Earth Science Informatics
https://doi.org/10.1007/s12145-024-01398-0

RESEARCH

Assessing the shear strength of sandy soil reinforced with polyethylene-terephthalate: an AI-based approach

Masoud Samaei[1] · Morteza Alinejad Omran[2] · Mohsen Keramati[2] · Reza Naderi[2] · Roohollah Shirani Faradonbeh[1]

Received: 17 April 2024 / Accepted: 24 June 2024

Abstract
This research aimed to investigate the effectiveness of Polyethylene-Terephthalate (PET) as a reinforcement material for sandy soils in enhancing the shear strength. To achieve this, different concentrations of PET were tested, and 118 sets of data were collected. Parameters such as relative density, normal stress in direct shear strength test, and types of PET elements (1×1, 1×5, and fiber) were also recorded. Subsequently, four decision tree-oriented machine learning (ML) methods—decision tree (DT), random forest (RF), AdaBoost, and XGBoost—were applied to construct models capable of forecasting enhancements in shear strength. The evaluation of these models' effectiveness was conducted using four established statistical metrics: R^2, RMSE, VAF, and A-10. The results showed that AdaBoost results in the highest prediction accuracy among other algorithms, representing the high modelling performance of the algorithm in dealing with complex nonlinear problems. The conducted sensitivity analysis also revealed that relative density is the most crucial parameter for all the algorithms in predicting the output, followed by PET percentage and normal stress. Furthermore, to make the developed model in this study more practical and easy to use, a Graphical User Interface (GUI) was created, enabling the engineers and researchers to perform the analysis straightforwardly.

Keywords Polyethylene terephthalate · Soil improvement · AdaBoost · XGBoost · Tree-based algorithms

Introduction

Overcoming the detrimental effects of non-recyclable materials has become a global concern in the last decade. Polyethylene-terephthalate (PET) is one of the non-recyclable materials used to produce liquid containers. On the other hand, using such materials in soil improvement has acquired significant importance during the last few years. According to prior studies by Peddaiah et al. (2018) and Consoli et al. (2002), using PET elements as a reinforcer for loose soils reduces its destructive effects on the environment, the cost of soil improvement projects, and increases the soil's shear strengths. PET, a prominent thermoplastic polymer resin in the polyester family, is extensively utilised across a wide range of applications. This includes fabricating packaging films, creating bottles and jars for consumer goods, in injection moulding processes for making durable items, and as a reinforcing material in certain composite products. Generally, this material is used in industry in textiles, rigid packaging, flexible packaging, photovoltaic modules, thermoplastic resins and waterproofing barriers. This material in geotechnical engineering is used with different shapes of fiber and/or crushed pieces with specific dimensions to improve soil specifications. Soil behaviour with the PET elements in the California bearing ratio (CBR) test was investigated by Sinha et al. (2019). Furthermore, Maher and Ho (1994) conducted the triaxial cyclic test to show

Communicated by: Hassan Babaie

Roohollah Shirani Faradonbeh
roohollah.shiranifaradonbeh@curtin.edu.au

Masoud Samaei
masoud.samaei@curtin.edu.au

Morteza Alinejad Omran
alinejad1993@gmail.com

Mohsen Keramati
keramati@shahroodut.ac.ir

Reza Naderi
r_naderi@shahroodut.ac.ir

[1] WA School of Mines: Minerals, Energy and Chemical Engineering, Curtin University, Kalgoorlie, WA 6430, Australia

[2] Department of Civil Engineering, Shahrood University of Technology, Shahrood, Iran

Published online: 10 July 2024

the behaviour of the PET content in cemented sand. In another study, the uniaxial compression test was undertaken on the reinforced cohesive soil with PET element and the stress–strain response of reinforced soil with 0% to 1% PET waste was investigated by Li and Ding (2002) and Babu and Chouksey (2011). The findings revealed a notable enhancement in the specimens' strength, friction angle, and the uniaxial strength of soil reinforced with 1% material, showing an increase of 73.8% compared to the unreinforced soil. Consoli et al. (2002) conducted a study on the engineering properties of sand reinforced with PET fibers. In their study, the dimension of PET elements was approximately 36 mm, and the percentage of PET was 0.9 per cent of soil weight. The results showed that PET increases the soil strength and soil specifications.

The behaviour of sandy and clayey soil with PET reinforcements was investigated by Acharyya et al. (2013). Their research determined that reinforcing sandy soil with PET fibers proves to be more effective compared to its application in clayey soil. Furthermore, the optimum percentage of the reinforcement material in the sand is higher than in clay, and the optimum wet percentage in clayey soil increases with the PET content. Alvarez et al. (2020) and Louzada et al. (2019) also investigated the behaviour of clayey soils. Botero et al. (2015) conducted an unconsolidated undrained triaxial test with different equivalent pressures (i.e., 2.5, 5, and 7.5 m) and with various PET contents (i.e., 0, 0.3, 0.6, and 1%). Their study showed that PET reinforcements result in a decrease in friction angle and an increase in cohesion. Necmi and Ekrem (2020) conducted uniaxial compressive tests to investigate the strength properties of clayey soils reinforced with PET fibers derived from waste plastic. Both exposed and unexposed samples were tested for their resistance to freeze–thaw cycles using a programmable cabinet under laboratory conditions. The findings of the experimental study indicate that the waste PET fibers enhance the strength of reinforced clayey soil samples compared to unreinforced sandy soil samples. Additionally, the PET fibers increase the resistance of reinforced clayey soil samples to the effects of freeze–thaw cycles.

Peddaiah et al. (2018) explored the effects of integrating PET strips into the mechanical characteristics of silty sand, employing a series of laboratory experiments such as compaction, direct shear, and California Bearing Ratio (CBR) tests on samples of silty sand blended with varied proportions of plastic strips and differing aspect ratios. The results indicated significant improvements in the maximum dry unit weight, shear strength parameters, and the CBR values of the soil upon the inclusion of plastic reinforcements. However, the degree of enhancement in soil properties was contingent upon the amount of plastic used, the dimensions of the strips, and the type of soil. The research concluded that a plastic content of 0.4% and a strip dimension of 15 mm × 15 mm markedly bettered the engineering properties of silty sand. In addition, the study examined the behavior of soil reinforced with PET through direct shear and triaxial tests, where Patil et al. (2016) found a notable augmentation in soil cohesion and strength due to PET incorporation.

Moghaddas Tafreshi et al. (2021) investigated the reinforced soil's cyclic behaviour. PET reinforcement was discovered to enhance all reinforcement configurations, with a slightly more significant improvement observed for PETs of smaller size. Unlike the unreinforced condition, settlement accumulation rates decreased, and resilient settlement increased in reinforced cases. With PET reinforcement, there was an average 42% improvement in lower cyclic loading but a 179% improvement in high cyclic loading, indicating that PET reinforcement is especially useful under critical loading. An extra layer of geogrid over the PET-reinforced zone reduced soil settlement by as much as 82% compared to the case without geogrid, as it increased the load distribution area and prevented PET distortion (Moghaddas Tafreshi et al. 2021). Table 1 summarises recent studies on PET and provides detailed information about each one.

There have been several research studies that have explored the impact of PET on reinforced soil's drainage conditions, pore water pressure, seismic behaviour, and stabilisation of soil used for pedestrian purposes (Shariatmadari et al. (2020); Fathi et al. (2020), Hafez et al. (2019) and Carvalho et al. (2019)). In addition, Mishra and Gupta (2018) showed the behaviour of the PET mixtures with Fly ash in clayey soil of flexible pedestrians. This material behaviour has also been investigated in several slope stability projects (Nadaf et al. 2019).

From the prior studies, it can be inferred that PET reinforcements significantly improve soil specifications as efficiently as other improvers. Furthermore, as the ratio of PET content increases, the soil strength increases. Therefore, this type of reinforcer results in an acceptable improvement in the maximum shear stress, friction angle and cohesion. Due to these materials' availability, knowing this reinforcer's accurate behaviour can be effective in geotechnical analysis. The literature review shows the influence of various parameters on improving soil's mechanical parameters. However, the degree of accuracy and influence of parameters on each other has yet to be well understood.

Recently, machine learning (ML) algorithms have been introduced to do deeper analyses and develop more accurate formulas and models for geotechnical engineering problems (Samaei et al. 2018, 2022; Momeni et al. 2023). According to the authors' awareness, ML methods have yet to be applied to forecast the shear strength of soils reinforced with PET. Predicting shear strength in soil strengthened by PET plastic elements involves complex, nonlinear interactions among various factors, including the type and content percentage of plastic elements, normal stress,

Table 1 Summary of recent studies on PET application for soil improvement

Results	Optimum PET Percentage	Soil Type	PET Content	Conducted Experiments	Reference
Increased peak compressive strength and ductility	1%	Kaolinite clay	PET fiber 0.5% to 4% by weight	Triaxial cyclic test	Maher and Ho (1994)
The unconfined compressive strength and the tensile strength of the cemented sand were significantly increased by fiber reinforcement	0.5%	Uncemented and artificially cemented sand	PET fiber 0.1 to 0.9% of soil dry weight	Unconfined compression tests, splitting tensile tests and drained triaxial compression tests	Consoli et al. (2002)
Significant improvement in the strength of the soil	1%	Red clayey soil and sand	PET chips 0.5%, 0.75% and 1.0% by dry weight of soil	Uniaxial compression test	Babu and Chouksey (2011)
Increase in the apparent cohesion	0.6%	Silty soil	PET fiber 0.3, 0.6, 1% of the soil dry weight	Unconsolidated undrained triaxial test	Botero et al. (2015)
Improvement of properties for clayey soil with the addition of plastic strips	0.4%	Silt, clay and sand	PET strips 0.2, 0.4, 0.6 and 0.8% of soil dry weight	Direct shear and California bearing ratio (CBR) tests	Peddaiah et al. (2018)
Reduce soil brittleness - increase in the damping ratio and decrease in shear modulus	1%	Sand	PET strips 0.5%, 0.75% and 1% by the soil weight	Shaking table tests	Fathi et al. (2020)
Mixtures containing higher PET content have almost identical drained and undrained behaviour	0.5 – 0.7%	Sandy soil (SP)	PET chips 0.6, 1, 1.5% of the total mixture weight	CD and CU triaxial tests	Shariatmadari et al. (2020)
Increase in cohesion and decrease in friction angle and adopt a better performance, more uniform and less vertical variation	1%	Clay soils of high plasticity	PET crushed (3–5 mm) 0.5, 1, 2.5, 3.5% by weight of the soil	Direct shear test	Alvarez et al. (2020)
Increase in the strength of the clayey soil	0.3%	Clayey soils	PET fiber 0.1, 0.2, 0.3% by the total weight of samples	Uniaxial compressive tests	Necmi and Ekrem (2020)
The achieved performance improvement is demonstrated here with the change in the plate settlement due to the reinforcement	Not measured	Sandy soil	PET bottles	Cyclic loads	Moghaddas Tafreshi et al. (2021)

relative density, and specific weight of soil (Acharyya et al. 2013, Tofigh Tabrizi et al. 2021). Traditional statistical methods often fall short in capturing these intricate relationships, whereas ML algorithms are designed to handle such complexity effectively, contributing to more sustainable and efficient engineering practices (Frank et al. 2020). ML techniques have demonstrated superior accuracy and generalisation capabilities in predictive modelling, which is crucial for ensuring the safety and reliability of engineering designs (Frank et al. 2020). Given the dataset derived from the current study's experimental results, ML algorithms are particularly advantageous due to their ability to process small volumes of data and extract meaningful patterns (Naghadehi et al. 2018).

