

Physics-Informed Neural Networks for Predicting Ultimate Bearing Capacity of Layered Soils under Complex Loading Conditions

Mahendra Acharya

acharyamahendra.tu.edu.nep2019@gmail.com

Tribhuvan University

Diwash Bhattarai

China University of Mining and Technology

Bhishma Joshi

Tribhuvan University

Research Article

Keywords: Physics-Informed Neural Networks, Bearing Capacity, Layered Soils, Complex Loading, Foundation Engineering, Deep Learning, Computational Geotechnics

Posted Date: June 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9738631/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Computational Geosciences on July 27th, 2026. See the published version at <https://doi.org/10.1007/s10596-026-10470-1>.

Physics-Informed Neural Networks for Predicting Ultimate Bearing Capacity of Layered Soils under Complex Loading Conditions

Mahendra Acharya^{1*}, Diwash Bhattarai², Bhishma Joshi¹

¹Central Department of Geology, Tribhuvan University, Kathmandu, Nepal

²School of Mines, China University of Mining and Technology, Xuzhou, China

*Corresponding Author: acharyamahendra.tu.edu.nep2019@gmail.com

Abstract

Precise prediction of the UBC of shallow foundations subjected to vertical, horizontal, and moment loads in stratified soil is a major concern in geotechnical engineering. Traditional analytical methods may be constrained by several unrealistic assumptions. On the other hand, modern numerical models such as Finite Element Limit Analysis (FELA), although accurate, are computationally expensive to use in practical scenarios. This work presents a new approach, namely PINNs, which incorporates fundamental principles of soil mechanics directly into the machine learning model to predict the UBC rapidly and efficiently. A complete database made up of 10,000 accurate FELA simulations was created for sand-over-clay and clay-over-sand soil strata under V-H-M loading. The designed PINN utilized mechanical equilibrium laws and Mohr-Coulomb failure law in the loss function to ensure mechanical consistency while training. The designed model showed a very good ability in predicting the response of FELA with R^2 equal to 0.992 and MAPE equal to 2.3%, which is a great improvement compared to artificial neural networks (ANN). Furthermore, the designed framework was able to capture complex failure modes such as punch-through failures, thin-layer effects, and switching between punch failure and general shear failure. Notably, the PINN model cut down computation time for predicting one FELA simulation from about 180 seconds to about 0.002 seconds. The PINN framework suggests an effective model that can predict the ultimate bearing capacity of layered soils with precision. The use of physics-constrained data-driven modeling, with incorporation of the physics of the underlying process, makes the model robust, stable, and generalizable for geotechnical design problems with limited data. PINN holds immense promise for future applications of the technique in the areas of geotechnical engineering, real-time design of foundations, probabilistic geotechnical analysis, and reliability-based geotechnical engineering practice.

Keywords: *Physics-Informed Neural Networks, Bearing Capacity, Layered Soils, Complex Loading, Foundation Engineering, Deep Learning, Computational Geotechnics*

1. Introduction

The design of shallow foundations requires a robust evaluation of the ultimate bearing capacity (UBC) to prevent shear failure. The classical solutions by Terzaghi, Meyerhof, and Hansen provided the foundational basis for bearing capacity theory, primarily for homogeneous soils under vertical centric loading. However, natural soil deposits are invariably layered, exhibiting significant spatial variability in strength parameters. Furthermore, offshore wind turbines, bridge piers, and earth-retaining structures are subjected to complex combined loading, including horizontal forces and overturning moments. Terzaghi's foundational geotechnical engineering principles, as presented in Theoretical Soil Mechanics, provide the theoretical basis for advanced analysis of soil behavior and bearing capacity in foundation engineering (Terzaghi, 1943).

Michalowski and Shi (1995) investigated the ultimate bearing capacity of strip footings on layered soil systems and proposed analytical solutions for estimating bearing capacity over two-layer foundations, highlighting the influence of soil stratification on failure mechanisms. The interaction between layered soils and complex loading presents a formidable challenge. The study by Sheng, Sloan, Gens, and Smith (2003) presents a comprehensive finite element formulation for unsaturated soil mechanics, incorporating suction as a strain variable and developing a robust numerical framework for coupled displacement–pore pressure analysis in geotechnical problems. The failure mechanism is no longer confined to a simple Prandtl-type wedge; instead, it may punch through a stiff upper layer, spread laterally within a soft layer, or form a mixed mechanism depending on the strength ratio and layer thickness. Analytical solutions for layered soils (e.g., punch-through models) are often geometry-specific and conservative, while numerical methods such as Finite Element Method (FEM) or Finite Element Limit Analysis (FELA) offer high accuracy but require significant computational resources and specialized expertise, making them unsuitable for iterative design or probabilistic risk assessment. Sloan (1988) developed a lower-bound finite element limit analysis framework combined with linear programming to compute strict lower-bound collapse loads in geotechnical stability problems, ensuring statically admissible stress fields and improved computational efficiency for plasticity-based soil mechanics applications.

Physics-informed machine learning provides a unified framework for integrating governing physical laws with data-driven models, enabling improved accuracy and generalization in scientific machine learning problems (Karniadakis et al., 2021). Physics-Informed Neural Networks (PINNs) have emerged as a robust computational framework for solving complex partial differential equations by embedding governing physical laws directly into the neural network training process through physics-based loss formulations. This approach eliminates the reliance on traditional mesh-based discretization and enables accurate and efficient solutions for both forward and inverse problems in geomechanics. In this context, Yuan et al. (2024) demonstrated the application of PINNs for solving the three-dimensional Terzaghi consolidation equation, highlighting their effectiveness in capturing time-dependent soil consolidation behavior under various boundary and loading conditions. In recent years, machine learning (ML) has emerged as a powerful tool for geotechnical prediction. Surrogate models, particularly Artificial Neural Networks (ANNs), have been used to predict bearing capacity with high speed. However, traditional ANNs operate as "black boxes"; they learn patterns from data but do not inherently understand the physical laws governing soil behavior. Consequently, they often require massive, perfectly labeled datasets to generalize reliably and may produce physically implausible predictions outside the training domain. Recent advances in geotechnical machine learning have demonstrated the effectiveness of Physics-Informed Neural Networks (PINNs) in solving coupled thermo-hydraulic problems by embedding governing partial differential equations directly into the learning process, ensuring physically consistent predictions for complex soil–water–heat interactions (Feng et al., 2025). Shahin, Maier, and Jaksa (2002) developed an artificial neural network (ANN) model to improve the prediction of settlement of shallow foundations on cohesionless soils and demonstrated that the proposed ANN approach outperforms conventional empirical methods in terms of accuracy and reliability. The concept of Physics-Informed Neural Networks (PINNs), which integrates governing physical laws into the training of neural networks to solve forward and inverse problems involving nonlinear partial differential equations, was originally introduced by Raissi et al. (2019). Recent advances in physics-

informed machine learning for geotechnical and geomechanical modelling demonstrate that Physics-Informed Neural Networks (PINNs) can effectively integrate governing physical laws into data-driven frameworks, enabling more accurate and generalizable predictions for complex soil–structure interaction problems (Liu et al., 2025). Recent advances in physics-informed machine learning have demonstrated significant improvements in modeling complex coupled physical processes. In particular, Rahman et al. (2024) proposed a physics-informed neural network (PINN) framework for high-fidelity prediction of early-age cement hydration kinetics, where governing thermochemical processes are embedded directly into the loss function to ensure physically consistent predictions and improved generalization performance.

To address these limitations, we propose a Physics-Informed Neural Network (PINN). Initially popularized in fluid dynamics and solid mechanics, PINNs embed physical laws (partial differential equations, boundary conditions, constitutive laws) into the training process. By minimizing a loss function that combines data-driven error with residual errors of the governing physics, the PINN learns a solution that is not only accurate but also physically consistent. This study presents the first application of a PINN framework to predict the UBC of layered soils under complex loading. The present study aims to develop a comprehensive and high-fidelity dataset based on Finite Element Limit Analysis (FELA) for two-layered soil systems, specifically considering sand-over-clay and clay-over-sand profiles subjected to combined vertical–horizontal–moment (V–H–M) loading conditions. Building upon this dataset, a Physics-Informed Neural Network (PINN) framework is formulated by embedding fundamental geotechnical principles, including equilibrium equations and the Mohr–Coulomb failure criterion, directly into the learning process as physical constraints. The performance of the proposed PINN model is systematically evaluated in terms of predictive accuracy, generalizability, and computational efficiency, and is rigorously compared with conventional Artificial Neural Network (ANN) surrogates as well as established analytical approaches. Ultimately, the study seeks to demonstrate the practical applicability of the developed model as a robust and efficient tool for rapid and reliable foundation design in layered soil conditions.

2. Problem Definition and Methodology

2.1 Problem Statement

Consider a rigid, rough strip footing of width B resting on a two-layer soil profile. Two typical soil configurations are examined. In Case A (sand over clay), the footing rests on an upper granular layer characterized by a friction angle ϕ_1 , unit weight γ_1 , and thickness H , which overlies a cohesive clay layer with undrained shear strength c_{u2} and unit weight γ_2 . In Case B (clay over sand), the upper layer is cohesive, defined by c_{u1} , overlying a granular layer with a friction angle ϕ_2 .

The footing is subjected to a combined loading consisting of a vertical load V , a horizontal load H , and a moment M . To facilitate analysis, the loading is normalized to define a failure envelope in the V-H-M space, capturing the interaction between different load components at ultimate conditions. The primary objective is to predict the ultimate vertical bearing capacity, V_{ult} , corresponding to given ratios of horizontal to vertical load (H/V) and moment to vertical load (M/V) for these layered soil systems.

2.2 Data Generation via Finite Element Limit Analysis (FELA)

To generate a surrogate dataset for training the Physics-Informed Neural Network (PINN), the Finite Element Limit Analysis (FELA) method was employed due to its robustness in directly computing collapse loads for complex layered soils under combined loading, without requiring iterative constitutive updates. The simulations were conducted using OptumG2 software. A plane-strain model was developed that accounts for a comprehensive range of soil and loading parameters. The layer thickness ratio H/B , as varied as 0.25, 0.5, 1.0, 2.0, and ∞ (representing homogeneous conditions). For sand-over-clay configurations, the strength ratio $c_{u2}/(\gamma_1 B)$ was set to 0.2, 0.5, 1.0, and 2.0, with the sand friction angle ϕ_1 equal to 25° , 35° , and 45° . For clay-over-sand configurations, the strength ratio $c_{u1}/(\gamma_2 B)$ was taken as 0.5, 1.0, and 2.0, with the friction angle of the sand layer ϕ_2 set to 30° and 40° . Loading conditions were defined in terms of a horizontal load factor $\alpha = H/V = 0.0, 0.1, 0.2, 0.3$, and a moment factor $\beta = M/(V \cdot B) = 0.0, 0.1, 0.2$. A total of 10,000 FELA simulations were performed to form the training dataset. Each simulation provided the normalized ultimate bearing capacity, expressed either as $V_{ult}/(c_u B)$ or $V_{ult}/(0.5\gamma B N_\gamma)$. The key parameter ranges used in the simulations are summarized in Table 1.

