A PROBABILISTIC TRANSFORMER-BASED MACHINE LEARNING MODEL FOR THE PREDICTION OF SETTLEMENT FOR SHALLOW FOUNDATION

Raied Ahmed Nishatt[1], Abdun Nur Tusar[2], Shriful Islam[3,*], Moriyam Akter Monni[4]

[1] *Undergraduate Thesis Student, Shahjalal University of Science and Technology, Bangladesh, e-mail:* raiedahmednishat@gmail.com
[2] *Former Thesis Student, Shahjalal University of Science and Technology, Bangladesh, e-mail:* nur.tusar.2020@gmail.com
[3] *Professor, Shahjalal University of Science and Technology, Bangladesh, e-mail:* sharif_sust_cee@yahoo.com
[4] *Former Thesis Student, Shahjalal University of Science and Technology, Bangladesh, e-mail:* moriyamaktermonni@gmail.com

***Corresponding Author**

ABSTRACT

Many civil engineering foundations can undergo significant settlement which are not within the allowable limit. Accurate and reliable prediction of foundation settlement is critical in geotechnical engineering as excessive settlement of foundations can compromise safety and serviceability of structures. Empirical correlation used in practice are often built using simplified assumptions and can fail due to the high variability present in soil. Machine learning models provide a promising alternative to these limitations, with improved prediction capabilities while reducing reliance on costly site investigations and empirical approximations. This study presents the first application of Tablular Prior-data Fitted Network (TabPFN), a probabilistic transformer-based meta-learning model, for predicting shallow foundation settlement on granular soils. TabPFN follows a transformer-based architecture and is pre-trained on a variety of synthetic tasks using a meta-learning framework. This allows the model to perform accurately without requiring preprocessing and hyperparameter tuning. A database of 189 SPT-based case histories was used for the model development. Model inputs were foundation width (B), length (L), embedment depth (Df), net applied pressure (q) and average SPT blow count (N). The model output was the measured settlement (Sc). The TabPFN model achieved an R^2 value of 0.9567 on the training set and 0.9521 for the testing set. The RMSE values were 5.19 mm and 6.45 mm, and MAE values were 2.96 mm and 4.14 mm for the training and testing set, respectively. The model's performance was compared with other traditional machine learning models, and it was tested under reduced data conditions as well. The model consistently showed superior performance with stable predictions. The findings demonstrate TabPFN's potential as a reliable tool for predicting shallow foundation settlement on granular soils. Its ability to produce accurate prediction with minimal configuration makes it particularly suitable for practical application, offering possibility for further adoption of transformer-based approaches in geotechnical engineering.

Keywords: *Geotechnical Engineering, Machine Learning, TabPFN, Foundation settlement, Probabilistic Transformer Model,*

1. INTRODUCTION

Shallow foundations are one of the most important components of a structure used for safely transferring structural loads to the near-surface ground. An important factor for consideration when designing these structures, is the settlement of the soil. Settlement under foundations, even in allowable limits can limit structural stability by inducing differential settlement, cracking or long-term performance degradation (Burland & Burbidge, 1985; Peck et al., 1974). If not taken into account, settlement under foundations in granular soil can pose significant safety and maintenance concerns, prompting the need for accurate determination of settlement (Meyerhof, 1965; Terzaghi & Peck, 2013).

Laboratory testing for determining settlement require an undisturbed sample, which is nearly impossible or very costly to obtain from the field for granular soils (Stokoe et al., 2013). Empirical relations based on Standard Penetration Test (SPT) or Cone Penetration Test (CPT) are popular for settlement calculation because of their availability and simplicity (Parry, 1972; Peck et al., 1974). However, these empirical relationships often fail due to the variation in soil conditions even requiring extensive local calibration which is time and resource intensive (Soulié et al., 1990).

In recent years, artificial intelligence and machine learning (ML) has rapidly gained popularity in geotechnical engineering as a means for addressing these limitations. M. Shahin et al. (2002) and Erzin & Gul (2013) demonstrated the effectiveness of Artificial Neural Network (ANN) for settlement prediction. ANNs trained on SPT-based datasets significantly outperform conventionally used empirical methods. Later more modern algorithms such as Genetic Programming, Adaptive Neuro-Fuzzy Inference System and Support Vector Machine (SVM) have been used for the prediction of settlements of soils (Mohammed et al., 2020; Rezania & Javadi, 2007; Samui, 2008). Hybrid algorithms integrating optimization algorithms have also shown promise when developing models for settlement prediction. Song et al. (2022) optimized SVM with particle swarm optimization to predict settlement of foundation pits. Zhai (2022) used artificial bee colony optimization with autoregressive models for the prediction of road foundation settlement and reported significant accuracy gains over conventional models. Ensemble algorithms such as Random Forest, Gradient Boosting Machines, etc. have also been explored, producing promising results in this field (Liu et al., 2024). However, a key limitation to these models is that they require a large amount of high-quality data, with extensive pre-processing and hyperparameter tuning to achieve good predictive accuracy.

