Multiscale and Multidisciplinary Modeling, Experiments and Design (2024) 7:3841–3864
https://doi.org/10.1007/s41939-024-00447-x

ORIGINAL PAPER

Cone penetration test-based assessment of liquefaction potential using machine and hybrid learning approaches

Jitendra Khatti[1] · Yewuhalashet Fissha[2,3] · Kamaldeep Singh Grover[1] · Hajime Ikeda[2] · Hisatoshi Toriya[2] · Tsuyoshi Adachi[2] · Youhei Kawamura[4]

Received: 1 March 2024 / Accepted: 29 March 2024 / Published online: 26 April 2024

Abstract

In soil mechanics, liquefaction is the phenomenon that occurs when saturated, cohesionless soils temporarily lose their strength and stiffness under cyclic loading shaking or earthquake. The present work introduces an optimal performance model by comparing two baselines, thirty tree-based, thirty support vector classifier-based, and fifteen neural network-based models in assessing the liquefaction potential. One hundred and seventy cone penetration test results (liquefied and non-liquefied) have been compiled from the literature for this aim. Earthquake magnitude, vertical-effective stress, mean grain size, cone tip resistance, and peak ground acceleration parameters have been used as input parameters to predict the soil liquefaction potential for the first time. Performance metrics, accuracy, an area under the curve (AUC), precision, recall, and F1 score have measured the training and testing performances. The comparison of performance metrics reveals that the model Runge–Kutta optimized extreme gradient boosting (RUN_XGB) has assessed the liquefaction potential with an overall accuracy of 99%, AUC of 0.99, precision of 0.99, recall value of 1, and F1 score of 1. Moreover, model RUN_XGB has a true negative rate of 0.98, negative predictive value of 1, Matthews correlation coefficient of 0.98, and average classification accuracy of 0.99, close to the ideal values and presents the robustness of the RUN_XGB model. Finally, the RUN_XGB model has been recognized as an optimal performance model for predicting the liquefaction potential. It has been noted that a low multicollinearity level affects the prediction accuracy of models based on conventional soft computing techniques, i.e., logistic regression. This research will help researchers choose suitable hybrid algorithms and enhance the accuracy of seismic soil liquefaction potential models.

Keywords Liquefaction potential · Optimization techniques · Xtreme-gradient boosting · Neural networks · Multicollinearity

1 Introduction

An earthquake is a phenomenon that creates significant destruction in any structure under fissures, unequal moments, and loss of strength or stiffness. Building settlements, landslides, failure of the earthen dams, and other hazardous phenomena occur due to loss of strength and stiffness (Das and Luo 2016). The process by which the strength of the soil decreases is known as liquefaction. The first liquefaction

✉ Jitendra Khatti
jitendrakhatti197@gmail.com

✉ Yewuhalashet Fissha
yowagaye@gmail.com

Kamaldeep Singh Grover
ksgrover@rtu.ac.in

Hajime Ikeda
ikeda@gipc.akita-u.ac.jp

Hisatoshi Toriya
toriya@gipc.akita-u.ac.jp

Tsuyoshi Adachi
adachi.t@gipc.akita-u.ac.jp

Youhei Kawamura
kawamura@eng.hokudai.ac.jp

[1] Department of Civil Engineering, Rajasthan Technical University, Kota, Rajasthan, India

[2] Department of Geosciences, Geotechnology, and Materials Engineering for Resources, Graduate School of International Resource Sciences, Akita University, Akita 0108502, Japan

[3] Department of Mining Engineering, Aksum University, 7080 Aksum, Tigray, Ethiopia

[4] Faculty of Engineering, Hokkaido University, Kita 8, Nishi 5, Kita-Ku, Sapporo 0608628, Japan

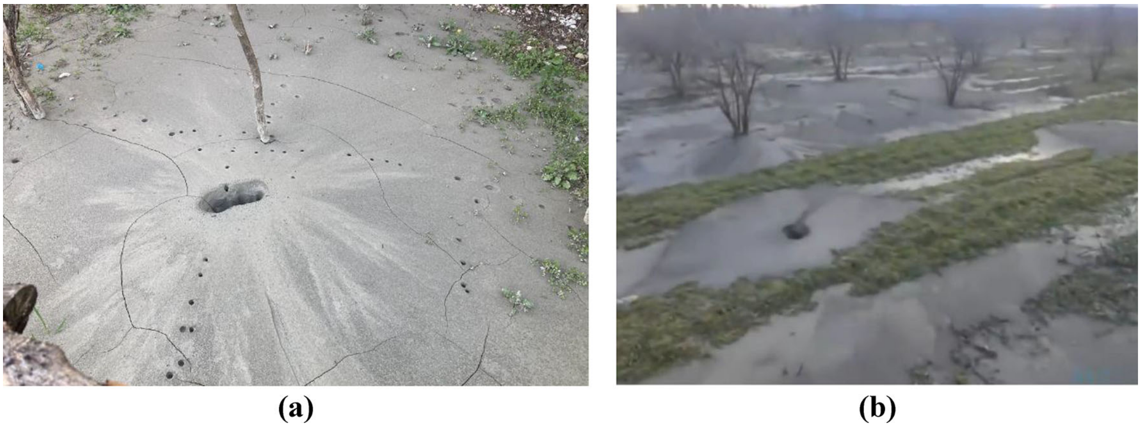


Fig. 1 After the Turkey–Syria earthquakes (2023): **a** liquefaction at the edge of Hatay city, and **b** soil liquefaction in northern Syria (Gazetas and Garini 2023)

occurred in Niigata (1964) and Alaska (1964) earthquakes due to medium to fine-grained saturated cohesionless soils. Casagrande (1936) introduced a concept for sandy soil based on a critical void ratio to understand the liquefaction phenomenon. Casagrande stated that (i) the dense sand behaves like dilated if it is subjected to shear and (ii) on the other side, the volume decreases in loose sand (Das and Luo 2016). The critical void ratio is "no volume or void ratio change during shear." Casagrande explained that the deposits of sand always have a void ratio that is more significant than the critical void ratio. The volume and void ratio of sand decreases under the vibration action. A deep degree of liquefaction, which occurs due to an increase in pore water pressure, was observed in Chile (2010), Wenchuan; China (2008), Kocaeli; Turkey (1999), Tangshan; China (1976), Niigata; Japan (1964), and Alaska; USA (1964) during the earthquakes (Sharma et al. 2021). The re-liquefaction occurred on the Pacific Coast of Tohoku (Japan) due to aftershocks (Yu et al. 2013). In 2019, ground shaking and slope failure occurred in the Poonch and Rajouri districts because of the Mirpur earthquake. Because of these earthquakes, the liquefaction phenomenon happened in different areas of Jammu and Kashmir (Ansari et al. 2022). Recently, two earthquakes, M7.8 and M7.5, hit Turkey–Syria on February 6, 2023 (Baltzopoulos et al. 2023). Significant liquefaction has been captured and caused severe damage to both surface buildings and underground infrastructure, as shown in Fig. 1.

Liquefaction is a drastic phenomenon that needs to be measured during an earthquake to decrease the liquefaction-induced damage. The empirical methods (based on in situ index tests) and laboratory methods (on undisturbed soil samples) are the main strategies for computing the liquefaction (Kumar et al. 2022a, 2023a). The lab procedures are challenging because of the need to keep the sample undisturbed and reconsolidated. The high-quality cycle simple shear testing is the most expensive and extremely difficult. The cyclic triaxial testing does not accurately represent the significant loading conditions for most seismic issues. Stress, strain, and energy-based methods determine the liquefaction potential. Youd and Idriss (2001) presented four methods for liquefaction analysis: (1) the Becker penetration test-BPT, (2) the in situ shear wave velocity test, (3) the cone penetration test-CPT, and (4) the standard penetration test-SPT. Robertson and Campanella (1985) introduced the CPT-based methods (an alternative approach to the SPT-based method) to determine soil liquefaction potential. It should be noted that each of these methods is fundamentally deterministic, calculating the factor of safety (FOS) against liquefaction while ignoring the numerous process uncertainties. The regional variance of many relevant parameters, such as unit weight and fine content, is the significant source of uncertainty in assessing liquefaction potential (Kumar et al. 2022a, 2023b). Mechanical and human errors affect the accuracy of estimating the liquefaction potential by the SPT/CPT test. In terms of the cyclic resistance ratio (CRR) and the cyclic stress ratio (CSR), the safety factor against liquefaction is computed. However, FOS > 1 may not always identify a non-liquefaction event, and conversely, FOS1 may not always result in liquefaction due to uncertainties associated with all these factors and the selected model. The boundary curve that separates liquefaction from non-liquefaction instances in the deterministic analysis is known as a “performance function” or "limit-state function." Many researchers have attempted to quantify the limits of the limit-state function. The probabilistic liquefaction reaction has been calculated in terms of liquefaction potential using probabilistic or statistical methods (Haldar and Tang 1979; Seed and Idriss 1971; Youd and Nobble 1997; and Toprak et al. 1999). The published models were developed using the post-liquefaction database. Still, these analyses did not consider the uncertainty of independent variables (inputs) and models. The output liquefaction potential may be inaccurate if significant uncertainties relate to the parameters and model. By estimating its probabilistic value and carrying out a reliability analysis while considering the model and parameter uncertainties, it is possible to avoid the difficulties brought on by the stochastic character of the input parameters used to determine liquefaction potential. Therefore, researchers and scientists applied soft computing techniques to assess the liquefaction potential using the in situ test database (Samui and Hariharan 2015). In recent decades, soft computing techniques have increased because of the availability of in situ databases (Kumar et al. 2023c). These techniques are more accurate and frequent than traditional methods in assessing the liquefaction potential.

Mase et al. (2023) assessed the liquefaction potential by applying the energy concept to a database collected from 38 Bengkulu City, Indonesia sites. Sahin and Demir (2023) evolved a greedy search-optimized stacking ensemble learning (G_SEL) approach using the Automated machine learning framework to estimate the liquefaction potential. The authors noted 98% accuracy of G_SEL for the SPT and CPT database and 99% accuracy of G_SEL for the shear wave velocity (SWV) database. Sun et al. (2023) reported that ground motion intensity measures are not able to predict the liquefaction potential because of rupture distance and earthquake magnitude. Kurnaz et al. (2023) employed support vector regression (SVR), decision tree (DT), random forest (RF), and artificial neural network (ANN) techniques to assess the factor of safety (FOS) against liquefaction. The researchers found the ANN technique more efficient (determination coefficient = 0.99) than other techniques in assessing the FOS. Jas and Dodagoudar (2023a) reviewed the articles published between 1994 and 2021. The reviewers concluded that a reliable and efficient assessment of liquefaction potential can be achieved by improving the database and re-configuring the existing soft computing techniques. Jas and Dodagoudar (2023b) employed the extreme gradient boosting-Shapley additive explanations (XGBoost-SHAP) approach for assessing the liquefaction potential using 253 data points. For this purpose, the researchers selected many input variables (Mag, σ_{v}, σ'_{v}, $a_{\max}$, d_{avg}, $(q_{\text{cN}})_{\text{avg}}$, $(F_{\text{sN}})_{\text{avg}}$, I_{c}, FC, q_{c1Ncs}, r_{d}, CSR, and K_{σ}) to employ the XGBoost-SHAP model. The research revealed that the SHAP technique (testing accuracy = 88.24%) minimizes the gap between traditional liquefaction field knowledge and soft computing methods. Ge et al. (2023) developed a region-specific fragility function using the Hierarchical Bayesian Model (HBM). The researchers noted that HBM performed better than the lumped parameter model (in case of a large database) and independent parameter model (in case of insufficient data). Demir and Sahin (2023) determined the liquefaction potential prediction capabilities of extreme gradient boosting (XGBoost), gradient boosting, and adaptive boosting techniques. The researchers found that the XGBoost technique is better than other techniques. Zhou et al. (2022) used CPT, SPT, and SWV databases to employ the genetic (GA) and grey wolf (GWO) optimized RF models. The authors noted that the GWO_RF model performed better for the CPT and SPT databases, while the GA_RF model performed better for the SWV database. Demir and Sahin (2022) compared random forest (RF), rational forest (ROTF), and canonical correlation forest (CCF) techniques in assessing the liquefaction potential using the CPT database. The researchers concluded that CCF and ROTF models are more precise in predicting the liquefaction potential than traditional RF models. Ozsagir et al. (2022) implemented ANN, RF, stochastic gradient descent (SGD), *k*-nearest neighbors (kNN), SVM, logistic regression, and DT techniques. It was noted that the DT model attained higher accuracy, i.e., 90%, in predicting the liquefaction potential than implemented models. The researchers also found that the assessment of liquefaction potential is affected by the mean grain size of soil, i.e., D_{50}. Ghani and Kumari (2022) performed first-order second-moment (FOSM) regression analysis (based on the reliability index) to assess liquefaction. Kumar et al. (2022a) found the best parameters for assessing the liquefaction potential. The authors reported that the FOS-based method is more accurate than the liquefaction potential index (LPI) and liquefaction severity index (LSI) methods. Furthermore, Kumar et al. (2022b) assessed the probability of liquefaction using artificial neural network optimized by teaching–learning (ANN_TL), shuffled complex evolution (ANN_SCE), imperialist competitive (ANN_IC) ant lion (ANN_AL), ant colony (AC), and artificial bee colony (ANN_ABC) algorithms. The ANN_AL model outperformed the other model by achieving determination coefficients of 0.682 and 0.723 in this study's training and testing phases. Kumar et al. (2022c) employed five ML models (XGBoost, RF, GBM, SVR, and GMDH) for predicting soil liquefaction potential. The author used SPT test data from 620 case records. In the reported study, the analysis was carried out with six input variables, i.e., d, $(N_1)_{60}$, FC, σ'_{v}, $a_{\max}$, and $M_{\text{w,}}$ and it was reported that the XGBoost model achieves the best prediction among all employed models. Ghani et al. (2022) implemented particle swarm (PS), genetic (GA), and firefly (FF) algorithms in an adaptive neuro-fuzzy inference system (ANFIS). The researchers found that the ANFIS_FF model assessed the liquefaction with residuals (RMSE) of 0.069 in the testing phase, which is excellent compared to other developed models. Zhang et al. (2021a) computed soil liquefaction using an extreme learning machine (ELM) via 226 CPT samples. The ELM model obtained a prediction accuracy of 87.50% in the testing phase, configured by a sign activation function and fifteen neurons. Zhang et al. (2021b) employed GWO-optimized SVM using the SPT database. The researchers noted that the GWO algorithm improved SVM accuracy by combining the SWV and SPT databases. Zhang et al. (2021c) used SWV and SPT databases to predict liquefaction utilizing deep neural networks (DNN). The authors trained and tested multilayer fully connected neural network models using SPT and combined SPT-SWV databases. The authors noted that the SPT-SWV database-tested model performed better than the SPT database-tested model, i.e., 93%. It was observed that including the SWV database improves the prediction accuracy of the computational models. Zhou et al. (2021a) predicted liquefaction potential using hybrid GWO_SVM and GA_SVM models. The investigators utilized SPT, CPT, and SWV databases and reported that model GWO_SVM predicted liquefaction potential with an accuracy of 98.25%, 90.32%, and 32.31%, respectively, higher than the GA_SVM model. Zheng et al. (2021) compared ANN and SVM models in assessing liquefaction and concluded that ANN is better than the SVM model. Zhou et al. (2021b) implemented a PSO-optimized kernel ELM model (PSO_KELM) using CPT and SWV databases. The combination of CPT and SWV gives a significant accuracy for the PSO_KELM model, i.e., 83.6%. Kumar et al. (2021) employed an emotional backpropagation neural network (EDNN) model for predicting liquefaction and concluded prediction accuracy was 89% higher than the deep learning model, i.e., 79%. Alizadeh et al. (2021), Hu (2021), Ahamd et al. (2021), Kim et al. (2021), Ghorbani and Eslami (2021), Demir and Şahin (2021), and Khan et al. (2021) noted that soft computing techniques are reliable and accurate in assessing the liquefaction of soil.