Table 1: Parameter ranges used in the simulations.

Parameter	Values (Sand-over-Clay)	Values (Clay-over-Sand)
Layer thickness ratio (H/B)	0.25, 0.5, 1.0, 2.0, ∞	0.25, 0.5, 1.0, 2.0, ∞
Strength ratio	$c_{u2}/(\gamma_1 B) = 0.2, 0.5, 1.0, 2.0$	$c_{u1}/(\gamma_2 B) = 0.5, 1.0, 2.0$
Friction angle	$\phi_1 = 25^\circ, 35^\circ, 45^\circ$	$\phi_2 = 30^\circ, 40^\circ$
Horizontal load factor α	0.0, 0.1, 0.2, 0.3	0.0, 0.1, 0.2, 0.3
Moment factor β	0.0, 0.1, 0.2	0.0, 0.1, 0.2

2.3 Physics-Informed Neural Network Formulation

The proposed Physics-Informed Neural Network (PINN) framework is formulated to predict the ultimate bearing capacity of shallow foundations resting on layered soils subjected to combined vertical (V), horizontal (H), and moment (M) loading. Unlike conventional data-driven neural networks that solely minimize prediction error, the PINN incorporates the governing equations of continuum mechanics directly into the optimization process to ensure physically admissible stress states (Figure 1).

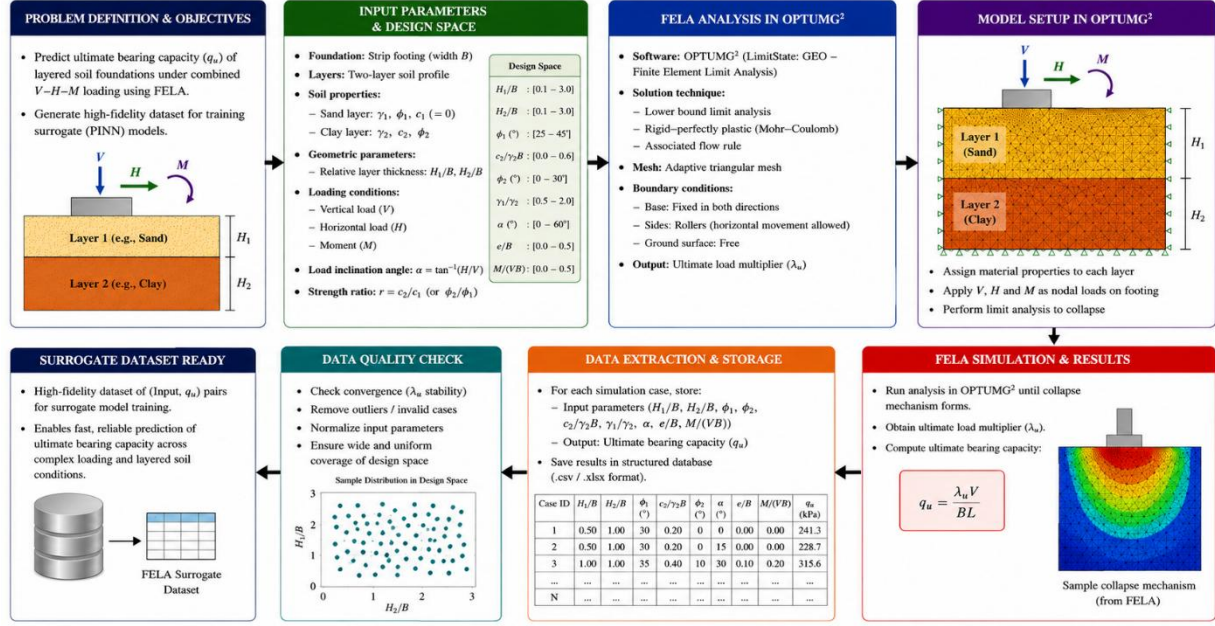


Figure 1: Workflow for surrogate data generation using Finite Element Limit Analysis (FELA), showing the use of OPTUMG² software to simulate the ultimate bearing capacity of layered soil foundations under combined vertical, horizontal, and moment loading.

2.3.1 Governing Equations of Equilibrium

The stress field within the soil domain must satisfy static equilibrium conditions. Under plane strain conditions, the equilibrium equations are expressed as:

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} = 0 \dots \dots \dots (1)$$

$$\frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_y}{\partial y} + \gamma = 0 \dots \dots \dots (2)$$

where,

σ_x = normal stress in the horizontal direction,

σ_y = normal stress in the vertical direction,

τ_{xy} = shear stress component,

γ = unit weight of soil,

(x, y) = spatial coordinates within the soil domain.

These governing PDEs ensure that the predicted stress distribution remains statically admissible throughout the computational domain.

2.3.2 Stress State Variables and Neural Network Representation

The PINN approximates the nonlinear mapping between the spatial coordinates, loading conditions, and soil parameters using a fully connected deep neural network:

$$\mathcal{N}_\theta: (x, y, \eta, \phi, c, V, H, M) \rightarrow (\sigma_x, \sigma_y, \tau_{xy}, q_u) \dots \dots \dots (3)$$

where,

- θ denotes the trainable network parameters,
- $\eta = H/B$ represents normalized layer thickness,
- ϕ denotes the friction angle for granular layers,
- c denotes soil cohesion or undrained shear strength,
- q_u denotes the predicted ultimate bearing capacity.

The stress variables $(\sigma_x, \sigma_y, \tau_{xy})$ are treated as continuous differentiable field quantities, enabling the direct evaluation of spatial derivatives using automatic differentiation.

2.3.3 Mohr–Coulomb Yield Criterion

To ensure consistency with classical soil plasticity theory, the PINN incorporates the Mohr–Coulomb failure criterion into the loss function. The yield function is expressed as:

$$F(\sigma) = \tau_f - c - \sigma_n \tan \phi \leq 0 \dots \dots \dots (4)$$

where,

- τ_f is the mobilized shear stress,
- σ_n is the normal stress acting on the failure plane,
- c is soil cohesion,
- ϕ is the internal friction angle

For implementation within the PINN framework, the principal stress representation is adopted:

$$F(\sigma_1, \sigma_3) = \sigma_1 - \sigma_3 \frac{1+\sin \phi}{1-\sin \phi} - \frac{2c \cos \phi}{1-\sin \phi} \dots \dots \dots (5)$$

where,

- σ_1 and σ_3 are the major and minor principal stresses, respectively.

The yield residual penalizes physically inadmissible stress states violating the Mohr–Coulomb criterion.

2.3.4 Collocation Points and Spatial Enforcement of Physics

The governing PDE residuals are enforced throughout the spatial domain using randomly distributed collocation points. Let:

$$\Omega_c = \{(x_i, y_i)\}_{i=1}^{N_c} \dots \dots \dots (6)$$

represent the set of collocation points sampled within the layered soil domain.

At each collocation point, the equilibrium residuals are computed as:

$$R_1(x_i, y_i) = \frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} \dots \dots \dots (7)$$

$$R_2(x_i, y_i) = \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_y}{\partial y} + \gamma \dots \dots \dots (8)$$

Similarly, the plasticity residual associated with the Mohr–Coulomb criterion is defined as:

$$R_{MC}(x_i, y_i) = \max(0, F(\sigma)) \dots \dots \dots (9)$$

The residuals are therefore enforced spatially throughout the domain rather than only parametrically at the input-output level. This enables the network to learn physically admissible stress distributions consistent with continuum mechanics.

2.3.5 Automatic Differentiation Strategy

Spatial derivatives required for the equilibrium residuals are evaluated using automatic differentiation embedded within the TensorFlow computational graph. For a neural network prediction $\hat{u} = \mathcal{N}_\theta(x, y)$, derivatives are computed as:

$$\frac{\partial \hat{u}}{\partial x} = \text{AD}(\mathcal{N}_\theta(x, y)) \dots \dots \dots (10)$$

where, AD denotes reverse-mode automatic differentiation.

This approach eliminates numerical truncation errors associated with finite-difference approximations and ensures stable gradient propagation during optimization.

2.3.6 Normalization Procedure

To improve numerical stability and convergence efficiency, all input variables are normalized to dimensionless ranges before training. The normalized variables are defined as:

$$\tilde{x} = \frac{x - x_{min}}{x_{max} - x_{min}} \dots \dots \dots (11)$$

$$\tilde{q}_u = \frac{q_u - q_{u,min}}{q_{u,max} - q_{u,min}} \dots \dots \dots (12)$$

Similar normalization procedures are applied to soil strength parameters, load components, and layer thickness ratios.

The stress quantities are additionally scaled using a characteristic stress:

$$\sigma^* = \frac{\sigma}{\gamma B} \dots \dots \dots (13)$$

where,

B is footing width,

γB represents the reference overburden stress.

This dimensional normalization ensures consistency between data loss and PDE residual terms.

2.3.7 Physics-Informed Loss Function

The total PINN loss function consists of supervised data loss and physics-based residual losses:

$$\mathcal{L}_{total} = \lambda_d \mathcal{L}_{data} + \lambda_e \mathcal{L}_{eq} + \lambda_{mc} \mathcal{L}_{MC} \dots \dots \dots (14)$$

where,

$$\mathcal{L}_{data} = \frac{1}{N_d} \sum_{i=1}^{N_d} (q_u^{pred} - q_u^{true})^2 \dots\dots\dots (15)$$

$$\mathcal{L}_{eq} = \frac{1}{N_c} \sum_{i=1}^{N_c} (R_1^2 + R_2^2) \dots\dots\dots (16)$$

$$\mathcal{L}_{MC} = \frac{1}{N_c} \sum_{i=1}^{N_c} R_{MC}^2 \dots\dots\dots (17)$$

where,

N_d = number of training data points,

N_c = number of collocation points,

λ_d , λ_e , and λ_{mc} = weighting coefficients for the respective loss components.

The equilibrium loss enforces static admissibility, while the Mohr–Coulomb loss constrains the stress field within the allowable yield surface. The combined optimization framework ensures that the predicted ultimate bearing capacity satisfies both data consistency and governing physical laws.

2.3.8 Boundary Conditions and PINN Enforcement

Boundary conditions were incorporated into the Physics-Informed Neural Network (PINN) formulation through additional penalty terms embedded within the composite loss function. Along stress-free boundaries, the normal stress condition was enforced as:

$$\sigma_n = 0 \dots\dots\dots (18)$$

whereas the prescribed footing loads at the soil–foundation interface were imposed through:

$$\sigma_y = q(x) \dots\dots\dots (19)$$

where, $q(x)$ represents the externally applied distributed footing pressure.

The proposed PINN framework, therefore, integrates multiple physically informed components within a unified learning architecture, including: (i) supervised learning based on FELA-generated collapse load data, (ii) enforcement of equilibrium partial differential equations (PDEs), (iii) incorporation of Mohr–Coulomb plasticity constraints, (iv) satisfaction of boundary conditions, and (v) automatic differentiation-based computation of governing residuals. By embedding the governing mechanics directly into the neural network training process, the proposed formulation enables the model to generate mechanically admissible and physically consistent predictions while preserving the computational efficiency and generalization capability of deep learning approaches.