Numerous studies in geotechnical engineering have successfully applied ML algorithms, underscoring their effectiveness in predicting soil properties and enabling engineers and researchers to leverage these techniques for future applications (Onyelowe et al. 2023). Several studies have explored different methodologies for predicting the shear strength of soil and understanding soil behaviour under various conditions (Rabbani et al. 2023a, 2023b, 2023c, 2023d, 2024). Tree-based models are widely used in machine learning due to their high accuracy and grey-box nature, which enhances interpretability and facilitates further research applications (Rabbani et al. 2024). In contrast, models like Artificial Neural Networks (ANNs) are often considered black-box models, making them less interpretable and, thus, less useful for certain applications (Samaei et al. 2018).

This research introduces advanced tree-based, non-destructive methods for accurately and reliably predicting shear strength in reinforced soils. These approaches are expected to significantly assist geotechnical engineers and researchers in evaluating soil behavior and performance. Specifically, XGBoost and AdaBoost are highlighted for their suitability in modeling the shear strength of soils enhanced with PET. XGBoost is noted for its high performance, regularization to prevent overfitting, efficient handling of missing data, and optimized computation (Chen and Guestrin 2016). AdaBoost, on the other hand, excels through its adaptive boosting mechanism, which focuses on the hardest-to-predict instances and iteratively adjusts weights, resulting in simpler models that generalize well and provide higher accuracy in certain scenarios (Dos Santos Aguiar and Paulo 2019). Other algorithms, such as neural networks, SVM, and linear regression, require larger datasets, extensive computational resources, or are too simplistic for capturing the complex relationships inherent in geotechnical data (Liu et al. 2022). Consequently, XGBoost and AdaBoost are preferred for their robustness and efficiency in this context. This study compares these advanced algorithms to traditional tree-based methods like Decision Tree (DT) and Random Forest (RF), demonstrating their capabilities in modelling the shear strength of PET-reinforced soils.

Materials and methods

This study started with data collection from soil samples extracted at Anzali Port, Iran. It proceeded by preparing soil samples combined with different percentages of PET materials, followed by direct shear tests. The data collected was then analyzed using machine learning models, including XGBoost, AdaBoost, Random Forest, and Decision Tree, with an emphasis on hyperparameter optimization. The models' performance was evaluated and ranked based on specific metrics. Finally, a graphical user interface (GUI) was developed to integrate the entire process for practical application. Figure 1 represents the flowchart of the methodology followed in this study. The foregoing stages are discussed in more detail in the following sections.

Studied soil

In this study, the material used is the sandy soil extracted from Anzali Port in Iran, which is reinforced with PET plastic waste material. First, several tests were carried out on the soil to identify its mechanical properties. The tests include sieve analysis, specific gravity (Gs) as well as a series of tests to determine the specific weight of soil (γ_d) to apply the desired relative density (Dr) on the soil and to determine the properties of PET, such as specific gravity, modulus of elasticity, and tensile strength. After combining the reinforcement with soil, the reinforced samples were subjected to the direct shear test.

The studied sandy soil in this research is carbonated sand without fine grains, rounded corners, and uniform and poorly grained (see Fig. 2), which falls into the poorly graded sand (SP) class according to the Unified Soil Classification System (USCS). It should be noted that all the experiments were conducted in dry conditions. The other mechanical properties of the soil, such as uniformity coefficient (Cu) and coefficient of curvature (Cc), are listed in Table 2, which were measured using the tests and based on the ASTM D422-43 standard (ASTM-D422 2016).

The distribution of particles larger than 0.075 mm (soil remaining on the 200 sieve) was determined by conducting a granulation test by sieve. Since the percentage passing through the 200 sieve was very low (less than 1% by weight of the soil), the soil was considered coarse grain material, and there was no need to conduct a hydrometric test. The sand used in this research is uniform, and the soil particles have a limited size range. This can be observed in the soil distribution curve shown in Fig. 3. As shown in Fig. 3 (the

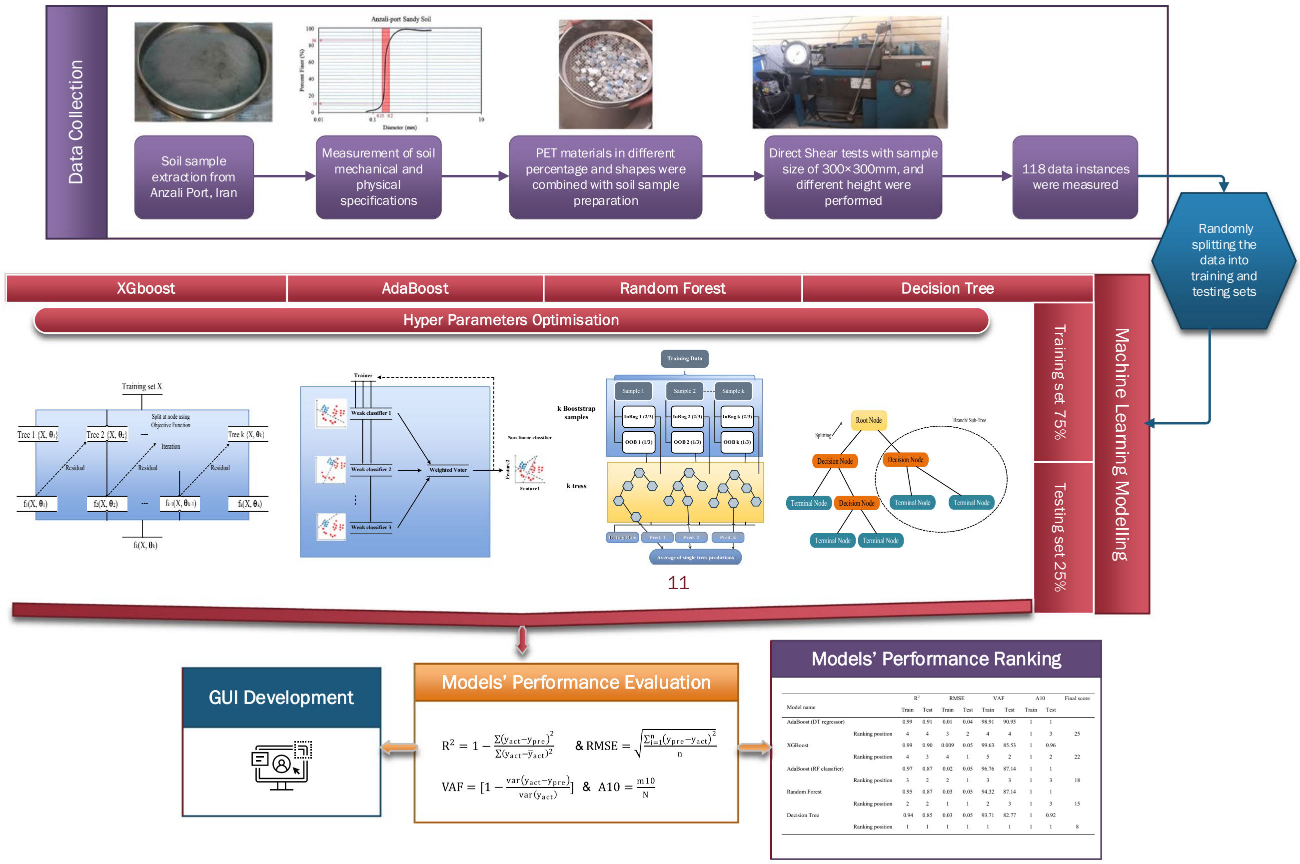


Fig. 1 Flowchart illustrating the methodology followed in this study

Fig. 2 The sandy soil used in this study

Table 2 Mechanical properties of sample soil

Properties	Unit	Value
Gs	-	2.65
$(\gamma d)_{min}$	gr/cm^3	1.6
$(\gamma d)_{max}$	gr/cm^3	1.25
D50	mm	0.17
Cu	-	1.25
Cc	-	1.07

red zone), the diameter range of grains is very narrow, and most of the grains are in almost the same dimension range. In other words, this is poorly-grained sand or the SP soil in the USCS. The preparation of the reinforced soil samples was carried out to achieve three distinct relative densities: 55%, 75%, and 95%, with the aim of examining how relative density influences the parameters of soil shear strength.

PET preparation

For this study, PET fragments were obtained by cleaning, drying, and then cutting discarded plastic bottles into chips measuring 1×1 cm and 1×5 cm with a thickness of 0.5 mm, or alternatively, into fibers (Fig. 4). The measured characteristics of the prepared PETs are listed in Table 3. In most prior studies, the optimal PET percentage was 1% (Malidarreh et al. 2018). However, in this study, four different percentages of PET (i.e., 0.1%, 0.5%, 1%, and 2%) were

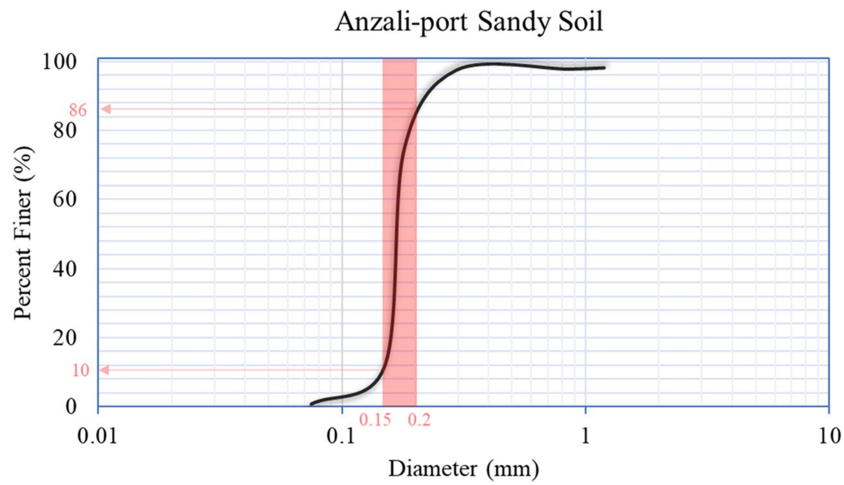


Fig. 3 The particle size distribution of the soil utilized in this study

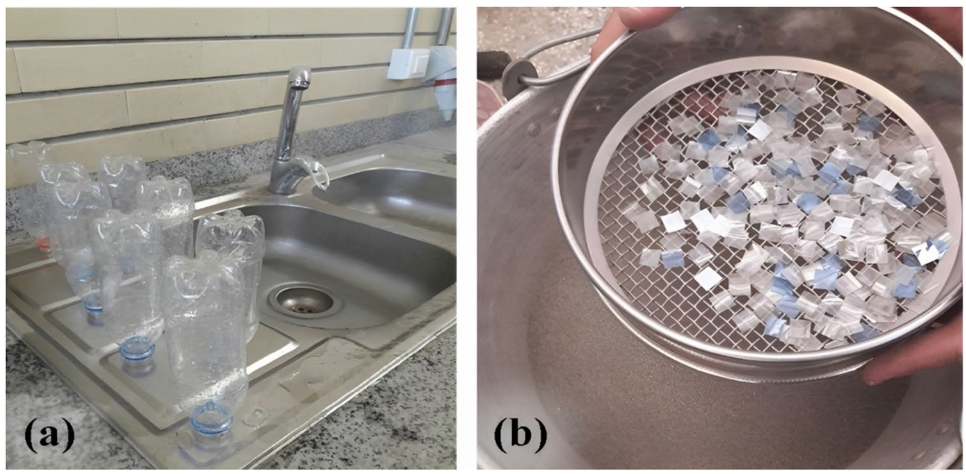


Fig. 4 Preparation of PET materials for the experiments: **a**) PET containers are washed and dried before being cut to the desired sizes to prevent adhesion errors, given that the containers are reused and recycled, and **b**) Shredded PET samples before adding to the soil

Table 3 Measured characteristics of PET elements

Length (mm) × width (mm)	Thickness (mm)	Gs(g/cm^3)	Chemical formula
10 × 10	0.5	1.38	$(C_{10}H_8O_4)_n$
10 × 50			
Fiber			

utilised to examine the impact of PET on enhancing shear strength in soil within a certain tolerance.