Zhang et al. (2020) assessed liquefaction potential using a calibrated SPT method via region-specific variability. Xing et al. (2020) computed the liquefaction using a multilayer perceptron neural network model. Rollins et al. (2020) performed a dynamic cone penetrometer (DCPT) test at Avasinis, Italy, for a sandy gravel alluvial soil profile with 20% to 40% gravel content. The researchers obtained corrected values at three sites, i.e., loose to medium–dense soil profiles. Park et al. (2020) predicted liquefaction-induced settlement at Pohang by applying the ANN model. The authors found that the ANN model developed by utilizing cyclic stress ratio corrected N-SPT and unit weight attains a higher performance in assessing the liquefaction-induced settlement, i.e., 86.01%. Nong et al. (2020) investigated the effect of cyclic loading frequency on sand liquefaction assessment. The investigators collected clean soil having relative densities of 80% and 40% for dense and loose sand, respectively. It was noted that sand liquefaction resistance increases with the loading frequency. The maximum increase in cyclic resistance was found at 19% and 15% for dense and loose sand, respectively, when the frequency becomes 0.1 to 0.5 or 1 Hz. Njock et al. (2020) applied a new evolutionary neural network (EVONN) with a t-distributed stochastic neighbor embedding algorithm. The proposed EVONN model attained 97% testing performance in this research. Mahmood et al. (2020) employed Bayesian neural network models using the CPT database to assess liquefaction potential. The researchers implemented the interpretive structural modelling (ISM) technique to construct the hybrid model and noted that the proposed hybrid model assessed the liquefaction potential with an accuracy of 97.248%. The liquefaction potential can be assessed by (i) evolutionary polynomial regression technique (Ghorbani et al. 2020); (ii) multi-objective feature selection algorithm (Das et al. 2020); (iii) multivariate adaptive regression splines (Chen et al. 2020); (iv) ICA-optimized multi-objective generalized feedforward neural network (Abbaszadeh and Maghsoudi 2020). Abbaszadeh and Maghsoudi (2020) developed ICA models with MOGFFN to predict soil LP using σ'$_v$ (kPa), a_{max}(g), CSR, FC (%), SPT value, CRR, r_d, V_s (m/s), and γ (KN/m^3), as an independent variable. The authors used 296 data points in the reported study. Based on the research findings, the ICA-MOGFFN ($R = 96\%$) is better than the MOGFFN (94%) approach for predicting soil liquefaction potential. Zhou et al. (2019) constructed a stochastic gradient boosting (SGB) model using SPT and CPT databases to predict liquefaction potential. It was noted that the SPT-based SGB model achieved an accuracy of 95.45% in the testing. Sabbar et al. (2019) compared the ANN and genetic programming models, and ANN outperformed the genetic programming model with a determination coefficient of 0.864. Rahbarzare and Azadi (2019) implemented hybrid GA and PSO-optimized fuzzy SVM models. The radial basis function (RBF) kernel-based model obtained 99.09% accuracy in assessing liquefaction. Kutanaei and Choobbasti (2019) assessed the liquefaction potential around a submarine pipeline during the earthquake. The researchers employed ANN with the local RBF differential quadrature method. The researchers found that liquefaction potential decreases with increasing hydraulic conductivity, deformation module, and Poisson ratio. Kurnaz and Kaya (2019a) employed an ensemble group method of data handling (E_GMDH) model using the CPT database to predict liquefaction. The accuracy of ANN, SVM, LR, RF, GMDH, and E_GMDH was compared. The authors noted that E_GMDH outperformed the other computational models with an accuracy of 98.36%. Furthermore, Kurnaz and Kaya (2019b) used the E_GMDH model for the SPT database to predict liquefaction. Model E_GMDH attained an accuracy of 99.30% higher than ANN, SVM, LR, RF, and GMDH. Johari et al. (2019) used a jointly distributed random variables (JDRV) method for liquefaction prediction using SWV and SPT databases. The researchers concluded that the JDRV method predicts liquefaction potential for SWV better than the SPT database. Youd (2018) assessed liquefaction-induced lateral spread displacement using multiple linear regression. Tang et al. (2018) used the SPT database to assess the liquefaction hazards by applying a Bayesian neural network (BNN). The authors compared the BNN model with ANN. It was noted that the BNN model assesses ground

Table 1 Details of soft computing models available in the literature

S. no	Soft computing approaches	Details of Input Parameters		Accuracy	References
		Variables	Count		
1	XGBoost_SHAP	Mag, σ_v, σ'_v, a_{max}, d_{avg}, $(q_{cN})_{avg}$, $(F_{sN})_{avg}$, I_c, FC, q_{c1Ncs}, r_d, CSR, and K_σ	253	88.24	Jas and Dodagoudar (2023)
2	RF	D, q_c, R_f, σ_v, σ'_v, a_{max}, M_w, d_w, FC, CSR	479	93.24	Demir and Sahin (2022)
3	XGBoost	Z, $(N_1)_{60}$, FC, σ'_v, a_{max}, and M_w	620	89.40	Kumar et al. (2022c)
4	PFEMS	$(N_1)_{60}$, FC, σ'_v, a_{max}, and M_w	834	91.20	Kumar et al. (2022a)
5	ICA_MOGFFN	σ'_v, a_{max}, CSR, FC, SPT value, CRR, r_d, V_s, and γ	296	96.00	Abbaszadeh and Maghsoudi (2020)
6	BN	$(N_1)_{60}$, FC, σ'_v, a_{max}, M_w, PGA, ST, GWT, and t_d	210	84.90	Hu and Liu (2019)
7	PSO_GA_FSVM	q_c, σ'_v, D_{50}, a_{max}, and M_w	109	100.00	Rahbarzare and Azadi (2019)
8	KFDA_LSSVM	D, σ_v, σ'_v, SC, V_s, a_{max}, and M_w	1029	88.56	Hoang and Bui (2018)
9	LR_MARS	M, h, R_f, σ_v, σ'_v, a_{max}; M, R_f, σ_v, σ'_v, q_c, D_{50}, a_{max}; M, h, σ_v, σ'_v, q_c	226 170 466	89.30 91.10 87.90	Zhang and Goh (2016)
10	PSO_SVM	M, d_w, d_s, σ_v, σ'_v, q_c, a_{max}, D_{50}, and τ_{av}, σ'_v	79	96.55	Xue and Yang (2016)
11	ANN	T and q_{c1N}	NA	99.99	Erzin and Ecemis (2015)
12	BOSVMs	M, d_w, d_s, σ_v, σ'_v, q_c, a_{max}, D_{50}, and τ_{av}, σ'_v	79	96.40	Zhang et al. (2022)
13	ANN	σ_v, σ'_v, D_{50}, $(N_1)_{60}$, a_{max}, M_w, and τ_{av})	370	99.60	Hsu et al. (2017)

cracks, sand boils, settlement, lateral spreading, and severity of liquefaction-induced hazards with high accuracies, i.e., 90.9%, 91.8%, 83.6%, 95.5%, and 93.6%, respectively. Nejad et al. (2018) implemented an RF model using the SWV database to evaluate the liquefaction potential. The researchers noted that the RF model achieved over 92% accuracy in assessing the liquefaction potential. Mola-Abasi et al. (2018) used the CPT database to assess the liquefaction using a triangular chart. The authors summarized that the triangular chart is more efficient for assessing the liquefaction. Hoang and Bui (2018) estimated liquefaction by employing hybrid kernel fisher discriminant analysis with the least square support vector machine (KFDA_LSSVM). It was noted that the KFDA improves the prediction accuracy of the LSSVM model in estimating liquefaction. Hsu et al. (2017) constructed ANN models using 370 case studies with input parameters (σ_v, σ'_v, D_{50}, $(N_1)_{60}$, a_{max}, M_w, and τ_{av},σ'_v) to predict the soil LP. The ANN model has a 99.6% success rate in correctly predicting when liquefaction and non-liquefaction will occur. Zhang and Goh (2016) mapped a comparative study between the LR-MARS model and ANN techniques to predict soil LP based on three datasets (75, 56, and 116 cases) for testing. In the 1st case study, the input parameters are M_w, h, R_f, σ_v, σ'_v, and a_{max}; however, M, R_f, σ_v, σ'_v, q_c, D_{50}, a_{max} and M_w, h, σ_v, σ'_v, q_c, and a_{max} are for the 2nd and 3rd case studies respectively. The LR-MARS approach performed better than the ANN model based on SR evaluation testing results, i.e., 89.3%, 91.1%, and 87.9% for the 1st, 2nd, and 3rd case studies. Xue and Yang (2016) constructed a PSO-SVM model to predict the LP of soils using M_w, d_w, d, σ_v, σ'_v, q_c, a_{max}, D_{50}, and τ_{av}, σ'_v for 79 CPT records. Study findings showed that the created PSO-SVM approach's classification accuracy rate (96.55%) is higher than those of a grid search, among other methods (i.e., simplified CPT and Seed techniques with an accuracy rate of 67% and 74%). Erzin and Ecemis (2015) mapped a comparative study among the ANN model, simulation, and experimental observations to estimate the liquefaction resistance. The authors used T and q_{c1N} for ANN models 1 and 2, respectively. The ANN model attained higher test performance ($R = 0.999$) than the field liquefaction screening chart ($R = 0.69$). Numerous investigators and scientists published several soft computing models to predict the liquefaction potential, as summarized in Table 1.

1.1 Gap identification

The published research reveals the superiority of soft computing models over conventional methods in terms of prediction accuracy. Still, these computational methods are more suitable if an extensive database is available. Each soft computing method has limitations, including machine and hybrid learning methods, and no accurate method is suggested to assess soil liquefaction potential. The ANN technique is the most famous for assessing liquefaction but doesn't establish the relationship between the independent and dependent variables. On the other side, the logistic regression (LR), k-nearest neighbor (kNN), support vector classifier (SVC)-based on linear, polynomial, and Gaussian kernels, least square support vector classifier (LSSVC)-based on linear, polynomial, and Gaussian kernels, decision tree (DT), random forest (RF), gradient boosting (GB), adaptive boosting (ADB), extreme gradient boosting (XGB), light gradient boosting (LGB), ensemble tree stacking (ETS), ensemble tree voting (ETV), and artificial neural network (ANN)-based on adam, stochastic gradient descent, Broyden–Fletcher–Goldfarb–Shanno models have not been employed and compared in assessing the liquefaction potential. Furthermore, the Slime Mold Algorithm (SMA), Harris Hawks Optimizer (HHO), Runge–Kutta Optimized (RUN), Particle Swarm Optimizer (PSO), and Genetic Algorithm (GA) have not been used to construct hybrid soft computing models. The hybrid models work efficiently and achieve high prediction accuracy in the case of large databases. Still, the impact of a small database on the accuracy of hybrid models has not been studied for hybrid computational models. Also, the literature study demonstrates that the impact of multicollinearity on the accuracy of conventional and hybrid soft computing models has not been investigated in predicting the liquefaction potential.

1.2 Novelty of the present work

The novelty of this research has been mapped based on the gap found in the literature and mentioned as follows:

- This research develops the logistic regression (LR), k-nearest neighbor (kNN), support vector classifier (SVC)-based on linear, polynomial, and Gaussian kernels, least square support vector classifier (LSSVC)-based on linear, polynomial, and Gaussian kernels, decision tree (DT), random forest (RF), gradient boosting (GB), adaptive boosting (ADB), extreme gradient boosting (XGB), light gradient boosting (LGB), ensemble tree stacking (ETS), ensemble tree voting (ETV), and artificial neural network (ANN)-based on adam, stochastic gradient descent, Broyden–Fletcher–Goldfarb–Shanno models and investigates the liquefaction potential prediction capabilities developed models.
- This research optimizes the models using Slime Mold Algorithm (SMA), Harris Hawks Optimizer (HHO), Runge–Kutta Optimized (RUN), Particle Swarm Optimizer (PSO), and Genetic Algorithm (GA). It maps a comparison among hybrid models, i.e., RUN_XGB, SMA_ADB, SMA_RF, SMA_GB, HHO_LGB, SMA_DT, SMA_LSSVC_RBF, SMA_ANN_Adam, PSO_SVC_RBF, HHO_SVC_L, SMA_SVC_P, SMA_LSSVC_P, SMA_LSSVC_L, SMA_ANN_IBFGS, and SMA_ANN_SGD to study the effect of the limited database on the accuracy of hybrid models.
- This research illustrates the impact of multicollinearity on the accuracy of the developed conventional and hybrid models for the first time.

1.3 Research significance

This research introduces an optimal performance model by comparing conventional and hybrid soft computing models. This work will help geotechnical engineers find the possibilities of liquefaction using soft computing techniques in case of a limited database.

