of the other steps. Another serious disadvantage in the decision tree algorithms would be that the continuous variables of the dataset could be discretised implicitly in the splitting process, such as losing their importance and some of the needed predictive information as the construction of the leaves progresses along.

Some of the setbacks mentioned above, especially as it relates to the decision tree ensembles, could be tremendously curtailed by aggregating multiple learners (both the stand-alone and tree ensembles) into what is referred to herein as meta-ensembles in the machine learning training. As could be observed in Fig. 8, the performance metrics of the meta-ensembles, most especially the VE model, are high (accuracy of 0.78, average precision of 0.78 and average recall of 0.76). As stated previously, the VE and SE work by a meta-heuristic combination of base algorithms' hyperparameters such that the outcomes of their prediction are even more accurate. It is quite interesting to note how the stacking technique of the model of algorithm performance (SE) is almost at par with that of the tree-ensembles. Further investigation of this behaviour shall be conducted subsequently through a sensitivity analysis.

4.2. Multiclass prediction uncertainties

Multiclass prediction probability distributions across each of the classes for the machine learning models are plotted in Fig. 9. Greater symmetry (or less skewness) in the distribution is exhibited by the ensemble models compared to the stand-alone ones. The stand-alone machine learning models generally possess nonsynchronous means and median scores across the class labels. Moreover, much greater biases in the prediction when using LR are

Table 4
Data type of machine learning features.

Feature	Data type
Carbonate content	Integer
Specific surface area	Integer
Cation exchange capacity	Integer
Clay fraction	Integer
Soil class	String

towards CH and CL classes, whereas the probability of predicting machine learning is higher compared to other classes as demonstrated by ANN. Overall, the probability of correctly predicting the classes is shown by the ensemble models given their relatively higher and balanced averages.

4.3. Evaluation of machine learning feature importance

Indicators of the importance of machine learning independent features can give insights into the dataset used while also contributing to an improvement of the machine learning model efficiency. Feature importance is a technique that enables an assignment of scores to input variables of a dataset and indicates the relative significance or usefulness of each variable to the goal of machine learning model prediction. For the sake of brevity, the influence of each input feature or variable across each of the classes is assessed in this study and given in Fig. 10 for the best meta-ensemble model. It is observed that the CF seems to have the greatest overall significance on the multiclass prediction of the natural soils' plasticity classes in accordance with the USCS.

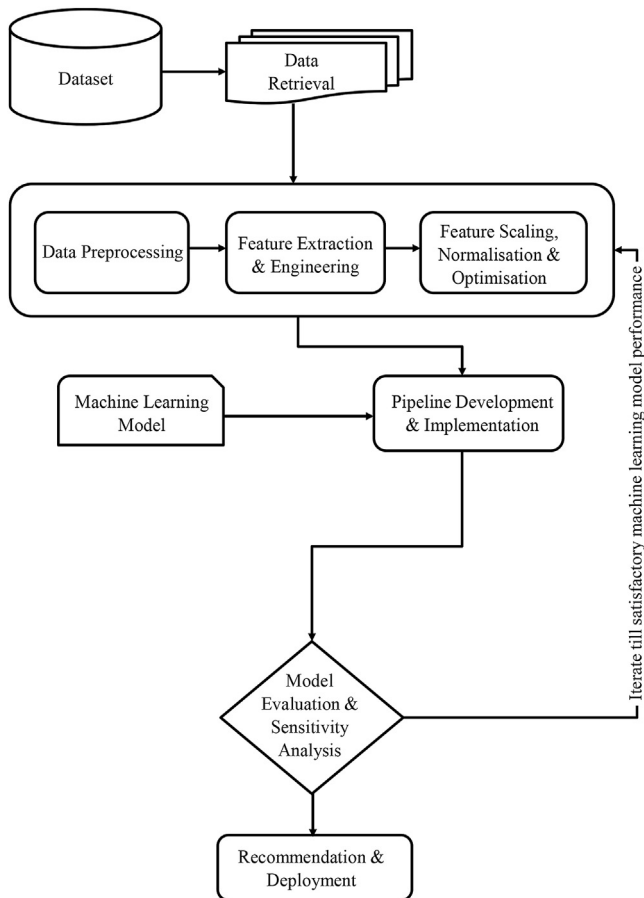


Fig. 6. Machine learning workflow diagram.

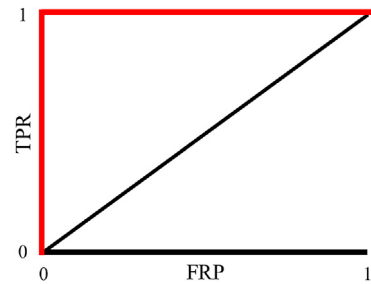


Fig. 7. A typical ROC curve.

Previous studies have established that CF, among other physico-chemical soil properties, does bear a strong relationship with the physical and mineral properties of soil, which in turn are known to affect soil plasticity (Mitchell and Soga, 2005; Nelson et al., 2015). The greatest influence of CEC on the MH soil class is also easily noticeable, along with the importance of SSA (although slightly lower than that of CF) in predicting CL soil classes. Moreover, it was recently confirmed by some probabilistic analyses that a soil's *LL* is most likely to be directly proportional to SSA and CEC (Spagnoli and Shimobe, 2019). The overall importance of SSA and CC in deciding the soil classes are not easily distinguishable, even though the influence of the later appears to be the least when evaluated cumulatively compared to the other features.