Direct shear tests were conducted to determine the soil's shear strength in an unreinforced and reinforced state with PET fragments. The tests were undertaken on 300 × 300 mm samples according to the ASTM D 3080–90 standard (ASTM-D3080 2011) and under three different normal stresses (Sn) of 50, 100, and 150 kPa on samples with a relative density of 55%, 75%, and 95% in the dry state, up to a strain of 10% and with a constant speed of 2.067 mm/min. For the purpose of creating predictive models, a dataset was compiled from 118 sample tests, with all relevant features such as PET type, PET percentage, Dr, Sn, and shear strength meticulously recorded for each test. Furthermore, as the PET type is a qualitative parameter, the values of 1 to 3 were replaced with the PET types 1 × 1 cm and 1 × 5 cm and fibers, respectively, to make all the independent variables quantitative. The descriptive statistics of the dataset are presented in Table 4.

Prior to predictive model development using ML algorithms, the variation in soil shear strength was evaluated relatively, given the utilisation of soils with distinct parameter ranges in the experiments. Consequently, the modelling process employed the ratio of increases. A part of the employed data in the study is presented in Table 5. As seen in this table, the percentage increase in soil shear strength without PET is indicated in the adjacent column, with a reference value of 1.

Table 4 Descriptive statistics of the recorded parameters

Category	Symbol	Unit	Min	Max	Avg	St deviation
Input	PET type	-	1	3	1.846	0.772
	Dr	%	0.55	0.95	0.75	0.164
	Sn	kPa	50	150	100	41
	PET percentage	%	0	2	0.6	0.687
Output	Shear Strength (τ_{max})	kPa	36.31	144.63	84.88	31.903
	Shear Strength increase ratio ($\tau_{max\ ratio}$)	-	0.86	1.42	1.10	0.128

Table 5 Sample dataset used in this study

PET type	Dr	Sn	PET percentage	τ_{max}	$\tau_{max\ ratio}$
1	0.55	50	0	36.313	1
1	0.55	50	0.1	37.371	1.0291
1	0.55	50	0.5	40.768	1.1226
2	0.55	50	0	39.313	1
2	0.55	50	0.1	47.077	1.197
2	0.55	50	0.5	43.680	1.111
3	0.95	150	0	110.657	1
3	0.95	150	0.1	144.631	1.307
3	0.95	150	0.5	130.070	1.175

Machine learning algorithms

Decision Tree

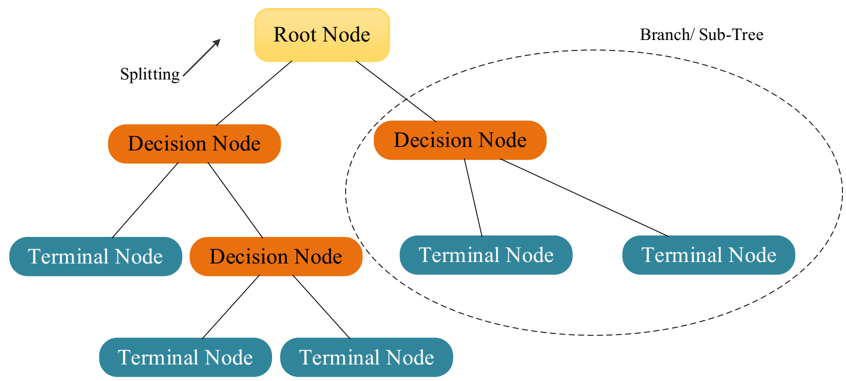
The decision tree (DT), a fundamental component of supervised learning in machine learning and artificial intelligence, has been deployed for tasks involving classification and prediction modeling. The DT methodology encompasses a variety of subsets, including chi-squared automatic interaction detection (CHAID), Quick, Unbiased, Efficient, Statistical Tree (QUEST), C5, and the Classification and Regression Tree (CART). Among these, CART and CHAID are unique in their ability to model and predict continuous variables, as noted by Lin and Fan (2019). The CART algorithm, in particular, is favoured for its "white-box" nature, offering straightforward interpretability and ease in understanding the connections between input and output variables. Additionally, CART's performance remains robust across large datasets, distinguishing it from other more opaque machine learning algorithms, especially in scenarios involving complex samples and a high variable count (Li and Sheu 2021). Samaei et al. (2018) highlighted that CART leverages principal component analysis (PCA) to pinpoint critical input variables while discarding those deemed insignificant. Depending on the nature of the output variable, a CART decision tree can be tailored as either a classification tree (CT) or a regression tree (RT). A CART model's effectiveness for a specific dataset is notably affected by key indices. For a CT, these critical indices encompass the Gini criterion, Entropy, and the Twoing criterion. Furthermore, the non-parametric nature of CART models means that assuming a distribution for the variables is unnecessary, allowing for flexibility in handling various types of data.

During the development of a CART model, implementing specific criteria is crucial to avoid overfitting and ensure the model's applicability to new datasets. These preventative measures include setting a minimum number of observations required for a node split, defining the maximum depth of the tree, and determining the smallest error reduction necessary for splitting a node. These parameters, which can be adjusted in the Scikit-Learn package during CART model development, help control the complexity of the tree and prevent overtraining. By applying these criteria, the training process can be guided to produce a simpler, more generalized tree. The structure of a decision tree comprises a root node from where the decision-making process starts, decision nodes (or interior nodes) that represent the points where choices are made, and terminal nodes (or leaf nodes) that signify the outcomes of those decisions (Samaei et al. 2018). Every tree originates from the root node, positioned at the top level of the tree, and bifurcates into left and right sub-branches. Each sample in the dataset is categorized starting from the root node, progressing through the decision nodes, until it can no longer be split, culminating in a terminal node (refer to Fig. 5) (Lin and Fan 2019).

Random forest

The Random Forest (RF) algorithm, established on the foundation of the CART decision tree algorithm by Breiman (2001), amalgamates numerous DTs to form a composite model. Zhou et al. (2020) demonstrated its versatility as both a classification and regression tool, highlighting its capability without necessitating prior assumptions about the relationship between predictor variables and the response variable. Initially, the RF algorithm generates samples through the bootstrap sampling technique from the dataset, with each bootstrap sample leading to the creation of an individual RF tree. Samples not selected during the bootstrap

Fig. 5 A Decision Tree consisting of the root nodes, decision nodes, and terminal nodes



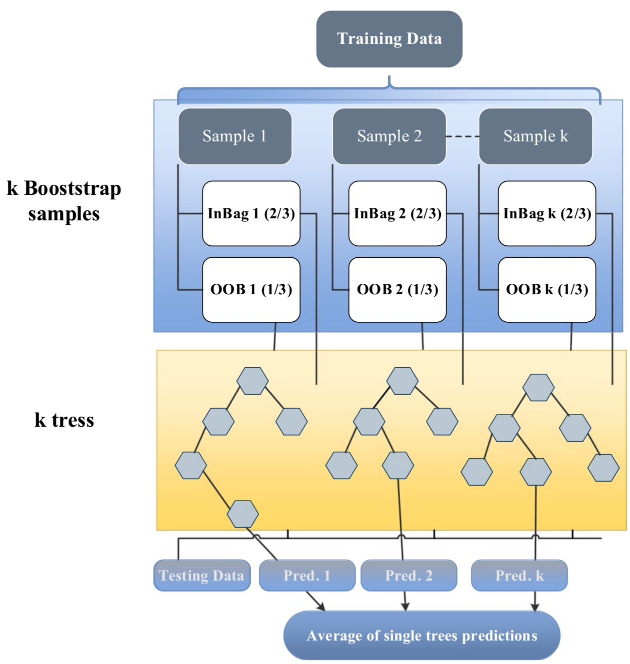


Fig. 6 The process of RF trees' growth

process, known as out-of-bag (OOB) samples, are utilized for validation purposes. The procedural framework of the RF algorithm is depicted in Fig. 6.

Probst et al. (2019) illustrated that several parameters can regulate the architecture of each tree within an RF: (i) the smallest allowable node size, (ii) the aggregate number of trees within the forest, and (iii) the level of randomness applied during the construction of trees. The RF algorithm offers several benefits, such as (a) the ability to save constructed trees for later reference, (b) resistance to overfitting issues, (c) shorter training times coupled with faster prediction capabilities, and (d) a built-in feature selection process that aids in prioritising parameters based on their significance. According to Witten et al. (2005), these attributes distinguish the RF algorithm from other machine learning methodologies.

In prior studies, the Gini-index splitting criterion has been used frequently as a measure to assess the level of inequality or discrepancy within a distribution. By using this criterion as a basis for splitting, decision trees can effectively divide up the feature space to enhance their predictive power. This has resulted in the Gini index becoming a widely utilised tool for evaluating theoretical concepts and practical applications (Charles et al. 2022). This research utilized the Scikit-Learn Python library (Pedregosa et al. 2011), which provided two criteria for splitting data: the Gini index and Entropy, both of which were applied in the development of the model. The Gini index assesses the purity of nodes to determine the lowest possible error rate for the chosen training data group. On the other hand, the Entropy method measures the uncertainty or randomness information in a data group, helps determine which feature to split on at each tree node, and splits the tree to give more information (Hannan and Anmala 2021). If the Entropy is extended, the subsets homogeneity improves (Kuhn and Johnson 2013).

XGBoost

Chen and Guestrin (2016) introduced extreme gradient boosting (XGBoost), a scalable approach to ML that enhances tree-boosting techniques. This algorithm has outperformed other ML algorithms across various engineering challenges, according to Zhu et al. (2021). Figure 7 illustrates that XGBoost merges weaker learners with stronger ones, utilising residuals from each iteration to refine earlier predictions. Unlike weak learners like DT or RF, which rely on random guesses and lack the capability to leverage errors for improving the ultimate model, the XGBoost algorithm employs gradient boosting. This enables it to use errors constructively to bolster the performance of the final model.