Deep learning architectures recurrent neural networks and transformers, though relatively new in the field of geotechnical engineering are showing promise in building more reliable and accurate predictive models. Hong et al. (2025) developed LSTM-transformer based hybrid models for the prediction of consolidation settlement in deep soft clay, with significant accuracy over empirical consolidation models. Li et al. (2024) applied LSTM-transformer models for railway foundation settlement prediction, which captured both temporal and cross-feature dependencies. Gao et al. (2024) developed "PCFFormer," a transformer-based model for embankment settlement prediction. These transformer-based models have strong potential for geotechnical application as they can perform very well with tabular data.

Despite their advantages, their adoption in predicting shallow foundation settlement remains limited. Existing research in this field has primarily focused on empirical correlation, neural network based models or ensemble models. But these models can lead to overfitting on small datasets often failing to capture the non-linear relationship present in granular soil.

This study addresses this gap by introducing the Tabular Prior-data Fitten Network (TabPFN), a probabilistic transformer-based meta-learning model, for settlement prediction of granular soil based on SPT-based data. By benchmarking against conventional ML models and commonly used empirical relations, and by evaluating the stability of the models under reduced data conditions this work provides a modern approach for shallow foundation settlement prediction.

2. METHODOLOGY

The research framework of this study is illustrated in Figure 1. The study starts with data preparation and preprocessing. The models are then trained and from the comparison of the four models, the best model is chosen. The best performing model is then validated through a detailed parametric study and benchmarked against previous published research to confirm the predictive superiority and reliability of the model.

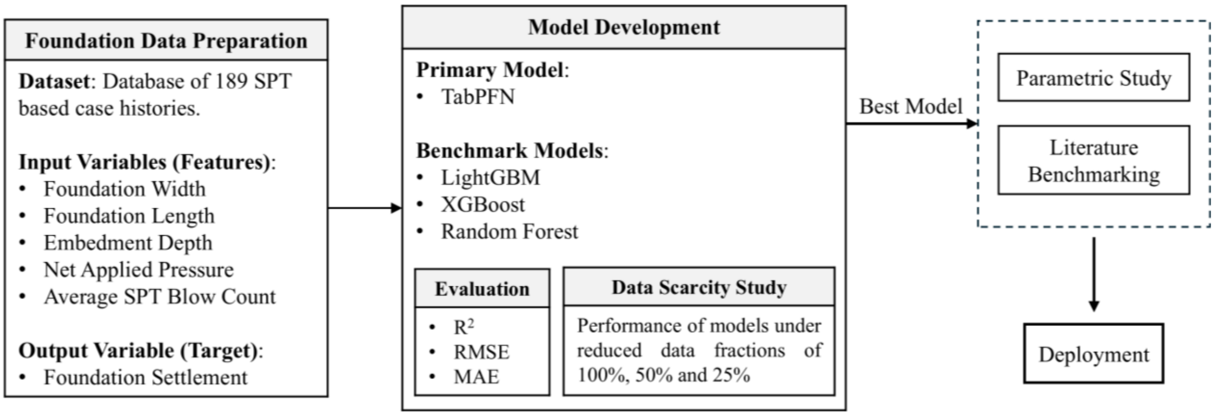

Figure 1. Methodological framework for model development and validation

2.1 Database Description

The SPT-based database used in this study was originally compiled by Shahin et al. (2002). It contains 189 individual case histories of shallow foundation settlement. The database contains a wide variety of soil types (silty sand, clayey sand, fine to coarse sand with gravel) and various foundation structures. The dataset was preprocessed by removing incomplete recording with missing values and the remaining 173 case histories were used for model development. The inputs in the model development were footing width (B), footing length (L), net applied pressure (q), embedment depth (D_f), and average Standard Penetration Test (SPT) blow count (N). The output variable was measured settlement (S_c) of the foundation. SPT-N values are representative of in-situ soil stiffness and strength, hence the use of SPT based dataset is highly relevant.

The statistical characteristics of the database are summarized in Table 1. The footing width ranges from 0.9 m to 55 m, with a mean of 7.62 m, while footing length varies between 0.9 m and 101 m (mean 16.13 m). The embedment depth spans from ground surface level to 10.7 m, with an average of 2.38 m. Applied pressures range between 18.3 kPa and 697 kPa, averaging 186.93 kPa, reflecting both lightly and heavily loaded footings. The SPT blow count varies widely, from 4 to 60, with a mean of 24.40, indicating soils from loose sands to dense granular deposits. The measured settlements display substantial variability, with values between 0.6 mm and 121 mm, and an average of 19.64 mm.

Table 1. Descriptive statistics of input and output variables

Parameter	Mean	Std. Dev.	Min	Median	Max
Footing width (m)	7.62	8.46	0.90	4.50	55.00
Footing length (m)	16.13	18.67	0.90	8.80	101.00
Applied pressure (kPa)	186.93	122.61	18.32	150.00	697.00
SPT blow count average (N)	24.40	13.57	4.00	20.00	60.00
Embedment depth (m)	2.38	1.88	0.00	2.10	10.70
Measured settlement (mm)	19.64	25.95	0.60	10.90	121.00

For model development, the database was partitioned into 141 training cases (81.5%) and 32 testing cases (18.5%). The split was performed randomly and preserved the statistical distributions of the variables.