Fig. 11 depicts the performance of the VE model by comparing the predicted values of the classes to their true or actual values. As could be observed, VE slightly over-predicted the CH (by approximately 10%) and CL (by approximately 17%) plasticity classes while under-predicting the machine learning (by approximately 30%) plasticity class. It is also clear from Fig. 11 that the MH class was most accurately predicted by the VE model. Fig. 12 depicts the density plot of feature importance in terms of the cumulative average of absolute values. The means of the distribution are higher for CF followed by CEC, CC and then SSA, hence, confirming the level of importance of these features when considered in that order. Also, a more symmetric distribution of the density plots about their mean or average value of importance is observed for the features with the highest level of influence in the multiclass prediction.

4.4. Sensitivity analysis of the best performing models

The meta-ensemble models used in the prediction did produce the best predictions as observed from the foregoing. However, given that a multiclass classification problem is being addressed in this study, it is imperative that the classifier boundaries between

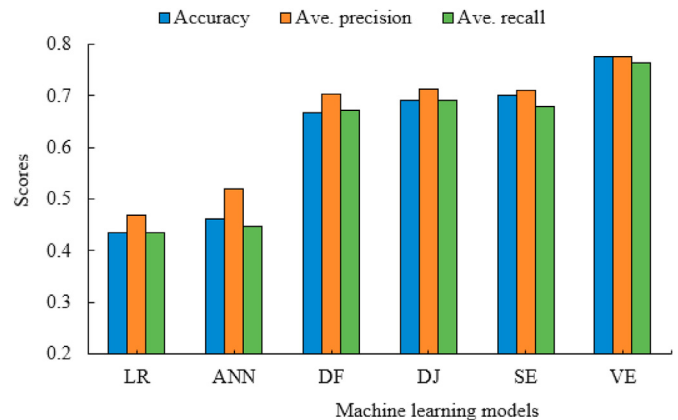


Fig. 8. Machine learning performance scores for all models.

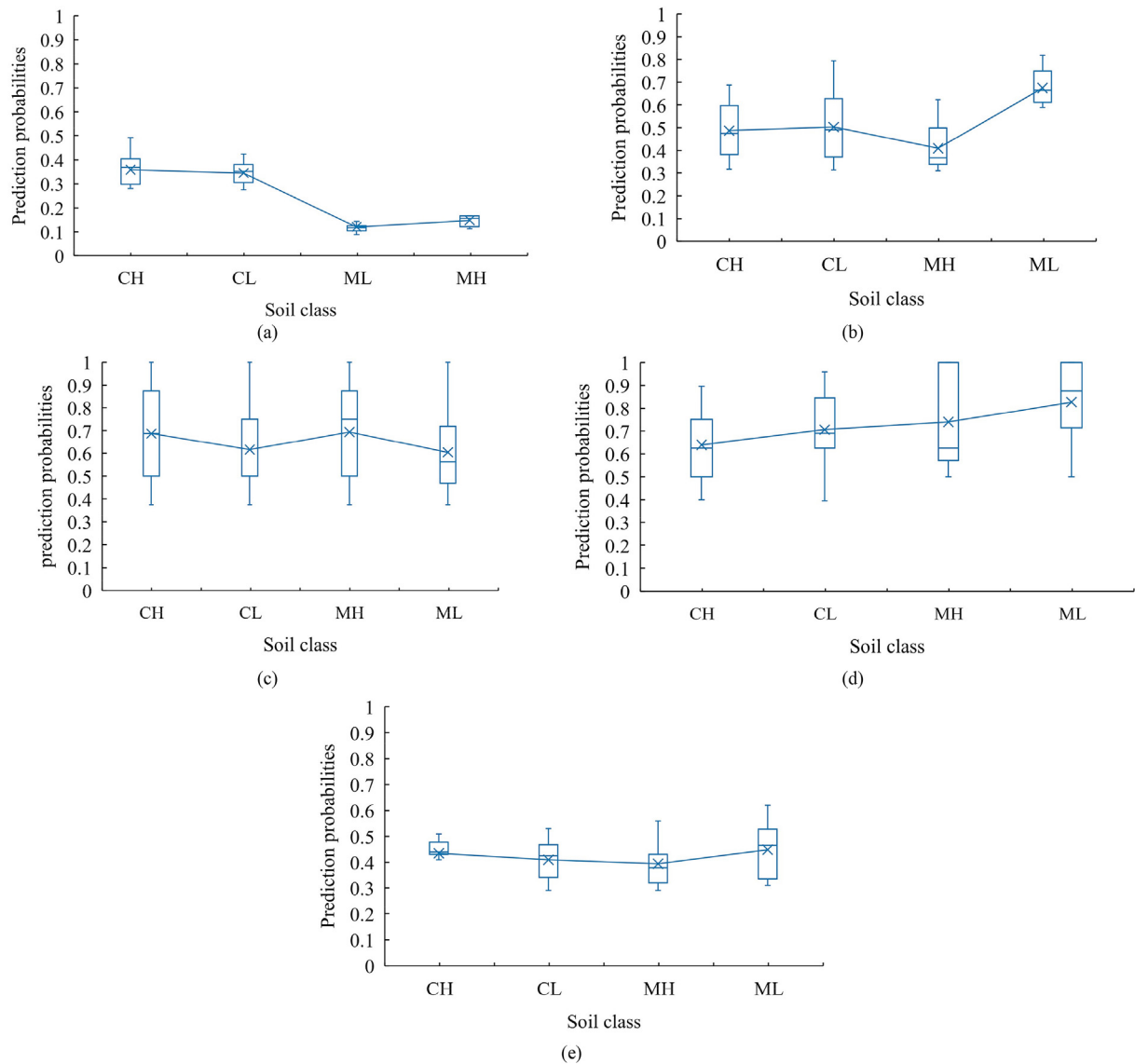


Fig. 9. Prediction probability scores for all models: (a) LR; (b) ANN; (c) DF; (d) DJ; (e) VE.