Dhaliwal et al. (2018) elucidate the salient advantages of employing the XGBoost algorithm, as enumerated below:

1. XGBoost exhibits a superior speed, being approximately ten times faster than other prevalent algorithms, which mitigates the challenges of temporal expenditure in

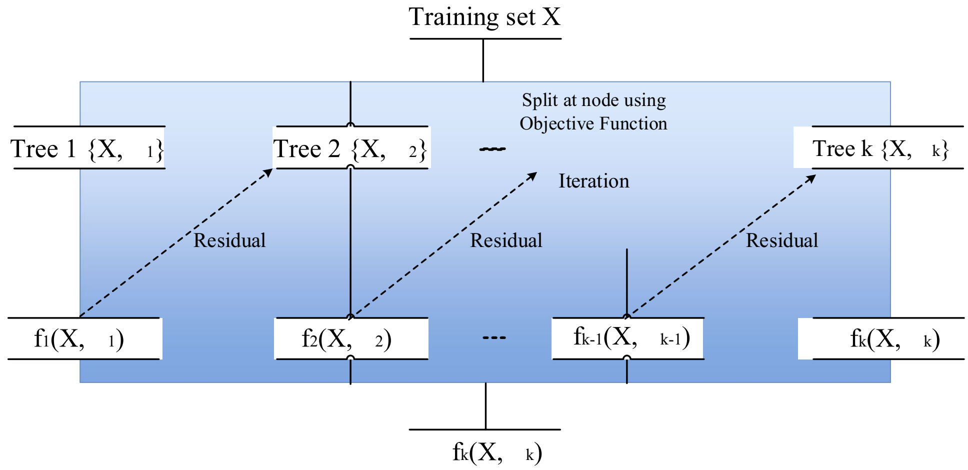


Fig. 7 The XGBoost algorithm involves a variable X and DT features denoted by Θ (such as tree depth, number of splits, etc.), which are utilised to construct a tree aimed at predicting X. Here, *f* represents the tree that has been learned

processing large datasets and computationally intensive tasks.

2. The algorithm's capability for parallel processing facilitates efficient scalability and the production of a copious number of instances while ensuring minimal resource consumption. This attribute is crucial for the efficacious classification and sophisticated preprocessing of data.
3. XGBoost's compatibility with a wide array of programming languages, including Java, Python, R, and C + +, extends its applicability across diverse computational environments.
4. Its strategic synthesis of weak learners into robust learners significantly diminishes the likelihood of overfitting and concurrently elevates the precision of predictive outcomes.
5. The feature of inherent cross-validation within the training dataset obviates the necessity for auxiliary cross-validation frameworks.
6. XGBoost undertakes extensive model evaluation processes to ascertain peak operational efficiency, meticulously adjusting parameters to avert overfitting and circumvent unnecessarily complex solutions.

To mitigate the risk of overfitting, it is imperative to adjust several internal parameters within the XGBoost framework, as delineated by (Chen and Guestrin 2016). These include the number of iterations, which denotes the quantity of trees integrated into the model; the maximum depth, indicative of the highest number of bifurcations; the subsample, representing the proportion of the dataset allocated for training purposes; the learning rate, which adjusts the model's weights to enhance performance; the colsample_bytree, reflecting the fraction of features used for tree construction; and lambda and alpha, which serve to regularise the model's weights, with higher values promoting a more cautious model approach.

AdaBoost

Studies from the 1990s revealed that combining multiple weak learners can result in a robust learner combination (Schapire 1990; Freund and Schapire 1997). Following this insight, Schapire (1990) laid down the initial theoretical framework for the boosting algorithm, which seeks to address this concept. The essence of boosting, a technique embedded in machine learning, lies in its ability to integrate a multitude of weak and less accurate predictors to create a highly precise prediction model. Freund and Schapire then improved this idea in 1997 by introducing Adaptive Boosting (AdaBoost) algorithm. Figure 8 shows the algorithm structure and the process of modelling. To train the AdaBoost algorithm, weights are equalised for all training examples. Then, weak models are trained on the training dataset. These weak models can be decision trees with low depth or logistic regression with few features. Next, the error of the weak models on the training dataset is calculated. The error is defined as the fraction of misclassified examples. The weights of the misclassified examples are increased so that they have more influence on the next weak model. This ensures that subsequent weak models focus more on correctly classifying these data points. The process of training a new weak model, calculating its error, and increasing the weights of the misclassified examples is repeated until a set number of weak models are created or the desired level of accuracy is reached. The final step is to combine the weak models by weighting their predictions according to their accuracy and outputting the final prediction (Azmi and Baliga 2020).

The application of the AdaBoost classifier leads to the removal of superfluous training data, including outlier features. This process enhances the classifier's accuracy by diminishing bias and variance errors through iterative training (Rajesh and Dhuli 2018). It is noteworthy to mention that while the

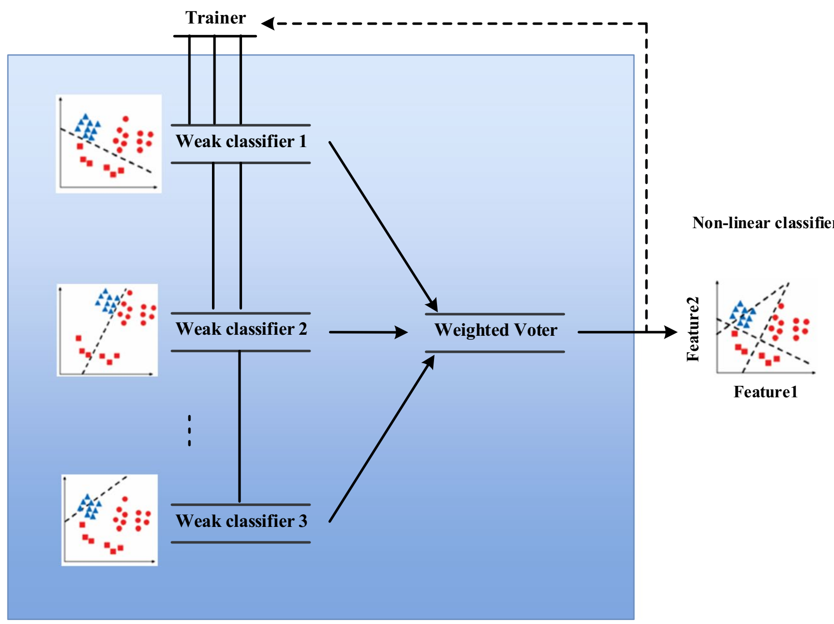


Fig. 8 Formulating a strong classifier through the aggregation of weak classifiers, each bearing unique weights, in the framework of the AdaBoost algorithm

computational time required for AdaBoost is comparatively lower than that of XGBoost due to its parsimonious set of hyperparameters, the former algorithm is known to be vulnerable to noisy data, which may have a detrimental impact on its overall model performance (Rätsch et al. 2001).

Goodness-of-fit evaluation indices

In this research, to forecast the enhancement of soil strength through the incorporation of PET elements, four distinct artificial intelligence models were constructed utilizing algorithms such as DT, RF, XGBoost, and AdaBoost. The datasets underwent a random partition into training (75%) and testing (25%) segments. The efficacy of each model was appraised using well-known metrics, namely the Determination Coefficient (DC, R^2), Root Mean Square Error (RMSE), Variance Accounted For (VAF), and the A-10 index. (Samaei et al. 2018; Naghadehi et al. 2019). The aforementioned metrics can be determined through the application of specific mathematical formulas:

$$R^2 = 1 - \frac{\sum \left(y_{act} - y_{pre}\right)^2}{\sum \left(y_{act} - \overline{y}_{act}\right)^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(y_{pre} - y_{act}\right)^2}{n}} \quad (2)$$

$$VAF = [1 - \frac{var(y_{act} - y_{pre})}{var(y_{act})}] \quad (3)$$

$$A - 10 = \frac{m10}{N} \quad (4)$$

Here, y_{act} and y_{pre} represent the actual and predicted values, respectively. N denotes the overall count of data points, while $m10$ refers to the number of instances where the ratio of actual to predicted value falls within the range of 0.9 to 1.1.

Model construction

In this section, the process of model development, parameter tuning, and the results for each algorithm are discussed in detail.

DT-based model development

The efficacy of the decision tree model hinges on the choice of parameters that enhance the learning procedure. In this study, the Scikit-learn Python library was utilized to build the decision tree model, and the Graphviz library was used to generate visualizations. During the feature selection process, the mean square error (MSE) serves as the criterion for variance reduction, which is pivotal in identifying the optimal split (Friedman 2001); Friedman's mean square error (MSE) utilises Friedman's improvement score to identify potential splits in the decision tree (Hastie et al. 2009); and Poisson uses Poisson deviance reduction to find splits (Hastie et al. 2009). Other important parameters, such as minimum_samples_split, minimum_leaf_samples, and tree_depth, are crucial in the decision tree modelling process. The selection of these parameters ought to be guided by the unique attributes of the dataset and the balance between model complexity and predictive accuracy (Mantovani et al. 2016). Therefore, a range of values was provided to the algorithm, and the most optimal values were selected through a trial and error approach, as shown in Table 6. When minimum_samples_split is set to 2, a node can be split if it contains at least two samples. This parameter setting can result in more flexible and complex decision trees, allowing the model to split nodes with fewer samples, leading to potentially more accurate predictions. It is particularly useful when working with smaller datasets, where setting a larger minimum sample split may result in underfitting and poorer model performance. The minimum leaf samples parameter is set at the final node of the tree to optimise for maximum accuracy, and a value of one was chosen for this parameter. Tree depth is also critical for controlling the decision tree's size and preventing overfitting. It can be used as a criterion to determine the best tree among a set of constructed trees. DTs were developed using the parameters mentioned above to predict reinforced soil behaviour. Table 6 displays the ideal settings for the DT parameters that yield the most effective model for forecasting the strength of reinforced soil. Additionally, Fig. 9 illustrates the configuration of the optimally trained DT dedicated to predicting the shear strength of reinforced soil.

Within DT models, input variables can be assigned ranks based on their relevance through an ad-hoc ranking technique. This approach ensures that variables essential for tree division and target prediction are situated closer to the root nod (Boutaba et al. 2018). The model operates accurately whereby any instance that involves a PET percentage lower than 0.5% is identified as soil that does not contain PET. In such cases, it is assigned the maximum shear strength ratio of 1. As the sample strength increases, the value increases and the assigned box for this specimen becomes darker. The algorithm assigns a brighter colour or sheer white colour to the three samples in which PET did not enhance their shear strength but also caused their looseness.

Table 6 The optimal values of DT parameters for the reinforced soil strength prediction

Parameter	Range	Optimum value
criterion	[MSE, friedman_ mse, poisson]	MSE
Minimum sample split	[2—6]	2
max depth	[2—10]	4
Minimum samples leaf	[1—∞]	1

RF-based model development

For the creation of an RF model aimed at predicting the strength of reinforced soil, Bootstrap sampling was initially utilized to prepare the training data. The RF algorithm then facilitated the generation of a diverse collection of DT samples, enhancing the precision of the forecast. An aggregation technique was subsequently employed to fine-tune the prediction accuracy. In the implementation with the Python SciKit package, the model's performance is evaluated using the out-of-bag (OOB) feature, which is enabled by setting it to True. This evaluation specifically considers the samples excluded from the bootstrap selection.