2 Research methodology

The present research introduces the optimal performance soft computing model for predicting the soil liquefaction potential. For this purpose, a database has been compiled from the published article by Ahmad et al. (2021). The data proportionality method has been applied, and training and testing databases have been created by randomly selecting 70% and 30% of the database. Furthermore, models based on logistic regression (LR), *k*-nearest neighbor (kNN), support vector classifier (SVC)-based on linear, polynomial, and Gaussian kernels, least square support vector classifier (LSSVC)-based on linear, polynomial, and Gaussian kernels, decision tree (DT), random forest (RF), gradient boosting (GB), adaptive boosting (ADB), extreme gradient boosting (XGB), light gradient boosting (LGB), ensemble tree stacking (ETS), ensemble tree voting (ETV), and artificial neural network (ANN)-based on adam, stochastic gradient descent, Broyden–Fletcher–Goldfarb–Shanno approaches have been developed and optimized by Slime Mould Algorithm (SMA), Runge–Kutta Optimizer (RUN), Particle Swarm Optimization (PSO), Harris Hawks Optimization (HHO), and Genetic Algorithm (GA). Thus, RUN_XGB, ET_voting, SMA_ADB, SMA_GB, SMA_RF, ST_Stacking, HHO_LGB, kNN, SMA_DT, SMA_LSSVC_RBF, SMA_ANN_Adam, PSO_SVC_RBF,

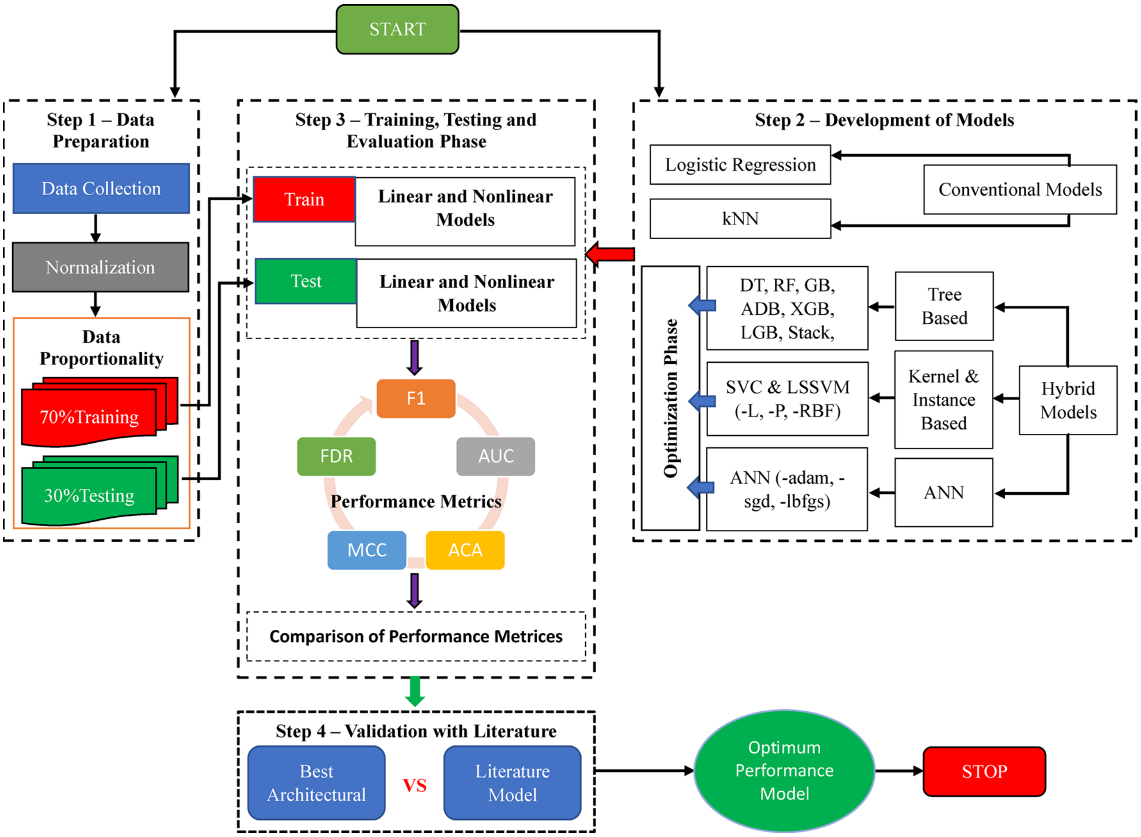

START
Step 1 – Data Preparation
Data Collection
Normalization
Data Proportionality
70%Training
30%Testing
Step 3 – Training, Testing and Evaluation Phase
Train
Linear and Nonlinear Models
Test
Linear and Nonlinear Models
F1
FDR
AUC
Performance Metrics
MCC
ACA
Comparison of Performance Metrices
Step 2 – Development of Models
Logistic Regression
kNN
Conventional Models
Optimization Phase
DT, RF, GB, ADB, XGB, LGB, Stack,
Tree Based
SVC & LSSVM (-L, -P, -RBF)
Kernel & Instance Based
Hybrid Models
ANN (-adam, -sgd, -lbfgs)
ANN
Step 4 – Validation with Literature
Best Architectural
VS
Literature Model
Optimum Performance Model
STOP

Fig. 2 Illustration of research methodology of the present work

HHO_SVC_L, SMA_SVC_P, SMA_LSSVC_P, SMA_LSSVC_L, SMA_ANN_IBFGS, logistic regression, and SMA_ANN_SGD models have been developed, trained, tested, and analyzed. The soft computing models have been developed using the Python platform. The metrics, such as recall, F1, TNR, NPV, FPR, FNR, FDR, FOR, ACA, and MCC, have measured the performance and accuracy of the models. Based on performance metrics, the optimum performance soft computing model has been recognized in predicting the soil liquefaction potential. Furthermore, the optimum performance model has been validated by comparing the models available in the literature (Ahmad et al. 2021). Figure 2 illustrates the research methodology of the present work.

3 Data analysis and computational approaches

3.1 Data analysis

The present investigation has been carried out using the database published by Ahmad et al. (2021). The database includes 170 case history records of both liquefied and non-liquefied sites. These records have been collected from field observations of nine distinct earthquakes that occurred around the world: the Niigata earthquake in 1964, the San Fernando Valley earthquake in 1971, the Haicheng earthquake in 1975, the Tangshan earthquake in 1976, the Vrancea earthquake in 1977, Imperial Valley earthquake in 1979, Nihonkai–Cho earthquake in 1983, Saguenay earthquake in 1988, and Loma Prieta earthquake in 1989. The database comprises 66 non-liquefaction and 104 liquefaction case history records. Five parameters, such as earthquake magnitude (M), peak ground acceleration ($a_{\max}$, in g), cone tip resistance ($\mathrm{q_c}$, in MPa), mean grain size (D_{50}), and vertical-effective stress (σ'_{v}, in kPa), have been selected out of six parameters for this research. The descriptive statistics of the five parameters obtained from the literature are mentioned in Table 2.

Table 2 presents that that database contains D_{50}, σ'_{v}, $a_{\max}$, M, and q_c in the range of 0.016 to 0.480 mm, 13.9 to 227.50 kPa, 0.1 to 0.6 g, 5.9 to 7.8, and 0.38 to 26.0 MPa, respectively. Several researchers have suggested that the database must be reliable and cover the full range of parameters to achieve better assessment (Khatti and Grover 2023a; Lu et al. 2020; Huang et al. 2020; Cavaleri et al. 2017). In addition, the frequency distribution of each parameter input variable is illustrated in Fig. 3.

Table 2 Statistical description of the features

Features	Unit	Count	Mean	StDev	Minimum	Maximum	IQ25	IQ50	IQ75
D_{50}	mm	170	0.168	0.102	0.016	0.480	0.07225	0.160	0.243
σ'_v	kPa	170	78.586	44.857	13.900	227.50	48.325	65.950	100.475
a_{max}	g	170	0.284	0.134	0.100	0.60	0.200	0.230	0.400
M	–	170	7.297	0.605	5.900	7.800	7.100	7.700	7.800
q_c	MPa	170	6.414	5.241	0.380	26.00	2.257	4.905	8.830

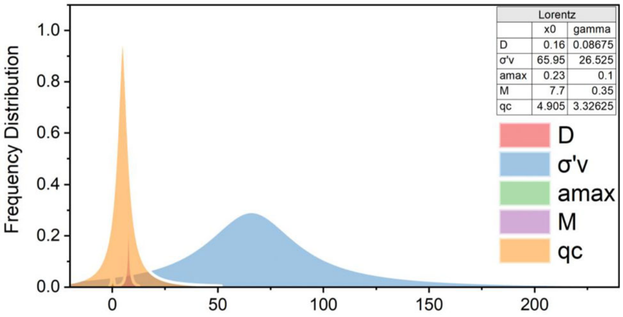


Fig. 3 Frequency distribution of features

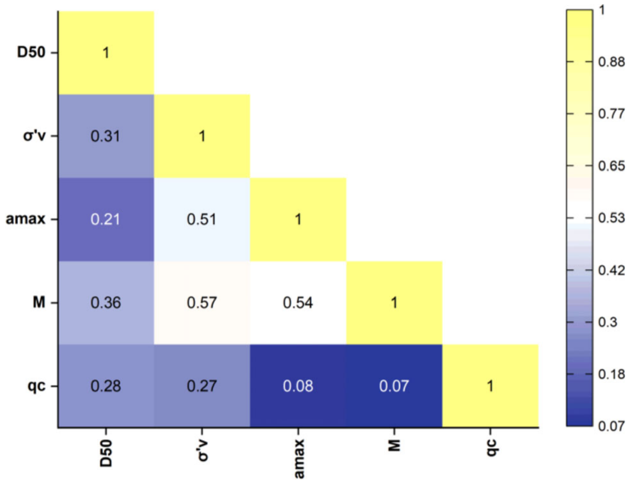


Fig. 4 Demonstration of relationship among variables in terms of the correlation matrix

3.1.1 Pearson product moment correlation coefficient

Pearson product-moment correlation coefficient (I) is a method to determine the relationship or correlation of the pair of variables. The relationship between variables is defined as very strong ($\pm 1.0 < I < \pm 0.81$), strong ($\pm 0.80 < I < \pm 0.61$), moderate ($\pm 0.60 < I < \pm 0.41$), weak ($\pm 0.40 < I < \pm 0.21$), and no ($\pm 0.20 < I < \pm 0.00$) correlation (Hair et al. 2013; Khatti and Grover 2021). However, five input variables, M, a_{max} (in g), q_c (in MPa), D_{50} (in mm), and σ'_v (in kPa), have been used to assess the soil liquefaction potential in the present research. The Pearson product-moment correlation coefficient has been computed and presented in Fig. 4.

Figure 4 demonstrates that (i) D_{50} weakly correlates with M, a_{max}, q_c, and σ'_v, (i) σ'_v moderately correlates with M and a_{max}, (iii) M moderately correlates with a_{max}, and (iv) q_c has no relationship with M and a_{max}. The correlation coefficient is the preliminary method to determine the multicollinearity of independent variables. However, the correlation coefficient method is one of the multicollinearity methods used to find the multicollinearity of input variables. Since the input variables have weak to moderate correlations. Therefore, it may be noted that the independent variables have multicollinearity. The variance inflation factor (VIF) is another method for determining the multicollinearity level, which has been discussed in the next section.

3.1.2 Multicollinearity analysis

When big data is utilized in artificial intelligence for more precise predictions, multicollinearity is a phenomenon that happens throughout the regression process (Chan et al. 2022). When performing a linear multivariable regression analysis, multicollinearity occurs (1) when dependent and independent variables do not significantly correlate (2) when independent and dependent variables are highly associated (3). Pairwise scatterplot, Pearson's correlation coefficient, and variance inflation factor (VIF) are utilized to find the multicollinearity in the database. (Shrestha 2020; Obite, 2020; Garg and Tai 2012; Mansfield and Helms 1982; Gunst and Webster 1975). This section uses the variance inflation factor to determine the multicollinearity. VIF's mathematical representation is:

$$VIF = \frac{1}{\left(1 - R^2\right)} \tag{1}$$

where R^2 is the coefficient of determination. The determination coefficient is estimated for each independent variable using the regression technique. The multivariable linear regression analysis is conducted for each independent variable to determine the VIF. One independent variable is

Table 3 Results of multicollinearity analysis

Particulars	Coefficients	Standard Error	*t* Stat	*P*-value	Lower 95%	Upper 95%	R^2	VIF
Intercept	– 0.4282	0.4566	– 0.9379	0.3497	– 1.3297	0.4733	–	–
D_{50}	0.8495	0.2912	2.9171	0.004	0.2745	1.4245	0.2689	1.3678
σ’v	– 0.0013	0.0008	– 1.7273	0.086	– 0.0029	0.0002	0.4676	1.8783
a_{max}	0.6113	0.2372	2.5776	0.0108	0.143	1.0796	0.3533	1.5464
M	0.1688	0.0567	2.9783	0.0033	0.0569	0.2808	0.4446	1.8004
q_c	– 0.0629	0.0055	– 11.4195	0	– 0.0738	- 0.052	0.2161	1.2757

retained among the others as the dependent or target independent variable. According to a number of experts, the troublesome multicollinearity of the input variables always has a variance inflation factor (VIF) of greater than 10. (Vittinghoff et al. 2006; Witten and James 2013; Menard 2002). Moreover, many investigators have reported that the 2.5 or 5 is considerable multicollinearity or collinearity of datasets (Menard 2002; Johnston et al. 2018). However, the level of multicollinearity hasn't been clearly defined. Therefore, Khatti and Grover (2023b) proposed a level of multicollinearity. In this work, the multicollinearity analysis has been performed for M, $a_{\max}$, q_c, D_{50}, σ'_v input variables and is given in Table 3.

Table 3 reveals that earthquake magnitude (M), peak ground acceleration ($a_{\max}$), cone tip resistance (q_c), mean grain size (D_{50}), and vertical-effective stress (σ'_v) have weak multicollinearity levels. Thus, the compiled database is suitable for the present study.

3.1.3 Hypothesis testing

The ANOVA and Z tests have been performed in this research to identify the hypothesis. ANOVA testing is a parametric statistical test used to determine the hypotheses of the different forms of research work. Always run the ANOVA test when there are more than 30 points. (Sawyer 2009; Larson 2008; Gelman 2005; Christensen 1996). The research hypotheses (H_R) for the present work are (i) q_c plays an important role in assessing the liquefaction potential than D_{50}, and (ii) all input parameters have an equal influence in assessing the liquefaction potential. The results of the ANOVA test are mentioned in Table 4.

Table 4 reveals that the input parameters satisfy the research hypothesis (H_R) and reject the null hypothesis (H_0). It can be seen that all input variables present F > F crit and $p \leq 0.005$. The z test is a statistical analysis used to determine the hypothesis and the statistical significance of a discovery or correlation. (Kim 2017; Lin and Mudholkar, 1980; Lawley 1938). The results of the z test for input parameters in assessing the liquefaction potential are given in Table 5.

Table 5 illustrates that input parameters have z one tail lower than z two tail, presenting the research hypothesis for the present research as true. To sum up, the ANOVA and z tests accept the research hypothesis (H_R) for this study, i.e., (i) qc plays an important role in assessing the liquefaction potential than D_{50} and (ii) all input parameters have an equal influence in assessing the liquefaction potential. However, Goh (1996) reported that qc is more sensitive than D_{50} in the liquefaction potential prediction.

3.2 Computational approaches

In this study, the conventional and hybrid approaches are arranged into four groups for comparison purposes: baseline, tree-based, support vector machine (SVM)-based, and artificial neural network (ANN) models. The adopted soft computing approaches are discussed as follows:

3.2.1 Baseline models

Logistic regression (LR) Logistic regression is a widely used statistical model for binary classification problems. It is a linear model that uses a logistic function to convert a linear combination of independent variables into a probability of the response variable (Hastie et al. 2009).