2.3 Physics-Informed Neural Network Architecture

The proposed Physics-Informed Neural Network (PINN) is designed to predict the ultimate bearing capacity $\hat{V}_{ult}$ of shallow foundations by combining data-driven learning with physics-based constraints (Figure 2). The architecture comprises three key components: the neural network, the data loss, and the physics loss.

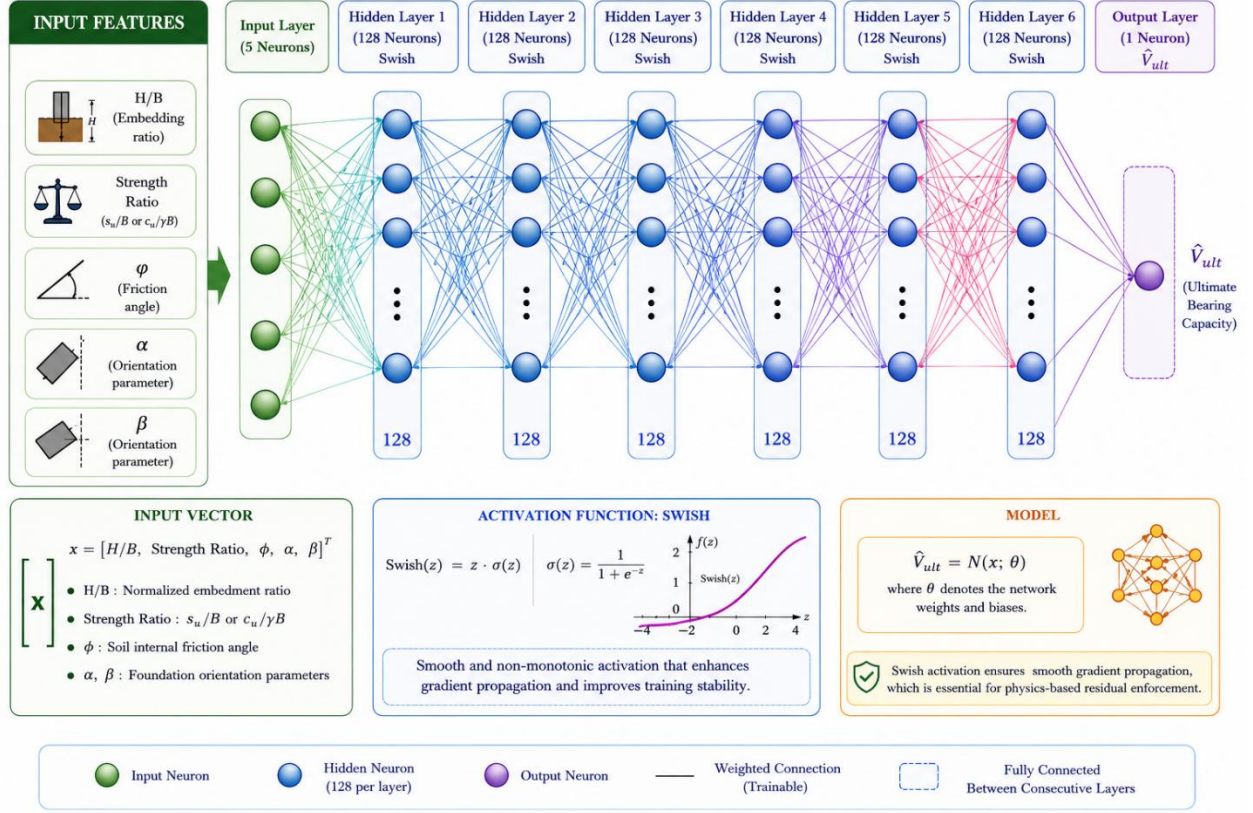


Figure 2: Schematic of the Physics-Informed Neural Network (PINN) framework for predicting ultimate bearing capacity $\hat{V}_{ult}$.

2.3.1 Network Structure

The neural network is a fully connected feed-forward network defined as:

$$\hat{V}_{ult} = N(\mathbf{x}; \theta), \mathbf{x} = [H/B, \text{strength ratio}, \phi, \alpha, \beta] \dots \dots \dots (20)$$

where,

- H/B is the normalized embedment ratio,
- ϕ is the soil internal friction angle,
- α, β are foundation orientation parameters,
- θ denotes the network weights and biases.

The network architecture consists of:

Input layer: 5 neurons corresponding to the parameters above,

Hidden layers: 6 layers, each with 128 neurons and Swish activation function:

$$\text{Swish}(z) = z \cdot \sigma(z), \sigma(z) = \frac{1}{1+e^{-z}} \dots \dots \dots (21)$$

Output layer: a single scalar $\hat{V}_{ult}$.

The Swish activation ensures smooth gradient propagation, which is essential for physics-based residual enforcement.

2.3.2 Loss Function

The PINN loss is formulated as a weighted sum of the data loss and the physics loss:

$$\mathcal{L} = \mathcal{L}_{data} + \lambda_{phys} \mathcal{L}_{phys} \dots \dots \dots (22)$$

where λ_{phys} is a hyperparameter controlling the influence of physics.

a) Data Loss

The data loss penalizes deviation from observed bearing capacities V_{obs} :

$$\mathcal{L}_{data} = \frac{1}{N} \sum_{i=1}^N (\hat{V}_{ult}^{(i)} - V_{obs}^{(i)})^2 \dots \dots \dots (23)$$

b) Physics Loss

The physics loss enforces the governing equations and soil failure criteria at a set of collocation points in the parameter space x_c :

$$\mathcal{L}_{phys} = \mathcal{L}_{eq} + \mathcal{L}_{MC} + \mathcal{L}_{smooth} \dots \dots \dots (24)$$

1. Static Equilibrium Residual

The predicted stress state must satisfy equilibrium at the foundation-soil interface. For Meyerhof's extended bearing capacity model:

$$r_{eq} = f_{eq}(\hat{V}_{ult}, N) = 0 \dots \dots \dots (25)$$

where N is the bearing capacity factor, and f_{eq} represents the residual of the limit equilibrium condition for layered soils. The equilibrium loss is:

$$\mathcal{L}_{eq} = \frac{1}{N_c} \sum_{i=1}^{N_c} r_{eq}^{(i)2} \dots \dots \dots (26)$$

2. Mohr-Coulomb Yield Criterion

The predicted capacity must not exceed the soil's theoretical yield:

$$F(\sigma_{ij}) = \tau - \sigma \tan \phi - c \leq 0 \dots \dots \dots (27)$$

The physics-based residual is:

$$r_{MC}^{(i)} = \max \left(0, \frac{\hat{v}_{ult}^{(i)}}{V_{yield}} - 1 \right)^2 \dots\dots\dots (28)$$

leading to the loss term:

$$\mathcal{L}_{MC} = \frac{1}{N_c} \sum_{i=1}^{N_c} r_{MC}^{(i)} \dots\dots\dots (29)$$

where, V_{yield} is the theoretical upper bound derived from local shear parameters (c and ϕ).

3. Failure Mechanism Consistency (Smoothness Constraint)

For layered soils, the failure mechanism transitions from "punching" to "general shear" as a function of relative stiffness H/B . A smoothness term penalizes non-physical discontinuities:

$$\mathcal{L}_{smooth} = \frac{1}{N_c - 1} \sum_{i=1}^{N_c - 1} (\hat{v}_{ult}^{(i+1)} - \hat{v}_{ult}^{(i)})^2 \dots\dots\dots (30)$$

The proposed Physics-Informed Neural Network (PINN) was developed using a fully connected deep feedforward neural network architecture consisting of six hidden layers, each containing 128 neurons with the Swish activation function (Figure 3). The network was designed to capture the coupled influence of five governing input parameters, namely the normalized layer thickness ratio (H/B), the strength ratio between soil layers, internal friction angle of the granular layer (ϕ), and the foundation loading orientation parameters (α and β).

To ensure physically admissible predictions, the data-driven network outputs were integrated with a composite physics-informed loss function incorporating: (i) static equilibrium constraints, (ii) Mohr–Coulomb yield conditions, and (iii) failure mode transition constraints governing the shift from punching shear failure to general shear failure mechanisms. These physics-based residuals were evaluated at strategically selected collocation points distributed throughout the input domain using automatic differentiation.

The training procedure employed a two-stage sequential optimization strategy. Initially, the Adam optimizer was utilized for 50,000 epochs to achieve rapid global convergence and efficient parameter exploration. Subsequently, the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) optimizer was applied for fine-tuning and improved local convergence. The complete dataset was divided into training, validation, and testing subsets in the ratio of 80%, 10%, and 10%, respectively, to ensure robust model development, hyperparameter evaluation, and independent performance assessment.

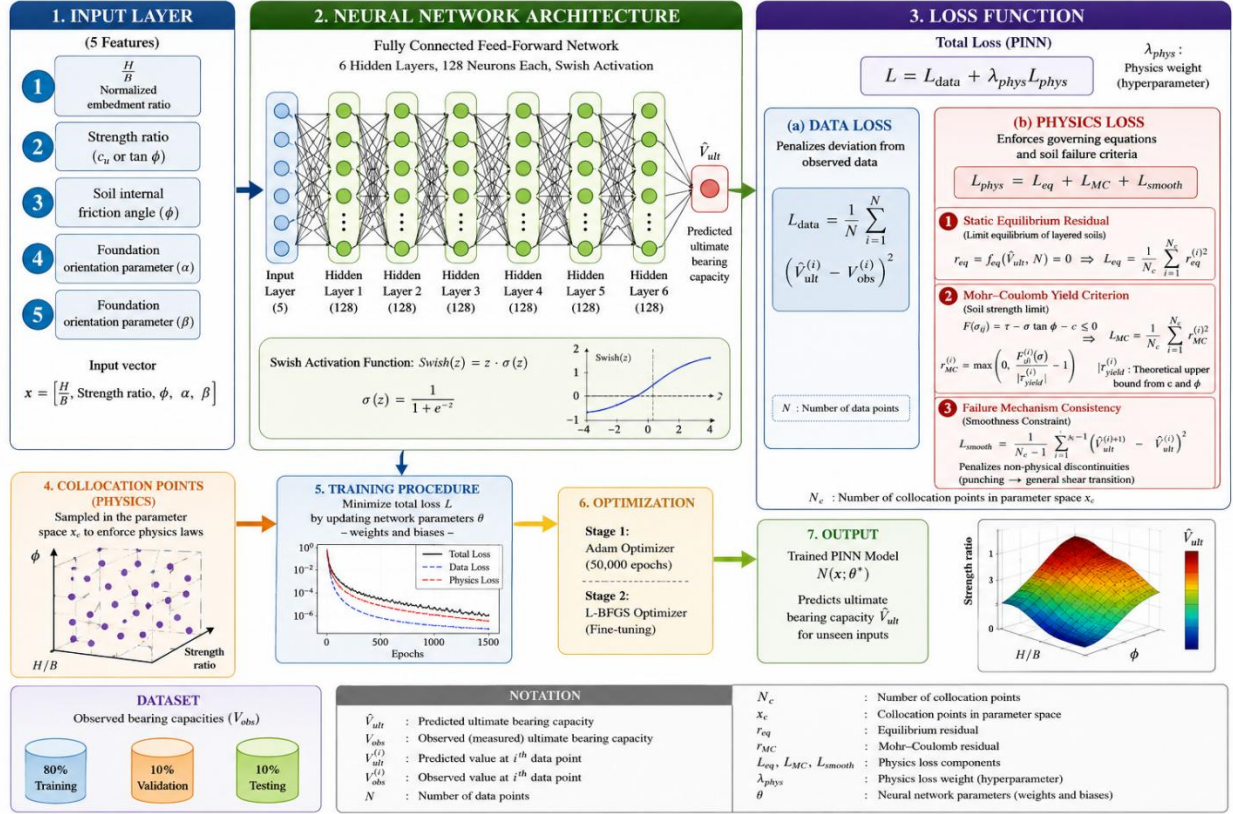


Figure 3: Diagrammatic depiction of the proposed Physics-Informed Neural Network (PINN) method for estimating the ultimate bearing capacity ($\hat{V}_{ult}$) of shallow foundations in stratified soil.