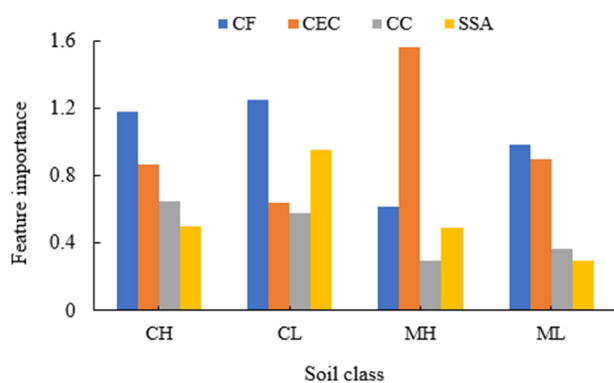


Fig. 10. Feature importance of multiclass prediction for VE model.

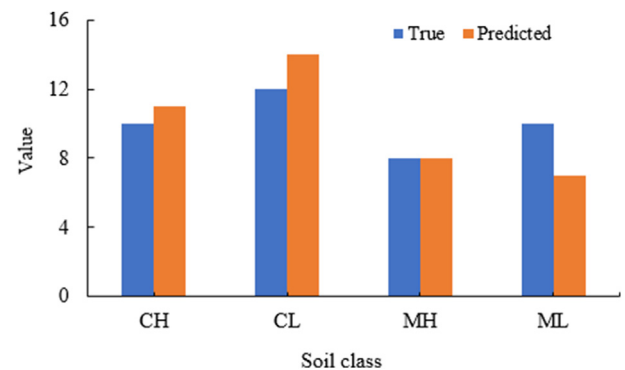


Fig. 11. True versus predicted values for each soil class for the VE model.

the plasticity categories of the natural soils are assessed with respect to a threshold value using the ROC curve.

An analysis of the sensitivity of the meta-ensemble models considered with all four input variables (SSA, CF, CEC, and CC) for

both VE and SE using the method of TVS is indicated in Fig. 13. An examination of Fig. 13 indicates the ability of the meta-ensemble models to distinguish between the soils' class labels. The ROC curve of VE model seems almost indistinguishable from that of SE

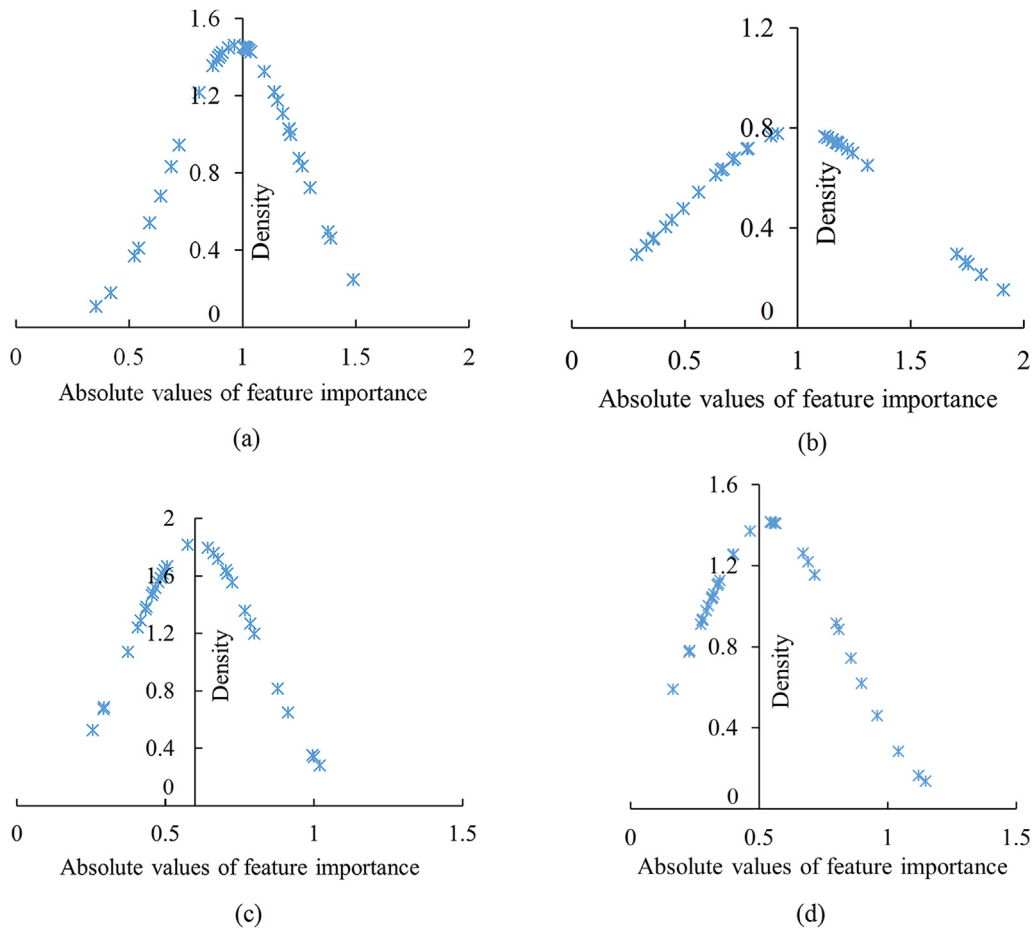


Fig. 12. Average absolute values of feature importance: (a) clay, (b) CEC, (c) CC, and (d) SSA.

model, especially when the multiclass AUC metric of both models is considered. However, when considered in terms of the machine learning score of accuracy, the VE model seemed to outperform SE model by about 10%, as observed previously. However, it is quite revealing through the ROC curve in Fig. 13 that reliance on a single metric score alone to determine the performance of the ensembles may be not entirely sufficient. Notwithstanding, further diagnostic analyses of the performance of the meta-ensembles across different CV methods used in the training and subsequent validation of the dataset are discussed below.