K-nearest neighbors (kNN) K-nearest neighbors (kNN) is a non-parametric instance-based learning algorithm that operates without making assumptions about the underlying data distribution. It utilizes the concept of proximity to classify or predict the grouping of a given data point. (Hastie et al. 2009).

Support vector machine (SVM) SVM is a popular machine learning algorithm for solving classification and regression problems. The selection of the best hyperplane that separates between classes is based on the width of the margin between classes (James et al. 2013). Kernel Support Vector Machine (K-SVM) is a machine learning algorithm that models non-linear relationships between input and output variables. In this study, three SVM models are implemented based on

Table 4 Results of the ANOVA test

Inputs	Source of variation	SS	df	MS	*F*	P-value	*F* crit
D_{50}	Between groups	16.66876	1	16.66876	133.6429	2.85E-26	3.869118
	Within groups	42.15744	338	0.124726			
σ'_v	Between groups	516,804.7	1	516,804.7	513.6187	8.35E-70	3.869118
	Within groups	340,096.6	338	1006.203			
a_{max}	Between groups	9.115144	1	9.115144	70.96825	1.05E-15	3.869118
	Within groups	43.41264	338	0.12844			
M	Between groups	3799.587	1	3799.587	12,561.78	2.2E-269	3.869118
	Within groups	102.2355	338	0.302472			
q_c	Between groups	2862.243	1	2862.243	206.5872	6.87E-37	3.869118
	Within groups	4682.953	338	13.85489			

Table 5 Results of z test for input parameters

Parameters	*D*	σ'v	a_{max}	*M*	qc	Liquefaction
Mean	0.1689	78.5865	0.2843	7.2976	6.4146	0.6118
Known variance	0.0105	2012.167	2012.167	2012.1671	2012.1671	0.2389
Observations	170	170	170	170	170	170
Hypothesized mean difference	0.0000	0.0000	0.0000	0.0000	0.0000	
z	− 11.5616	22.6632	− 0.0952	1.9432	1.6866	
P (Z< = *z*) one tail	0.0000	0.0000	0.4621	0.0260	0.0458	
z Critical one tail	1.6449	1.6449	1.6449	1.6449	1.6449	
P (Z< = *z*) two tail	0.0000	0.0000	0.9242	0.0520	0.0917	
z Critical two tail	1.9600	1.9600	1.9600	1.9600	1.9600	

the kernel function, namely: linear kernel (SVC_L), polynomial kernel (SVC_P), and Gaussian kernel (SVC_RBF). The Least Squares Support Vector Machine (LSSVM) is a type of SVM algorithm that aims to minimize classification error rather than maximize the margin between the hyperplane and the data points. LSSVM aims to find a hyperplane that separates the data points into two classes while minimizing the classification error. LSSVM minimizes the sum of squared errors between the predicted and actual class labels, subject to a constraint that ensures that all data points are correctly classified (Suykens and Vandewalle 1999). In this study, three kernel functions have been implemented, i.e., linear, polynomial, and Gaussian function kernels, and related models are designated as LSSVC_L, LSSVC_P, and LSSVC_RBF.

3.2.2 Tree-based models

Decision tree (DT) A decision tree is a non-parametric hierarchical model commonly used for solving classification and regression problems. It is a tree-like structure that partitions the input space based on the values of the input variables to make predictions about the output variable (Hastie et al. 2009). The present research has implemented entropy as a measure of information gain, defined as the amount of information needed to describe data accurately. If the data is homogeneous, the entropy equals 0. If not, the entropy will increase towards 1.

Random forest (RF) Random Forest is an ensemble method that combines multiple decision trees to make predictions by aggregating their results. It uses randomness in data sampling and feature selection processes to improve the model's robustness and generalization capabilities (Hastie et al. 2009).

Gradient boosting (GB) Gradient boosting is a popular ensemble learning method that combines multiple weak models to make more accurate predictions. The Gradient Boosting algorithm minimizes a loss function using gradient descent (Hastie et al. 2009). By adding weak learners to the model, the algorithm gradually reduces the residual errors, thereby improving the model's accuracy.

Adaptive boosting (ADB) Adaptive boosting, also known as AdaBoost, is an ensemble learning method that combines multiple weak models to make more accurate predictions.

The basic idea of AdaBoost is to iteratively train a series of weak classifiers on weighted versions of the training data (Hastie et al. 2009). The final classifier is a weighted combination of these weak classifiers, where the weights are determined by their overall performance in the training process.

Extreme gradient boosting (XGB) XGBoost (eXtreme Gradient Boosting) is a popular decision tree-based ensemble learning algorithm for classification, regression, and ranking tasks. XGB is based on the gradient boosting framework (Chen and Guestrin 2016). XGB calculates the negative gradient of the loss function concerning the current predictions and fits a new tree to the negative gradient. This process continues until a stopping criterion is reached or a predefined number of trees has been added.

Light gradient boosting (LGB) The light gradient-boosting machine (LGB) algorithm is based on gradient boosting, which involves iteratively adding weak learners to improve the overall model performance. LGB does not grow a tree level-wise. Instead, it grows tree leaf-wise. It chooses the leaf it believes will yield the most significant decrease in loss (Ke et al. 2017).

Ensemble tree stacking (ETS) Stacking is an ensemble learning method that involves using the predictions of the individual models as input to a second-level model, which learns to combine these predictions to make a final prediction. This second-level model is typically a linear model, such as logistic regression or a support vector machine, but it can also be another tree-based model (Rokach 2019). Combining random forest and gradient boosting algorithms allows the ensemble stacking tree classifier to produce more accurate and stable predictions while maintaining computational efficiency.

Ensemble tree voting (ETV) Voting is an ensemble learning method that combines the predictions of multiple models to make more accurate predictions. This combination is done by selecting the most frequent prediction. Voting can be performed in different ways, such as simple majority voting or weighted voting, where each model's vote is weighted based on its performance on the validation set (Rokach 2019). Like the stacking approach, the most accurate models, i.e., RF, GB, and XGB, have been compiled using voting techniques.

3.2.3 Neural network-based models

Artificial Neural Network (ANN) is an approach to deep learning inspired by the structure and function of the human brain. ANNs consist of interconnected layers of nodes, or neurons, that process input data and produce output predictions. One of the key components of ANNs is the optimization algorithm used to update the weights and biases of the model during training. The choice of an optimization algorithm can have a significant impact on the performance of the model. Three commonly used optimization algorithms are Adam, Stochastic Gradient Descent (SGD), and Quasi-Newton-based optimizer (lbfgs).

Adam is an adaptive learning rate optimization algorithm that adjusts the learning rate of each weight based on the magnitude of the gradients and the second moments of the gradients. It allows Adam to adapt to different data types and converge faster than other optimization algorithms. Adam also includes bias-correction terms to improve performance on small datasets (Kingma and Ba 2015). SGD is a stochastic optimization algorithm that updates the model weights using the gradient of the loss function concerning the weights. SGD updates the weights in small batches, which reduces the computational cost and can lead to faster convergence (Goodfellow et al. 2016). However, SGD can suffer from slow convergence and be sensitive to the choice of learning rate and batch size (Lin et al. 2020). Quasi-Newton methods (lbfgs) are optimization algorithms that use an approximation of the Hessian matrix to update the model's weights. Quasi-Newton methods can converge faster than SGD and are less sensitive to the choice of learning rate and batch size (Goodfellow et al. 2016). However, they can be computationally expensive.

3.3 Hyperparameters tuning

Hyperparameter tuning is a critical task in machine learning that aims to optimize the hyperparameters of a given ML model. In this study, the hyperparameters of the tree-based models were the tree depth, the number of estimators for ensemble models, and the learning rate for boosting-based models. For the SVM-based models, the hyperparameters are the scale and regularization parameters. Finally, the hyperparameters refer to the number of layers and neurons in each layer for the ANN-based models. Five hyperparameter tuning algorithms, i.e., Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Slime Mold Algorithm (SMA), Harris Hawks Optimization (HHO), and RUNge Kutta optimizer (RUN), to examine the optimization performance of swarm-based, bio-based, and evolutionary-based algorithms. Each optimization algorithm was run for each model during the cross-validation until no further improvement in the accuracy score was noticed. The optimization algorithms are briefly discussed as follows:

3.3.1 Particle swarm optimization (PSO)

Particle Swarm Optimization (PSO) is a metaheuristic optimization approach that mimics the collective behavior of social animals, such as flocks of birds (Clerc 2010). The PSO algorithm maintains a population of potential solutions or particles that traverse through a search space to find the optimal solution. Each particle represents a potential solution to the optimization problem. Its motion is governed by its previous best position and the best position discovered by the swarm so far (Clerc 2010). The iterative process continues until a stopping criterion, such as a maximum number of iterations or a satisfactory solution, is achieved.

3.3.2 Genetic algorithm (GA)

The genetic algorithm is a technique based on natural selection, the process that propels biological evolution, and it is used to solve both limited and unconstrained optimization issues. In binary classification, binary chromosomes are implemented. Then, two parent chromosomes are randomly selected, and their fitness is evaluated by measuring the error metrics. More error means less fit, and crossover and mutation are to be applied to the original solutions for better fitness and less error. This procedure is repeated over generations of solutions until the required fitness is achieved.

3.3.3 Slime mold algorithm (SMA)

The Slime Mold Algorithm (SMA) is a bio-inspired optimization technique miming slime mold organisms' growth and foraging behaviors. Slime mold, also known as *Physarum polycephalum*, exhibits remarkable problem-solving capabilities despite lacking a centralized nervous system. SMA leverages the principles of this biological organism to guide the search for optimal solutions in complex optimization problems. Inspired by slime mold's self-organization and adaptability, SMA employs positive feedback, negative feedback, and random explorations to explore the search space efficiently and converge toward optimal solutions (Li et al. 2020).

3.3.4 Harris hawks optimization (HHO)

Harris Hawks Optimization (HHO) is an optimization algorithm inspired by the hunting behavior of Harris's hawks, a species of raptors known for their exceptional group cooperation during hunts. HHO introduces a hierarchical leadership structure among the hawks, where a dominant leader coordinates and guides the group's hunting strategy. This hierarchical leadership structure fosters effective exploration and exploitation of the search space. The algorithm employs diverse mechanisms, such as position updating, exploration through randomization, and exploitation through information sharing to optimize solutions. In HHO, the hunting process is simulated through mathematical equations that guide the movement of the hawks in the search space. By emulating the cooperative hunting behavior of Harris's hawks, HHO aims to enhance the efficiency and effectiveness of optimization processes, particularly in handling complex and multimodal problems (Heidari et al. 2019).

3.3.5 Runge–kutta optimizer (RUN)

The Runge–Kutta optimizer (RUN) is an optimization algorithm based on the popular Runge–Kutta numerical integration method commonly used in solving differential equations. The Runge–Kutta method approximates the solution of a differential equation by iteratively calculating the function values at different points within a given interval. RUN adapts this concept to the optimization domain by treating the hyperparameter optimization as a dynamic system and utilizing numerical integration to find optimal solutions. In RUN, candidate solutions are treated as trajectories in the search space, and the optimization process is viewed as a trajectory optimization problem. By iteratively updating the positions of the trajectories using Runge–Kutta integration, RUN explores and exploits the search space to locate promising regions and converge toward optimal solutions. Using the Runge–Kutta method in optimization allows for efficient search space exploration, particularly in problems with complex dynamics and interdependencies between hyperparameters (Ahmadianfar et al. 2021).

3.4 Configuration of hybrid models

In this study, the SMA, HHO, RUN, PSO, and GA optimization algorithms have been implemented, and RUN_XGB, SMA_ADB, SMA_RF, SMA_GB, HHO_LGB, SMA_DT, SMA_LSSVC_RBF, SMA_ANN_Adam, PSO_SVC_RBF, HHO_SVC_L, SMA_SVC_P, SMA_LSSVC_P, SMA_LSSVC_L, SMA_ANN_IBFGS, and SMA_ANN_SGD hybrid models have been employed. These hybrid models have been configured with the following hyperparameters, as mentioned in Table 6.

3.5 Evaluation metrics

A confusion matrix is the most important tool for evaluating the performance of a classification model. It comprehensively summarizes the model's performance by comparing predicted and actual labels. True positives, true negatives, false positives, and false negatives are used to calculate other evaluation metrics, such as precision, recall, and F1 score. Table 7 shows a typical confusion matrix for a binary classification problem (Hastie et al. 2009).

Precision is the proportion of true positives among all the positive predictions made by the model. It measures the model's ability to identify positive examples correctly. High precision means that the model makes fewer false positive

Table 6 Configuration of hybrid models

Hybrid Models	Hyperparameters
RUN-XGB	Learning rate = 0.57967808; Estimators = 153; Max depth = 2
SMA-ADB	Estimators = 185; Learning rate = 0.76809871
SMA-GB	Learning rate = 0.62206498; Estimators = 30; Max depth = 4
SMA-RF	Max depth = 9; Estimators = 27
HHO-LGB	Learning rate = 0.32571603; Estimators = 69
SMA-DT	Max depth = 2
SMA-LSSVC-RBF	Sigma = 0.1; Gamma = 0.64497407
SMA-ANN-Adam	Learning rate = 0.19395991; Hidden layer sizes = (3,1,1)
PSO-SVC-RBF	Gamma = 3.53716106e-02; C = 7.06565338e + 02
HHO-SVC-L	Gamma = 4.85396728e-01; C = 5.23750445e + 02
SMA-SVC-P	Gamma = 0.59147854; C = 121.92240537
SMA-LSSVC-P	Sigma = 3.14099178e + 03; Gamma = 9.06860154e-01
SMA-LSSVC-L	Sigma = 3.98097589e + 03; Gamma = 9.86913073e-01
SMA-ANN-lbfgs	Learning rate = 0.5936825; Hidden layer sizes = (5,3,3)
SMA-ANN- SGD	Learning rate = 0.01; Hidden layer sizes = (3,1,1)

Table 7 Typical confusion matrix for binary classification

		Actual Classes	
		Positive (1)	Negative (0)
Predicted Classes	Positive (1)	True Positive (TP)	False Positive (FP)
	Negative (0)	False Negative (FN)	True Negative (TN)

predictions. The mathematical expression for precision is (Ahmad et al. 2021).

$$precision = \frac{TP}{TP + FP} \tag{2}$$

The recall is the proportion of true positives among all the actual positive examples in the data, as mentioned in Eq. 3. It measures the model's ability to correctly identify all the positive examples. The high recall means that the model makes fewer false negative predictions.

$$recall = \frac{TP}{TP + FN} \tag{3}$$

F1 score is the harmonic mean of precision and recall, as mentioned in Eq. 4, and it provides a single measure of the model's overall performance. It is a useful metric when the number of positive and negative examples in the data is imbalanced. A high F1 score indicates that the model has a good balance of precision and recall.

$$F1 = \frac{2 * TP}{2 * TP + FP + FN} \tag{4}$$

The other classification measure implemented in this study is the area under the curve (AUC). The AUC is a metric that measures the model's ability to distinguish between positive and negative classes. It is computed as the area under the receiver operating characteristic (ROC) curve, which plots the true positive rate against the false positive rate for different classification thresholds. A model with an AUC of 1.0 has perfect discrimination between the two classes, while a model with an AUC of 0.5 has no discrimination ability and is essentially random. The accuracy is also implemented to evaluate the model performance. The accuracy measures the proportion of correctly classified instances out of the total number of instances, as mentioned in Eq. 5. It is a simple and intuitive performance measure, but it can be misleading in cases where the class distribution is imbalanced. Evaluating the classification performance based on other metrics is also important.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

Some more performance metrics have been used to measure the performance of the models in the training and testing phases (Kumar et al. 2022a).

$$TNR = \frac{TN}{TN + FP} \tag{6}$$

$$NPV = \frac{TN}{TN + FN} \tag{7}$$

$$FPR = \frac{FP}{FP + FN} \tag{8}$$

$$FNR = \frac{FN}{TP + FN} \tag{9}$$

$$FDR = \frac{FP}{FP + TP} \tag{10}$$

$$FOR = \frac{FN}{FN + TN} \tag{11}$$

$$ACA = 1 - \frac{FPR + FNR}{2} \tag{12}$$

$$MCC = \frac{TP * TN - FP * FN}{(TP + TN)(TP + FN)(TN + FP)(TN + FN)} \quad (13)$$

where TN is the true negative, TP is the true positive, FP is the false positive, FN is the false negative, TNP is the true negative rate, NPV is the negative predictive value, FPR is the false positive rate, FNR is the false negative rate, FDR is the false discovery rate, FOR is the false omission rate, ACA is the average classification accuracy, and MCC is the Matthews correlation coefficient.