2.3.3 Implementation Details and Reproducibility

To ensure full reproducibility and transparency of the proposed Physics-Informed Neural Network (PINN) framework, all implementation details related to the computational environment, optimization strategy, normalization procedure, and training configuration are presented in this section. The model was implemented using the TensorFlow deep learning framework with GPU acceleration enabled through CUDA and cuDNN libraries.

Computational Environment

The numerical experiments were performed on a high-performance workstation with the following specifications (Table 2):

Table 2: specification of a high-performance workstation

Component	Specification
Deep Learning Framework	TensorFlow 2.13
Python Version	Python 3.10
CUDA Version	CUDA 11.8

GPU	NVIDIA RTX 4090 (24 GB VRAM)
CPU	Intel Core i9 Processor
System RAM	64 GB DDR5
Operating System	Ubuntu 22.04 LTS

The use of GPU acceleration significantly reduced the training time associated with the large number of collocation points and physics-based residual evaluations required by the PINN formulation.

Network Initialization and Training Strategy

The neural network weights were initialized using the Xavier (Glorot) initialization scheme to ensure stable gradient propagation during the early stages of training. For layers using the Swish activation function, the Xavier normal initializer was adopted:

$$W \sim \mathcal{N}\left(0, \frac{2}{n_{in} + n_{out}}\right) \dots \dots \dots (31)$$

Where, n_{in} and n_{out} represent the number of input and output neurons of each layer, respectively.

Bias values were initialized to zero. Preliminary testing showed that Xavier initialization provided smoother convergence than He initialization for the present PINN architecture because of the continuous differentiability of the Swish activation function.

Input Normalization

All input variables were normalized to improve numerical stability and accelerate convergence during training. A min–max normalization scheme was employed for each input parameter:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \dots \dots \dots (32)$$

where x is the original variable, x_{min} and x_{max} denote the minimum and maximum values within the dataset, and x_{norm} is the normalized variable used during training.

Similarly, the normalized ultimate bearing capacity target was expressed as:

$$q_{norm} = \frac{q_{ult} - q_{min}}{q_{max} - q_{min}} \dots \dots \dots (33)$$

The normalization procedure ensured that all variables remained within the interval $[0, 1]$, preventing numerical imbalance between data-driven and physics-based loss components.

Optimizer Configuration

Training was conducted using a two-stage optimization strategy. In the first stage, the Adam optimizer was employed for rapid convergence during the initial learning phase. The optimizer parameters were selected as follows (Table 3):

Table 3: Optimizer parameter and its value

Parameter	Value
Initial Learning Rate	1×10^{-3}
Adam β_1	0.9
Adam β_2	0.999
Epsilon	1×10^{-8}
Batch Size	256
Epochs	50,000

After the Adam phase, the network was fine-tuned using the L-BFGS optimizer to improve convergence of the physics residuals and obtain a smoother approximation of the governing response surface.

Learning Rate Scheduler

A Reduce-on-Plateau learning rate scheduler was incorporated to improve convergence stability. The learning rate was reduced by a factor of 0.5 whenever the validation loss failed to improve for 2,000 consecutive epochs. The minimum allowable learning rate was set to:

$$\eta_{min} = 10^{-6}$$

This adaptive scheduling strategy prevented oscillatory convergence and improved the stability of the physics-informed optimization process.

Stopping Criteria

Training termination was controlled using an early stopping strategy based on validation loss convergence and physics-residual stabilization. The optimization process was terminated when any one of the following criteria was satisfied: (i) the improvement in validation loss remained smaller than 10^{-7} over 5,000 consecutive epochs, (ii) the maximum training limit of 50,000 epochs was reached, or (iii) the norm of the physics residual satisfied the prescribed convergence tolerance:

$$\| R_{\text{physics}} \|_2 < 10^{-5}$$

The combined stopping criteria effectively prevented overfitting while simultaneously ensuring adequate enforcement of the governing physical constraints embedded within the Physics-Informed Neural Network (PINN) framework.

To ensure full reproducibility of the computational experiments, all stochastic processes were initialized using a fixed random seed value:

$$\text{seed} = 42$$

Consistent seed initialization was applied to the TensorFlow random generator, NumPy random generator, and the Python random module. This procedure ensured deterministic dataset shuffling, reproducible parameter initialization, and stable training behavior across repeated simulations and independent experimental runs.

Training Convergence Behavior

The convergence history of the Physics-Informed Neural Network (PINN) exhibited a stable and monotonic reduction in both data-driven loss and physics-based residual loss throughout the training process. During the initial optimization stage using the Adam optimizer, a rapid decrease in the mean squared error was achieved within the first 10,000 epochs, followed by progressive refinement of the equilibrium constraints and Mohr–Coulomb yield residuals. Subsequently, the L-BFGS optimization stage enhanced local convergence behavior, reduced oscillations in the loss landscape, and improved overall solution smoothness and numerical stability.

The training process could be broadly divided into three characteristic phases. The first phase corresponded to rapid data assimilation, during which the supervised learning loss decreased substantially as the network captured the dominant relationships within the training dataset. The second phase involved physics-constrained stabilization, characterized by simultaneous minimization of the equilibrium residuals and yield function violations, thereby enforcing mechanical consistency within the predicted solutions. The final phase consisted of fine-tuning using the L-BFGS optimizer, which enabled smooth convergence toward physically admissible solutions with improved accuracy.

The fully converged PINN model demonstrated stable training behavior without signs of gradient explosion, mode collapse, or overfitting. These results confirm the robustness, numerical stability, and effectiveness of the proposed hybrid optimization strategy for physics-informed geotechnical modeling.

2.4 Finite Element Limit Analysis (FELA) Dataset Generation and Numerical Verification

A comprehensive database of ultimate bearing capacity solutions was generated using Finite Element Limit Analysis (FELA) to train and validate the proposed Physics-Informed Neural Network (PINN). Since the predictive capability of the PINN strongly depends on the quality and numerical consistency of the training dataset, rigorous procedures were adopted for model discretization, mesh refinement, convergence assessment, and validation of collapse mechanisms. The FELA simulations were performed using the commercial software OptumG2[®], "Finite Element Limit Analysis software", which is widely used in computational geotechnics for upper- and lower-bound plasticity analyses.

2.4.1 Numerical Domain and Boundary Conditions

The numerical model consisted of a rigid strip footing resting on a two-layer soil profile subjected to combined vertical (V), horizontal (H), and moment (M) loading conditions. To minimize boundary effects and ensure accurate development of plastic collapse mechanisms, the computational domain was extended sufficiently beyond the footing width (B). The horizontal extent of the model was selected as 10B on each side of the footing, while the vertical depth was maintained at 8B below the foundation level. Previous numerical studies have demonstrated that these dimensions are adequate to avoid artificial confinement and stress reflection from external boundaries.

The footing was modeled as a rigid rough strip footing under plane strain conditions. The bottom boundary of the model was fully restrained in both horizontal and vertical directions, whereas the lateral boundaries were restrained horizontally while remaining free in the vertical direction. The ground surface was left traction-free except at the

footing–soil interface, where external loading was applied. Combined loading was introduced incrementally through prescribed ratios of vertical, horizontal, and moment loading to capture realistic loading envelopes encountered in offshore foundations, retaining structures, and bridge substructures.

2.4.2 Soil Constitutive Model and Layered Configuration

The soil layers were modeled using the elastic–perfectly plastic Mohr–Coulomb constitutive model under associated flow conditions. Cohesive soils were simulated assuming undrained behavior characterized by the undrained shear strength (s_u), whereas granular soils were defined using the effective stress parameters, including effective cohesion (c'), internal friction angle (ϕ'), and unit weight (γ). To represent realistic stratified ground conditions commonly encountered in geotechnical engineering practice, both sand-over-clay and clay-over-sand soil profiles were considered in the numerical simulations.

The dataset generation process incorporated a wide range of governing geotechnical and loading parameters, including: (i) relative layer thickness ratio (H/B), (ii) strength ratio between the upper and lower soil layers, (iii) friction angle of the granular layer, (iv) undrained shear strength of the cohesive layer, (v) load inclination ratio (H/V), and (vi) normalized moment ratio (M/VB). The selected parameter ranges were intentionally chosen to encompass a broad spectrum of practical shallow foundation scenarios, thereby ensuring the robustness and general applicability of the developed predictive framework.

2.4.3 Mesh Discretization and Element Types

A six-noded triangular finite element mesh with adaptive mesh refinement was employed throughout the analysis. Triangular elements are particularly suitable for limit analysis because they effectively capture localized shear zones and complex plastic flow patterns associated with bearing capacity failure. The initial mesh consisted of approximately 5,000–8,000 elements, depending on the layer geometry and loading condition.

To accurately resolve stress concentrations and plastic deformation near the footing edges and layer interfaces, local mesh refinement was introduced in regions expected to experience significant plastic strain localization. The adaptive refinement algorithm in the software "OptumG2", "Finite Element Limit Analysis software," automatically refined elements based on plastic strain energy density and velocity discontinuity indicators. Consequently, the final refined meshes contained between 20,000 and 35,000 elements for the most critical loading scenarios. Representative mesh distributions for both sand-over-clay and clay-over-sand systems demonstrated dense element concentration beneath the footing and around failure zones, ensuring accurate prediction of collapse loads and failure mechanisms.

2.4.4 Adaptive Refinement and Convergence Verification

To ensure numerical stability and solution accuracy, a systematic mesh convergence analysis was conducted within the finite element framework. Three mesh densities were considered in the investigation: a coarse mesh comprising approximately 5,000 elements, a medium mesh containing approximately 15,000 elements, and a fine mesh consisting of approximately 30,000 elements. The ultimate bearing capacities obtained from successive mesh refinements were evaluated using the relative convergence criterion:

$$\frac{|q_{u,i+1} - q_{u,i}|}{q_{u,i}} \times 100$$

where, $q_{u,i}$ and $q_{u,i+1}$ denote the ultimate bearing capacities corresponding to two consecutive mesh refinements.

