XLIII CILAMCE
21-25, November, 2022
FOZ DO IGUAÇU - PR - BRAZIL

Machine Learning Applied to Predict Pile Bearing Capacity

Yago F. Gomes[1], Dimas B. Ribeiro[1]

[1]*Dept. of Geotechnics, Aeronautics Institute of Technology*
Praça Marechal Eduardo Gomes, 50, Vila das Acácias, 12228-900, São José dos Campos, SP, Brazil
yago@ita.br, dimasbetioli@gmail.com

Abstract. The present work presents the application of machine learning algorithms to the problem of predicting load capacity of piles. A dataset was compiled from the literature, composed by 165 load tests associated with SPT results performed in different regions of Brazil. From this raw base, 5 datasets were generated based on well-known semi-empirical methods. Such sets were then applied to six ML algorithms and a linear regression. The performances, measured by R^2 and $RMSE$, were compared to those achieved by semi-empirical methods. The RF technique stood out from the others, with a maximum R^2 of 0.77. A case study was then carried out and its results reinforced the good performance of ML algorithms against semi-empirical methods. Despite the limitations of the work regarding the dataset, the conclusions point to the use of ML tools as a good alternative to the classical methods of calculating load capacity.

Keywords: Machine learning, SPT, Precast concrete piles, Bearing capacity

1 Introduction

Foundation elements are a fundamental part of engineering projects, and the calculation of its bearing capacity is essential to design safe constructions. The best method to obtain the bearing capacity of piles is to measure the capacity of one or more foundation elements in the construction site through load tests, using standards such as ABNT NBR 6122[1]. However, as this method has a high cost of time and money, some approaches have been developed to predict pile capacity with lower costs. Some of these alternatives are the semi-empirical methods, which are based on basic soil information such as SPT results. Some of the most known and used semi-empirical methods are the ones from Decourt and Quaresma (1978)[2], Teixeira (1996)[3], Aoki and Velloso (1975)[4] and Meyerhof (1976)[5]. Although these methods are widespread and present good results, new technologies that are emerging may be able to better meet the needs of designers.

For the past years, Machine Learning (ML) techniques have been applied in several geotechnical problems such as slope stability (Ferentinou and Sakellariou, 2007[6]), soil classification (Carvalho and Ribeiro, 2019[7]) and foundation settlement (Samui, 2008 [8]. In addition to these applications, one can find in the literature studies exploring pile capacity prediction using these computational tools. However, there are still many gaps to be filled in this research field. Some of these gaps refer to the reduced size of datasets, poor variety of soil types and the application of few ML techniques.

This work seeks to explore some of these gaps through the application of ML techniques for predicting the bearing capacity of precast concrete piles. The study uses a dataset composed by 165 static load tests associated with SPT soundings and taken from several Brazillian regions. Using this raw data, five datasets were obtained based on the inputs used by well known semi-empirical methods. These datasets were used to train different ML algorithms and test the resulting models. The leave-one-out cross validation approach was used to split data into train and test sets, for its ability to better explore the information within the dataset.

The results obtained from ML models are first compared with the ones from the semi-empirical methods. This comparison was made using the coefficient of determination (R^2) and root-mean-square error ($RMSE$) as metrics to evaluate performance. After this analysis, the top three ML techniques and all semi-empirical methods are used for a case study, using examples not included in the initial dataset. This sought to evaluate the method's performance when exposed to new information. Next, an extra analysis was performed only with the Random Forest technique, which was the one with best performance. Information about the soil type obtained from SPT test was added to the initial dataset, with the objective of evaluating the sensitivity of the performance to this new information.

2 Semi-Empirical Methods

Semi-empirical methods are based on empirical correlations between in situ test results and values obtained from load tests. According to Cintra (2010)[9], these methods present better results than purely theoretical methods which are based on equilibrium equations, like Terzaghi (1943)[10]. For its calculation, the geometric characteristics of the foundation, its type, executive method and soil parameters are considered. The latter can be obtained through tests such as the SPT.