In the RF run, the number of estimators shows how many DTs were generated. Figure 10 represents some of the developed DTs in this study. Variable importance evaluations increase along with the number of estimators. The investigation conducted by Lunetta et al. (2004) showed that running the algorithm with more estimators increases the algorithm's run time. However, this method leads to more stable results. In the process of constructing an RF model to predict the increase in shear strength, 200 estimators were chosen as the optimal number after a series of trial-and-error iterations. By investigating deeper trees to maximise performance, this method may cause the model to become overfitted. A better result was obtained using the Entropy criterion than the Gini criterion when it came to gaining information. Table 7 shows the optimum values obtained for the RF-based model.

XGBoost-based model development

Although RF is a proficient prediction technique, it still has some limitations. One of these is that if the algorithm makes an error in the initial stages, it is likely to repeat this mistake in the following iterations (Javeed et al. 2019). On the other hand, both the XGBoost and AdaBoost algorithms do not have this limitation. These boosting techniques can learn from the errors made by individual trees and use this knowledge to improve the overall model performance significantly (Demir and Sahin 2023).

In this research, the development of an XGBoost model utilized the Scikit-learn and XGBoost Python libraries to train a comprehensive tree. Fine-tuning hyperparameters plays a vital role in crafting an effective XGBoost model. The key parameters requiring optimization encompass:

In the creation of an XGBoost-based model for this investigation, parameters such as the quantity of features, trees, maximum tree depth, the decision on bootstrap sampling, the minimal number of samples required in a node pre-split, and the minimal samples needed in a terminal leaf node were meticulously adjusted. Increasing these hyperparameter settings typically simplifies the model and lowers the risk of overfitting. A systematic trial-and-error approach

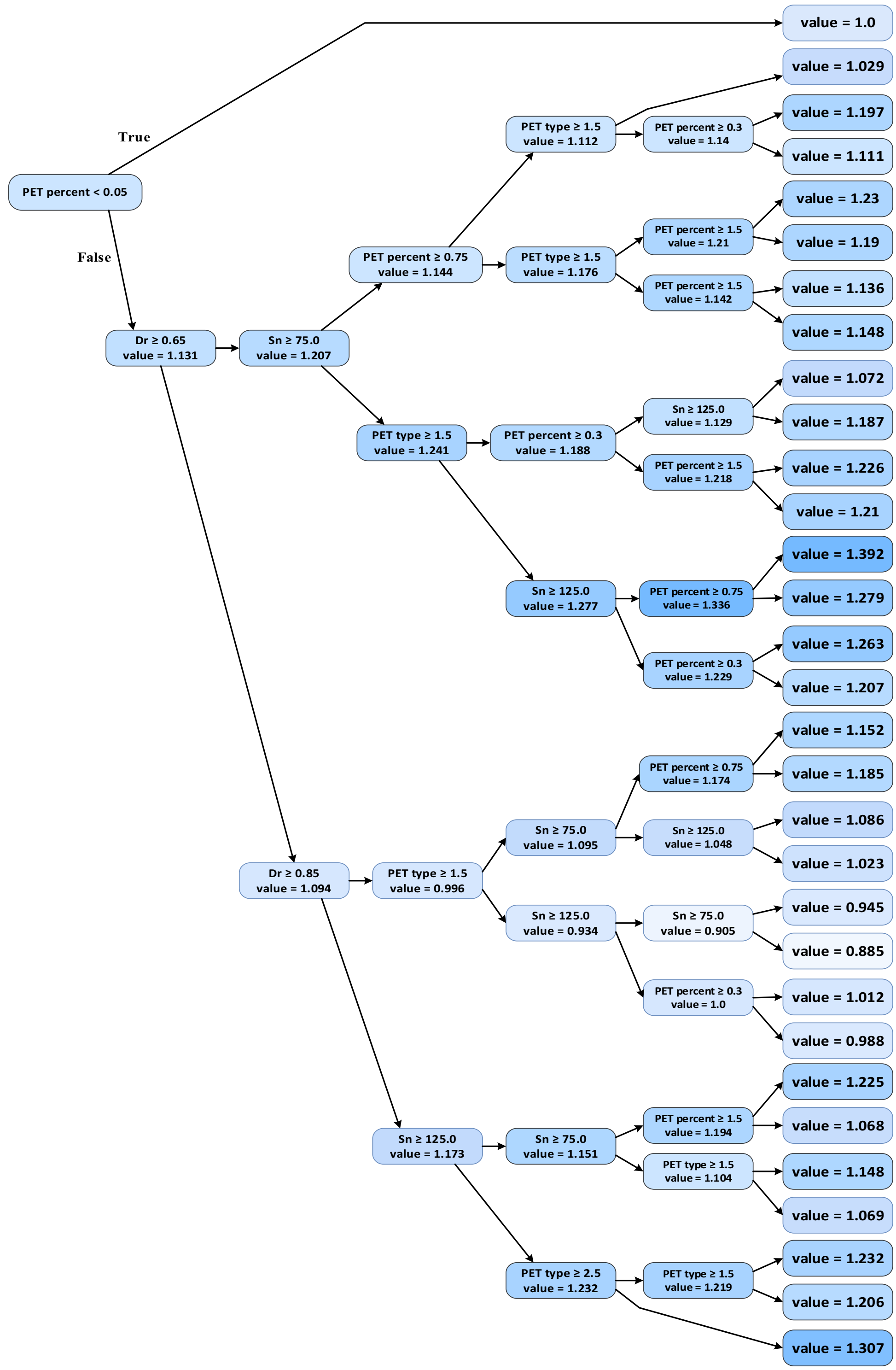


Fig. 9 The configuration of the optimal DT model's tree structure

was utilized to derive the optimal values for these XGBoost parameters. The optimal settings for the XGBoost parameters, aimed at forecasting the performance of reinforced soil, are detailed in Table 8. A limitation observed with XGBoost is the inconsistency in the speed of tree generation and learning, which tends to undervalue newer trees relative to older ones (Ahn et al. 2023). To address the inconsistency in XGBoost between generating trees and learning from them, Vinayak and Gilad-Bachrach (2015) introduced the DART (Dropout Additive Regression Trees) booster for XGBoost. This approach involves randomly dropping (or "dropping out") certain trees during the boosting process, which encourages diversity in the final ensemble of trees and prevents overfitting. In the study using the DART booster for

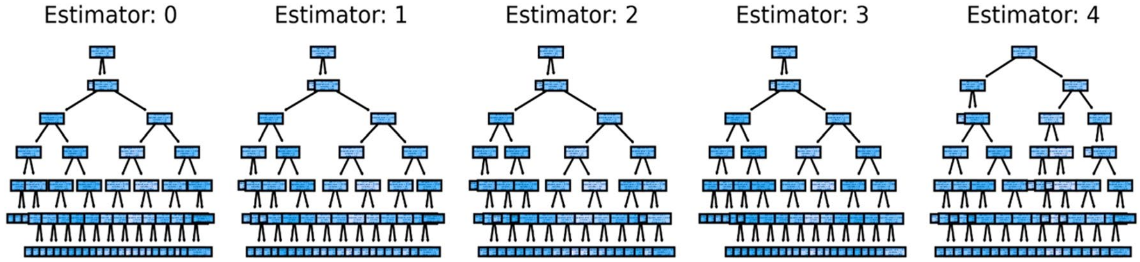


Fig. 10 Five different estimators (DTs) developed in this study

Table 7 Hyperparameters of the RF model for predicting the shear strength in soils reinforced with PET

Parameter	Range	Optimum value
Criterion	[Entropy, Gini]	Entropy
Estimators number	[50, 100, 200, 300]	200
Bootstrap	[True, False]	True
Max depth	[3—7]	4
Max features	[sqrt, log2]	Sqrt
OOB score	[True, False]	True

Table 8 The best XGBoost algorithm parameters for predicting the shear strength of soils treated with PET

Parameter	Range	Optimum value
Learning rate (eta)	[0.1, 0.2, 0.3, 0.5, 1, 1.5, 2, 3]	0.3
number of estimators	[50, 100, 200, 300]	100
max depth	[2, 3, 4, 5, 6]	3
Gamma	[0.0001, 0.001, 0.01, 0.1,0.5, 1]	0.001
min child weight	[0.2, 0.5, 0.8, 0.9, 1, 1.5, 2]	0.9
max delta step	[0.2, 0.5, 0.8, 0.9, 1, 1.5, 2]	0.9
Booster	[gbtree, gblinear, DART]	DART

XGBoost, it was discovered that the model's performance was improved compared to other boosting algorithms. Specifically, the DART booster achieved better accuracy and required fewer trees to be trained to reach that level of accuracy. These findings demonstrate the efficacy of the DART booster for XGBoost, which offers a promising solution to the issue of inconsistency between generating and learning from trees in this algorithm. Figure 11 illustrates several of the predictors that XGBoost generated for forecasting the shear strength of soils improved with PET.

AdaBoost-based model development

In the development of an AdaBoost model, adjustments were made to some features through a trial-and-error process. Initially, both CART and RF learners were tested with the AdaBoost algorithm, revealing that the RF regressor achieved higher accuracy. As previously noted, the entropy criterion consistently outperformed the Gini-index criterion for this dataset. Therefore, the entropy criterion was chosen for the splitting process. In order to achieve the desired tree depth (TD), a parametric analysis was conducted, revealing that reducing the TD value results in a less complex model. Nevertheless, accuracy was also a crucial consideration, and decreasing the TD led to a decrease in accuracy. Based on the experience gained while developing other tree-based models, the initial values of [2, 3, 4, 5] were chosen for TD. Subsequently, modifications were made to other features in an attempt to achieve the lowest possible TD number while maintaining the highest level of accuracy. At last, a TD value of 3 was opted for model creation. Table 9 displays the other essential factors needed to construct the AdaBoost model.