2.2 TabPFN Model

Tabular Prior-data Fitted Network or TabPFN is a tabular foundation model that approaches supervised prediction as an approximate Bayesian inference problem solved by a single pretrained transformer. Instead of learning model parameters from each new dataset via iterative optimization, TabPFN is trained offline on a very large, diverse family of synthetic tabular tasks sampled from an explicit prior; during application it performs in-context learning by ingesting the training examples and producing predictions for the test set in a single forward pass (i.e., no further gradient updates) (Hollmann et al., 2022, 2025). Formally, the ideal Bayesian posterior predictive for a new input $x^*$is,

$$p(y^* \mid x^*, D) = \int p(y * \mid x *, \theta) p(\theta \mid D) d\theta \quad (1)$$

TabPFN learns a function $g_\omega(D, x^*)$ parameterized by pretrained weights ω that approximates this integral by mapping the entire dataset D and query x^* to a predictive distribution $\hat{p}(y^* \mid x^*, D)$, thereby approximating Bayesian model averaging without costly posterior sampling (Hollmann et al., 2022). The transformer backbone uses self-attention to encode sets of (feature, label) pairs so that structural patterns across examples are exploited during inference, allowing interval estimates and uncertainty quantification (Hollmann et al., 2022, 2025).

TabPFN differs from conventional supervised learners in three practical ways. First, it is a meta-learner: its inductive bias is specified by the synthetic prior used during pretraining rather than by a dataset-specific regularizer or architecture search. This enables strong performance on small to moderate tabular datasets with minimal or no hyperparameter tuning (Hollmann et al., 2022). Second, TabPFN's in-context strategy produces probabilistic outputs directly, which facilitates calibrated prediction intervals and principled uncertainty estimates that are valuable for engineering decision making. Third, inference is extremely fast where predictions for an entire test set occur in one forward pass, providing substantial computational advantages over iterative training pipelines and AutoML solutions (Hollmann et al., 2022, 2025).

2.3 Benchmark ML Models

2.3.1 XGBoost

XGBoost (eXtreme Gradient Boosting) is a high-performance implementation of Gradient Boosted Decision Trees (GBDT), designed with explicit regularization and system optimizations to reduce overfitting, improve scalability and performance (Chen & Guestrin, 2016). Like GBDTs, its prediction is of the form,

$$\hat{y}(x) = \sum_{k=1}^{K} f_k(x) \quad (2)$$

with each leaf tree f_k fitted to residuals of the loss plus regularization. The objective function in XGBoost contains both loss term and a regularization penalty $\Omega(f\_k)$, which includes terms for number of leaves, leaf weights, and optionally L1/L2 penalties on weights. This regularization distinguishes XGBoost from simpler boosting algorithms (Chen & Guestrin, 2016).

XGBoost uses second-order (gradient and Hessian) information to approximate the optimal leaf weight or split gain, which enhances convergence and supports complex models with better control over overfitting. Additional features include sparse feature handling, approximate split finding, shrinkage, and subsample/colsample parameters to reduce correlation between trees. In geotechnical engineering and related soil behavior modeling, XGBoost has been widely adopted and shown to outperform conventional empirical and simpler ML methods especially when the dataset includes complex nonlinear interactions among input parameters (Han et al., 2025; Huang et al., 2024).

2.3.2 LightGBM

LightGBM (Light Gradient Boosting Machine) is another tree-based ensemble algorithm belonging to the GBDT family. It builds models by sequentially adding decision trees, each attempting to correct errors made by preceding trees, using a differentiable loss function and gradient descent on that loss (Ke et al., 2017). LightGBM distinguishes itself from level-wise tree growth by using a leaf-wise (best-first) strategy: at each iteration, the leaf with the greatest potential gain is split, allowing more aggressive reduction in loss compared to standard level-by-level expansion, though this requires careful regularization to avoid overfitting. LightGBM undergoes several optimizations for enhancing model efficiency. One is histogram-based splitting, which bins continuous feature values into discrete buckets to accelerate split gain calculations. Another is Exclusive Feature Bundling (EFB), which reduces the number of effective features by combining mutually exclusive. A further technique is Gradient-Based One-Side Sampling (GOSS), which retains instances with large gradients (where errors are larger) and randomly subsamples from those with smaller gradients, to focus learning on harder cases (Ke et al., 2017). For regression tasks such as settlement prediction, the generic model form is,

$$\hat{y}(x) = \sum_{m=1}^{M} f_m(x) \tag{3}$$

where each f_m is a regression tree fitted to pseudo-residuals of the loss function. Regularization via limiting tree depth, minimum data per leaf, subsampling, and penalizing tree complexity help prevent overfitting. LightGBM has proven successful in many geotechnical and civil engineering prediction tasks for its speed, ability to handle large feature dimensions, and strong predictive performance when tuned (Demir & Sahin, 2023; Tran & Do, 2021).

2.3.3 Random Forest

Random Forest (RF) is a tree-based supervised ensemble learning algorithm designed to address the instability and overfitting issues of single decision trees (Breiman, 2001). It creates a collection of decision trees during training and aggregates their predictions to improve generalization. For regression tasks, the prediction $\hat{y}(x)$ for an input vector x is given by the average of the predictions from all trees:

$$\hat{y}(x) = \frac{1}{T}\sum_{t=1}^{T} h_t(x) \tag{4}$$

where T is the number of trees and $h_t(x)$ denotes the prediction of tree t.