4 Results and discussion

In the present research, seventy-seven computing models (2 baseline, 30 tree-based, 30 SVC-based, and 15 NN-based) have been developed to determine the optimum performance model for assessing the soil liquefaction potential. The training and validation phase results have been discussed in the following sections and summarized in Table 8.

4.1 Simulation of soft computing models

4.1.1 Baseline models

The training and testing phase results of logistic regression and kNN models have been presented in terms of the AUC curve in (Fig. 5a, b). Figure 5 reveals that the logistic regression model has attained 0.51 and 0.50 AUC in the training and validation phases. Conversely, the kNN model has attained better AUC than logistic regression models in both phases, i.e., 0.90.

Tree-based models Five optimization algorithms have been implemented for hyperparameters tuning of tree-based models. In this study, DT, RF, GB, ADB, XGB, and LGB tree-based soft computing approaches have been employed and optimized by each SMA, RUN, PSO, HHO, and GA algorithm. Figure 6 Illustrates the validation accuracy matrix for the optimized tree-based models.

Figure 6 presents that models SMA_DT, RUN_DT, PSO_DT, HHO_DT, and GA_DT obtained equal validation accuracy, i.e., 0.92 in assessing soil liquefaction potential. Furthermore, models SMA_RF, RUN_RF, PSO_RF, HHO_RF, and GA_RF achieved the validation performance of 0.96, 0.94, 0.94, 0.88, and 0.88, respectively. It is found that the SMA_RF model outperformed the SMA_DT, RUN_DT, PSO_DT, HHO_DT, and GA_DT as well as RUN_RF, PSO_RF, HHO_RF, and GA_RF models. In addition, the comparison of SMA_GB, RUN_GB, PSO_GB, HHO_GB, and GA_GB models showed that model SMA_GB (= 0.96) assessed the liquefaction potential better than RUN_GB (= 0.94), PSO_GB (= 0.94), HHO_GB (= 0.94), and GA_GB (= 0.92) models in the validation phase. Conversely, the comparison of SMA_XGB, RUN_XGB, PSO_XGB, HHO_XGB, and GA_XGB models showed that the SMA_XGB (= 0.98) model can predict the liquefaction potential of soil than RUN_XGB (= 0.96), PSO_XGB (= 0.96), HHO_XGB (= 0.94), and GA_XGB (= 0.92) models. The SMA, RUN, PSO, HHO, and GA-optimized LGB models didn't perform well in predicting soil liquefaction potential. Model SMA_LGB performed better than RUN_LGB (= 0.88), PSO_LGB (= 0.88), HHO_LGB (= 0.86), and GA_LGB (= 0.86) models with an accuracy of 0.90 only. Thus, models SMA_DT, SMA_RF, and SMA_GB attained higher validation accuracy because the SMA algorithm has a high exploration and exploitation ability. In the case of the ADB and LGB soft computing approaches, models HHO_ADB and HHO_LGB have higher accuracy because the HHO algorithm consists of fast convergence and strong local search capability. Interestingly, it has been observed that model XGB has outperformed because of the stability of the RUN optimization algorithm. Based on the validation accuracy comparison, models SMA_DT, SMA_RF, SMA_GB, HHO_ADB, RUN_XGB, and HHO_LGB have been identified as the best architectural model and compared with stacking and voting soft computing approaches. Figure 7(a) and (b) compares training and validation accuracies for the best architectural tree model, stacking, and voting soft computing models in terms of the AUC curve. Figure 7 reveals that the RUN-XGB model has an AUC value of 0.98, close to the voting model in the validation phase.

SVC-based models Support vector classifier models have been employed to assess the liquefaction potential using SVC and LSSVC approaches. Linear, polynomial, and radial basis function kernels have been implemented to employ the SVC and LSSVC models. Thus, SVC_L, SVC_P, SVC_RBF, LSSVC_L, LSSVC_P, and LSSVC_RBF models have been employed and optimized by each SMA, RUN, PSO, HHO, and GA algorithm. Figure 8 demonstrates the validation accuracy matric for the SVC-optimized models.

Figure 8 demonstrates that the HHO-optimized linear kernel function-based SVC model has outperformed with 0.82 accuracy in the validation phase. The performance comparison of linear kernel function-based LSSVC models shows that the optimization algorithms have not impacted accuracy in the validation phase. Thus, all optimized linear kernel function-based LSSVC models have attained 0.55 validation accuracy, comparatively less than the HHO-optimized linear kernel function-based SVC model. Also, Fig. 8 reveals that the SMA-optimized polynomial kernel function-base SVC model has obtained higher accuracy in

Table 8 Performance results of developed models for liquefaction and non-liquefaction cases

Model	Phase	Accuracy	AUC	Liquefaction			Non-Liquefaction			Rank
				Precision	Recall	F1 score	Precision	Recall	F1 score	
RUN_XGB	Training	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	1
	Validation	**0.98**	**0.98**	**0.96**	**1**	**0.98**	**1**	**0.96**	**0.98**	
	all	**0.99**	**0.99**	**0.99**	**1**	**1**	**1**	**0.98**	**0.99**	
ET_Voting	Training	1	1	1	1	1	1	1	1	2
	Validation	0.98	0.98	0.96	1	0.98	1	0.96	0.98	
	all	0.99	0.99	0.99	1	1	1	0.98	0.99	
SMA_ADB	Training	1	1	1	1	1	1	1	1	3
	Validation	0.96	0.96	1	0.92	0.96	0.93	1	0.96	
	All	0.988	0.99	1	0.98	0.99	0.97	1	0.99	
SMA_GB	Training	1	1	1	1	1	1	1	1	4
	Validation	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	
	All	0.98	0.98	0.99	0.99	0.99	0.98	0.98	0.98	
SMA_RF	Training	0.99	0.99	1	0.99	0.99	0.98	1	0.99	5
	Validation	0.96	0.96	0.93	1	0.96	1	0.92	0.96	
	all	0.98	0.98	0.98	0.99	0.99	0.98	0.97	0.98	
ST_Stacking	training	1	1	1	1	1	1	1	1	5
	Validation	0.94	0.94	0.9	1	0.95	1	0.88	0.94	
	All	0.98	0.977	0.97	1	0.99	1	0.95	0.98	
HHO_LGB	Training	1	1	1	1	1	1	1	1	5
	Validation	0.9	0.9	0.86	0.96	0.91	0.95	0.84	0.89	
	All	0.97	0.96	0.96	0.99	0.98	0.98	0.94	0.96	
KNN	Training	0.92	0.89	0.91	0.99	0.94	0.97	0.8	0.88	6
	Validation	0.88	0.88	0.81	1	0.9	1	0.76	0.86	
	All	0.91	0.89	0.88	0.99	0.93	0.98	0.79	0.87	
SMA_DT	Training	0.86	0.86	0.93	0.85	0.89	0.75	0.88	0.81	7
	Validation	0.92	0.92	0.89	0.96	0.93	0.96	0.88	0.92	
	All	0.87	0.87	0.92	0.88	0.9	0.82	0.88	0.85	
SMA_LSSVC_RBF	Training	0.87	0.845	0.88	0.94	0.91	0.86	0.76	0.81	8
	Validation	0.86	0.86	0.81	0.96	0.88	0.95	0.76	0.84	
	All	0.87	0.85	0.86	0.94	0.9	0.89	0.76	0.82	
SMA_ANN_Adam	Training	0.8	0.84	0.97	0.73	0.83	0.65	0.95	0.77	9
	Validation	0.88	0.88	0.92	0.85	0.88	0.85	0.92	0.88	
	All	0.83	0.85	0.95	0.76	0.84	0.71	0.94	0.81	
PSO_SVC_RBF	Training	0.85	0.81	0.84	0.96	0.9	0.9	0.66	0.76	10
	Validation	0.78	0.78	0.7	1	0.83	1	0.56	0.72	
	All	0.835	0.79	0.8	0.97	0.88	0.93	0.62	0.75	
HHO_SVC_L	Training	0.83	0.8	0.85	0.9	0.88	0.78	0.71	0.74	11
	Validation	0.82	0.82	0.77	0.92	0.84	0.9	0.72	0.8	
	All	0.83	0.8	0.83	0.9	0.87	0.82	0.71	0.76	
SMA_SVC_P	Training	0.85	0.81	0.84	0.96	0.9	0.9	0.66	0.76	12
	Validation	0.76	0.76	0.69	0.96	0.81	0.93	0.56	0.7	

Table 8 (continued)

Model	Phase	Accuracy	AUC	Liquefaction			Non-Liquefaction			Rank
				Precision	Recall	F1 score	Precision	Recall	F1 score	
	All	0.83	0.79	0.8	0.96	0.87	0.91	0.62	0.74	
SMA_LSSVC_P	Training	0.83	0.75	0.8	1	0.89	1	0.51	0.68	13
	Validation	0.72	0.72	0.65	1	0.79	1	0.44	0.61	
	All	0.8	0.74	0.75	1	0.86	1	0.48	0.65	
SMA_LSSVC_L	training	0.74	0.63	0.72	1	0.84	1	0.27	0.42	14
	Validation	0.54	0.54	0.53	1	0.69	1	0.08	0.15	
	All	0.68	0.59	0.66	1	0.8	1	0.2	0.33	
SMA_ANN_lbfgs	Training	0.65	0.5	0.66	1	0.79	0	0	0	15
	Validation	0.5	0.5	0.51	1	0.68	0	0	0	
	All	0.61	0.5	0.61	1	0.76	0	0	0	
Logistic regression	Training	0.65	0.5	0.66	1	0.79	0	0	0	16
	Validation	0.51	0.5	0.51	1	0.68	0	0	0	
	all	0.61	0.5	0.61	1	0.76	0	0	0	
SMA_ANN_SGD	Training	0.65	0.5	0.66	1	0.79	0	0	0	17
	Validation	0.5	0.5	0.51	1	0.68	0	0	0	
	All	0.61	0.5	0.61	1	0.76	0	0	0	

Bold values correspond to an optimum performance model

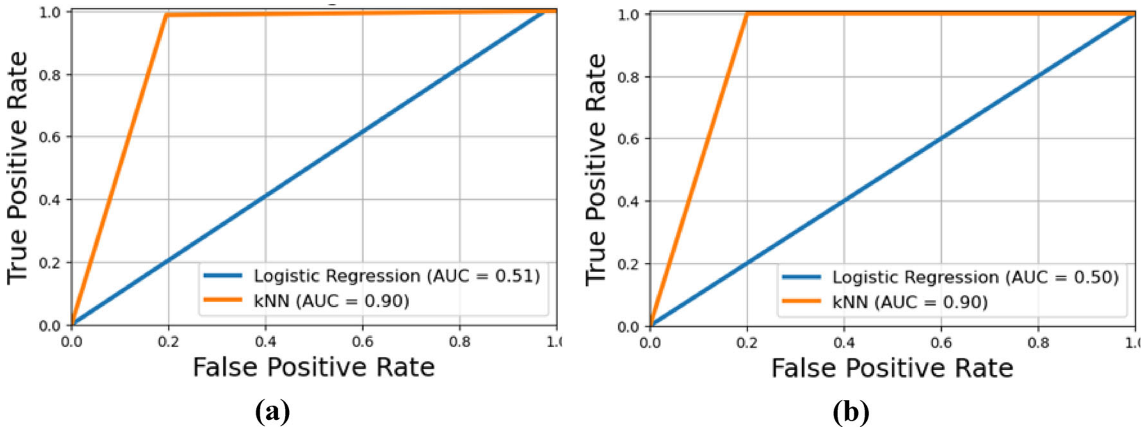


Fig. 5 Illustration of AUC curve for baseline models in **a** training and **b** validation phase

the validation phase, i.e., 0.76. Moreover, Fig. 6 presents that all optimized polynomial kernel function-based LSSVC models have gained 0.73 validation accuracy. Again, the optimized SVC models, HHO_SVC_L and SMA_SVC_P, have outperformed the optimized LSSVC models in assessing the liquefaction potential. Figure 8 demonstrates that the PSO-optimized RBF kernel function-based SVC model has achieved 0.78 validation accuracy, which is comparatively higher than other optimized RBF kernel function-based SVC models. Interestingly, it has been observed that the impact of the optimization algorithm has been determined for LSSVC models after implementing the RBF kernel function. Conversely, Fig. 8 reveals that the SMA-optimized RBF kernel function-based LSSVC model has obtained 0.86 validation accuracy, followed by RUN_LSSVC_RBF, PSO_LSSVC_RBF, HHO_LSSVC_RBF. Model GA_LSSVC_RBF has attained the least validation accuracy, i.e., 0.51. Thus, models HHO_SVC_L, SMA_SVC_P, PSO_SVC_RBF, SMA_LSSVC_L, SMA_LSSVC_P, and SMA_LSSVC_RBF have been recognized as the best architectural model in predicting the liquefaction potential of soil. Figure 9(a) and (b) demonstrates the AUC curve

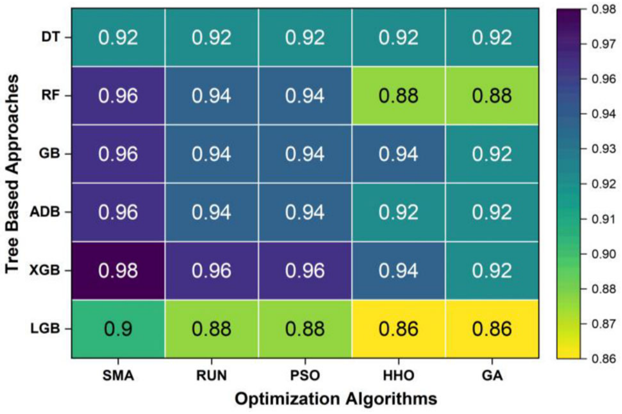


Fig. 6 Illustration of the accuracy of tree-based optimized models

property for the best architectural SVC-based models in the training and validation phases.