Basically, the total resistance is the sum of lateral friction (R_l) and tip resistance (R_p). These will depend on the foundation geometry such as its cross section perimeter (U_l), length (L), tip area (A_p) and soil parameters. Equation 1 presents this sum.

$$R_t = R_l + R_p = U_l L r_l + A_p r_p \tag{1}$$

The r_l and r_p terms are the lateral unit resistance and the tip unit resistance, respectively. These variables are not the same for different semi-empirical methods. Some of the best known and most used methods are those of Decourt and Quaresma (1978)[2], Teixeira (1996)[3], Aoki and Velloso (1975)[4] and Meyerhof (1976)[5], which are explored in this work. Each of these methods have different equations and ways of using the SPT test result. As these methods are well known and explored in the literature, details about each of them will be omitted in this work.

3 ML Techniques

According to Alpaydin (2020) [11], ML techniques seek to optimize the performance of a certain function through examples or known experiences, using statistical theories to create mathematical models. This approach is widely used for cases in which the relationship between input and output variables of a problem is not known, but there are data and examples about it.

In the present work, techniques of this type were used to generate a regression capable of relating the inputs extracted from the raw dataset to the load capacity of piles. Six ML algorithms were used: k nearest neighbors (KNN), kernel k nearest neighbors (KKNN), decision trees (DT), random forests (RF), artificial neural networks (ANN) and support vector machines (SVM). These were selected because of their wide application and reasonable performance in the literature, as well as their different biases. An multiple linear regression (LR) was also included in the analysis as a baseline for the ML techniques performance.

A brief description of each method used in this work is presented below. More details about these techniques are widely explored in the literature and are not part of the scope of this work.

3.1 KNN and KKNN

KNN uses a space where each dimension corresponds to an input of the problem. The inputs of an example are interpreted as coordinates of a point that represents the example. In this context, KNN assumes that points that are close to each other are similar. Thus, one way of estimating the output of an unknown example within a cloud of known examples is to use the output of its closest neighbors (Song et al., 2017).

The KKNN technique, on the other hand, uses this same concept, but maps the examples in a n-high dimension, with the objective of improving the performance of the technique for more complex problems (Kuo et al., 2008[12]). A kernel (concept from the SVM technique) is used to map points into a higher dimension space.

3.2 DT and RF

The DT is a flow-chart-like model with nodes that create ramifications and divide the dataset using rules. It starts with the complete dataset and splits it to other nodes through rules that are usually inequalities applied to the inputs. This process continues until the last nodes, called leafs, assign outputs for the example.

RFs are a collection of DTs randomly build, by selecting different subsets for the trees and nodes. This process minimizes the overfit problem that DTs can present. Each DT can make a different prediction for the example, and the mean of these predictions gives the final value for RF (Ho, 1995 [13]).

3.3 ANN

This technique is based on the functioning of the neurological system of intelligent organisms that acquires knowledge through experience.

It structure is composed by a number of nodes (neurons) arranged in layers: an input layer, an output layer and one or more hidden layers. The nodes from the first layer receive the inputs (signals), multiply and adjust them with weights. Then, a threshold is added and a activation function is applied to produce an output. This output can be the ANN final prediction or the input for another neuron (Moselhi et al., 1992 [14]).

3.4 SVM

SVMs use statistical learning principles to minimize the errors associated with the training dataset and maximize the model generalization (Vapnik, 1999 [15]). This model uses a space where the inputs define coordinates and points represent examples. Then, a hyperplane with margins is proposed as the model prediction result. This hyperplane is optimized by the distance to the known examples and smaller margins means smaller errors.

This technique uses the kernel trick to map points into a higher dimensional space, making possible to achieve better performances for non linear problems.

3.5 LR

A multiple linear regression was used as a baseline for the ML techniques performance. For inputs x_1, x_2, ..., x_n and an output y, a linear function calibrates the coefficients $\beta_{0...n}$ as presented by Equation 2:

$$y = \beta_0 + x_1.\beta_1 + x_2.\beta_2 + ... + x_n.\beta_n \ . \tag{2}$$

4 Dataset

The raw data used in this work is composed by 165 precast concrete pile load test associated with SPT sounding results collected from the works of Lobo (2005)[16], Vianna (2000)[17] and Santos (1988)[18]. These load tests were carried out in different locations in Brazil and were all perfomed according to the Brazilian Standard ABNT NBR 6122/2010 [1].

Figure 1 shows the sounding locations and, as can be seen, they were taken mostly from the southeast and south regions of Brazil. The country has a tropical climate and high temperatures predominance, and 65% of its territory is composed by non homogeneous lateritic soils. Thus, the clay-ferruginous soil is the most common soil type (Morais, 2020 [19]). It is also worth mentioning that the most predominant soil from the dataset soundings is the Silty-Clay, representing 45% of the total.