Random Forest uses bootstrap aggregation and random selection which improves its generalization capabilities. In bootstrap aggregation (bagging), training data is divided into random subsets and decision trees are constructed on each of the subsets. This reduces variance and enhances the overall stability of the ensemble by averaging predictions across multiple trees. The predictor variable at each node split is selected at random. This introduces diversity among the trees, which reduces correlation effects and improves predictive accuracy (Biau & Scornet, 2016; Hastie et al., 2009). Unlike single decision trees, Random Forest models grow without pruning and the ensemble structure of the model helps in reducing overfitting.

2.4 Model Development and Evaluation

Models were developed in the Google Colab environment using Python. Required data analysis and machine learning libraries were used. A consistent train-test split was applied when training the models, ensuring a fair comparison. For hyperparameter optimization of benchmark models, Bayesian optimization using the Hyperopt library was used. For the TabPFN model, no hyperparameter tuning was required and the out of box performance was reported.

For model performance evaluation, Coefficient of Determination (R^2), Root Mean Squared Error (RMSE), Mean Squared Error (MSE) and Mean Absolute Error (MAE) were employed. These metrics quantify the agreement between model predictions and observed values on the test set.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{7}$$

For all metrics, y_i represents the actual observed value, $\hat{y}_i$ denotes the predicted value, $\bar{y}$ is the mean of the actual values, and n is the total number of observations.

3. RESULTS AND DISCUSSION

3.1 Performance Evaluation of all Models

The predictive performance of the developed models are summarized in Table 2. The TabPFN model demonstrated best predictive performance among all the trained models, achieving R^2 values of 0.9567 and 0.9521 for the training and testing set respectively. The corresponding error matrices were low as well, with RMSE of 5.19 mm for training and 6.45 mm for testing and MAE values of 2.96 mm and 4.15 mm.

Table 2. Performance evaluation of developed models

Model	**Train R²**	**Test R²**	**Train RMSE**	**Test RMSE**	**Train MAE**	**Test MAE**
TabPFN	0.9567	0.9521	5.19	6.45	2.96	4.15
LightGBM	0.9831	0.9439	3.24	6.98	1.92	4.59
XGBoost	0.9675	0.9299	4.50	7.80	2.74	5.19
Random Forest	0.8723	0.7832	8.91	13.72	5.07	8.02

In comparison, the tree-based ensemble models used for benchmarking showed mixed behavior. LightGBM achieved the high accuracy on training data an R^2 of 0.9831 and the very low errors with RMSE of 3.24 mm and MAE of 1.92 mm. However, its performance deteriorated on the test set with R^2 of 0.9439, RMSE of 6.98 mm and MAE score of 4.59 mm, indicating that the model was overfitting on the training data and did not generalize well. A similar trend was observed for XGBoost, which maintained high training accuracy with R^2 of 0.9675 and RMSE of 4.50 mm but suffered from a large drop in testing performance with R^2 value reduced to 0.9299 and RMSE of 7.80 mm. These highlight the tendency of boosting-based methods to overfit on small datasets, particularly in problems with high soil variability. Random Forest exhibited significantly lower predictive performance compared to the other models with training R^2 of 0.8723 and high errors (RMSE = 8.91 mm, MAE = 5.07 mm). The testing performance of RF declined even further showing R^2 of 0.7832, RMSE of 13.72 mm and MAE of 8.02 mm. This suggests that the averaging mechanism of Random Forest may be unstable for complex nonlinear interactions in the input variables.