Figure 9(a) and (b) illustrates that model SMA_LSSVC_RBF has attained 85% (training = 0.85) and 86% (validation = 0.86) AUC values in training and validation phase, respectively, comparatively higher than the other best architectural SVC-based models. The results show that the RBF kernel function is better if the database has complex and nonlinear patterns or clusters. The radial basis kernel function, consisting of one hidden layer, helps generalize the targets. The comparison based on kernel functions presents that RBF kernel function-based SVC and LSSVC models have higher AUC values than linear and polynomial kernel-based SVC and LSSVC models.

Neural network-based models The artificial neural network models have been employed with Adam, sgd, and lbfgs algorithms to predict the liquefaction potential. Each SMA, RUN, GA, HHO, and PSO algorithm has optimized each ANN_adam, ANN_sgd, and ANN_lbfgs model. The validation accuracy of the employed models has been drawn, as shown in Fig. 10.

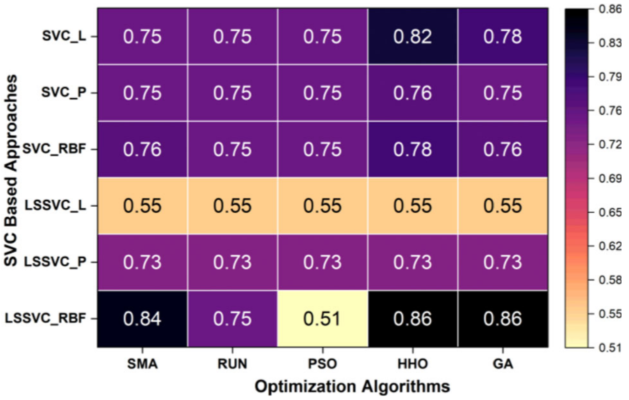
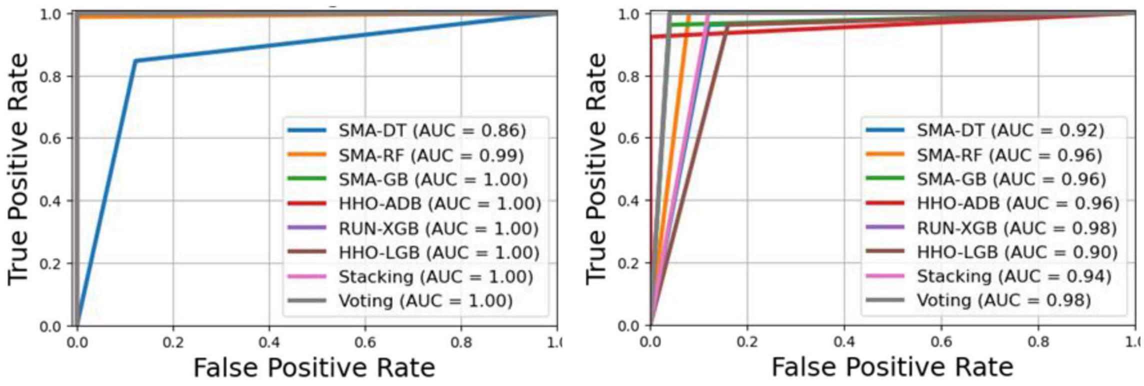


Fig. 8 illustration of validation accuracy matric for the optimized SVC-based models

From the study of Fig. 10, the following Observations have been mapped: (i) model SMA_ANN_adam has gained higher validation accuracy, i.e., 0.88, than other optimized ANN_adam models; and (ii) all optimized ANN_sgd and ANN_lbfgs models have attained equal accuracies in the validation phase. Finally, models SMA_ANN_adam, SMA_ANN_sgd, and SMA_ANN_lbfgs have been determined as the best architectural models for predicting liquefaction potential. The area under the curve (AUC) has been calculated for the best architectural ANN-based models, as shown in Fig. 11. Figure 11(a) and (b) illustrate that model SMA_ANN_adam has gained higher AUC in training (= 0.84) and validation (= 0.88) than SMA_ANN_sgd and SMA_ANN_lbfgs models.



(a) **(b)**

Fig. 7 Illustration of AUC curve for tree-based models in **a** training and **b** validation phase

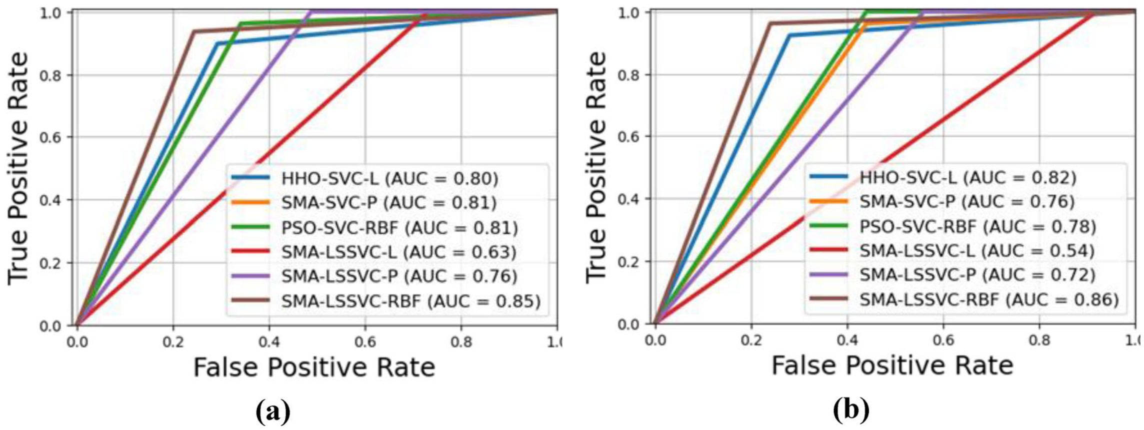


Fig. 9 Illustration of AUC curve for SVC-based models in **a** training and **b** validation phase

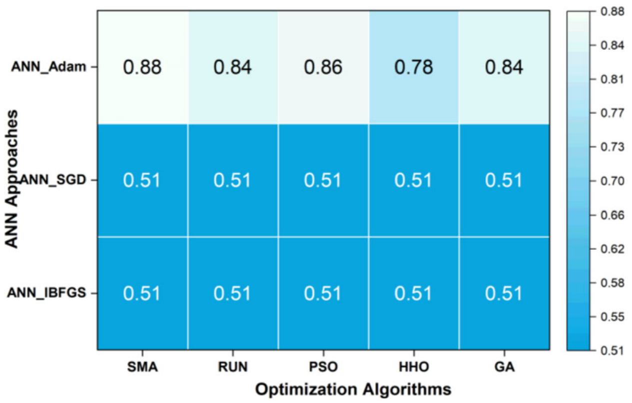


Fig. 10 Illustration of validation accuracies of ANN models

4.2 Discussion on results

The overall comparison demonstrates that models kNN, SMA_DT, SMA_RF, SMA_GB, HHO_ADB, RUN_XGB, HHO_LGB, stacking, voting, HHO_SVC_L, SMA_SVC_P, PSO_SVC_RBF, SMA_LSSVC_L, SMA_LSSVC_P, SMA_LSSVC_RBF, SMA_ANN_Adam, SMA_ANN_sgd, and SMA_ANN_lbfgs have obtained higher AUC, i.e., 0.5, 0.9, 0.92, 0.96, 0.96, 0.96, 0.98, 0.90, 0.94, 0.98, 0.82 0.76, 0.78, 0.54, 0.72, 0.86, 0.88, 0.5, 0.5. Finally, models kNN, RUN_XGB, SMA_LSSVC_RBF, and PSO_ELM have been identified as the best architectural models for predicting liquefaction potential. However, the AUC has been plotted to summarize the receiver operating characteristic (ROC) curve. The ROC curve has five degrees of ratings: not discriminating (0.5–0.6), poor (0.6–0.7), fair (0.7–0.8), good (0.8–0.9), and excellent (0.9–1.0) (Bradley 1997). Table 8 compares the accuracies of all models developed in the present research. Table 8 shows that (a) model kNN has attained overall accuracy of 91% and an AUC of 0.89, presenting a good degree of rating; (b) model kNN has gained 0.88 and 0.98 precisions in the training and validation phase; (c) model RUN-XGB has achieved accuracy, and AUC of 0.99; (d) the precision (liquefaction = 0.99, non-liquefaction = 1.00), recall (liquefaction = 1.00, non-liquefaction = 0.98), and F1 score (liquefaction = 1.00, non-liquefaction = 0.99) of model RUN_XGB demonstrates that model RUN_XGB is highly capable of predicting non-liquefaction and liquefaction cases; (d) AUC value for model RUN_XGB reveals that model has predicted liquefaction potential with an excellent degree; (e) model SMA_LSSVC_RBF has predicted liquefaction potential with higher precision (liquefaction = 0.86, non-liquefaction = 0.89), recall (liquefaction = 0.94, non-liquefaction = 0.76), and F1 score (liquefaction = 0.90, non-liquefaction = 0.82) than other SVC-based soft computing models.; and (f) model SMA_ANN_adam has predicted the liquefaction potential with better precision (liquefaction = 0.95, non-liquefaction = 0.71), recall (liquefaction = 0.76, non-liquefaction = 0.94), and F1 score (liquefaction = 0.84, non-liquefaction = 0.81) than other neural network-based models. To sum up, the RUN_XGB model has been identified as the optimum performance soft computing model for predicting soil liquefaction potential. The XGB approach utilizes more accurate approximations to find the best tree model. The RUN algorithm involves slope calculations at multiple steps at or between the current and following discrete time values. Therefore, model RUN_XGB has predicted liquefaction potential with higher accuracy in the training and validation phases. Furthermore, the prediction accuracy of model RUN_XGB has been cross-validated by comparing the published model by Ahmad et al. (2021), as shown in Table 9.

Table 9 demonstrates that the GB model has also outperformed the published model by Ahmad et al. (2021). Hence, it can be stated that the GB model is the best architectural model for assessing soil liquefaction potential.

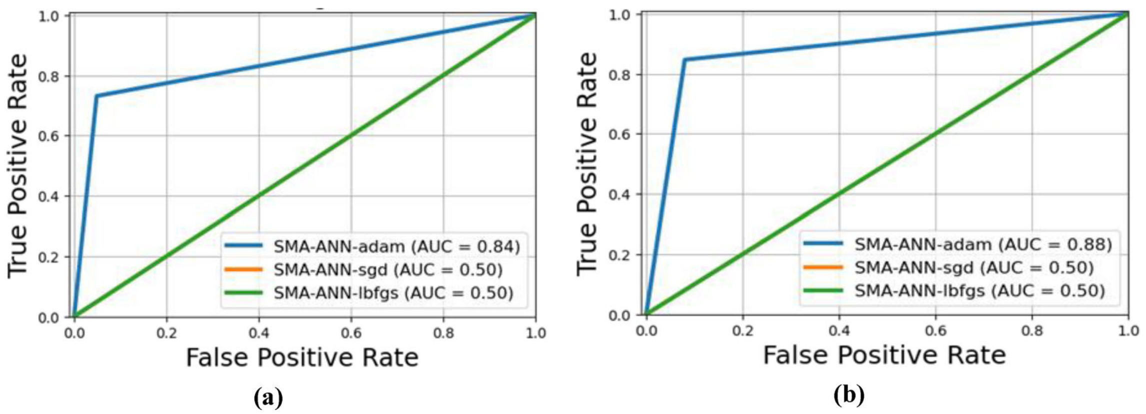


Fig. 11 Illustration of AUC curve for ANN-based models in **a** training and **b** validation phase

Table 9 Details for validation of the GB model

References	Phase	Liquefaction			Non-Liquefaction		
		Precision	Recall	F1	Precision	Recall	F1
Ahmad et al. (2021)	Testing	0.7647	1.00	0.8667	1.00	0.68	0.8095
Present study	Testing	0.9800	0.98	0.9600	1.00	0.98	1

Table 10 Results of performance metrics for developed models

Model	Phase	FPR	FNR	TNR	NPV	FDR	FOR	MCC	ACA
Ideal values	Ideal	0.00	0.00	1.00	1.00	0.00	0.00	− 1 to + 1	1.00
kNN	Training	0.22	0.05	0.78	0.89	0.11	0.11	0.75	0.86
	Testing	0.16	0.00	0.84	1.00	0.13	0.00	0.85	0.92
	All	0.20	0.04	0.80	0.93	0.12	0.07	0.79	0.88
RUN_XGB	**Training**	**0**	**0**	**1.00**	**1.00**	**0**	**0**	**1.00**	**1.00**
	Testing	**0.04**	**0**	**0.96**	**1.00**	**0.03**	**0**	**0.96**	**0.98**
	All	**0.01**	**0**	**0.98**	**1.00**	**0.009**	**0**	**0.98**	**0.99**
SMA_LSSVC_RBF	Training	0.24	0.06	0.75	0.86	0.12	0.13	0.71	0.84
	Testing	0.24	0.038	0.76	0.95	0.19	0.05	0.74	0.86
	All	0.24	0.05	0.75	0.89	0.14	0.1	0.72	0.85
SMA_ANN_Adam	Training	0.048	0.27	0.95	0.65	0.033	0.35	0.65	0.84
	Testing	0.08	0.15	0.92	0.85	0.08	0.15	0.76	0.88
	All	0.06	0.24	0.94	0.71	0.05	0.28	0.68	0.85

*Bold values correspond to the best architectural model

4.3 Analysis of soft computing models

After a thorough analysis of AUC and the accuracies of the best architectural model, the capabilities of baseline, tree-based, SVC-based, and ANN-based models have been analyzed by implementing TNR, NPV, FPR, FNR, FDR, FOR, ACC, and MCC performance metrics. The results of TNR, NPV, FPR, FNR, FDR, FOR, ACA, and MCC metrics for all soft computing models are given in Table 10. The comparative study of baseline, tree-based, SVC-based, and ANN-based models reveals that models kNN, RUN_XGB, SMA_LSSVC_RBF, and SMA_ANN_adam have attained higher AUC values in the training and validation phase. Therefore, the TNR, NPV, FPR, FNR, FDR, FOR, ACA, and MCC performance metrics of models kNN, RUN_XGB, SMA_LSSVC_RBF, and SMA_ANN_adam have been compared and found that model RUN_XGB has attained 0.98 TNR, 1.0 NPV, 0.01 FPR, 0.0 FNR, 0.009 FDR, 0.0 FOR, 0.99 ACA, and 0.98 MCC, comparatively better than other models and close to the ideal values. Hence, the RUN_XGB model has been recognized as an optimum performance model for predicting soil liquefaction potential.