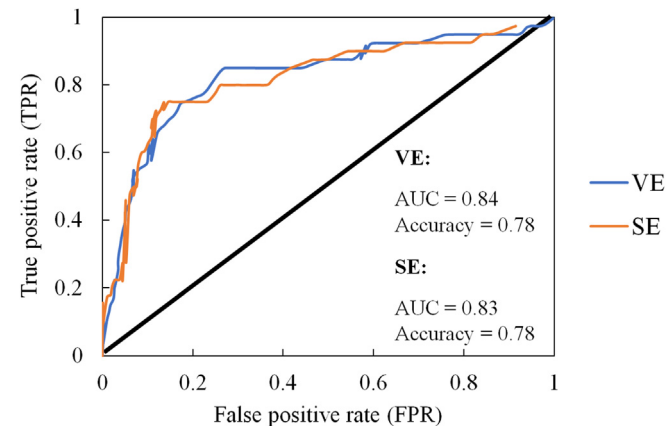


Fig. 13. ROC for VE and SE models using TVS method.

4.5. Comparison between different CV techniques

As mentioned previously, both KFCV and MCCV methods could be applied to improve the prediction of machine learning analyses while also serving as fine-tuning mechanisms to the more conventional TVS method. Although, when considered in terms of their bias–variance trade-offs, the MCCV is mostly deemed as having greater biases than the KFCV but with the former seeming to provide slightly more confidence in machine learning predictions given that it is more repeatable than the later due to its capacity to provide results with lower degrees of variance. As indicated in Fig. 14, the predictions offered using all three methods of training, testing, and validation of the dataset used in this study by adopting the meta-ensemble models do not appear to show any difference in the resulting ROC curves of the VE model. Hence, this does confirm that the multiclass machine learning models used predominantly for training and testing under TVS did not overfit the data, although the resulting metrics for using this technique are slightly higher according to Table 5. On the other hand, when using the SE model for the multiclass machine learning prediction, Fig. 14 and Table 5 demonstrate that the outcome is only slightly better when the dataset is trained, tested, and validated using MCCV and KFCV methods, albeit with the former showing a greater ability to distinguish between the positive classes. Table 5 also shows that the assignment of ‘weights’ to cater for any class imbalance or under-represented class instances does improve the multiclass accuracy metric though only slightly.

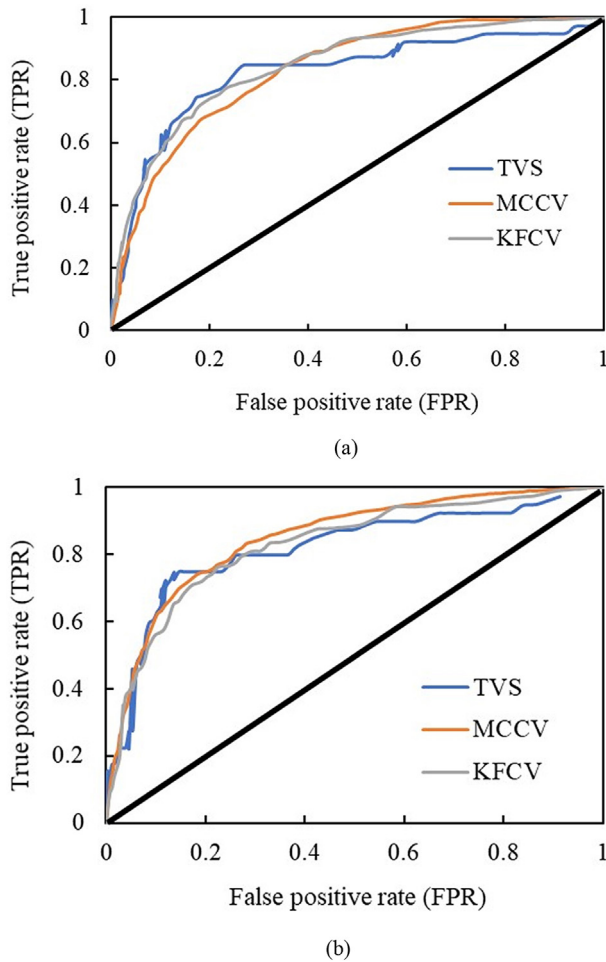


Fig. 14. ROC curves for (a) VE and (b) SE models.

4.6. Lift curve

The lift curve (LC) could provide another means for checking the sensitivity of the meta-ensemble models by lending more credence to their effectiveness in the multiclass prediction of the soil's plasticity class. Again, this is because the prediction that could be achieved from a somewhat 'random' model or through an act of guessing may be incorrect compared to a very good model with a greater percentage of datasets. LC is visually represented as a curve of cumulative gain ratio for any random model. Fig. 15 indicates no difference in the LC when applying the KFCV, MCCV, and TVS techniques to predict the multiclass soil labels, confirming no form of overfitting for dataset training carried out using the VE model.

Table 5
Machine learning metric scores for the meta-ensemble models.