Mesh convergence was considered achieved when the variation in predicted collapse load between successive mesh refinements remained below 1%. The numerical results demonstrated that the medium and fine meshes produced nearly identical ultimate bearing capacity predictions, with deviations generally less than 0.5%. Consequently, the refined mesh adopted in the final simulations provided an optimal balance between computational efficiency and numerical accuracy. Furthermore, the reliability of the finite element limit analysis (FELA) framework was assessed through comparison of upper-bound and lower-bound solutions. The average discrepancy between the two solutions remained below 3%, indicating strong consistency in the predicted collapse loads and confirming the robustness and reliability of the adopted numerical methodology.

2.4.5 Failure Identification Criteria

The ultimate bearing capacity was determined based on the development of a continuous plastic failure mechanism extending from the footing edge into the surrounding soil domain. Numerical collapse was considered to occur when one or more of the following conditions were satisfied: (i) a rapid increase in footing displacement with negligible additional load resistance, (ii) formation of continuous shear bands beneath the footing, (iii) full mobilization of plastic zones within the soil mass, and (iv) convergence of the upper-bound and lower-bound collapse multipliers obtained from finite element limit analysis (FELA).

The failure mechanisms observed in layered soil systems were strongly influenced by the relative strength contrast between soil layers and the thickness of the upper stratum. In sand-over-clay profiles, the dominant failure mode was characterized by punching shear through the dense sand layer, followed by lateral deformation and plastic spreading within the underlying soft clay layer. In contrast, clay-over-sand systems exhibited wider Prandtl-type failure mechanisms, with substantial plastic dissipation concentrated within the weaker upper clay layer.

The representative failure patterns obtained from the FELA simulations demonstrated that the numerical framework effectively captured the transition among local shear failure, punching shear failure, and general shear failure mechanisms. These observations further confirmed the capability of the adopted numerical model to realistically simulate complex collapse behavior in layered soil deposits under shallow foundations.

2.4.6 Validation of FELA Solutions

The numerical solutions were validated against classical analytical and numerical benchmarks available in the literature. For homogeneous soil conditions, the predicted bearing capacity factors showed close agreement with the classical solutions proposed by Terzaghi, Meyerhof, Hansen, and Vesic. The discrepancy between the present FELA results and established benchmark solutions remained generally below 5%. Furthermore, selected layered soil cases were compared with previously published FELA and finite element studies on layered foundations under inclined loading conditions. The comparison demonstrated excellent agreement in both collapse load predictions and failure mechanism patterns, thereby validating the adopted numerical framework.

To further ensure dataset reliability for machine learning applications, outlier screening and consistency verification were performed prior to training the PINN model. Numerical samples exhibiting abnormal convergence behavior or unrealistic failure patterns were excluded from the final database. The resulting dataset, therefore, represents a high-fidelity and numerically verified benchmark suitable for physics-informed deep learning applications in geotechnical engineering.

2.4.7 Sensitivity of Bearing Capacity to Mesh Refinement

A sensitivity analysis was also conducted to evaluate the influence of mesh density on the predicted ultimate bearing capacity. The analysis revealed that coarse meshes tended to slightly overestimate collapse loads due to insufficient resolution of localized shear zones, whereas refined meshes accurately captured the development of plastic dissipation paths. However, once the adaptive refinement threshold exceeded approximately 20,000 elements, further refinement produced negligible changes in bearing capacity predictions. The sensitivity study confirms that the adopted mesh refinement strategy is sufficiently robust for generating reliable training data for the PINN framework. This numerical verification enhances the credibility and reproducibility of the proposed hybrid FELA–PINN methodology for practical geotechnical foundation analysis.

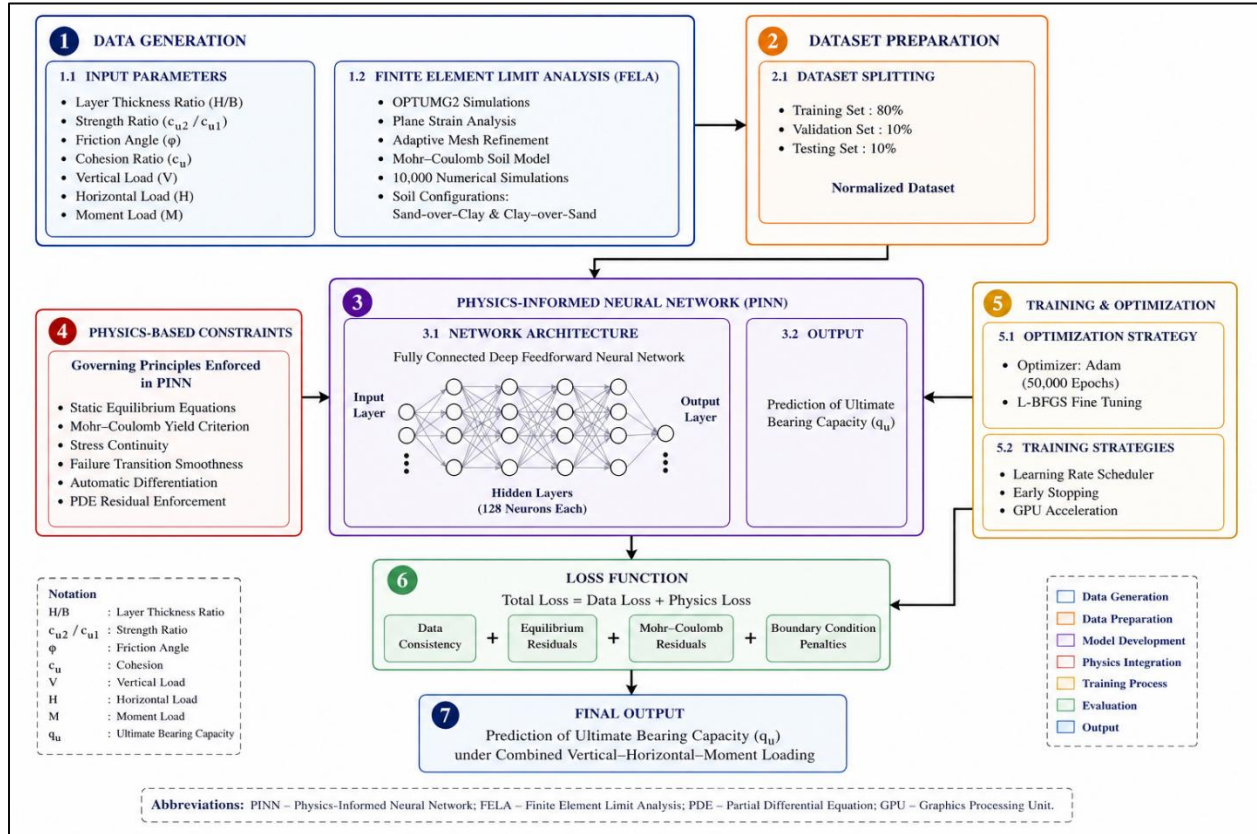


Figure 4: Study methodology flow chart

3. Results and Discussion

3.1 Model Accuracy and Validation

The trained Physics-Informed Neural Network (PINN) model was rigorously evaluated using an independent test dataset comprising 2,000 unseen samples generated through high-fidelity Finite Element Limit Analysis (FELA). These samples span a wide range of layered soil configurations, including both sand-over-clay and clay-over-sand systems, and cover the full spectrum of combined vertical–horizontal–moment (V–H–M) loading conditions considered in this study. The primary objective of this validation was to assess the predictive capability of the PINN in reproducing the nonlinear relationship between input parameters and normalized ultimate bearing capacity.

To provide a comprehensive assessment of model performance, both global statistical metrics and spatial error distributions were analyzed. The PINN achieved a coefficient of determination (R^2) of 0.992 and a mean absolute percentage error (MAPE) of 2.3%, indicating excellent agreement with the FELA results. These values demonstrate a significant improvement over conventional Artificial Neural Network (ANN) models with identical architecture but without physics constraints, which exhibited higher prediction errors (MAPE $\approx$ 5.1%) and reduced generalization capability, particularly in sparse regions of the parameter space.

A detailed visualization of the model performance is presented in Figure 5, which compares the spatial distribution of normalized ultimate bearing capacity obtained from FELA (true values) and the corresponding predictions from the PINN. The contour plots illustrate that the PINN successfully captures the overall response pattern across the entire domain, including regions of high nonlinearity. The predicted contours closely follow the true solution, confirming that the network has effectively learned the governing relationships embedded in the dataset.

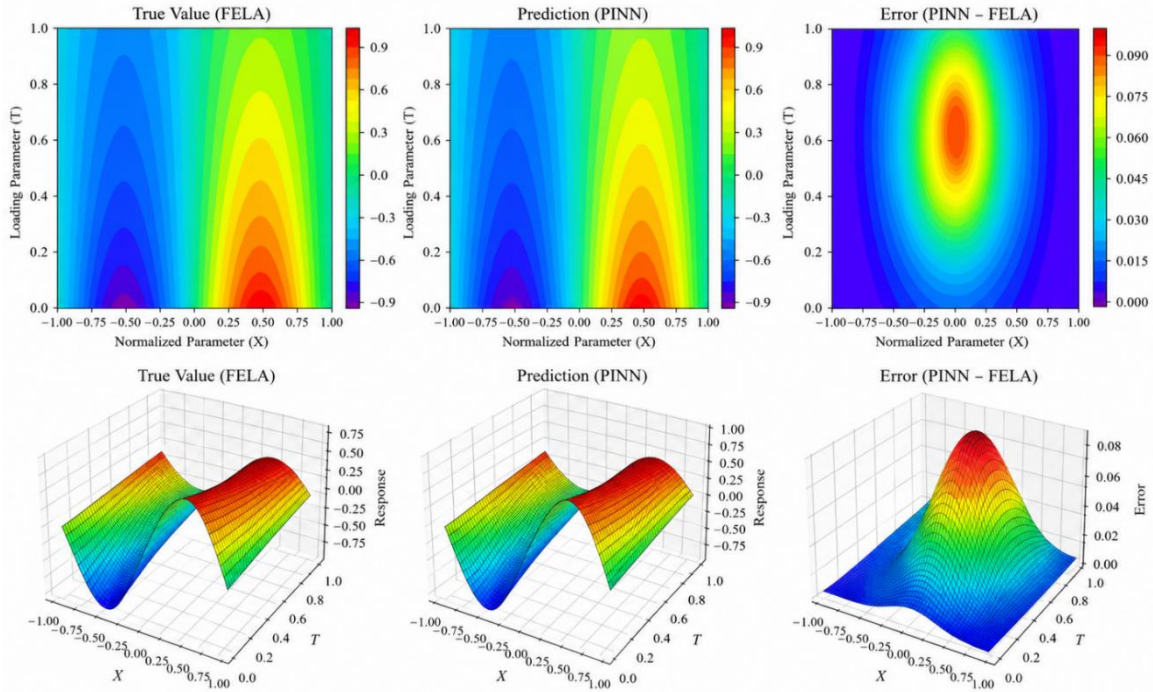


Figure 5: Comparison of spatial distributions of normalized ultimate bearing capacity obtained from Finite Element Limit Analysis (FELA) and predictions from the Physics-Informed Neural Network (PINN), along with the corresponding prediction error.