In addition, cross-validation has been performed for the best architectural models, i.e., kNN, RUN_XGB, SMA_LSSVC_RBF, and SMA_ANN_adam, using 10 k-fold. For the cross-validation, the computation cost of the best architectural models, developed by a k-fold value of 5, has been compared with the k-fold value of 10, as given in Table 11.

Table 11 illustrates that model RUN_XGB (based on $k = 5$) has attained an 11.242 s computational cost, higher than other best architectural models. In the case of a 10 k-fold value, the model RUN_XGB obtained 21.997 s computational cost, which is comparatively higher than the 5 k-fold and the best architectural models (based on $k = 10$). The comparison of the computational cost of RUN_XGB (based on 5 and 10) cross-validates and introduces the RUN_XGB model as an optimum performance model in predicting the soil liquefaction potential.

5 Summary and conclusions

The present research has been carried out to introduce an optimal performance model for predicting the liquefaction potential using 170 results of cone penetration testing. For that aim, two baselines, thirty tree-based, thirty SVC-based, and fifteen NN-based models, have been employed, trained, tested, and analyzed. Accuracy, AUC, precision, recall, and F1 score metrics have been implemented to measure the prediction capabilities. The following significant outcomes of the research have been drawn:

- Capabilities of Soft Computing Models – The kNN has assessed liquefaction potential better than logistic regression. Conversely, models based on tree, kernel & instance, and artificial neural network approaches attained higher accuracy. Still, models RUN_XGB, SMA_LSSVC_RBF, and SMA_ANN_Adam have outperformed with an accuracy of over 80%. The accuracy of the comparison between conventional and hybrid models reveals that the hybrid model efficiently predicts the liquefaction potential compared to conventional models.
- Optimum Performance Model – Model RUN_XGB has been identified as an optimum performance model with an overall accuracy of 0.99 and AUC of 0.99. Also, model RUN_XGB has higher precision (0.99 for liquefaction, 1 for non-liquefaction), recall (1 for liquefaction, 0.98 for non-liquefaction), and F1 score (1 for liquefaction, 0.99 for non-liquefaction). The analysis of FPR, FNR, TNR, NPV, FDR, FOR, MCC, and ACA metrics presents the robustness of the RUN_XGB model in predicting the liquefaction potential.

Table 11 Comparison of computational cost for cross-validation

Model	Computation cost for $k = 5$ models (in sec)				Computation cost for $k = 10$ models (in sec)			
	Training cost	Validation cost	Optimization algorithm convergence cost	Overall cost	Training Cost	Validation Cost	Optimization algorithm convergence cost	Overall cost
kNN	0.00049	0.00141	–	1.9e-3	0.0008	0.0008	–	1.6e-3
SMA_LSSVC_RBF	0.0018	0.0006	0.55	0.5524	0.0018	0.0004	0.55	0.5522
RUN_XGB	0.0297	0.0028	11.21	**11.242**	0.02955	0.0027	21.965	**21.997**
SMA_ANN_Adam	0.0553	0.0008	6.75	6.8061	0.0533	0.0005	5.705	5.7588

*Bold values correspond to the higher computational cost of an optimum performance model

- Impact of Multicollinearity – The multicollinearity analysis reveals that the input variables, such as M, a_{max}, q_c, D_{50}, σ'_v, have weak multicollinearity (VIF < 2.5). However, the impact of multicollinearity has been observed for conventional soft computing models, i.e., logistic regression. In the presence of multicollinearity, the conventional models estimate the unreliable coefficient with less interpretation capability; these models are less stable and have an overfitting issue. Conversely, the optimized soft computing models have performed better in weak multicollinearity because of the optimization technique.

To sum up, the present study introduces the RUN_XGB model as the optimum performance soft computing model for predicting soil liquefaction potential. The prediction capabilities of RUN_XGB reveal that the RUN_XGB model may be employed for slope stability analysis. The present study may also be extended by creating an artificial database to reduce the multicollinearity of the database used in this study. One of the limitations of the employed models in this work is determining the optimal structure using different analyses. Therefore, it is suggested to optimize the coefficients/weights of the employed models using metaheuristic optimization algorithms, i.e., squirrel search algorithm (SSA), improved squirrel search algorithm (ISSA), grey wolf optimizer (GWO) algorithm, random walk grey wolf optimizer (RW_GWO) algorithm, sailfish optimizer (SAO) algorithm, sandpiper optimization algorithm (SOA). Also, the dimensionality analysis may be carried out by creating different combinations of the significant features to analyze the performance of ML models. This research will help earthquake and geotechnical engineers decide the liquefaction phenomenon during an earthquake.

Author contributions Jitendra Khatti and Yewuhalashet Fissha: Main author, conceptualization, literature review, manuscript preparation, methodological development, statistical analysis, detailing, and overall analysis; Kamaldeep Singh Grover, Hajime Ikeda, Hisatoshi Toriya, Tsuyoshi Adachi, and Youhei Kawamura: Main author, detailing, overall analysis; conceptualization, comprehensive analysis, manuscript finalization, detailed review, and editing.

Funding Not applicable.

Data availability All data, models, and code generated or used during the study appear in the submitted article.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Abbaszadeh Shahri A, Maghsoudi Moud F (2020) Liquefaction potential analysis using hybrid multi-objective intelligence model. Environ Earth Sci 79(19):441. https://doi.org/10.1007/s12665-020-09173-2

Ahmad M, Tang XW, Qiu JN, Ahmad F, Gu WJ (2021) Application of machine learning algorithms for the evaluation of seismic soil liquefaction potential. Front Struct Civ Eng 15:490–505. https://doi.org/10.1007/s11709-020-0669-5

Ahmadianfar I, Heidari AA, Gandomi AH, Chu X, Chen H (2021) RUN beyond the metaphor: an efficient optimization algorithm based on Runge Kutta method. Expert Syst Appl 181:115079. https://doi.org/10.1016/j.eswa.2021.115079

Alizadeh Mansouri M, Dabiri R (2021) Predicting the liquefaction potential of soil layers in Tabriz city via artificial neural network analysis. SN Appl Sci 3:1–31. https://doi.org/10.1007/s42452-021-04704-3

Ansari A, Zahoor F, Rao KS, Jain AK (2022) Deterministic approach for seismic hazard assessment of Jammu Region, Jammu and Kashmir. In Geo-Congress 2022 (pp. 590–598)

Baltzopoulos G, Baraschino R, Chioccarelli E, Cito P, Vitale A, Iervolino I (2023) Near-source ground motion in the M7.8 Gaziantep (Turkey) earthquake. Earthq Eng Struct Dyn. https://doi.org/10.1002/eqe.3939

Bradley AP (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recogn 30(7):1145–1159. https://doi.org/10.1016/S0031-3203(96)00142-2

Cavaleri L, Chatzarakis GE, Di Trapani F, Douvika MG, Roinos K, Vaxevanidis NM, Asteris PG (2017) Modeling of surface roughness in electro-discharge machining using artificial neural networks. Adv Mater Res 6(2):169. https://doi.org/10.12989/amr.2017.6.2.169

Chan JYL, Leow SMH, Bea KT, Cheng WK, Phoong SW, Hong ZW, Chen YL (2022) Mitigating the multicollinearity problem and its machine learning approach: a review. Mathematics 10(8):1283. https://doi.org/10.3390/math10081283

Chen Z, Li H, Goh ATC, Wu C, Zhang W (2020) Soil liquefaction assessment using soft computing approaches based on capacity energy concept. Geosciences 10(9):330. https://doi.org/10.3390/geosciences10090330

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). https://doi.org/10.1145/2939672.2939785

Christensen R (1996) Analysis of variance, design, and regression: applied statistical methods. CRC Press

Clerc M (2010) Particle swarm optimization (Vol. 93). John Wiley & Sons

Das BM, Luo Z (2016) Principles of soil dynamics. Cengage Learning

Das SK, Mohanty R, Mohanty M, Mahamaya M (2020) Multi-objective feature selection (MOFS) algorithms for prediction of liquefaction susceptibility of soil based on in situ test methods. Nat Hazards 103:2371–2393. https://doi.org/10.1007/s11069-020-04089-3

Demir S, Şahin EK (2021) Assessment of feature selection for liquefaction prediction based on recursive feature elimination. Avrupa Bilim Ve Teknol Derg 28:290–294. https://doi.org/10.31590/ejosat.998033

Demir S, Sahin EK (2022) Comparison of tree-based machine learning algorithms for predicting liquefaction potential using canonical correlation forest, rotation forest, and random forest based on CPT data. Soil Dyn Earthq Eng 154:107130. https://doi.org/10.1016/j.soildyn.2021.107130

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(4):3173–3190. https://doi.org/10.1007/s00521-022-07856-4

Erzin Y, Ecemis N (2015) The use of neural networks for CPT-based liquefaction screening. Bull Eng Geol 74:103–116. https://doi.org/10.1007/s10064-014-0606-8

Garg A, Tai K (2012) Comparison of regression analysis, artificial neural network and genetic programming in handling the multicollinearity problem. In 2012 proceedings of international conference on modelling, identification and control (pp. 353–358). IEEE

Garini E, Gazetas G (2023) The 2 earthquakes of February 6th 2023 in Turkey & Syria – Second Preliminary Report. NTUA, Greece

Ge Y, Zhang Z, Zhang J, Huang H (2023) Developing region-specific fragility function for predicting probability of liquefaction induced ground failure. Probab Eng Mech 71:103381. https://doi.org/10.1016/j.probengmech.2022.103381

Gelman A (2005) Analysis of variance—why it is more important than ever. Ann Stat 33(1):1–53. https://doi.org/10.1214/009053604000001048

Ghani S, Kumari S (2022) Prediction of liquefaction using reliability based regression analysis. In: Choudhary AK (ed) Advances in geo-science and geo-structures : select Proceedings of GSGS 2020. Springer Singapore, Singapore, pp 11–23

Ghani S, Kumari S, Ahmad S (2022) Prediction of the seismic effect on liquefaction behavior of fine-grained soils using artificial intelligence-based hybridized modeling. Arab J Sci Eng 47(4):5411–5441. https://doi.org/10.1007/s13369-022-06697-6

Ghorbani A, Eslami A (2021) Energy-based model for predicting liquefaction potential of sandy soils using evolutionary polynomial regression method. Comput Geotech 129:103867. https://doi.org/10.1016/j.compgeo.2020.103867

Ghorbani A, Jahanpour R, Hasanzadehshooiili H (2020) Evaluation of liquefaction potential of marine sandy soil with piles considering nonlinear seismic soil–pile interaction; A simple predictive model. Mar Georesour Geotechnol 38(1):1–22. https://doi.org/10.1080/1064119X.2018.1550543

Goh AT (1996) Neural-network modeling of CPT seismic liquefaction data. J Geotech Eng 122(1):70–73. https://doi.org/10.1061/(ASCE)0733-9410(1996)122:1(70)

Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT press

Gunst RF, Webster JT (1975) Regression analysis and problems of multicollinearity. Commun Stat-Theory Methods 4(3):277–292. https://doi.org/10.1080/03610927308827246

Hair J Jr, Wolfnibarger MC, Ortinau DJ, Bush RP (2013) Essentials of marketing. McGraw Hill, New York, USA

Haldar A, Tang WH (1979) Probabilistic evaluation of liquefaction potential. J Geotech Eng Div 105(2):145–163. https://doi.org/10.1061/AJGEB6.0000765

Hastie T, Tibshirani R, Friedman J (2009) Boosting and additive trees. In: Hastie T, Tibshirani R, Friedman J (eds) The elements of statistical learning: data mining, inference, and prediction. Springer, NY, pp 337–387

Heidari AA, Mirjalili S, Faris H, Aljarah I, Mafarja M, Chen H (2019) Harris hawks optimization: Algorithm and applications. Futur Gener Comput Syst 97:849–872. https://doi.org/10.1016/j.future.2019.02.028

Hoang ND, Bui DT (2018) Predicting earthquake-induced soil liquefaction based on a hybridization of kernel Fisher discriminant analysis and a least squares support vector machine: a multi-dataset study. Bull Eng Geol Env 77:191–204. https://doi.org/10.1007/s10064-016-0924-0

Hsu SC, Yang MD, Chen MC, Lin, JY (2017) Artificial neural network of liquefaction evaluation for soils with high fines content. In The 2006 IEEE International Joint Conference on Neural Network Proceedings (pp. 2643–2649). IEEE. https://doi.org/10.1109/IJCNN.2006.247143

Hu J (2021) A new approach for constructing two Bayesian network models for predicting the liquefaction of gravelly soil. Comput Geotech 137:104304. https://doi.org/10.1016/j.compgeo.2021.104304

Hu J, Liu H (2019) Bayesian network models for probabilistic evaluation of earthquake-induced liquefaction based on CPT and Vs databases. Eng Geol 254:76–88. https://doi.org/10.1016/j.enggeo.2019.04.003

Huang J, Asteris PG, Pasha SMK, Mohammed AS, Hasanipanah M (2020) A new auto-tuning model for predicting the rock fragmentation: a cat swarm optimization algorithm. Eng Comput. https://doi.org/10.1007/s00366-020-01207-4

James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning. Springer, New York

Jas K, Dodagoudar GR (2023a) Liquefaction potential assessment of soils using machine learning techniques: a state-of-the-art review from 1994–2021. Int J Geomech 23(7):03123002. https://doi.org/10.1061/IJGNAI.GMENG-7788

Jas K, Dodagoudar GR (2023b) Explainable machine learning model for liquefaction potential assessment of soils using XGBoost-SHAP. Soil Dyn Earthq Eng 165:107662

Johari A, Khodaparast AR, Javadi AA (2019) An analytical approach to probabilistic modeling of liquefaction based on shear wave velocity. Iran J Sci Technol Trans Civ Eng 43:263–275. https://doi.org/10.1007/s40996-018-0163-7

Johnston R, Jones K, Manley D (2018) Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. Qual Quant 52:1957–1976

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu TY (2017) Lightgbm: A highly efficient gradient boosting decision tree. Adv Neural Inform Process Syst 30

Khan S, Sasmal SK, Kumar GS, Behera RN (2021) Assessment of liquefaction potential based on SPT data by using machine learning approach. In: Sitharam TG (ed) Seismic Hazards and Risk Select Proceedings of 7th ICRAGEE 2020. Springer Singapore, Singapore, pp 145–156. https://doi.org/10.1007/978-981-15-9976-7_14

Khatti J, Grover KS (2023a) Prediction of UCS of fine-grained soil based on machine learning part 1: multivariable regression analysis, gaussian process regression, and gene expression programming. Multiscale Multidiscip Model Exp Des. https://doi.org/10.1007/s41939-022-00137-6