Model	Score		CV technique		
			TVS	MCCV	KFCV
VE	Accuracy	Balanced	0.763	0.631	0.688
		Weighted	0.789	0.612	0.65
	AUC	Weighted	0.84	0.832	0.848
SE	Accuracy	Balanced	0.685	0.651	0.644
		Weighted	0.716	0.637	0.605
	AUC	Weighted	0.83	0.835	0.834

Overall, even though the highest lifting as noticed throughout the entire percentile range occurred with training, testing, and validation carried out with the SE model using the MCCV method, the same amount of lifting (at four) was observed for the three dataset validation methods.

5. Research significance and recommendations for future study

The need to appropriately classify soils, especially from the point of view of geotechnics field practice, cannot be over-emphasised. This is in most parts because an inaccurate classification of a silty soil as a plastic clay soil could result to undesirable cost overruns that are associated with a remedial measure such as stabilisation or even utter disregard of soils by incorrectly assuming they do not meet required specifications. Consequently, this study has utilised an intensive data-driven decision-making approach through the concept of machine learning to rigorously classify soils. It is believed that the procedures adopted in this study have provided a succinct basis for categorising soils because unlike previous techniques, which relied mainly on the activities of first remoulding the soil and then testing or laborious physical measurements to indirectly infer soil classes, the method followed herein is based on properties that intrinsically affect the natural soil behaviour under in-situ conditions or without having to physically work the soil. Granted, not every conceivable physico-chemical property of the soils is considered in this study. Nevertheless, the fundamental basis for future research has been provided. Accordingly, it is

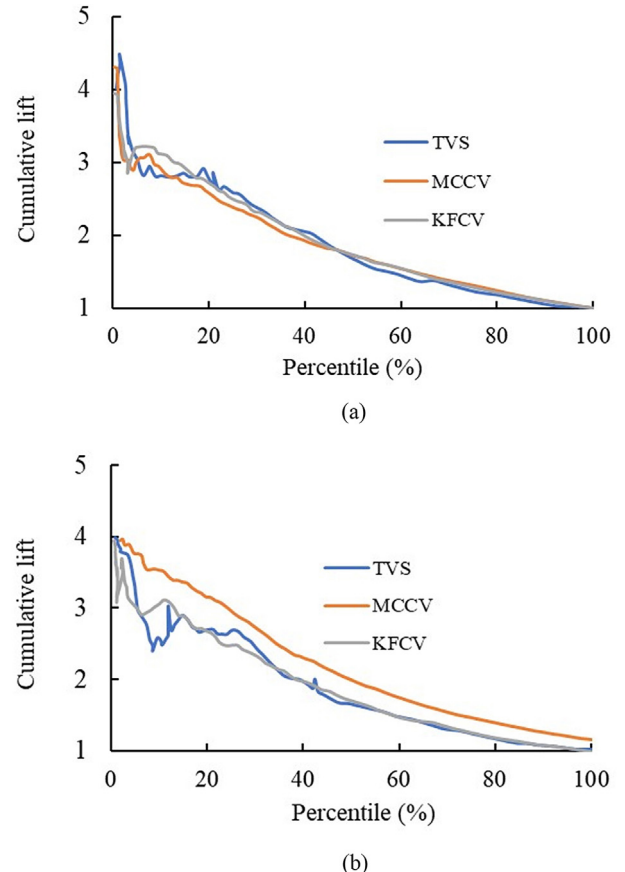


Fig. 15. Machine learning cumulative LCs for (a) VE and (b) SE models.

recommended that soil's inherent properties not covered in this study should be used to further evaluate and forecast soil classes. Besides, statistical comparisons are also suggested between class predictions based entirely on physical, electrical, or mechanical properties and those relying on soil's chemical components.

Meanwhile, for practical deployment and application of the concepts and ideas advanced in the present research, the resources, models, and their predictions, including the developed machine learning background codes utilised, can be carefully saved on an organisation's server, software and hardware assets and the best models reloaded, and predictions made for new soil data for in the classification problem. This procedure can be best implemented during the preliminary phases of a geotechnical site investigation and design.

6. Conclusions

In this study, a novel approach of soil classification based on their physico-chemical characteristics was carried out using the concept of machine learning. Multiclass elements of stand-alone machine learning algorithms, decision tree ensembles and meta-ensemble models were employed to classify soils according to their plastic behaviours. Machine learning predictions were also performed on the best models across multiple CV methods. The following are the important highlights from this study:

- (1) CF had the highest influence on the multiclass prediction of the natural soils' plasticity in accordance with the USCS, while SSA and CC possessed the least significance within the context of this study and the nature of the dataset used.
- (2) Stand-alone machine learning models (LR and ANN) produced less accurate predictive performance compared to tree-based and meta-ensemble models. Overall, the least performing model is the LR (accuracy of approximately 0.436), with the ANN clearly having the highest accuracy of the two stand-alone models (accuracy of approximately 0.462).
- (3) Tree-based ensemble models (DJ and DF) gave less accuracy (accuracy of 0.68, average precision of 0.71, and recall rate of 0.68) when compared to the meta-ensemble models (SE and VE). Overall, the meta-ensemble models, as could be observed, outperformed the rest of the models (accuracy of 0.75, average precision of 0.74, and average recall rate of 0.72).
- (4) Sensitivity analysis of the meta-ensemble models proved their capacities to discriminate between soil classes across different CV models considered. Machine learning training and validation using MCCV and KFCV methods enabled better prediction while also ensuring the dataset were not overfitted by the machine learning models.
- (5) Further confirmation of this phenomenon was depicted by the continuous rise of the cumulative LC of the best performing models when using the MCCV technique. Overall, this study demonstrated that soil's physico-chemical properties do have a direct influence on plastic behaviour and, therefore, can be relied upon to classify soils.