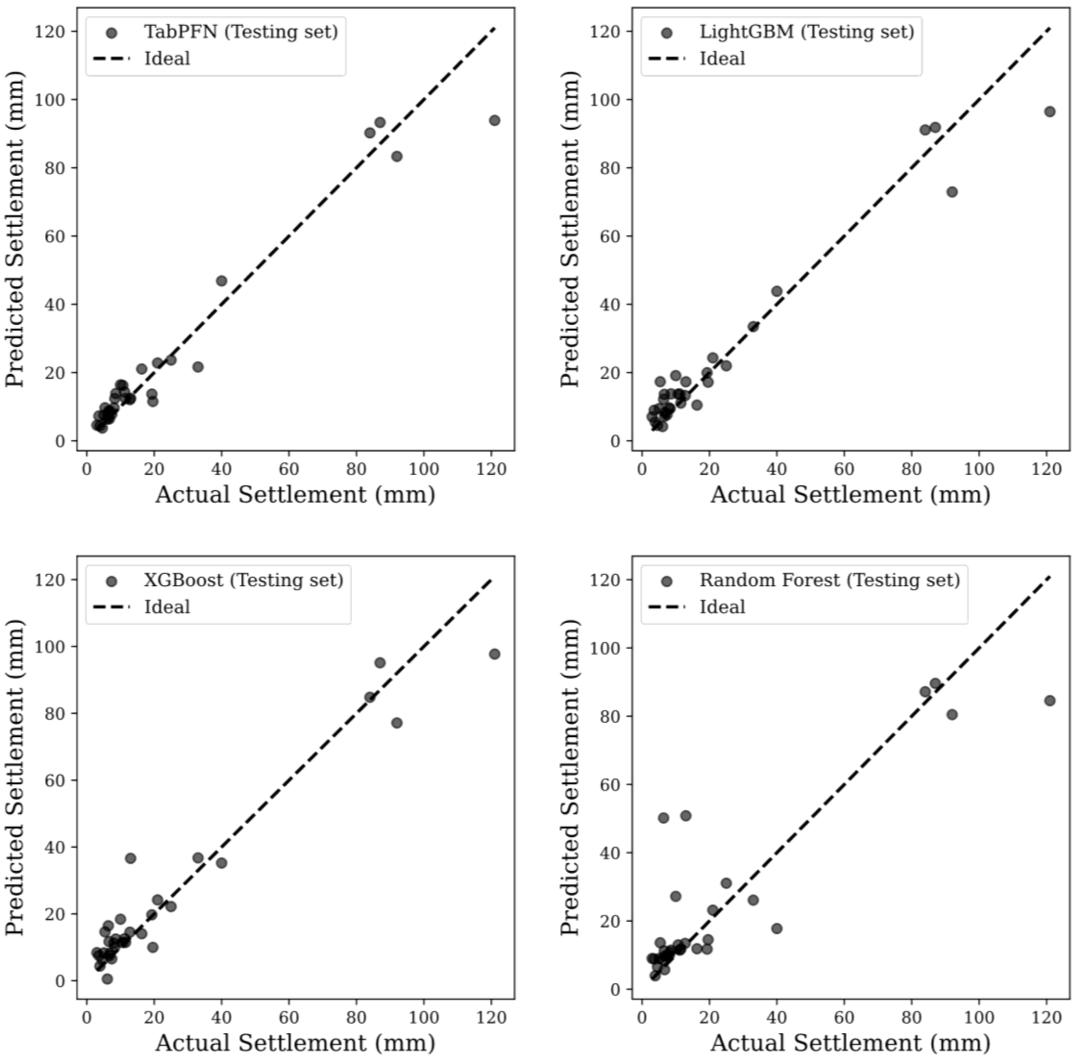

Figure 2. Actual vs. Predicted scatter plots for developed models on testing set.

The actual vs. predicted scatter plots for all trained models on the testing set is shown in Figure 2. TabPFN model had the best fit with points closely clustered around the ideal line. LightGBM, XGBoost and Random Forest all showed progressively greater dispersion in the scatter, which is consistent with their lower R^2 value and higher error matrices. The overall generalization capability of TabPFN was higher than the tree-based benchmarking models. This can be attributed to it's probabilistic transformer-based architecture which uses meta-learned prior, reducing its reliance on hyperparameter tuning and even preprocessing steps. Such characteristics are useful for geotechnical data where datasets are usually limited in size and are expensive to obtain. Unlike the conventional models, TabPFN doesn't require careful parameter calibration, highlighting its suitability in practical application in settlement prediction.

3.2 Effect of Training Data Size

The stability of the trained models under reduced data conditions were checked by systematically reducing the amount of training data used to train the model. The comparison of the R^2 and RMSE scores of the models trained with 100%, 50% and 25% of the training data are shown in Figure 3. The error bars represent the 95% confidence interval after 15 trial runs by randomly sampling the data.

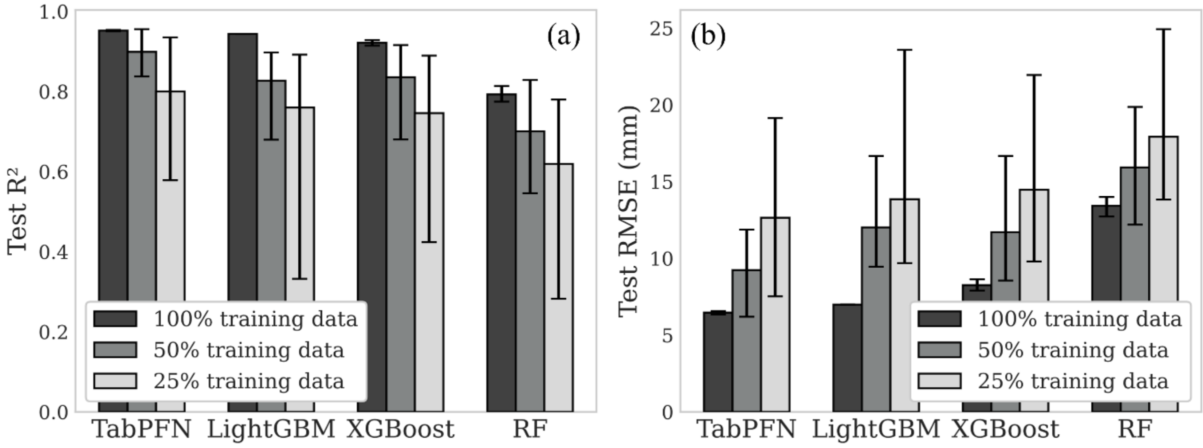

Figure 3. Bar chart of (a) R^2 and (b) RMSE scores of all models trained at different percentages of training data evaluated on the test set.

TabPFN consistently maintains high predictive accuracy with narrow error ranges across all training percentages. Even when 25% of the training data was used, TabPFN showed an average R^2 value of 0.8, with smaller error bars than the other models. This stability under limited data conditions is particularly useful in geotechnical engineering as field and laboratory datasets are often not available and expensive to obtain.