Results and discussion

The objective of this research was to forecast the rise in shear strength of sandy soils bolstered by PET elements employing four tree-based ML methodologies: DT, RF, XGB, and AdaBoost. The effectiveness of the models was assessed through four widely recognized performance metrics (R^2, RMSE, VAF, and A-10). The results of the training and testing phases for these tree-based models are shown in Table 10. Zorlu et al. (2008) introduced a ranking system to compare the accuracy of the developed models. This system makes it feasible to rank all the models based on their performance across all performance indices, including train and test scores. As shown in Table 10, AdaBoost, leveraging the DT regressor as its primary methodology, attained the highest prediction accuracy, earning a final ranking of 25, while RF and DT displayed acceptable results. Furthermore, it was found that the XGBoost model performs better than RF and DT in predicting the output parameter; however, its prediction performance is relatively lower than the AdaBoost model. Figure 12 illustrates scatter plots depicting the performance of the AdaBoost model during both the training

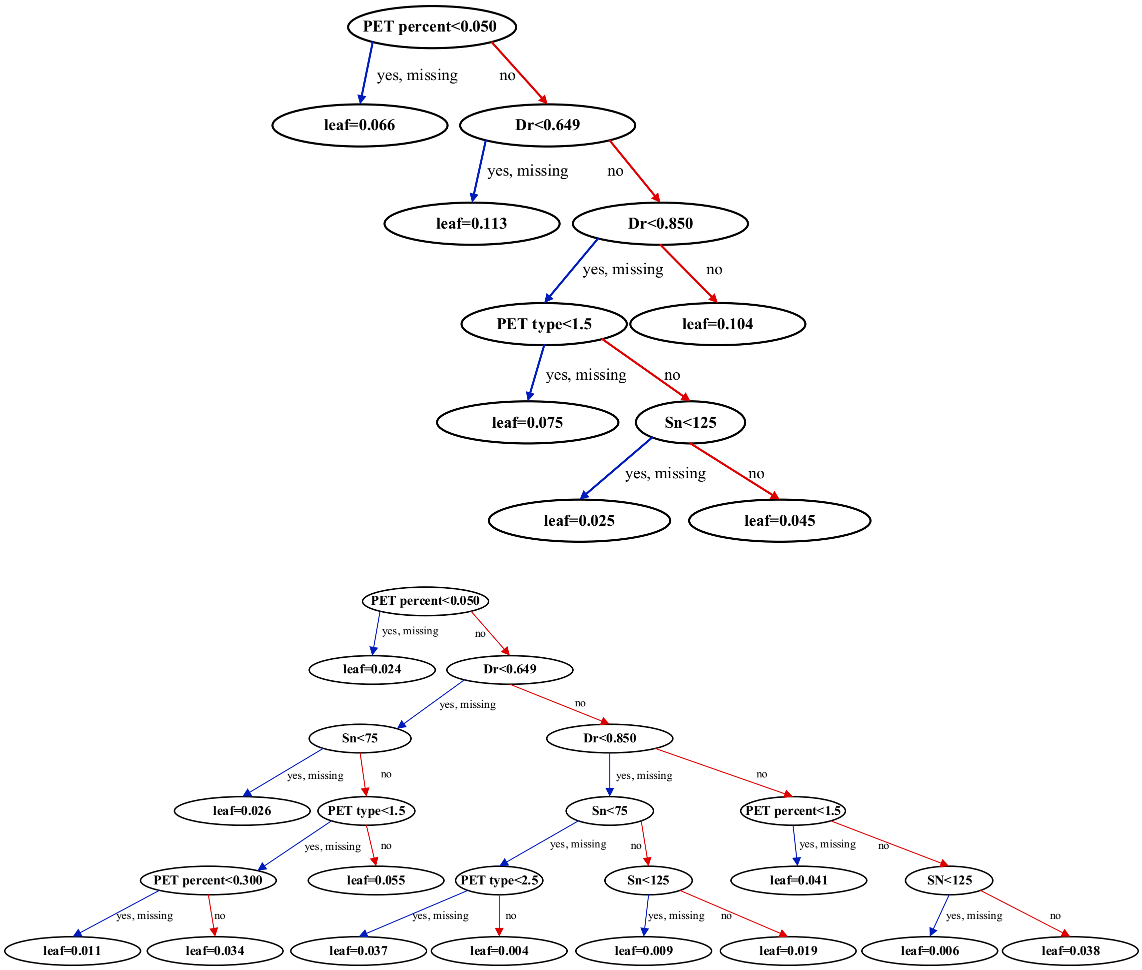


Fig. 11 XGBoost devised four tree predictors for forecasting enhanced shear strength

and testing phases, highlighting the robust predictive capabilities of the model.

The XGBoost model outperformed the RF and DT models in predicting the output parameter due to several key advantages. Firstly, as mentioned in Section "XGBoost", XGBoost is built on the gradient boosting framework, which combines the predictions of multiple weak learners to create a strong predictive model, effectively reducing bias and variance compared to single decision trees or ensemble methods that do not employ boosting (Kavzoglu and Teke 2022). Additionally, XGBoost incorporates regularisation techniques, such as L1 and L2 regularisation, which prevent overfitting and enhance the model's ability to generalise to unseen data, a feature not inherently present in RF (Dhaliwal et al. 2018). Furthermore, XGBoost is designed to handle missing data efficiently during the training process, improving its predictive power and accuracy (Dhaliwal et al. 2018). These combined features make XGBoost a robust and powerful tool for predictive modelling, surpassing the performance of RF and DT in this study.

However, the prediction performance of XGBoost was relatively lower than that of the AdaBoost model, and several factors could be attributed to this discrepancy. Firstly, the adaptive boosting mechanism of AdaBoost focuses on the hardest-to-predict instances by adjusting the weights of misclassified examples and iteratively improving the model (Dos Santos Aguiar and Paulo 2019; Rajesh and Dhuli 2018). This approach allows AdaBoost to handle datasets with noise and outliers more effectively, leading to higher accuracy in specific scenarios (Rätsch et al. 2001). Lastly, the specific characteristics of the used dataset, including low noise and the nature of variable relationships, may have favoured AdaBoost, compared to XGBoost. Its ability to emphasize difficult-to-predict samples likely provided an advantage in achieving higher predictive accuracy.

Figure 13 also presents the importance of input parameters for each of the four AdaBoost, XGB, RF, and DT algorithms during model development. The reduction in the model's criterion attributable to that feature is calculated and normalized to determine the importance of a feature.

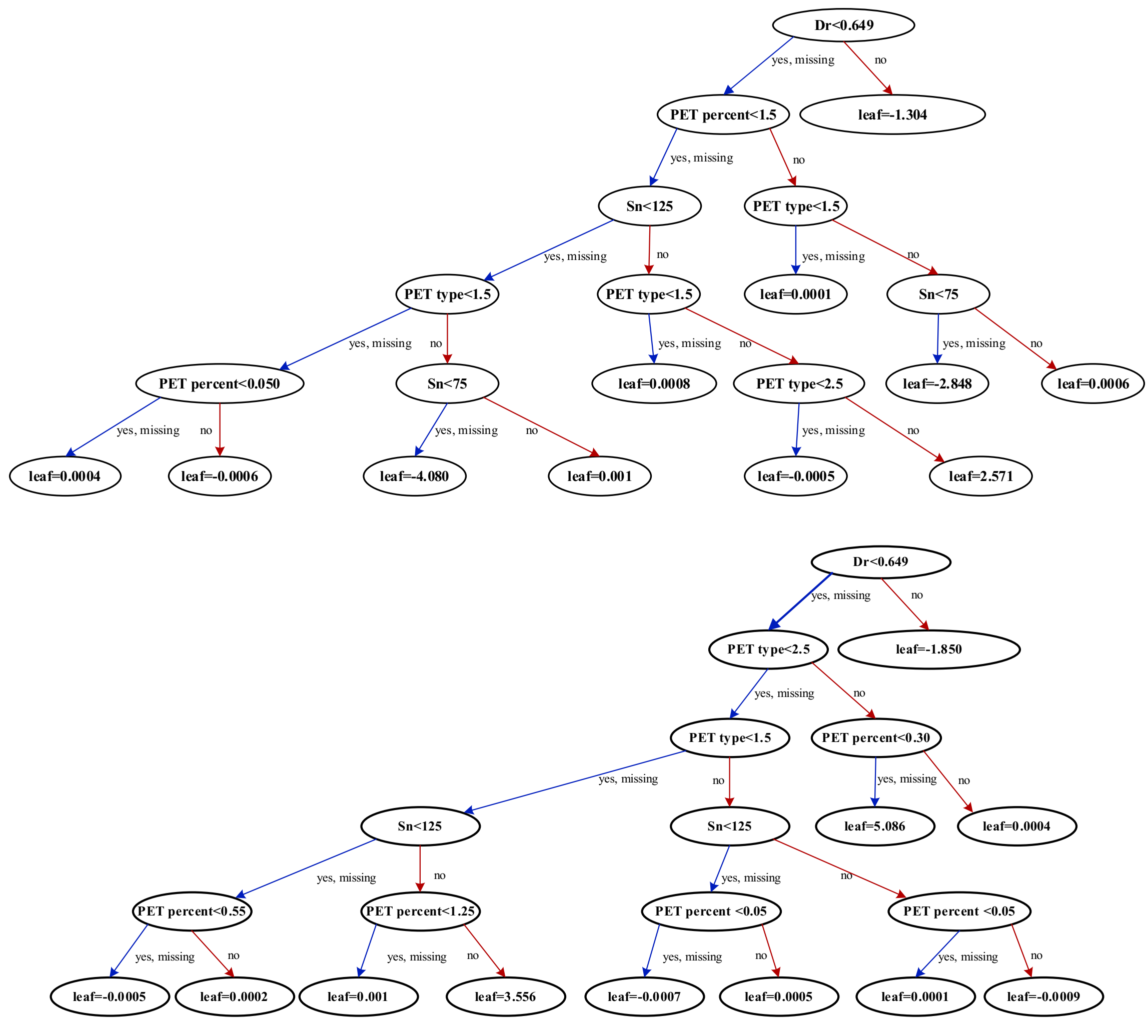


Fig. 11 (continued)

Table 9 The optimum values of the parameters for the best AdaBoost model to predict the shear strength of soils enhanced with PET

Attribute	Scale	Optimum value
Weak learner algorithm	Random Forest, Decision Tree	Random Forest Regressor
Tree depth	[2—6]	3
Algorithm	SAMME, SAMME.R	SAMME
Criterion	Entropy, Gini	Entropy
Number of estimators	[50, 100, 200, 300]	100
Learning rate	[0.5, 1, 1.5, 2]	1.0

This measure can be referred to as either Gini importance or Mean Decrease in Impurity (MDI). When using Gini importance or MDI, a feature's importance is determined by the number of splits it contributes to (across all trees) concerning the number of samples it splits (Gwetu et al. 2019). However, the Scikit-learn Python package was utilised in this study, which calculates feature importance by averaging the actual decrease in node impurity across all trees and weighting it based on the number of sample splits (Pedregosa et al. 2011). From this figure, it can be seen that relative density is the most important parameter for all the algorithms. Following this, the PET percentage is detected as the second important parameter for the algorithms except for the XGBoost model. On the other hand, the normal stress was also detected as a more important parameter than the PET type in all the algorithms.