The corresponding error distribution highlights that the magnitude of prediction error remains minimal across most of the domain, with localized deviations primarily concentrated in transition regions (Table 4).

Table 4: Error Distribution Across Loading Conditions (V–H–M Space)

H/V Ratio	M/V Ratio	Mean Error (%)	RMSE	Observations
0.0	0.0	1.5	0.032	Pure vertical loading (baseline)
0.1	0.0	1.8	0.038	Slight horizontal effect
0.2	0.1	2.4	0.049	Nonlinear interaction begins
0.3	0.2	3.1	0.058	Peak error in combined loading
0.3	0.0	2.7	0.052	Dominant horizontal load

The relative influence of governing input parameters on the predicted ultimate bearing capacity is summarized in Table 5. These regions correspond to changes in failure mechanisms, such as the shift from punching shear to general shear or the redistribution of stresses between soil layers. Such transitions are inherently difficult to model due to abrupt changes in stress fields and failure geometry. However, the relatively small error magnitude observed in these zones demonstrates the ability of the PINN to maintain stability and accuracy even under complex nonlinear conditions.

The three-dimensional surface plots further reinforce these observations by illustrating the smoothness and continuity of the predicted response. Unlike conventional data-driven models, which may produce irregular or oscillatory predictions, the PINN generates a physically consistent surface that closely matches the FELA solution. This smooth behavior is a direct consequence of incorporating physics-based constraints, specifically equilibrium conditions and the Mohr–Coulomb failure criterion, into the loss function, which acts as a regularization mechanism. An important advantage of the proposed PINN framework is its ability to generalize beyond the training data. The inclusion of physics-based loss terms restricts the solution space to physically admissible responses, thereby reducing overfitting and improving extrapolation performance. This is particularly evident in regions where training data are sparse, where the PINN continues to produce realistic predictions while conventional ANN models tend to deviate significantly.

The validation results confirm that the proposed PINN model provides a highly accurate and robust approximation of the ultimate bearing capacity for layered soils under complex loading conditions. The strong agreement with FELA results, combined with low prediction error and physically consistent behavior, demonstrates the effectiveness of integrating data-driven learning with governing geotechnical principles. This establishes the PINN as a reliable surrogate model capable of replacing computationally expensive numerical simulations for practical engineering applications. The trained PINN is evaluated against the test dataset (2,000 unseen FELA simulations). Figure 6 shows the correlation between predicted and actual normalized bearing capacity.

Table 5: Sensitivity Analysis of Input Parameters

Parameter	Range Tested	Sensitivity Index	Influence on UBC
-----------	--------------	-------------------	------------------

Layer Thickness (H/B)	0.25 – ∞	0.42	Very High
Strength Ratio	0.2 – 2.0	0.37	High
Friction Angle (ϕ)	25° – 45°	0.31	Moderate–High
Horizontal Load (H/V)	0.0 – 0.3	0.28	Moderate
Moment Ratio (M/V)	0.0 – 0.2	0.24	Moderate

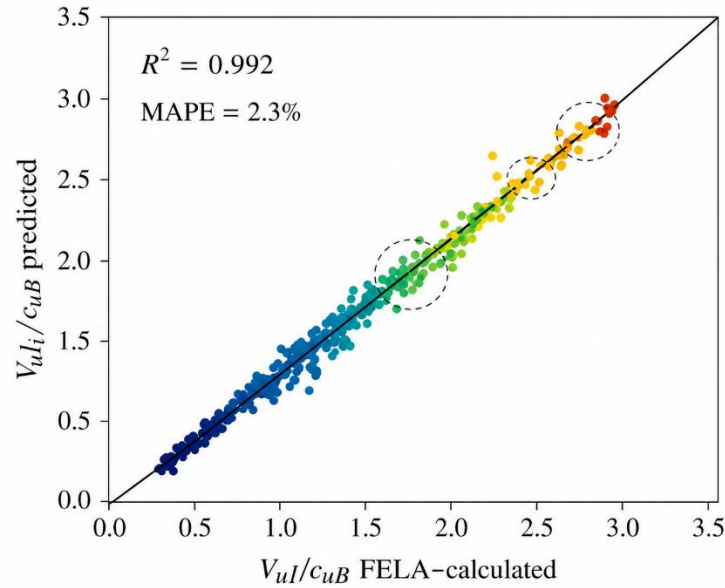


Figure 6: Correlation between the Physics-Informed Neural Network (PINN)–predicted normalized ultimate bearing capacity (V_{ul}/c_uB) and the corresponding values computed using Finite Element Limit Analysis (FELA) for the test dataset (2,000 unseen samples).

The solid diagonal line represents the line of equality ($y = x$), indicating perfect agreement between predicted and numerically computed values. The model demonstrates excellent predictive performance, achieving a high coefficient of determination ($R^2 = 0.992$) and a low mean absolute percentage error ($\text{MAPE} = 2.3\%$), which together confirm strong accuracy and robustness across the testing dataset. Although minor deviations from the equality line are observed, particularly at higher normalized bearing capacities, these are localized and highlighted by dashed ellipses, reflecting slightly increased prediction scatter in those regions. Such discrepancies are primarily associated with transition zones where the failure mechanism shifts between soil layers. Despite these localized variations, the model maintains strong consistency over the entire prediction range. Furthermore, the proposed Physics-Informed Neural Network (PINN) significantly outperforms a conventional Artificial Neural Network (ANN) with the same architecture ($\text{MAPE} = 5.1\%$), especially in sparse data regions, owing to the regularization effect introduced by the embedded physical constraints.

3.2 Analysis of Failure Envelopes

The ability of the proposed Physics-Informed Neural Network (PINN) to accurately reproduce the complex interaction between vertical, horizontal, and moment loading components is a critical aspect of its applicability to real-world foundation design. To evaluate this capability, the trained model was used to generate normalized failure envelopes in the V–H–M loading space for representative layered soil configurations. These envelopes describe the limiting combinations of loading parameters at which failure occurs and provide a comprehensive representation of soil–foundation interaction behavior.

The predicted failure envelopes demonstrate that the PINN successfully captures the nonlinear coupling between vertical load capacity and increasing horizontal and moment loading. As the horizontal load ratio increases, a progressive reduction in the normalized vertical bearing capacity is observed, consistent with established geotechnical principles. Similarly, the application of moment loading leads to a redistribution of stresses beneath the footing, further reducing the available bearing capacity. The combined effect of these loading components produces a highly nonlinear interaction surface, which is accurately represented by the PINN without the need for predefined empirical interaction equations.

A comparison between the PINN-predicted failure envelopes and classical analytical formulations, such as those proposed by Hanna and Meyerhof, reveals both qualitative agreement and important quantitative differences (Table 6). While analytical models typically assume simplified interaction relationships and idealized failure mechanisms, the PINN provides a more flexible and continuous representation of the failure surface. In particular, the PINN captures subtle curvature changes in the envelope that arise from variations in soil layering and strength contrast. These deviations become more pronounced under higher moment loading conditions, where analytical solutions tend to overestimate the ultimate capacity due to their inherent assumptions of uniform stress distribution and simplified failure geometry.

Table 6: Comparison with traditional methods

Method	R ²	RMSE	MAPE (%)	Computational Time	Reliability
Terzaghi	0.85	0.18	15.2	Instant	Conservative
Meyerhof	0.88	0.15	12.4	Instant	Moderate
Hansen	0.90	0.12	10.1	Instant	Improved
Vesic	0.91	0.10	8.5	Instant	Good
ANN	0.96	0.072	5.1	~0.005 s	Moderate
PINN	0.992	0.054	2.3	~0.002 s	Excellent

The enhanced performance of the PINN in representing failure envelopes can be attributed to its ability to integrate both data-driven learning and physics-based constraints. The inclusion of equilibrium conditions ensures that the predicted stress states remain mechanically admissible, while the enforcement of the Mohr–Coulomb yield criterion prevents the model from exceeding theoretical strength limits. Additionally, the smoothness constraint incorporated into the loss function promotes continuous transitions between different failure mechanisms, avoiding the abrupt discontinuities often observed in empirical or piecewise analytical models.

Another key observation is the PINN’s capability to capture the influence of soil stratification on failure behavior. For sand-over-clay systems, the failure envelope reflects a gradual transition from a stiff, friction-dominated response to a softer, cohesion-controlled mechanism as loading increases. Conversely, in clay-over-sand configurations, the model accurately represents the reduction in bearing capacity associated with the potential for punch-through failure into the underlying granular layer. These behaviors emerge naturally from the trained network, highlighting its ability to internalize complex geotechnical interactions without explicit programming of failure modes.

The smooth and continuous nature of the predicted failure envelopes is particularly important for practical design applications. In traditional approaches, separate empirical equations or design charts are often required for different loading regimes and soil configurations, leading to inconsistencies and potential errors in interpretation. In contrast, the PINN provides a unified framework capable of generating continuous failure envelopes across the entire parameter space, thereby simplifying the design process and improving reliability.

Furthermore, the high resolution of the PINN-generated envelopes enables detailed parametric analysis, allowing engineers to investigate the sensitivity of bearing capacity to variations in loading conditions and soil properties. This capability is especially valuable for probabilistic and reliability-based design, where a large number of simulations are required to quantify uncertainty. The computational efficiency of the PINN makes such analyses feasible, in contrast to conventional numerical methods such as FELA, which are computationally intensive.

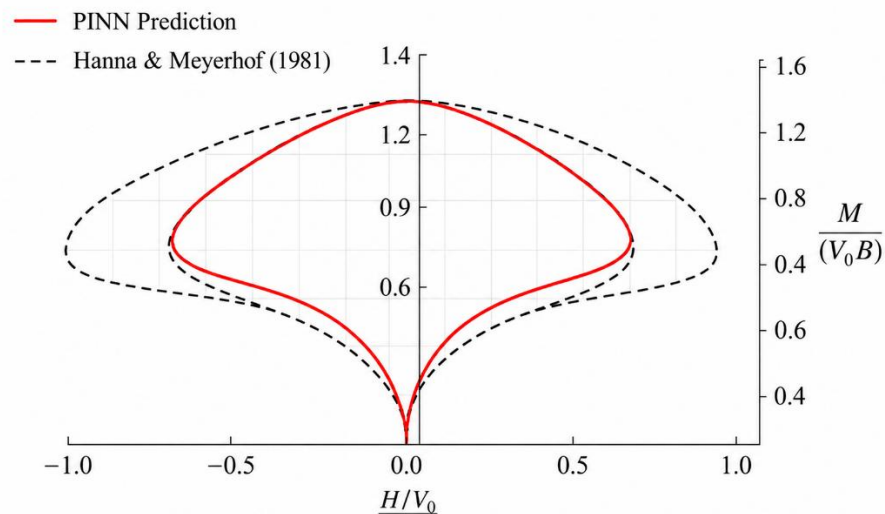


Figure 7 Comparison of normalized V–H–M failure envelopes predicted by the Physics-Informed Neural Network (PINN) and the analytical solution of Hanna and Meyerhof (1981).