LighGBM achieved performance close to TabPFN but the error ranges for this model trained on 50% and 25% training data were significantly higher than that of TabPFN. This suggests that LightGBM has greater sensitivity to data reduction. This can be attributed to LightGBM's leaf-wise tree growth and hyperparameter sensitivity, which can amplify its variance under reduced data conditions. XGBoost demonstrated a lower performance than LightGBM but it had more stable predictions in reduced data conditions with the error ranges smaller than LightGBM. Random Forest performed the weakest with sharp decline in the R^2 score and highest RMSE as well as broader error range as data percentage was reduced. This degradation in performance is due to its ensemble averaging, which is less effective when training data is limited.

3.3 Parametric Study of TabPFN Model

A parametric study was conducted on the trained TabPFN model to investigate how settlement responds to variations in the input features. To ensure comparability among variables with differing units and scales, all inputs were normalized to a [0,1] range prior to analysis. Figure 3 presents the variation in predicted settlement as a function of normalized input features for the TabPFN model. Settlement decreases sharply with increasing SPT blow count (N) in the low-N range, after which the curve flattens which is consistant with realistic soil behavior. Loose soils (low N) are highly compressible, whereas the marginal benefit of increasing stiffness diminishes in dense soils.

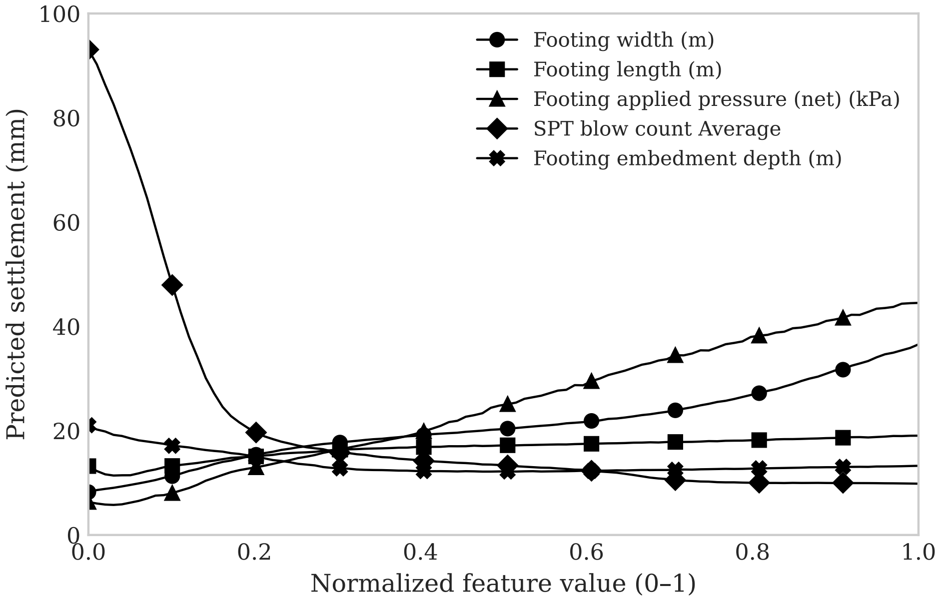

Figure 4. Results of parametric study of TabPFN: effect of feature variables on foundation settlement.

Applied footing pressure (q) shows a near-linear positive relationship with settlement across its normalized range. Footing width and length both exhibit positive relation with settlement, although their slopes are less pronounced compared to N or q. Embedment depth shows noticeable settlement reduction at lower normalized values but plateaus beyond mid-range, suggesting diminishing benefits with deeper embedment. These parametric trends align with prior findings in the literature (Ngo & Tran, 2024; Rezania & Javadi, 2007; Shahin et al., 2002). The nonlinear trends captured here, particularly the sensitivity of settlement to SPT-N in weak soils, highlights the capability of the TabPFN model to learn complex patterns in soil, that are less apparent in conventional models.

3.4 Comparison with Literature Benchmarks

Table 3 presents a comparison of the proposed TabPFN model with previously published machine learning and empirical approaches for shallow foundation settlement prediction using this dataset.

Table 3. Comparison of test performance of current study with existing literature

Reference	Model	R^2	RMSE	MAE
Current Study	TabPFN	0.95	6.45	4.14
Mohammed et al. (2020)	ANFIS-PSO	0.86	9.02	6.50
Shahnazari et al. (2014)	EPR	0.87	9.53	6.88
Rezania & Javadi (2007)	GP	0.95	6.86	4.92
Shahin et al. (2002)	ANN	0.92	9.20	6.65
Schmertmann et al. (1978)	Formula	0.34	26.72	18.2
Schultze & Sherif (1973)	Formula	0.61	20.56	10.62
Meyerhof (1965)	Formula	0.34	26.64	19.49

The proposed TabPFN model achieved a test R^2 of 0.95, with RMSE and MAE of 6.45 mm and 4.14 mm, respectively. This performance is superior to all the published machine learning models and substantially better than conventional empirical correlations. Compared to the Genetic Programming (GP) model reported by Rezania & Javadi (2007), which achieved a similar R^2 value of 0.95, TabPFN demonstrated lower error magnitudes with RMSE of 6.45 mm and MAE of 4.92 mm.