Declaration of competing interest

The authors wish to confirm that there are no known conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgments

The authors hereby wish to offer their thanks and deep appreciation to the reviewers for their thoughtful comments, suggestions and efforts towards improving our article.

References

- Abbey, S.J., Olubanwo, A.O., Ngambi, S., Eyo, E.U., Adeleke, B., 2019. Effect of organic matter on swell and undrained shear strength of treated soils. *J. Civ. Constr. Environ. Eng.* 4, 48–58.
- Abbey, S.J., Eyo, E.U., Oti, J., Amakye, S.Y., Ngambi, S., 2020. Mechanical properties and microstructure of fibre-reinforced clay blended with by-product cementitious materials. *Geoscience* 2020 (10), 241.
- Abbey, S.J., Eyo, E.U., Okeke, C.A.U., Ngambi, S., 2021. Experimental study on the use of RoadCem blended with by-product cementitious materials for stabilisation of clay soils. *Construct. Build. Mater.* 280, 122476.
- Abduljawwad, S.N., 1994. Swelling behaviour of calcareous clays from the eastern Province of Saudi Arabia. *Q. J. Eng. Geol. Hydrogeol.* 27, 333–351.
- Akgün, H., Günel, A., Arslan, A., Youse, K., Öner, G., Kerem, M., 2018. Assessment of the effect of mineralogy on the geotechnical parameters of clayey soils: a case study for the Orta County, Çankırı, Turkey. *Appl. Clay Sci.* 164, 44–53.
- Al-Rawas, A.A., 1999. The factors controlling the expansive nature of the soils and rocks of Northern Oman. *Eng. Geol.* 53, 327–350.
- Arnepalli, D.N., Shanthakumar, S., Rao, H.B., Singh, D.N., 2008. Comparison of methods for determining specific-surface area of fine-grained soils. *Geotech. Geol. Eng.* 26, 121–132.
- Bayat, H., Ebrahimi, E., Ersahin, S., Hepper, E.N., Narain, D., Amer, A.M., Yukselenaksoy, Y., 2015. Analyzing the effect of various soil properties on the estimation of soil specific surface area by different methods. *Appl. Clay Sci.* 116–117, 129–140.
- Bekhor, S., Livneh, M., 2014. Using the artificial neural networks methodology to predict the vertical swelling percentage of expansive clays. *J. Mater. Civ. Eng.* 26, 06014007.
- Bol, E., 2013. The influence of pore pressure gradients in soil classification during piezocone penetration test. *Eng. Geol.* 157, 69–78.
- Cai, G., Liu, S., Puppala, A.J., 2011. Comparison of CPT charts for soil classification using PCPT data: Example from clay deposits in Jiangsu Province, China. *Eng. Geol.* 121, 89–96.
- Casagrande, A., 1948. Classification and identification of soils. *Trans. Am. Soc. Civ. Eng.* 113, 901–930.
- Cerato, A.B., 2001. Influence of Specific Surface Area on Geotechnical Characteristics of Fine-Grained Soils. MSc Thesis. University of Massachusetts, MA, USA.
- Christidis, G.E., 1998. Physical and chemical properties of some bentonite deposits of Kimolos Island, Greece. *Appl. Clay Sci.* 13, 79–98.
- Das, S.K., 2013. Artificial neural networks in geotechnical engineering: modeling and application issues. *Metaheuristics in Water, Geotechnical and Transport Engineering*, first ed. Elsevier, Amsterdam, The Netherlands.
- DeRousseau, M.A., Laftchiev, E., Kasprzyk, J.R., Rajagopalan, B., Srubar, W.V., 2019. A comparison of machine learning methods for predicting the compressive strength of field-placed concrete. *Construct. Build. Mater.* 228, 116661.
- Dexter, A.R., 1990. Changes in the matric potential of soil water with time after disturbance of soil by moulding. *Soil Tillage Res.* 16, 35–50.
- Dreiseitl, S., Ohno-Machado, L., 2002. Logistic regression and artificial neural network classification models: a methodology review. *J. Biomed. Inf.* 35, 352–359.
- Ermias, B., Vishal, V., 2020. Application of artificial intelligence for prediction of swelling potential of clay-rich soils. *Geotech. Geol. Eng.* 38, 6189–6205.
- Erzin, Y., Gunes, N., 2013. The unique relationship between swell percent and swell pressure of compacted clays. *Bull. Eng. Geol. Environ.* 72, 71–80.
- Eyo, E.U., Ng'ambi, S., Abbey, S.J., 2019. Effect of intrinsic microscopic properties and suction on swell characteristics of compacted expansive clays. *Transp. Geotech.* 18, 124–131.
- Eyo, E.U., Ng'ambi, S., Abbey, S.J., 2020. Incorporation of a nanotechnology-based additive in cementitious products for clay stabilisation. *J. Rock Mech. Geotech. Eng.* 12 (5), 1056–1069.
- Eyo, E., Abbey, S., Oti, J., Ng'ambi, S., Ganjian, E., Coakley, E., 2021. Microstructure and physical-mechanical characteristics of treated kaolin-bentonite mixture for application in compacted liner systems. *Sustainability* 13 (4), 1617.
- Eyo, E.U., Abbey, S.J., 2021. Machine learning regression and classification algorithms utilised for strength prediction of OPC/by-product materials improved soils. *Construct. Build. Mater.* 284, 122817.
- Gaidzinski, R., Dualibi Fh, J., Tavares, L.M., 2011. Influence of aging on the technological properties of clays in the ceramic industry. *Appl. Clay Sci.* 54, 47–52.
- Gajurel, A., Mukherjee, P.S., Chittoori, B., 2019. Estimating optimal additive content for soil stabilization using machine learning methods. In: *Proceedings of the Geo-Congress*, pp. 662–672, 2019.
- Goh, A.T.C., Zhang, Y., Zhang, R., Zhang, W., Xiao, Y., 2017. Evaluating stability of underground entry-type excavations using multivariate adaptive regression splines and logistic regression. *Tunn. Undergr. Space Technol.* 70, 148–154.