After assembling the raw dataset a pre-processing was carried out, in order to create datasets to be submitted to the ML models. This work proposes five different datasets: four based on the semi-empirical methods from section 2 and one compiling the inputs from the previous four. This choice of inputs was made in order to use the experience of the authors of semi-empirical methods to select a set of suitable inputs to be applied to ML techniques. In addition, it is worth comparing the same algorithm being trained and tested with different selections of inputs.

Each dataset (except for the last one) is composed by four inputs: the pile diameter (D), its length (L), a SPT average for the pile shaft (SPT_l) and for the pile tip (SPT_p). The difference between the datasets lies in the SPT_l and SPT_p values.This is because the pile geometry is the same for all, but each semi-empirical method calculates its average shaft and tip SPT in a different way, as mentioned in Section 2.

To simplify, the names of the datasets created from the semi-empirical methods were summarized to acronyms that refer to their authors. For the sets created from Teixeira, Decourt and Quaresma, Aoki and Velloso and Mayerhof, the acronyms Tx, Dq, Av and Mey were used, respectively. Furthermore, All refers to the set resulting from the joining of all inputs. The name of an ML technique with one of the above acronyms refers to the technique trained with that dataset

Next step was to split the datasets into training and testing sets. For this, the present work used the leave-one-out cross-validation approach, for its ability to better explore the dataset information. In this technique the full

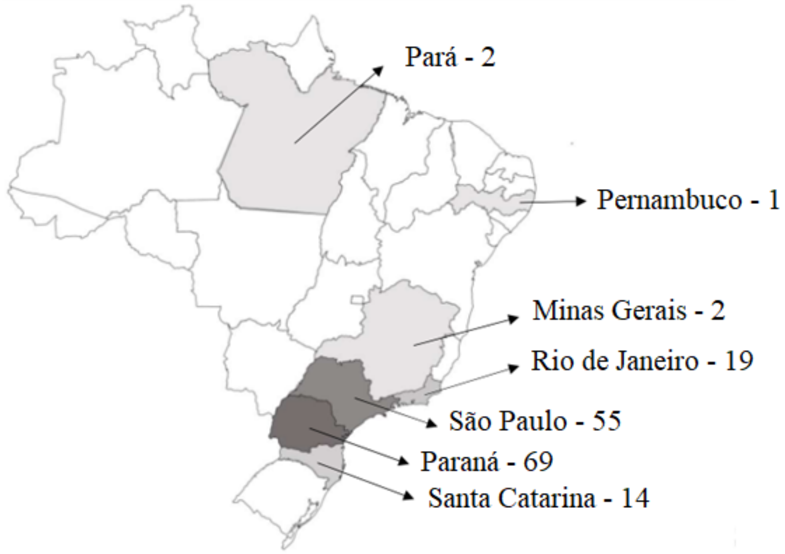

Figure 1. Pile load test locations.

dataset is used for training and one example is separated for testing. This process is repeated for each example until all examples are used as test. Finally, the average accuracy results in the model overall performance.

5 Results

After building the datasets from the collected raw dataset, they were submitted to the ML techniques and linear regression. In addition, predicted values were also calculated according to the semi-empirical methods. The performances were measured as a function of the coefficient of determination R^2 and $RMSE$.

Initially, the load capacities were calculated according to semi-empirical methods, in order to obtain baseline performances to be compared to the ML techniques. Table 1 presents the R^2 and $RMSE$ values for these techniques.

Table 1. Semi-empirical methods performance

Método	R²	RMSE (kN)
Tx	0.755	1431.9
Dq	0.749	897.3
Mey	0.660	888.7
Av	0.614	915.1

Among these classic models, the ones that stood out were those of Teixeira and Decourt and Quaresma, with R^2 close to 0.75.

Then, each of the five generated datasets were subjected to each of the ML techniques and LR. Table 2 presents the results obtained. The columns represent the semi-empirical model used as a basis for the dataset and the rows the ML techniques applied to these sets. Shaded regions highlight the best performing algorithm for each dataset.

The RF technique stood out for presenting the best performance for each of the data sets, with the exception of Aoki and Velloso dataset. The highest value of R^2 reached by this technique was 0.77, when using the set of inputs based on Teixeira.