Khatti J, Grover KS (2023b) Prediction of compaction parameters for fine-grained soil: Critical comparison of the deep learning and standalone models. J Rock Mech Geotech Eng. https://doi.org/10.1016/j.jrmge.2022.12.034

Khatti J, Grover KS (2021) relationship between index properties and CBR of soil and prediction of CBR. In: Muthukkumaran K et al (eds) Indian Geotechnical Conference. Springer, Singapore, pp 171–185

Kim HY (2017) Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. Restor Dent Endod 42(2):152–155. https://doi.org/10.5395/rde.2017.42.2.152

Kim HS, Kim M, Baise LG, Kim B (2021) Local and regional evaluation of liquefaction potential index and liquefaction severity number for liquefaction-induced sand boils in Pohang, South Korea. Soil Dyn Earthq Eng 141:106459. https://doi.org/10.1016/j.soildyn.2020.106459

Kingma D, Ba J (2015) Adam: A method for stochastic optimization in: Proceedings of the 3rd international conference for learning representations (iclr'15). San Diego, 500

Kumar D, Samui P, Kim D, Singh A (2021) A novel methodology to classify soil liquefaction using deep learning. Geotech Geol Eng 39:1049–1058. https://doi.org/10.1007/s10706-020-01544-7

Kumar DR, Samui P, Burman A (2022a) Determination of best criteria for evaluation of liquefaction potential of soil. Transp Infrastruct Geotechnol. https://doi.org/10.1007/s40515-022-00268-w

Kumar DR, Samui P, Burman A (2022b) Prediction of probability of liquefaction using hybrid ANN with optimization techniques. Arab J Geosci 15(20):1587. https://doi.org/10.1007/s12517-022-10855-3

Kumar DR, Samui P, Burman A (2022c) Prediction of probability of liquefaction using soft computing techniques. J Inst Eng (india) Series A 103(4):1195–1208. https://doi.org/10.1007/s40030-022-00683-9

Kumar DR, Samui P, Burman A, Wipulanusat W, Keawsawasvong S (2023a) Liquefaction susceptibility using machine learning based on SPT data. Intell Syst Appl 20:200281. https://doi.org/10.1016/j.iswa.2023.200281

Kumar DR, Samui P, Burman A, Kumar S (2023b) Seismically induced liquefaction potential assessment by different artificial intelligence procedures. Transp Infrastruct Geotechnol. https://doi.org/10.1007/s40515-023-00327-w

Kumar DR, Samui P, Burman A (2023c) Suitability assessment of the best liquefaction analysis procedure based on SPT data. Multiscale Multidiscip Model Exp Des 6(2):319–329. https://doi.org/10.1007/s41939-023-00148-x

Kurnaz TF, Kaya Y (2019a) A novel ensemble model based on GMDH-type neural network for the prediction of CPT-based soil liquefaction. Environmental Earth Sciences 78(11):339. https://doi.org/10.1007/s12665-019-8344-7

Kurnaz TF, Kaya Y (2019b) SPT-based liquefaction assessment with a novel ensemble model based on GMDH-type neural network. Arab J Geosci 12:1–14. https://doi.org/10.1007/s12517-019-4640-5

Kurnaz TF, Erden C, Kökçam AH, Dağdeviren U, Demir AS (2023) A hyper parameterized artificial neural network approach for prediction of the factor of safety against liquefaction. Eng Geol 319:107109. https://doi.org/10.1016/j.enggeo.2023.107109

Kutanaei SS, Choobbasti AJ (2019) Prediction of liquefaction potential of sandy soil around a submarine pipeline under earthquake loading. J Pipeline Syst Eng Pract 10(2):04019002. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000349

Larson MG (2008) Analysis of variance. Circulation 117(1):115–121. https://doi.org/10.1161/CIRCULATIONAHA.107.654335

Lawley DN (1938) A generalization of Fisher's z test. Biometrika 30(1/2):180–187. https://doi.org/10.2307/2332232

Li S, Chen H, Wang M, Heidari AA, Mirjalili S (2020) Slime mould algorithm: a new method for stochastic optimization. Futur Gener Comput Syst 111:300–323. https://doi.org/10.1016/j.future.2020.03.055

Lin T, Stich SU, Patel KK, Jaggi M (2020) Don't Use Large Mini-Batches, Use Local SGD. s.l., ICLR 2020

Lin CC, Mudholkar GS (1980) A simple test for normality against asymmetric alternatives. Biometrika 67(2):455–461. https://doi.org/10.1093/biomet/67.2.455

Lu S, Koopialipoor M, Asteris PG, Bahri M, Armaghani DJ (2020) A novel feature selection approach based on tree models for evaluating the punching shear capacity of steel fiber-reinforced concrete flat slabs. Materials 13(17):3902. https://doi.org/10.3390/ma13173902

Mahmood A, Tang XW, Qiu JN, Gu WJ, Feezan A (2020) A hybrid approach for evaluating CPT-based seismic soil liquefaction potential using Bayesian belief networks. J Cent South Univ 27(2):500–516. https://doi.org/10.1007/s11771-020-4312-3

Mansfield ER, Helms BP (1982) Detecting multicollinearity. Am Stat 36(3a):158–160. https://doi.org/10.1080/00031305.1982.10482818

Mase LZ, Agustina S, Hardiansyah MF, Supriani F, Tanapalungkorn W, Likitlersuang S (2023) Application of simplified energy concept for liquefaction prediction in Bengkulu City Indonesia. Geotech Geol Eng 41(3):1999–2021. https://doi.org/10.1007/s10706-023-02388-7413(1999),pp.10,2021.1007/s10706-023-02388-7

Menard S (2002) Applied logistic regression analysis (No. 106). Sage.

Mola-Abasi H, Kordtabar B, Kordnaeij A (2018) Liquefaction prediction using CPT data by triangular chart identification. Int J Geotech Eng 12(4):377–382. https://doi.org/10.1080/19386362.2017.1282399

Nejad AS, Güler E, Özturan M (2018) Evaluation of liquefaction potential using random forest method and shear wave velocity results. In 2018 International Conference on Applied Mathematics & Computational Science (ICAMCS. NET) (pp. 23–233). IEEE. https://doi.org/10.1109/ICAMCS.NET46018.2018.00012

Njock PGA, Shen SL, Zhou A, Lyu HM (2020) Evaluation of soil liquefaction using AI technology incorporating a coupled ENN/t-SNE model. Soil Dyn Earthq Eng 130:105988. https://doi.org/10.1016/j.soildyn.2019.105988

Nong Z, Park SS, Jeong SW, Lee DE (2020) Effect of cyclic loading frequency on liquefaction prediction of sand. Appl Sci 10(13):4502. https://doi.org/10.3390/app10134502

Obite CP, Olewuezi NP, Ugwuanyim GU, Bartholomew DC (2020) Multicollinearity effect in regression analysis: a feed forward artificial neural network approach. Asian J Probab Stat 6(1):22–33

Ozsagir M, Erden C, Bol E, Sert S, Özocak A (2022) Machine learning approaches for prediction of fine-grained soils liquefaction. Comput Geotech 152:105014. https://doi.org/10.1016/j.compgeo.2022.105014

Park SS, Ogunjinmi PD, Woo SW, Lee DE (2020) A simple and sustainable prediction method of liquefaction-induced settlement at Pohang using an artificial neural network. Sustainability 12(10):4001. https://doi.org/10.3390/su12104001

Rahbarzare A, Azadi M (2019) Improving prediction of soil liquefaction using hybrid optimization algorithms and a fuzzy support vector machine. Bull Eng Geol Env 78(7):4977–4987. https://doi.org/10.1007/s10064-018-01445-3

Robertson PK, Campanella RG (1985) Liquefaction potential of sands using the CPT. J Geotech Eng 111(3):384–403. https://doi.org/10.1061/(ASCE)0733-9410(1985)111:3(384)

Rokach L (2019) Ensemble learning: pattern classification using ensemble methods. World Sci. https://doi.org/10.1142/9789811201967_0001

Rollins KM, Amoroso S, Milana G, Minarelli L, Vassallo M, Di Giulio G (2020) Gravel liquefaction assessment using the dynamic cone penetration test based on field performance from the 1976 Friuli earthquake. J Geotech Geoenviron Eng 146(6):04020038. https://doi.org/10.1061/(ASCE)GT.1943-5606.0002252

Sabbar AS, Chegenizadeh A, Nikraz H (2019) Prediction of liquefaction susceptibility of clean sandy soils using artificial intelligence techniques. Indian Geotech J 49:58–69. https://doi.org/10.1007/s40098-017-0288-9

Sahin EK, Demir S (2023) Greedy-AutoML: A novel greedy-based stacking ensemble learning framework for assessing soil liquefaction potential. Eng Appl Artif Intell 119:105732. https://doi.org/10.1016/j.engappai.2022.105732

Samui P, Hariharan R (2015) A unified classification model for modeling of seismic liquefaction potential of soil based on CPT. J Adv Res 6(4):587–592. https://doi.org/10.1016/j.jare.2014.02.002

Sawyer SF (2009) Analysis of variance: the fundamental concepts. J Man Manip Ther 17(2):27E-38E. https://doi.org/10.1179/jmt.2009.17.2.27E

Seed HB, Idriss IM (1971) Simplified procedure for evaluating soil liquefaction potential. J the Soil Mech Found Div 97(9):1249–1273. https://doi.org/10.1061/JSFEAQ.0001662

Sharma M, Satyam N, Reddy KR (2021) State of the art review of emerging and biogeotechnical methods for liquefaction mitigation in sands. J Hazard Toxic Radioact Waste 25(1):03120002. https://doi.org/10.1061/(ASCE)HZ.2153-5515.0000557

Shrestha N (2020) Detecting multicollinearity in regression analysis. Am J Appl Math Stat 8(2):39–42

Sun P, Huang D, Du S (2023) Improving soil liquefaction prediction through an extensive database and innovative ground motion characterization: a case study of Port Island liquefied site. Soil Dyn Earthq Eng 165:107696. https://doi.org/10.1016/j.soildyn.2022.107696

Suykens JA, Vandewalle J (1999) Least squares support vector machine classifiers. Neural Process Lett 9:293–300. https://doi.org/10.1023/A:1018628609742

Tang XW, Bai X, Hu JL, Qiu JN (2018) Assessment of liquefaction-induced hazards using Bayesian networks based on standard penetration test data. Nat Hazard 18(5):1451–1468

Toprak S, Holzer TL, Bennett MJ, Tinsley III JC (1999) CPT-and SPT-based probabilistic assessment of liquefaction. In Proc., 7th US–Japan Workshop on Earthquake Resistant Design of Lifeline Facilities and Countermeasures against Liquefaction (pp. 69–86). Buffalo, NY: Multidisciplinary Center for Earthquake Engineering Research

Vittinghoff E, Glidden DV, Shiboski SC, McCulloch,CE (2006) Regression methods in biostatistics: linear, logistic, survival, and repeated measures models. Springer New York, NY. https://doi.org/10.1007/978-1-4614-1353-0

Witten D, James G (2013) An introduction to statistical learning with applications in R. springer publication

Xing DU, Yongfu SU, Yupeng SO, Binghui SO, Xiaolong ZH, Shasha SO, Yue WA (2020) Multilayer perception neural network for assessment and prediction of earthquake-induced sand liquefaction. 工程地质学报 28(6):1425–1432. https://doi.org/10.13544/j.cnki.jeg.2019-321

Xue X, Yang X (2016) Seismic liquefaction potential assessed by support vector machines approaches. Bull Eng Geol Environ 75:153–162. https://doi.org/10.1007/s10064-015-0741-x

Youd TL (2018) Application of MLR procedure for prediction of liquefaction-induced lateral spread displacement. J Geotech Geoenviron Eng 144(6):04018033. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001860

Youd TL, Noble SK (1997) Liquefaction criteria based on statistical and probabilistic analyses. In Proc., NCEER Workshop on Evaluation of Liquefaction Resistance of Soils. State Univ. of New York, Buffalo, NY, pp. 201–205

Yu H, Miao YU, Subhamoy B (2013) Review on liquefaction-induced damages of soils and foundations during 2011 of the Pacific Coast of Tohoku Earthquake (Japan). Chin J Geotech Eng 35(5):834–840

Zhang W, Goh AT (2016) Evaluating seismic liquefaction potential using multivariate adaptive regression splines and logistic regression. Geomech Eng 10(3):269–284. https://doi.org/10.12989/gae.2016.10.3.269

Zhang J, Xiao S, Huang H, Zhou J (2020) Calibrating a standard penetration test based method for region-specific liquefaction potential assessment. Bull Eng Geol Env 79:5185–5204. https://doi.org/10.1007/s10064-020-01815-w

Zhang Y, Qiu J, Zhang Y, Xie Y (2021a) The adoption of a support vector machine optimized by GWO to the prediction of soil liquefaction. Environ Earth Sci 80:1–9. https://doi.org/10.1007/s12665-021-09648-w

Zhang Y, Xie Y, Zhang Y, Qiu J, Wu S (2021b) The adoption of deep neural network (DNN) to the prediction of soil liquefaction based on shear wave velocity. Bull Eng Geol Env 80:5053–5060. https://doi.org/10.1007/s10064-021-02250-1

Zhang YG, Qiu J, Zhang Y, Wei Y (2021c) The adoption of ELM to the prediction of soil liquefaction based on CPT. Nat Hazards 107(1):539–549. https://doi.org/10.1007/s11069-021-04594-z

Zhang X, He B, Sabri MMS, Al-Bahrani M, Ulrikh DV (2022) Soil liquefaction prediction based on Bayesian optimization and support vector machines. Sustainability 14(19):11944. https://doi.org/10.3390/su141911944

Zheng G, Zhang W, Zhang W, Zhou H, Yang P (2021) Neural network and support vector machine models for the prediction of the liquefaction-induced uplift displacement of tunnels. Undergr Space 6(2):126–133. https://doi.org/10.1016/j.undsp.2019.12.002

Zhou J, Li E, Wang M, Chen X, Shi X, Jiang L (2019) Feasibility of stochastic gradient boosting approach for evaluating seismic liquefaction potential based on SPT and CPT case histories. J Perform Constr Facil 33(3):04019024. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001292

Zhou J, Huang S, Wang M, Qiu Y (2021a) Performance evaluation of hybrid GA–SVM and GWO–SVM models to predict earthquake-induced liquefaction potential of soil: a multi-dataset investigation. Eng Comput. https://doi.org/10.1007/s00366-021-01418-3

Zhou J, Huang S, Zhou T, Armaghani DJ, Qiu Y (2022) Employing a genetic algorithm and grey wolf optimizer for optimizing RF models to evaluate soil liquefaction potential. Artif Intell Rev 55(7):5673–5705. https://doi.org/10.1007/s10462-022-10140-5

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