- Han, T., Siddique, A., Khayat, K., Huang, J., Kumar, A., 2020. An ensemble machine learning approach for prediction and optimization of modulus of elasticity of recycled aggregate concrete. *Construct. Build. Mater.* 244, 118271.
- Hanandeh, S., Ardah, A., Abu-Farsakh, M., 2020. Using artificial neural network and genetics algorithm to estimate the resilient modulus for stabilized subgrade and propose new empirical formula. *Transp. Geotech.* 24, 100358.
- Ikizler, S.B., Aytekin, M., Vekli, M., Kocabaş, F., 2010. Prediction of swelling pressures of expansive soils using artificial neural networks. *Adv. Eng. Software* 41, 647–655.
- Joshi, A., 2020. *Machine Learning and Artificial Intelligence*, first ed. Springer International Publishing. <https://doi.org/10.1007/978-3-030-26622-6>.
- Kaliakin, V.N., 2017. Example problems related to soil identification and classification. In: *Soil Mechanics: Calculations, Principles, and Methods*. Elsevier, Amsterdam, The Netherlands. <https://doi.org/10.1016/b978-0-12-804491-9.00002-1>.
- Kalkan, E., Akbulut, S., 2004. The positive effects of silica fume on the permeability, swelling pressure and compressive strength of natural clay liners. *Eng. Geol.* 73, 145–156.
- Kang, M.C., Yoo, D.Y., Gupta, R., 2021. Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete. *Construct. Build. Mater.* 266, 121117.
- Kayadelen, C., Taşkıran, T., Günaydin, O., Fener, M., 2009. Adaptive neuro-fuzzy modeling for the swelling potential of compacted soils. *Environ. Earth Sci.* 59, 109–115.
- Kim, S.Y., Upneja, A., 2021. Majority voting ensemble with a decision tree for business failure prediction during economic downturns. *J. Innov. Knowl.* 6, 112–123.
- Liao, K.W., Fan, J.C., Huang, C.L., 2011. An artificial neural network for groutability prediction of permeation grouting with microfine cement grouts. *Comput. Geotech.* 38, 978–986.
- Mahmoudi, S., Bennour, A., Meguebli, A., Srasra, E., Zargouni, F., 2016. Characterization and traditional ceramic application of clays from the Douiret region in South Tunisia. *Appl. Clay Sci.* 127–128, 78–87.
- Mehta, B., Sachan, A., 2017. Effect of mineralogical properties of expansive soil on its mechanical behavior. *Geotech. Geol. Eng.* 35, 2923–2934.
- Mitchell, J.K., Soga, K., 2005. *Fundamentals of Soil Behaviour*. John Wiley & Sons, Hoboken, NJ, USA.
- Moreno-Maroto, J., Alonso-Azcárate, J., 2017. Plastic limit and other consistency parameters by a bending method and interpretation of plasticity classification in soils. *Geotech. Test J.* 40, 467–482.
- Moreno-Maroto, J.M., Alonso-Azcárate, J., 2018. What is clay? A new definition of “clay” based on plasticity and its impact on the most widespread soil classification systems. *Appl. Clay Sci.* 161, 57–63.
- Moreno-Maroto, J.M., Alonso-Azcárate, J., O’Kelly, B.C., 2021. Review and critical examination of fine-grained soil classification systems based on plasticity. *Appl. Clay Sci.* <https://doi.org/10.1016/j.clay.2020.105955>, 200.
- Mozumder, R.A., Laskar, A.I., 2015. Prediction of unconfined compressive strength of geopolymer stabilized clayey soil using Artificial Neural Network. *Comput. Geotech.* 69, 291–300.
- Mozumder, R.A., Laskar, A.I., Hussain, M., 2017. Empirical approach for strength prediction of geopolymer stabilized clayey soil using support vector machines. *Construct. Build. Mater.* 132, 412–424.
- Nelson, J.D., Chao, K.C.G., Overton, D.D., Nelson, E.J., 2015. *Foundation Engineering for Expansive Soils*. John Wiley & Sons, Hoboken, NJ, USA.
- Ngun, B.K., Mohamad, H., Sulaiman, S.K., Okada, K., Ahmad, Z.A., 2011. Some ceramic properties of clays from central Cambodia. *Appl. Clay Sci.* 53, 33–41.
- Okeke, C., Abbey, S.J., Oti, J., Eyo, E.U., Johnson, A., Ngambi, S., 2021. Appropriate use of lime in the study of the physicochemical behaviour of stabilised lateritic soil under continuous water ingress. *Sustainability* 2021 (13), 257.
- Pham, B.T., Nguyen, M.D., Nguyen-Thoi, T., Ho, L.S., Koopalipoor, M., Kim Quoc, N., Armaghani, D.J., Le, H. Van, 2021. A novel approach for classification of soils based on laboratory tests using Adaboost, Tree and ANN modeling. *Transp. Geotech.* 27, 100508.
- Polidori, E., 2007. Relationship between the Atterberg limits and clay content. *Soils Found.* 47, 887–896.
- Polidori, E., 2009. Reappraisal of the activity of clays. *Activ. Chart. Soils Found.* 49, 431–441.
- Polidori, E., 2015. Proposal for a new classification of common inorganic soils for engineering purposes. *Geotech. Geol. Eng.* 33, 1569–1579.
- Saito, T., Miki, G., 1975. Swelling and residual strength characteristics of soils based on a newly proposed “plastic ratio chart”. *Soils Found.* 15, 61–68.