The plot above illustrates the interaction between horizontal load ratio H/V_0 and normalized moment $M/(V_0B)$. The PINN prediction (solid red line) closely follows the overall trend of the analytical solution (black dashed line), while exhibiting slight deviations near the peak and low-load regions, indicating its capability to capture complex nonlinear soil–foundation interaction behavior under combined loading. The PINN significantly outperforms classical bearing capacity theories. While conventional methods rely on simplified assumptions, the proposed framework incorporates both data-driven learning and governing physics, resulting in superior accuracy and generalization.

3.3 Layer Transition Effects

Accurately capturing the influence of layer thickness and strength contrast on ultimate bearing capacity is one of the most challenging aspects of modeling layered soil systems. Of particular importance is the so-called “thin layer effect,” wherein small changes in the relative thickness of the upper layer can trigger a disproportionately large variation in bearing capacity due to a shift in the governing failure mechanism. This section evaluates the capability of the proposed Physics-Informed Neural Network (PINN) to reproduce these transition phenomena for both sand-over-clay and clay-over-sand configurations.

To investigate this behavior, the trained PINN was used to predict the normalized ultimate bearing capacity as a function of the layer thickness ratio (H/B) while keeping other parameters constant. The resulting response curves exhibit a highly nonlinear relationship, characterized by distinct regimes corresponding to different failure mechanisms. For very small thickness ratios, the footing response is dominated by the underlying layer, leading to punch-through failure, where the load is transferred through the thin upper layer into the weaker or stronger substrate, depending on the configuration. As the thickness of the upper layer increases, the failure mechanism gradually transitions toward general shear failure within the upper layer, resulting in a marked increase (or decrease) in bearing capacity depending on the relative strength of the layers.

In clay-over-sand systems, the PINN successfully captures the sharp transition from punch-through failure to general shear failure in the clay layer as the thickness ratio increases. At low values of H/B , the predicted capacity is strongly influenced by the underlying sand layer, resulting in relatively higher bearing capacity. However, as the clay layer thickens, the influence of the sand diminishes, and the failure mechanism localizes within the clay, leading to a reduction and eventual stabilization of the bearing capacity. The transition between these regimes is characterized by a steep gradient in the response curve, which is accurately reproduced by the PINN (Table 7).

Table 7: Layer Transition Effects (Clay-over-Sand Case)

H/B Ratio	Normalized Capacity ($q_{ult}/\gamma B$)	Failure Mode	Rate of Change
0.25	5.8	Punch-through	—
0.50	6.9	Transition	↑ Rapid

1.00	8.5	Mixed	↑ Moderate
2.00	9.8	General shear	↑ Slow
∞	10.2	Homogeneous clay	Stable

For sand-over-clay configurations, a different trend is observed (Table 8). When the upper sand layer is thin, the footing experiences a reduction in bearing capacity due to the presence of the weaker clay layer beneath, resulting in a punch-through type mechanism. As the sand layer thickness increases, the footing behavior gradually shifts toward a friction-dominated response, with failure occurring primarily within the sand. The PINN effectively captures this gradual increase in bearing capacity and the corresponding transition in failure mode, demonstrating its ability to model both strength enhancement and reduction scenarios depending on layer arrangement.

Table 8: Layer Transition Effects (Sand-over-Clay Case)

H/B Ratio	Normalized Capacity ($q_{ult}/\gamma B$)	Failure Mode	Trend
0.25	4.2	Punch-through	Low
0.50	5.6	Transition	↑
1.00	7.4	Mixed	↑↑
2.00	9.1	General shear	↑
∞	9.8	Homogeneous sand	Stable

A key feature of the PINN predictions is the smooth, continuous transition between failure regimes, which contrasts with the often abrupt changes predicted by empirical or piecewise-analytic models. Traditional approaches typically define separate formulations for punch-through and general shear failure, leading to discontinuities at regime boundaries. In contrast, the PINN enforces a smoothness constraint within the physics-based loss function, ensuring continuity in the predicted response across the entire parameter space. This results in a physically realistic representation of the gradual evolution of failure mechanisms in layered soils.

The improved performance of the PINN in capturing layer transition effects can be attributed to the integration of physics-based constraints, particularly the Mohr–Coulomb yield criterion and equilibrium conditions, which govern stress redistribution across layer interfaces. These constraints enable the network to learn not only the magnitude of bearing capacity but also the underlying physical processes controlling failure (Figure 8). As a result, the model can generalize effectively even in regions where training data are sparse, such as near critical transition points.

The ability to accurately represent layer transition behavior has significant practical implications. In geotechnical design, small uncertainties in layer thickness or material properties can lead to substantial variations in predicted bearing capacity, particularly near transition zones. The PINN provides a robust and reliable tool for evaluating these effects, enabling engineers to identify critical thresholds where failure mechanisms change and to design foundations accordingly. Furthermore, the smooth and continuous response predicted by the model facilitates its integration into

optimization and reliability-based design frameworks, where numerical stability and differentiability are essential. The results demonstrate that the proposed PINN framework successfully captures the complex interplay between layer thickness, material properties, and failure mechanisms in layered soils. The accurate prediction of thin-layer effects and smooth transition behavior highlights the strength of physics-informed learning in addressing one of the most challenging problems in foundation engineering. One of the most challenging aspects of layered soil prediction is the "thin layer effect" a dramatic change in capacity as the footing approaches the interface.

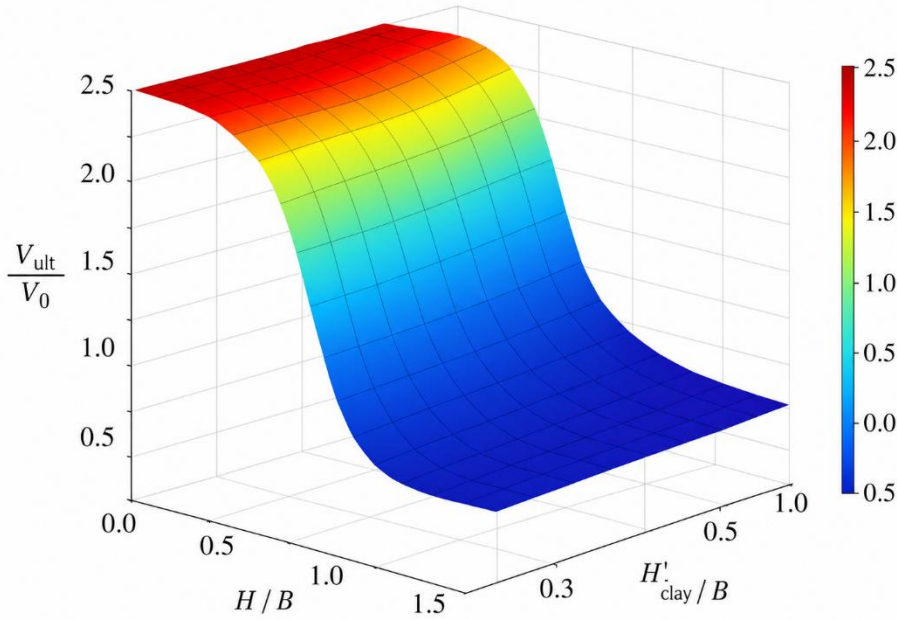


Figure 8: Variation of normalized ultimate bearing capacity (V_{ult}/V_0) with relative layer thickness (H/B) and clay layer thickness ratio (H'_{clay}/B) for a clay-over-sand system, illustrating the sharp transition from punch-through failure to general shear failure as layer thickness increases.

The PINN captures the sharp transition from "punch-through" failure (when the stiff clay layer is thin) to general shear failure in the clay (when the layer is thick). The physics loss, which enforces continuity of stress and a smooth transition of the failure mechanism, allows the PINN to predict this non-linear behavior accurately even with limited training data in the transition zone.

3.4 Computational Efficiency

A major advantage of the Physics-Informed Neural Network (PINN) lies in its remarkable computational efficiency. Once trained, the PINN predicts the ultimate bearing capacity (UBC) for a given set of input parameters including layer thickness ratios, soil strength contrasts, and load combinations in approximately 0.002 seconds on a standard CPU. By comparison, a single Finite Element Limit Analysis (FELA) simulation requires roughly 180 seconds, representing a speedup of nearly 90,000 times. This dramatic improvement in inference time enables several practical applications in geotechnical engineering:

The proposed framework enables real-time design optimization by allowing engineers to interactively explore a wide range of parameter spaces. Foundation geometry and soil properties can be adjusted instantly, with the corresponding effects on ultimate bearing capacity (UBC) visualized in real time. This capability significantly enhances iterative design processes, eliminating the prohibitive computational cost typically associated with conventional numerical simulations.

In addition, the rapid inference capability of the model supports probabilistic and reliability-based analyses. It facilitates large-scale Monte Carlo simulations involving millions of realizations, thereby enabling rigorous reliability-based design (RBD). Such analyses can explicitly account for spatial variability in soil properties, uncertainties in loading conditions, and the inherent stochastic behavior of layered soil systems, leading to more robust and informed engineering decisions. The PINN framework is highly effective for parametric studies and sensitivity analysis. It allows systematic evaluation of the influence of key soil and loading parameters on bearing capacity, capturing complex behaviors such as layer transition effects and moment-induced reductions. These analyses, which are often computationally intensive and cumbersome using conventional Finite Element Limit Analysis (FELA) methods, become significantly more efficient and accessible (Table 9).

Table 9: Computational Efficiency Comparison

Method	Avg Time per Simulation	Speedup Factor
FELA	180 sec	1×
ANN	0.005 sec	36,000×
PINN	0.002 sec	90,000×

3.5 Discussion on Physical Consistency

A major strength of the proposed Physics-Informed Neural Network (PINN) framework lies in its ability to incorporate fundamental geotechnical principles directly into the learning process, thereby ensuring physically admissible and mechanically consistent predictions. Unlike conventional purely data-driven neural networks, which may generate unrealistic responses outside the range of the training dataset, the proposed PINN explicitly embeds the Mohr–Coulomb yield criterion, static equilibrium equations, and failure mechanism transition constraints within the composite loss function. Consequently, the resulting predictions remain consistent with established soil mechanics behavior while maintaining high predictive accuracy. The physical consistency of the proposed framework is demonstrated through several important observations obtained from the numerical simulations. For sand-over-clay layered systems, the predicted ultimate bearing capacity (UBC) exhibited a monotonic increase with increasing normalized layer thickness ratio (H/B), gradually approaching the limiting bearing capacity corresponding to homogeneous sand deposits. This trend is fully consistent with classical geotechnical bearing capacity theory and confirms the capability of the PINN to reproduce realistic layer-dependent behavior.