Other models such as ANFIS-PSO, EPR and ANN, reported R^2 values consist between the range of 0.86 and 0.92, with high error compared to the TabPFN model. Empirical relations proposed by Schmertmann et al. (1978) and Meyerhof (1965) showed significantly weaker prediction accuracy with R^2 values as low as 0.34 and high RMSE values. These models are widely adopted in practice but fail to take into account some key aspects of soil behavior, which explain their poor behavior. This comparison makes it clear that machine learning models significantly outperform the empirical relations, prompting the need for these intelligent methods but they still have some limitations. These models lose predictive accuracy when data is limited and high variability in the geotechnical data is present. In contrast, TabPFN excellently captures these limitations of low data availability and site-specific data variability, returning error matrices that are meaningful to foundation design. The model is able to predict settlement within a few millimeters of observed values, showing high reliability and offering engineers a probabilistic transformer based alternative to overcome the limitations posed by traditional methods.

4. CONCLUSIONS

The study introduces TabPFN, a probabilistic transformer-based machine learning model for settlement prediction of shallow foundations on granular soil. The model was benchmarked against modern ensemble models, previously published machine learning performances in literature and empirical relations. TabPFN consistently outperformed all the compared models with testing R^2 score of 0.95 and low error matrices with RMSE of 6.45 mm and MAE of 4.14 mm.

The model is able to deliver stable performance even when trained on 50% and 25% of the training data. The benchmarking models all exhibit performance degradation with large error margin when trained under this reduced data scenario. This stability under reduced data further highlights the applicability of this model in practice where high-quality data is limited and hard to obtain.

A parametric study was conducted on the trained model and it reported a geotechnically consistent trend among the parameters. TabPFN also outperformed previously reported models found in literature proposing a significant improvement over traditional methods.

Despite these strengths, there are some limitations to this study. The SPT-based dataset used in this study is of moderate size and focused on granular soils. A broader validation integrating diverse soil types and loading conditions is required to confirm the generalization of the model across different field conditions. Furthermore, TabPFN provides probabilistic predictions. Full calibration of these uncertainty estimates for more reliable predictions remains open for research. Future work should focus on incorporating larger datasets and developing more optimized transformer-based frameworks.

ACKNOWLEDGEMENTS

The authors thank Dr. Mohammad Rezania for giving access to the dataset used in this study, originally prepared by Dr. Mohamed Shahin. We also appreciate the scientific research community that is working to make data resources available for continuing transparent scientific progress.

DECLERATION OF USE OF AI

No AI tools were used for the preparation of this manuscript.

REFERENCES

Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, *25*(2), 197–227. https://doi.org/10.1007/S11749-016-0481-7/FIGURES/4

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Burland, J. B., & Burbidge, M. C. (1985). Settlement of foundations on sand and gravel. *Proceedings of the Institution of Civil Engineers (London)*, *78*(pt 1), 1325–1381. https://doi.org/10.1680/IICEP.1985.1058

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Demir, S., & Sahin, E. K. (2023). Predicting occurrence of liquefaction-induced lateral spreading using gradient boosting algorithms integrated with particle swarm optimization: PSO-XGBoost, PSO-LightGBM, and PSO-CatBoost. *Acta Geotechnica*, *18*(6), 3403–3419. https://doi.org/10.1007/S11440-022-01777-1/TABLES/8

Erzin, Y., & Gul, T. O. (2013). The use of neural networks for the prediction of the settlement of pad footings on cohesionless soils based on standard penetration test. *Geomechanics and Engineering*, *5*(6), 541–564. https://doi.org/10.12989/gae.2013.5.6.541

Gao, S., Chen, C., Jiang, X., Zhu, S., Cai, H., & Li, W. (2024). Transformer-based settlement prediction model of pile composite foundation under embankment loading. *Computers and Geotechnics*, *176*, 106783. https://doi.org/10.1016/j.compgeo.2024.106783

Han, X., Yi, J., Li, X., Li, S., Tang, H., Wang, Z., & Ran, J. (2025). Application of extreme gradient boosting for predicting standard penetration test N-values from cone penetration test data. *Bulletin of Engineering Geology and the Environment*, *84*(4), 1–18. https://doi.org/10.1007/s10064-025-04219-w

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning* (2nd ed.). Springer New York. https://doi.org/10.1007/978-0-387-84858-7

Hollmann, N., Müller, S., Eggensperger, K., & Hutter, F. (2022). TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second. *11th International Conference on Learning Representations, ICLR 2023*. https://arxiv.org/pdf/2207.01848

Hollmann, N., Müller, S., Purucker, L., Krishnakumar, A., Körfer, M., Hoo, S. Bin, Schirrmeister, R. T., & Hutter, F. (2025). Accurate predictions on small data with a tabular foundation model. *Nature*, *637*(8045), 319–326. https://doi.org/10.1038/s41586-024-08328-6

Hong, S., Chung, T. K., Yoo, B. S., & Kim, S. R. (2025). Innovative LSTM-transformer hybrid models for predicting consolidation settlement. *Acta Geotechnica*, 1–27. https://doi.org/10.1007/S11440-025-02752-2/FIGURES/21

Huang, G. ;, Liu, Z. ;, Wang, Y. ;, Yang, Y., Huang, G., Liu, Z., Wang, Y., & Yang, Y. (2024). A Multi-Objective Prediction XGBoost Model for Predicting Ground Settlement, Station Settlement, and Pit Deformation Induced by Ultra-Deep Foundation Construction. *Buildings 2024, Vol. 14, Page 2996*, *14*(9), 2996. https://doi.org/10.3390/BUILDINGS14092996

Huong Ngo, T. T., & Tran, V. Q. (2024). Predicting and evaluating settlement of shallow foundation using machine learning approach. *Science Progress*, *107*(4). https://doi.org/10.1177/00368504241302972

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Advances in Neural Information Processing Systems 30 (NIPS 2017)*. https://github.com/Microsoft/LightGBM.