- Senol, A., Edil, T.B., Bin-Shafique, M.S., Acosta, H.A., Benson, C.H., 2006. Soft sub-grades’ stabilization by using various fly ashes. *Resour. Conserv. Recycl.* 46, 365–376.
- Shahri, A.A., Malehmir, A., Juhlin, C., 2015. Soil classification analysis based on piezocone penetration test data – a case study from a quick-clay landslide site in southwestern Sweden. *Eng. Geol.* 189, 32–47.
- Shotton, J., Nowozin, S., Sharp, T., Winn, J., Kohli, P., Criminisi, A., 2013. Decision jungles: compact and rich models for classification. In: *Proceedings of the 26th International Conference on Neural Information Processing Systems*, pp. 234–242.
- Smith, C.W., Hadas, A., Dan, J., Koyumdjisky, H., 1985. Shrinkage and Atterberg limits in relation to other properties of principal soil types in Israel. *Geoderma* 35, 47–65.
- Soleimani, S., Rajaei, S., Jiao, P., Sabz, A., Soheilinia, S., 2018. New prediction models for unconfined compressive strength of geopolymer stabilized soil using multi-genetic programming. *Meas.* 113, 99–107.
- Spagnoli, G., Sridharan, A., Oreste, P., Di Matteo, L., 2017. A probabilistic approach for the assessment of the influence of the dielectric constant of pore fluids on the liquid limit of smectite and kaolinite. *Appl. Clay Sci.* 145, 37–43.
- Spagnoli, G., Shimobe, S., 2019. A statistical reappraisal of the relationship between liquid limit and specific surface area, cation exchange capacity and activity of clays. *J. Rock Mech. Geotech. Eng.* 11, 874–881.
- Sridharan, A., Gurtug, Y., 2004. Swelling behaviour of compacted fine-grained soils. *Eng. Geol.* 72, 9–18.
- Tekin, E., Akbas, S.O., 2011. Artificial neural networks approach for estimating the groutability of granular soils with cement-based grouts. *Bull. Eng. Geol. Environ.* 70, 153–161.
- Tinoco, J., Gomes Correia, A., Cortez, P., 2014. Support vector machines applied to uniaxial compressive strength prediction of jet grouting columns. *Comput. Geotech.* 55, 132–140.
- Venkat, B., J. P.A., Sekhar, M.R., Richard, W., 2008. Laboratory procedure to obtain well-mixed soil binder samples of medium stiff to stiff expansive clayey soil for deep soil mixing simulation. *Geotech. Test J.* 31 (3), 225–238.
- Yazdandoust, F., Yasrobi, S.S., 2010. Effect of cyclic wetting and drying on swelling behavior of polymer-stabilized expansive clays. *Appl. Clay Sci.* 50, 461–468.
- Yilmaz, I., 2004. Relationships between liquid limit, cation exchange capacity, and swelling potentials of clayey soils. *Eurasian Soil Sci.* 37, 506–512.
- Yilmaz, I., Kaynar, O., 2011. Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. *Expert Syst. Appl.* 38, 5958–5966.
- Zhang, W., Goh, A.T.C., 2016. Evaluating seismic liquefaction potential using multivariate adaptive regression splines and logistic regression. *Geomech. Eng.* 10, 269–284.
- Zhang, W., Zhang, R., Wu, C., Goh, A.T.C., Lacasse, S., Liu, Z., Liu, H., 2020. State-of-the-art review of soft computing applications in underground excavations. *Geosci. Front.* 11, 1095–1106.
- Zhang, W., Li, H., Li, Y., Liu, H., Chen, Y., Ding, X., 2021. Application of deep learning algorithms in geotechnical engineering: a short critical review. *Artif. Intell. Rev.* <https://doi.org/10.1007/s10462-021-09967-1>.



Dr. Eyo Eyo is a geotechnical and civil engineer. He is currently a lecturer in geotechnical engineering at the University of the West of England, UK. He has also previously worked as a supervising quality control and quality assurance engineer at Fugro GB Marine Ltd., Wallingford, UK. He obtained his MSc degree (with distinction) in Civil and Structural Engineering from Coventry University, UK. Following his MSc degree, Eyo subsequently obtained a scholarship through the European Union/United Kingdom partnership SME KEEN project in 2016 to enable him to conduct PhD research in Geotechnical Engineering at Coventry University. His research interest lies in ground improvement and, more specifically, soil mixing technology (SMT). Eyo has applied artificial intelligence (AI) through machine learning (machine learning) paradigms to ground improvement that has enabled precise and accurate design/data-driven decisions and predictions of the mechanical properties of stabilised soils. He has had about two years of prior experience working as a graduate civil engineer/PA to the project manager at D & S Construction Company Ltd. during the construction of a 66 km road and bridge project at Auch, Nigeria. Before joining Fugro, Eyo worked as an Assistant Lecturer at Coventry University and Associate Lecturer at the University of the West of England, UK. He is currently a registered member (GMICE) of the Institute of Civil Engineers (ICE), UK.