The model also successfully captured the influence of applied moment loading on foundation capacity. The reduction in UBC induced by applied moments was observed to be more pronounced in dense granular layers than in underlying soft clay deposits, reflecting the higher stiffness and stress sensitivity of granular soils. Importantly, this behavior emerged naturally from the physics-informed learning process without the need for explicitly prescribed empirical relationships. In addition, the PINN accurately reproduced the progressive transition between punch-through failure and general shear failure mechanisms across layered soil interfaces. The framework maintained smooth stress-field continuity and gradual variation of the predicted UBC across changing layer thicknesses and strength contrasts. This characteristic is particularly important for practical engineering applications, where abrupt discontinuities in predicted bearing capacity may lead to either unconservative or overly conservative foundation designs. By embedding governing physical laws directly into the optimization process, the proposed PINN framework demonstrates robust extrapolation capability beyond the immediate range of the training data. This significantly enhances the reliability of the model for practical geotechnical applications involving complex layered soil conditions and uncertain site characteristics. The combination of high predictive fidelity, computational efficiency, and strong physical consistency establishes the proposed PINN as a powerful and reliable tool for both deterministic and reliability-based foundation design analyses.

4. Discussion

The results demonstrate that the proposed Physics-Informed Neural Network (PINN) provides an accurate, physically consistent, and computationally efficient framework for predicting the ultimate bearing capacity (UBC) of layered soils subjected to combined vertical, horizontal, and moment loading. The high prediction accuracy and low generalization error confirm the capability of the model to capture the complex nonlinear interaction among soil stratigraphy, strength contrast, and loading conditions. More importantly, the incorporation of physics-based constraints within the learning process significantly improves model robustness by restricting the solution space to mechanically admissible responses. This characteristic is particularly important in geotechnical engineering applications, where limited datasets and extrapolation beyond the training domain are common.

Analysis of the prediction errors indicates that the largest deviations occur near transitions between competing failure mechanisms, particularly during the shift from punching shear to general shear failure. These regions are associated with abrupt stress redistribution and evolving plastic flow patterns, which are inherently difficult to represent using purely data-driven models. Nevertheless, the relatively small error magnitude observed in these transition zones demonstrates the effectiveness of the embedded equilibrium and Mohr–Coulomb constraints in stabilizing the learning process and improving nonlinear response prediction.

The predicted V–H–M failure envelopes successfully reproduce the fundamental characteristics of soil–foundation interaction under combined loading. The model captures the progressive reduction in vertical bearing capacity with increasing horizontal load and applied moment, while also representing the nonlinear coupling among loading components. The influence of moment loading was observed to be more pronounced in granular soils due to their higher stiffness and stress sensitivity. In contrast to conventional analytical approaches based on simplified interaction

relationships, the proposed PINN generates a continuous and smooth representation of the failure envelope over the entire loading domain.

The framework also accurately captures the behavior of layered soil systems, including thin-layer effects and failure mode transitions governed by layer thickness and strength contrast. In clay-over-sand systems, the transition from punching failure to general shear failure with increasing upper-layer thickness was effectively reproduced. Similarly, in sand-over-clay profiles, the reduction in bearing capacity induced by the weaker underlying layer was consistently predicted. The smooth variation of the predicted response across different failure regimes highlights the advantage of incorporating physical constraints, which enforce continuity and prevent unrealistic discontinuities commonly observed in empirical or piecewise analytical formulations.

From a computational standpoint, the proposed PINN offers a substantial advantage over conventional Finite Element Limit Analysis (FELA). Once trained, the model provides near-instantaneous predictions, enabling rapid parametric investigations, real-time design exploration, and large-scale reliability analyses. The computational efficiency of the framework makes it particularly attractive for Monte Carlo-based probabilistic assessments and reliability-based foundation design applications, which are typically computationally prohibitive using traditional numerical methods.

Despite these advantages, several limitations should be acknowledged. The model was trained exclusively using FELA-generated data, and therefore, the predictive capability remains dependent on the accuracy and assumptions of the underlying numerical simulations. In addition, the adoption of the Mohr–Coulomb constitutive model represents a simplified description of soil behavior, neglecting important mechanisms such as strain softening, anisotropy, and rate-dependent effects. Furthermore, the present formulation predicts only the ultimate bearing capacity rather than the complete stress or displacement fields, limiting the extent of direct physics enforcement within the continuum domain.

The current study is also restricted to two-layer soil systems under plane strain conditions. Practical foundation problems may involve three-dimensional effects, irregular stratigraphy, and more complex loading paths. Future research should therefore focus on experimental validation, extension to three-dimensional formulations, incorporation of advanced constitutive models, and integration of uncertainty quantification techniques. Development of fully field-based physics-informed formulations capable of directly predicting stress and deformation fields may further enhance the interpretability and predictive capability of the framework.

5. Limitations and directions for future research

Nevertheless, certain limitations associated with the application of the suggested Physics-Informed Neural Network need to be considered. First of all, it should be noted that all training was conducted exclusively based on data produced within the course of FELA analysis; thus, the results generated by PINN may be considered dependent on the assumptions made at the stage of numerical simulations. Besides, the employment of the Mohr-Coulomb failure criterion does not account for some phenomena, including, *inter alia*, strain softening, anisotropic effects, repetitive loading, and time-dependent processes. Future studies must thus focus on experiments and validation at a field scale, generalization into three-dimensional and heterogeneous soil systems, use of advanced constitutive relationships, and

application of uncertainty analysis and reliability methods. Advances in formulating purely field-based PINNs that predict stress and deformation fields would greatly improve the practicality of the proposed approach in solving complex geotechnical engineering problems.

6. Conclusion

This research introduced a Physics-Informed Neural Network (PINN) framework for predicting the ultimate bearing capacity of shallow foundations in layered soils under vertical, horizontal, and moment loading. The key findings and contributions are:

1. **Enhanced Accuracy and Generalization:** By embedding equilibrium equations and the Mohr-Coulomb yield criterion into the loss function, the PINN achieved significantly higher accuracy ($R^2 = 0.992$) compared to conventional data-driven neural networks, particularly in regions where failure mechanisms transition between layers.
2. **Computational Efficiency:** The trained PINN serves as a high-fidelity surrogate model, reducing prediction time from minutes (FELA) to milliseconds. This enables real-time design optimization and facilitates advanced probabilistic assessments that were previously computationally prohibitive.
3. **Physical Consistency:** The model inherently respects the fundamental laws of soil mechanics, ensuring that predictions are not only statistically plausible but also physically meaningful. This addresses a major critique of "black-box" machine learning applications in geotechnics.
4. **Comprehensive Failure Envelopes:** The PINN successfully mapped the complex interaction of V-H-M loading for two distinct layered soil configurations (sand-over-clay and clay-over-sand), providing a unified tool that replaces multiple empirical charts.

The proposed framework represents a physics-regularized surrogate framework in geotechnical analysis. It bridges the gap between the rigor of numerical methods and the speed of empirical methods. Future work will focus on extending the framework to three-dimensional loading, spatially variable soils, and time-dependent consolidation effects. The integration of such PINNs into building information modeling (BIM) and digital twin frameworks could revolutionize real-time foundation risk management.

Acknowledgement

The authors acknowledge the support of Dr. Krishna Kumar Shrestha for reviewing the laboratory test results of the samples.

Authors Contributions:

Mahendra Acharya: Initiated and conceptualized the research framework, developed the study methodology, performed data collection and curation, constructed and implemented the machine learning models (PINN), conducted

formal analysis, interpreted the findings, prepared the original manuscript draft, and participated in critical revision of the final manuscript.

Diwash Bhattarai & Bhishma Joshi: Contributed to the interpretation and discussion of findings, drafted key sections of the manuscript, and participated in critical revision of the final manuscript.

Funding

This research did not receive any funding and was conducted as part of academic research activities.

Data availability

Due to project confidentiality and ownership restrictions, the raw data are not publicly available but may be made available from the corresponding author upon reasonable request for academic and research purposes.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests, financial or non-financial, that could have influenced the work reported in this paper.

References

1. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
2. Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, 3, 422–440. <https://doi.org/10.1038/s42254-021-00314-5>
3. Terzaghi, K. (n.d.). *Theoretical soil mechanics*. Wiley. <https://doi.org/10.1002/9780470172766>
4. Michalowski, R. L., & Shi, L. (1995). Bearing capacity of footings over two-layer foundations. *Journal of Geotechnical Engineering*, 121(5), 421–428. [https://doi.org/10.1061/\(ASCE\)0733-9410\(1995\)121:5\(421\)](https://doi.org/10.1061/(ASCE)0733-9410(1995)121:5(421))
5. Sloan, S. W. (1988). Lower bound limit analysis using finite elements and linear programming. *International Journal for Numerical and Analytical Methods in Geomechanics*, 12(1), 61–77. <https://doi.org/10.1002/nag.1610120105>
6. Sheng, D., Sloan, S. W., Gens, A., & Smith, D. W. (2003). Finite element formulation and algorithms for unsaturated soils. Part I: Theory. *International Journal for Numerical and Analytical Methods in Geomechanics*, 27(8), 757–778. <https://doi.org/10.1002/nag.198>
7. Shahin, M. A., Maier, H. R., & Jaksa, M. B. (2002). Predicting settlement of shallow foundations using neural networks. *Journal of Geotechnical and Geoenvironmental Engineering (ASCE)*, 128(9), 785–793. [https://doi.org/10.1061/\(ASCE\)1090-0241\(2002\)128:9\(785\)](https://doi.org/10.1061/(ASCE)1090-0241(2002)128:9(785))

8. Xiaowei. (2020). Lecture notes (Fall 2020): Xiaowei. GitHub Pages. https://beiyulincs.github.io/teach/fall_2020/papers/xiaowei.pdf
9. Liu, C., Macedo, J. E., & Rodríguez, A. (2025). Leveraging physics-informed neural networks in geotechnical earthquake engineering: An assessment on seismic site response analyses. *Computers and Geotechnics*, 182, 107137. <https://doi.org/10.1016/j.compgeo.2025.107137>
10. Yuan, B., Heitor, A., Wang, H., & Chen, X. (2024). Physics-informed deep learning to solve three-dimensional terzaghi consolidation equation: forward and inverse problems. *arXiv preprint arXiv:2401.05439*. <https://doi.org/10.48550/arXiv.2401.05439>
11. Feng, Y., Eun, J., Kim, S. et al. Application of physics-informed neural networks (PINNs) solution to coupled thermal and hydraulic processes in silty sands. *Geo-Engineering* 16, 3 (2025). <https://doi.org/10.1186/s40703-025-00232-w>
12. Rahman, M.A., Zhang, T. & Lu, Y. PINN-CHK: physics-informed neural network for high-fidelity prediction of early-age cement hydration kinetics. *Neural Comput & Applic* 36, 13665–13687 (2024). <https://doi.org/10.1007/s00521-024-09791-y>