Table 10 Ranking of developed models based on performance indices

Model name		R^2		RMSE		VAF		A10		Final score
		Train	Test	Train	Test	Train	Test	Train	Test	
AdaBoost (DT regressor)		0.99	0.91	0.01	0.04	98.91	90.95	1	1	
	Ranking position	4	4	3	2	4	4	1	3	25
XGBoost		0.99	0.90	0.009	0.05	99.63	85.53	1	0.96	
	Ranking position	4	3	4	1	5	2	1	2	22
AdaBoost (RF classifier)		0.97	0.87	0.02	0.05	96.76	87.14	1	1	
	Ranking position	3	2	2	1	3	3	1	3	18
Random Forest		0.95	0.87	0.03	0.05	94.32	87.14	1	1	
	Ranking position	2	2	1	1	2	3	1	3	15
Decision Tree		0.94	0.85	0.03	0.05	93.71	82.77	1	0.92	
	Ranking position	1	1	1	1	1	1	1	1	8

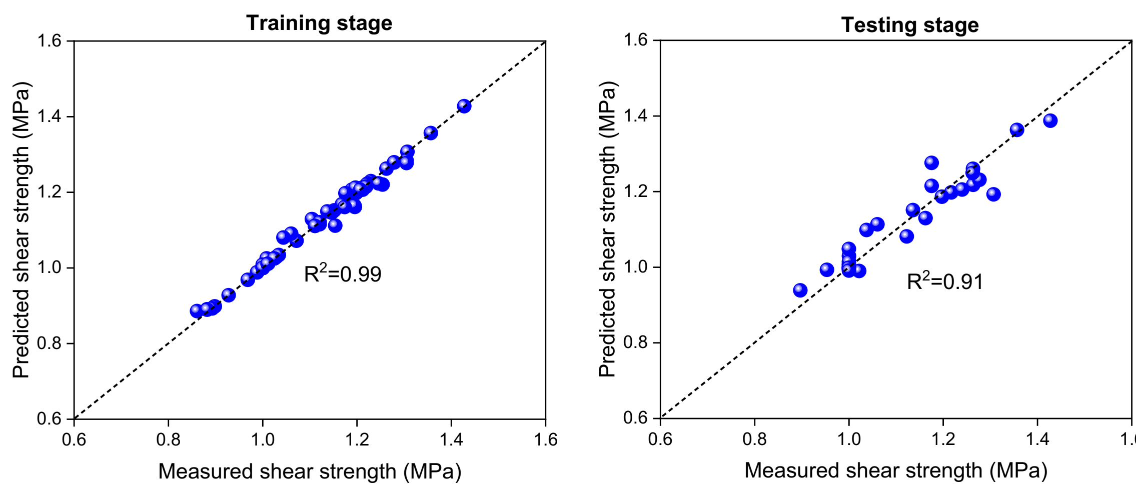


Fig. 12 The correlation between the actual and predicted values of shear strength in reinforced soil by the AdaBoost predictor in the training and testing stages

Graphical user interface

In response to concerns from civil engineers about the practicality of using machine learning (ML) techniques in real-world applications, a graphical user interface (GUI) was developed in this study (see Fig. 14). This GUI can help overcome some of the perceived challenges of implementing ML in practice by providing users with an easy-to-use interface. The GUI facilitates the input of essential features, enables users to run the toolbox with ease, and allows them to view the predictions generated by the ensemble models developed in the study. By providing access to the ensemble models and associated results, this interface enables users to better understand the application of ML in civil engineering and the potential benefits of using such techniques. The program is available online through the https://github.com/Faradonbeh/Soil-Shear-Strength-Improvement-using-PET.git website, which hosts the software and relevant coding. The program allows users to modify various parameters to better understand the impact of PET elements on shear strength enhancement. With these capabilities, users can conduct in-depth analyses of the underlying data and test different scenarios to assess the robustness of the results.

Conclusions

This study developed four decision tree-based predictive models (DT, RF, XGB, and AdaBoost) to assess the improvement of soil shear strength using PET elements. The advanced use of AI-based models effectively detected

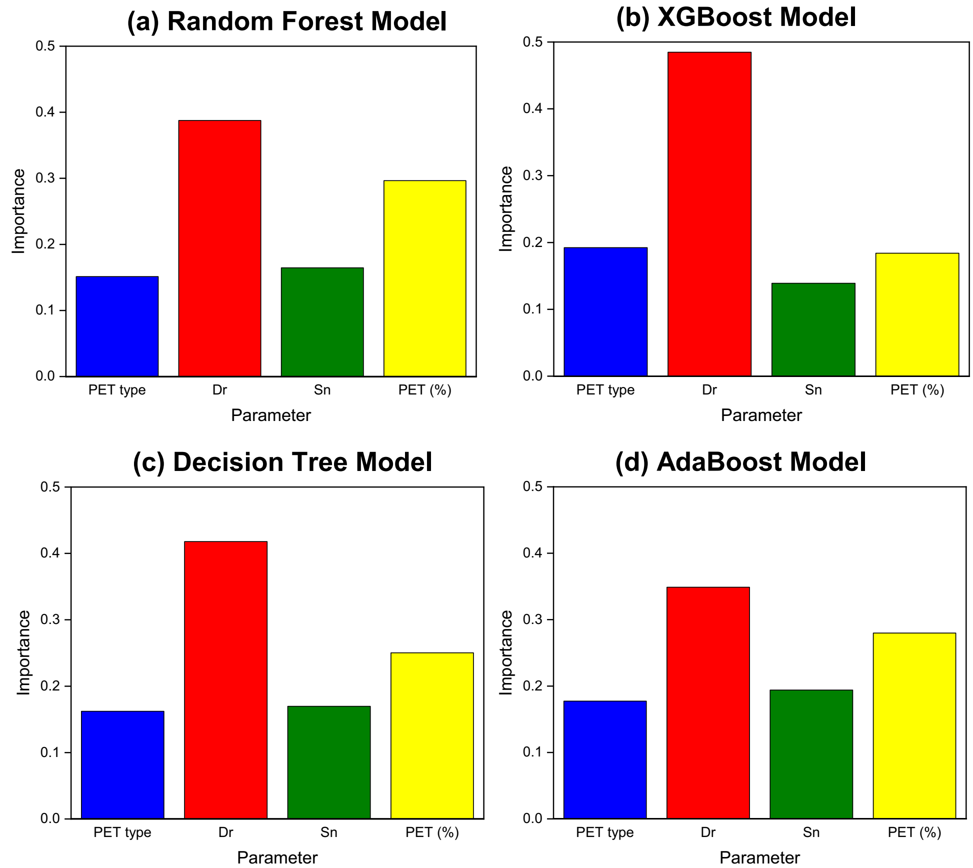


Fig. 13 Analysis of the significance of features (**a**) RF, (**b**) XGB, (**c**) DT, and (**d**) AdaBoost

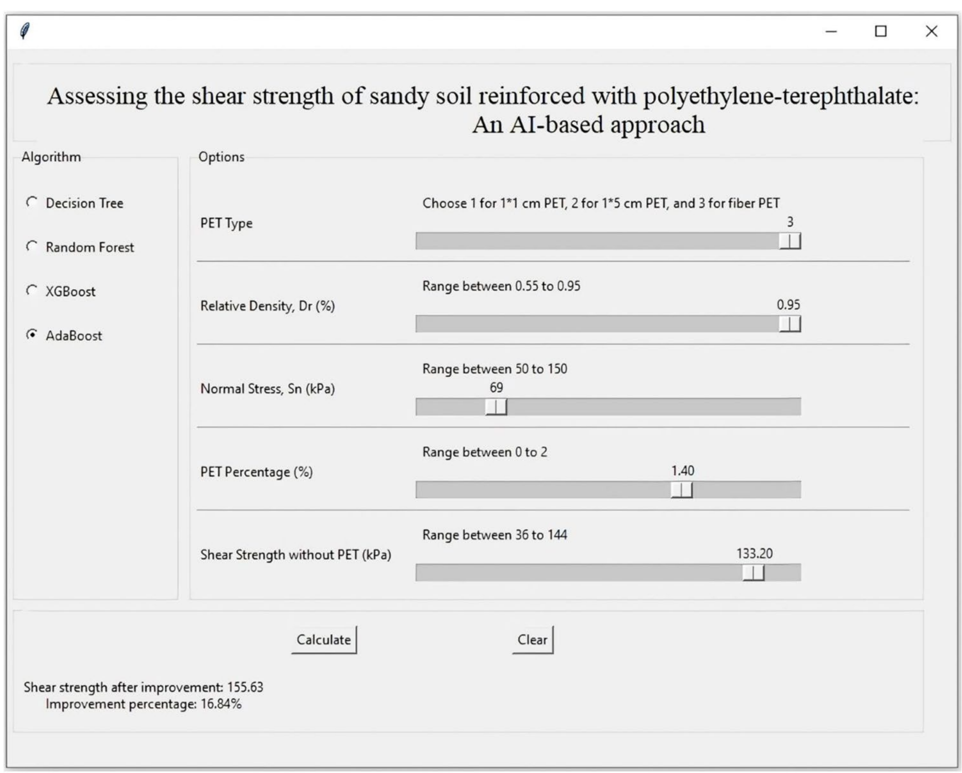


Fig. 14 The GUI developed in this study for soil's shear strength improvement prediction using PET elements

complex nonlinear relationships between variables, simplifying and accelerating the shear strength measurement process through a nondestructive method. These models account for the variability of soil properties, offering valuable predictions for practical engineering applications. A comprehensive dataset of 118 shear tests on soil samples with varying PET contents was used to develop these models. Extensive parametric analyses were conducted to optimize the algorithms. Among the models, AdaBoost demonstrated superior performance, with R2 scores of 0.99 for training and 0.91 for testing, outperforming other models. The models' prediction performance was further validated using RMSE, VAF, and A10-index, highlighting AdaBoost's remarkable capability in predicting shear strength.

The XGBoost model, employing DT as a weak learner, effectively mitigated errors by learning from previous iterations, refining its predictions without exacerbating error rates. This study confirmed that AdaBoost excels in low-noise datasets, making it suitable for applications where computational resources and immediacy of results are not critical concerns. The sensitivity analysis identified relative density as the most crucial parameter, followed by PET percentage and normal stress. To enhance practical applicability, a Graphical User Interface (GUI) was developed, allowing researchers and engineers to easily utilize the models. Future enhancements could include incorporating more data and addressing the public availability of ML models.

Additionally, the study promoted sustainable engineering practices by using PET elements, a non-recyclable material, for soil reinforcement. This effective use of waste materials not only improves soil properties but also addresses environmental concerns related to plastic waste.

Moreover, the introduced methods significantly aid geotechnical engineers and researchers in evaluating soil behavior and performance, providing a robust, practical tool for enhancing the safety and efficiency of engineering designs.

Author contributions Conceptualization: Masoud Samaei, Roohollah Shirani Faradonbeh, Morteza Alinejad Omran, Mohsen Keramati, Reza naderi; Methodology: Masoud Samaei, Roohollah Shirani Faradonbeh; Formal analysis and investigation: Masoud Samaei, Roohollah Shirani Faradonbeh, Morteza Alinejad Omran; Writing—original draft preparation: Masoud Samaei and Morteza Alinejad Omran; Writing—review and editing: Masoud Samaei, Roohollah Shirani Faradonbeh, Morteza Alinejad Omran, Mohsen Keramati, Reza Naderi; Resources: Morteza Alinejad Omran, Mohsen Keramati, Reza Naderi; Supervision: Roohollah Shirani Faradonbeh, Mohsen Keramati, Reza Naderi.