Li, Z., Peng, Y., Li, J., & Tang, Z. (2024). Composite Foundation Settlement Prediction Based on LSTM–Transformer Model for CFG. *Applied Sciences*, *14*(2), 732. https://doi.org/10.3390/app14020732

Liu, H., Su, H., Sun, L., & Dias-da-Costa, D. (2024). State-of-the-art review on the use of AI-enhanced computational mechanics in geotechnical engineering. *Artificial Intelligence Review*, *57*(8), 196. https://doi.org/10.1007/s10462-024-10836-w

Meyerhof, G. G. (1965). Shallow Foundations. *Journal of the Soil Mechanics and Foundations Division*, *91*(2), 21–31. https://doi.org/10.1061/JSFEAQ.0000719

Mohammed, M., Sharafati, A., Al-Ansari, N., & Yaseen, Z. M. (2020). Shallow Foundation Settlement Quantification: Application of Hybridized Adaptive Neuro-Fuzzy Inference System Model. *Advances in Civil Engineering*, *2020*(1). https://doi.org/10.1155/2020/7381617

Parry, R. H. (1972). A direct method of estimating settlements in sand from SPT values. *Symposium on Interaction of Structure and Foundations, 1972, Midlands Soil Mechanics and Foundation Engineering Society, United Kingdom*.

Peck, R. B. ., Hanson, W. Edmund., & Thornburn, T. H. . (1974). *Foundation engineering*. Wiley.

Rezania, M., & Javadi, A. A. (2007). A new genetic programming model for predicting settlement of shallow foundations. *Canadian Geotechnical Journal*, *44*(12), 1462–1473. https://doi.org/10.1139/T07-063

Samui, P. (2008). Support vector machine applied to settlement of shallow foundations on cohesionless soils. *Computers and Geotechnics*, *35*(3), 419–427. https://doi.org/10.1016/j.compgeo.2007.06.014

Schmertmann, J. H., Hartman, J. P., & Brown, P. R. (1978). Improved Strain Influence Factor Diagrams. *Journal of the Geotechnical Engineering Division*, *104*(8), 1131–1135. https://doi.org/10.1061/AJGEB6.0000683

Schultze, E., & Sherif, G. (1973). Prediction of settlements from evaluated settlement observations for sand. *Proceedings Eighth International Conference on Soil Mechanics and Foundation Engineering*, *1*(3), 225–230.

Shahin, M. A., Maier, H. R., & Jaksa, M. B. (2002). Artificial Neural Network–Based Settlement Prediction Formula for Shallow Foundations on Granular Soils. *Australian Geomechanics*, *37*, 45–52.

Shahin, M., Maier, H. R., & Jaksa, M. B. (2002). Predicting Settlement of Shallow Foundations using Neural Networks. *Journal of Geotechnical and Geoenvironmental Engineering*, *128*(9), 785–793. https://doi.org/10.1061/(ASCE)1090-0241(2002)128:9(785)

Shahnazari, H., Shahin, M., & Tutunchian, M. (2014). Evolutionary-based approaches for settlement prediction of shallow foundations on cohesionless soils. *International Journal of Civil Engineering*, *12*, 55–64.

Song, Z., Liu, S., Jiang, M., & Yao, S. (2022). Research on the Settlement Prediction Model of Foundation Pit Based on the Improved PSO-SVM Model. *Scientific Programming*, *2022*, 1–9. https://doi.org/10.1155/2022/1921378

Soulié, M., Montes, P., & Silvestri, V. (1990). Modelling spatial variability of soil parameters. *Canadian Geotechnical Journal*, *27*(5), 617–630. https://doi.org/10.1139/t90-076

Stokoe, K. H., Kacar, O., & Van Pelt, J. (2013). Predicting settlements of shallow footings on granular soil using nonlinear dynamic soil properties. *Proceedings of 18th International Conference on Soil Mechanics and Geotechnical Engineering*, 3467–3470.

Terzaghi, K., & Peck, R. B. (2013). *Soil mechanics in engineering practice*. Read Books Ltd.

Tran, V. Q., & Do, H. Q. (2021). Prediction of California Bearing Ratio (CBR) of Stabilized Expansive Soils with Agricultural and Industrial Waste Using Light Gradient Boosting Machine. *Journal of Science and Transport Technology*, *1*(1), 1–8. https://doi.org/10.58845/JSTT.UTT.2021.EN.1.1.1-8

Zhai, J. (2022). Machine-Learning-Based Road Soft Soil Foundation Treatment and Settlement Prediction Method. *Scientific Programming*, *2022*, 1–7. https://doi.org/10.1155/2022/3463413