Funding No funding was obtained for this study.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

Acharyya R, Lahiri A, Mukherjee SP, Raghu PV (2013) Improvement of undrained shear strength of clayey soil with pet bottle strips. In Proceedings of Indian Geotechnical Conference, pp 22–24

Ahn JM, Kim J, Kim K (2023) Ensemble machine learning of gradient boosting (XGBoost, LightGBM, CatBoost) and attention-based CNN-LSTM for harmful algal blooms forecasting. Toxins 15:608

Alvarez A, Sosa J, Duran G, Pacheco L (2020) Improved mechanical properties of a high plasticity clay soil by adding recycled PET. In IOP Conf Ser: Mater Sci Eng 758:012075

Astm-D3080 (2011) Standard Test Method for Direct Shear Test of Soils Under Consolidated Drained Conditions. ASTM, USA

Astm-D422 (2016) Standard Test Method for Particle-Size Analysis of Soils. ASTM, USA

Azmi SS, Baliga S (2020) An overview of boosting decision tree algorithms utilizing AdaBoost and XGBoost boosting strategies. Int Res J Eng Technol 7:6867–6870

Babu GL, Chouksey SK (2011) Stress-strain response of plastic waste mixed soil. Waste Manag 31:481–488

Botero E, Ossa A, Sherwell G, Ovando-Shelley E (2015) Stress-strain behavior of a silty soil reinforced with polyethylene terephthalate (PET). Geotext Geomembr 43:363–369

Boutaba R, Salahuddin MA, Limam N, Ayoubi S, Shahriar N, Estrada-Solano F, Caicedo OM (2018) A comprehensive survey on machine learning for networking: evolution, applications and research opportunities. J Int Serv Appl 9:1–99

Breiman L (2001) Random forests. Mach Learn 45:5–32

Carvalho B, Casagrande M, Farias M (2019) Evaluation of resilient behavior of a clayey soil with polyethylene terephthalate (PET) insertion for application in pavements base. In Geotechnical engineering in the XXI century: lessons learned and future challenges, pp 1510–1517

Charles V, Gherman T, Paliza JC (2022) The Gini index: a modern measure of inequality. In: Charles V, Emrouznejad A (eds) Modern Indices for International Economic Diplomacy. Palgrave Macmillan, Cham, pp 55–84

Chen T, Guestrin C (2016) Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp 785–794

Consoli NC, Montardo JP, Prietto PDM, Pasa GS (2002) Engineering behavior of a sand reinforced with plastic waste. J Geotech Geoenviron Eng 128:462–472

Demir S, Sahin EK (2023) An investigation of feature selection methods for soil liquefaction prediction based on tree-based ensemble algorithms using AdaBoost, gradient boosting, and XGBoost. Neural Comput Appl 35:3173–3190

Dhaliwal S, Nahid A, Abbas R (2018) Effective intrusion detection system using XGBoost. Information 9(7):149

Fathi H, Chenari RJ, Vafaeian M (2020) Shaking Table Study on PET Strips-Sand Mixtures Using Laminar Box Modelling. Geotech Geol Eng 38:683–694

Frank M, Drikakis D, Charissis V (2020) Machine-learning methods for computational science and engineering. Computation 8:15

Freund Y, Schapire RE (1997) A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 55:119–139

Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29:1189–1232
Gwetu MV, Tapamo J-R, Viriri S (2019) Exploring the impact of purity gap gain on the efficiency and effectiveness of random forest feature selection. Computational Collective Intelligence: 11th International Conference, ICCCI 2019, Hendaye, France, September 4–6, 2019, Proceedings, Part I 11. Springer
Hafez M, Mousa R, Awed A, El-Badawy S (2019) Soil Reinforcement using recycled plastic waste for sustainable pavements. In: El-Badawy S, Valentin J (eds) Sustainable solutions for railways and transportation engineering. GeoMEast 2018. Sustainable Civil Infrastructures. Springer, Cham
Hannan A, Anmala J (2021) Classification and Prediction of Fecal Coliform in Stream Waters Using Decision Trees (DTs) for Upper Green River Watershed, Kentucky, USA. Water 13:2790
Hastie T, Rosset S, Zhu J, Zou H (2009) Multi-class adaboost. Statistics and its interface 2:349–60
Javeed A, Zhou SJ, Liao YJ, Qasim I, Noor A, Nour R (2019) An Intelligent Learning System Based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection. Ieee Access 7:180235–180243
Kavzoglu T, Teke A (2022) Predictive Performances of ensemble machine learning algorithms in landslide susceptibility mapping using random forest, extreme gradient boosting (XGBoost) and natural gradient boosting (NGBoost). Arab J Sci Eng 47:7367–7385
Kuhn M, Johnson K (2013) Applied predictive modeling, New York: Springer 613
Li J, Ding DW (2002) Nonlinear elastic behavior of fiber-reinforced soil under cyclic loading. Soil Dyn Earthq Eng 22:977–983
Li H, Sheu PC-Y (2021) A scalable association rule learning heuristic for large datasets. J Big Data 8:86
Lin CL, Fan CL (2019) Evaluation of CART, CHAID, and QUEST algorithms: a case study of construction defects in Taiwan. J Asian Archit Build Eng 18:539–553
Liu L, Zhou W, Gutierrez M (2022) Effectiveness of predicting tunneling-induced ground settlements using machine learning methods with small datasets. J Rock Mech Geotech Eng 14:1028–1041
Louzada NDL, JaC Malko, Casagrande MD (2019) Behavior of Clayey Soil Reinforced with Polyethylene Terephthalate. J Mater Civil Eng 31:04019218
Lunetta KL, Hayward LB, Segal J, Van Eerdewegh P (2004) Screening large-scale association study data: exploiting interactions using random forests. BMC Genet 5:32
Maher MH, Ho YC (1994) Mechanical-Properties of Kaolinite Fiber Soil Composite. J Geotech Eng-Asce 120:1381–1393
Malidarreh NR, Shooshpasha I, Mirhosseini SM, Dehestani M (2018) Effects of reinforcement on mechanical behaviour of cement treated sand using direct shear and triaxial tests. Int J Geotech Eng 12:491–499
Mantovani RG, Horváth T, Cerri R, Vanschoren J, De Carvalho AC (2016) Hyper-parameter tuning of a decision tree induction algorithm. 2016 5th Brazilian Conference on Intelligent Systems (BRACIS). IEEE
Mishra B, Gupta MK (2018) Use of randomly oriented polyethylene terephthalate (PET) fiber in combination with fly ash in subgrade of flexible pavement. Constr Build Mater 190:95–107
Moghaddas Tafreshi SN, Parvizi Omran M, Rahimi M, Dawson A (2021) Experimental investigation of the behavior of soil reinforced with waste plastic bottles under cyclic loads. Transp Geotech 26:100455
Momeni E, Samaei M, Hashemi A, Dowlatshahi MB (2023) A review on the application of soft computing techniques in foundation engineering. Artificial intelligence in mechatronics and civil engineering: Bridging the Gap 16:111–133
Nadaf MB, Dutta S, Mandal J (2019) Fly ash as backfill material in slopes using waste pet bottles as reinforcement. Springer, Waste Management and Resource Efficiency
Naghadehi MZ, Samaei M, Ranjbarnia M, Nourani V (2018) Corrigendum to "State-of-the-art predictive modeling of TBM performance in changing geological conditions through gene expression programming" [Measurement 126 (2018) 46–57] (Measurement (2018) 126 (46–57), (S0263224118304366). Measurement: Journal of the International Measurement Confederation. https://doi.org/10.1016/j.measurement.2018
Naghadehi MZ, Samaei M, Ranjbarnia M (2019) Superior modeling of hard rock TBM performance using novel predictive analytics methodologies. Proceedings of the 3rd International conference on applied researches in structural engineering and construction management, pp 1–11
Necmi Y, Ekrem K (2020) The Mechanical Performance of Clayey Soils Reinforced with Waste PET Fibers. Int J Earth Sci Knowl Appl 2:19–26
Onyelowe KC, Mojtahedi FF, Ebid AM, Rezaei A, Osinubi KJ, Eberemu AO, Salahudeen B, Gadzama EW, Rezazadeh D, Jahangir H (2023) Selected AI optimization techniques and applications in geotechnical engineering. Cogent Engineering 10:2153419
Patil A, Waghere G, Inamdar N, Gavali P, Dhore R, Shah S (2016) Experimental review for utilisation of waste plastic bottles in soil improvement techniques. Int J Eng Res Appl 6:25–31
Peddaiah S, Burman A, Sreedeep S (2018) Experimental Study on Effect of Waste Plastic Bottle Strips in Soil Improvement. Geotech Geol Eng 36:2907–2920
Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: Machine Learning in Python. J Mach Learn Res 12:2825–2830
Probst P, Wright MN, Boulesteix AL (2019) Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery 9**:**e1301
Rabbani A, Samui P, Kumari S (2023a) Implementing ensemble learning models for the prediction of shear strength of soil. Asian J Civil Eng 24:2103–2119
Rabbani A, Samui P, Kumari S (2023b) A novel hybrid model of augmented grey wolf optimizer and artificial neural network for predicting shear strength of soil. Model Earth Syst Environ 9:2327–2347
Rabbani A, Samui P, Kumari S (2023c) Optimized ANN-based approach for estimation of shear strength of soil. Asian J Civil Eng 24:3627–3640
Rabbani A, Samui P, Kumari S, Saraswat., Tiwari M, Rai A (2023d) Optimization of an artificial neural network using three novel meta-heuristic algorithms for predicting the shear strength of soil. Transp Infrastruct Geotechnol 1–22
Rabbani A, Muslih JA, Saxena M, Patil SK, Mulay B. N, Tiwari M, Usha A, Kumari S, Samui P (2024) Utilization of tree-based ensemble models for predicting the shear strength of soil. Transp Infrastruct Geotechnol 1–24
Rajesh KNVPS, Dhuli R (2018) Classification of imbalanced ECG beats using re-sampling techniques and AdaBoost ensemble classifier. Biomed Signal Process Control 41:242–254
Rätsch G, Onoda T, Müller KR (2001) Soft margins for AdaBoost. Mach Learn 42:287–320
Samaei M, Ranjbarnia M, Zare NM (2018) Prediction of the Rock Brittleness Index Using Nonlinear Multivariable Regression and the CART Regression Tree. J Civil Environ Eng 48:33–40
Samaei M, Massalow T, Abdolhosseinzadeh A, Yagiz S, Sabri MMS (2022) Application of Soft Computing Techniques for Predicting Thermal Conductivity of Rocks. Applied Sciences-Basel 12:9187

Dos Santos Aguiar MJ, Paulo M (2019) Learning to classify a subject-line quality for email marketing using data mining techniques, Master Dissertation, University of Porto, p 113

Schapire RE (1990) The strength of weak learnability. Mach Learn 5:197–227

Shariatmadari N, Karimpour-Fard M, Hasanzadehshooiili H, Hoseinzadeh S, Karimzadeh Z (2020) Effects of drainage condition on the stress-strain behavior and pore pressure buildup of sand-PET mixtures. Construc Build Mater 233:117295

Sinha AK, Jha JN, Choudhary AK (2019) A study on cbr behaviour of waste pet strip reinforced stone dust. In Proceedings of the 1st international conference on sustainable waste management through design: ICSWMD 2018. Lecture Notes in Civil Engineering, vol 21. Springer, Cham, pp 302–312

Tofigh Tabrizi M, Keramati M, Ramesh A (2021) Investigation of dynamic behavior of Anzali Port sandy soil reinforced with PET fibers. J Mar Eng 37–48

Vinayak RK, Gilad-Bachrach R (2015) Dart: Dropouts meet multiple additive regression trees. In Artificial intelligence and statistics, pp 489–497

Witten IH, Frank E, Hall MA, Pal CJ, Data M (2005) Practical machine learning tools and techniques. Data mining. Elsevier Amsterdam, The Netherlands

Zhou J, Asteris PG, Armaghani DJ, Pham BT (2020) Prediction of ground vibration induced by blasting operations through the use of the Bayesian Network and random forest models. Soil Dyn Earthq Eng 139:106390

Zhu X, Chu J, Wang KD, Wu SF, Yan W, Chiam K (2021) Prediction of rockhead using a hybrid N-XGBoost machine learning framework. J Rock Mech Geotech Eng 13:1231–1245

Zorlu K, Gokceoglu C, Ocakoglu F, Nefeslioglu HA, Acikalin S (2008) Prediction of uniaxial compressive strength of sandstones using petrography-based models. Eng Geol 96:141–158

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.