In addition, the ANN and KNN techniques also had outstanding performance, reaching R^2 close to 0.75. In addition to the relatively good performance compared to semi-empirical methods, they have a very low computational cost, which is an advantage over the RF technique.

Table 2. ML algorithms performance

	Tx		Dq		Mey		Av		All	
algoritmo	R²	RMSE	R²	RMSE	R²	RMSE	R²	RMSE	R²	RMSE
RF	0.770	636.0	0.765	642.4	0.760	649.9	0.738	677.3	0.762	645.6
KNN	0.732	687.3	0.762	651.2	0.741	679.9	0.749	670.2	0.756	664.3
ANN	0.748	664.6	0.752	659.1	0.751	660.4	0.750	662.7	0.730	690.3
KKNN	0.724	698.1	0.750	665.4	0.742	676.3	0.730	696.0	0.722	712.6
SVM	0.732	696.7	0.744	686.2	0.728	704.6	0.703	735.8	0.731	701.9
LR	0.723	695.9	0.728	690.3	0.724	696.0	0.726	692.7	0.715	712.6
DT	0.737	679.6	0.717	704.4	0.715	706.0	0.710	713.4	0.733	684.2

As expected, the worst performance was achieved by linear regression. This may suggest that this technique is equivalent to classical methodologies. Nevertheless, among regression models, those based on ML may still be better alternatives to the problem.

It is interesting to note that the best performing semi-empirical methods were those whose datasets applied to ML models achieved better results. This may indicate that the selection of inputs made by Teixeira and Decourt and Quaresma has a greater capacity to explain the behavior of the pile, or at least for those in this work's dataset.

In a second moment, a case study was carried out with three new examples. None of these examples had been used before in the training or testing phase of the models. This analysis sought to evaluate the behavior of the techniques against a real example of application.

First, the load capacities were calculated according to semi-empirical methods. These results are presented in Table 3, which also contains the load capacity obtained by the load test and the average error of these methods.

Table 3. Semi-empirical method's predicted values and absolute mean error for case study

Pile ID	Qu	Tx	Dq	Av	Mey
120	680	285,5	488,2	646,9	922,0
121	815	200,6	452,3	437,15	558,2
122	710	143,8	399,0	294,7	477,8
Absolute mean error		71,0%	38,8%	36,2%	33,3%

As for the ML models, the top three best techniques were applied: RF, ANN and KNN. The results obtained for the five data sets submitted to these three techniques are presented in Table 4, as well as the resulting mean errors.

In a last moment, information about the predominant soil type in the tip and lateral regions of the pile was added to the dataset. The intervals considered to obtain the predominant type of soil were the same used by the authors of the semi-empirical methods to obtain the mean SPT values for the tip and shaft.

As this new input is categorical, that is, non-numerical, the only ML technique (among those used in this work) capable of using these two types of input data at the same time is the RF.

The results for the case study and R^2 for each new dataset obtained by RF are presented in Table 5

6 Conclusions

This work applied ML algorithms to predict the load capacity of precast concrete piles, using data collected from the literature. Inputs were selected from semi-empirical methods, considering the most popular ones in Brazil and worldwide. At first, some ML techniques were applied and their performances were compared to those of classical calculation methods. The algorithm that stood out was the RF, reaching the highest R^2, surpassing all semi-empirical methods. The ANN and KNN models also achieved reasonable performances, comparable to the best-performing semi-empirical method, that of Teixeira. These results indicate that ML algorithms can be a good

Table 4. ML techniques' predicted values and absolute mean error for case study

Algorithm		N° Estaca 120	121	122	Absolute mean error
Tx	rf	928,8	901,8	650,1	18,6 %
	knn	787,2	672,6	621,4	15,2 %
	brnn	983,7	834,9	705,1	15,9%
Dq	rf	858,6	900,6	692,5	13,1%
	knn	844,0	656,7	650,4	17,3%
	brnn	1062,0	818,0	733,0	19,9%
Mey	rf	921,6	896,7	638,4	18,5%
	knn	844,0	656,7	650,4	17,3%
	brnn	1046,8	805,0	707,1	18,5%
Av	rf	901,2	913,2	631,1	18,6%
	knn	1121,3	571,1	627,0	35,3%
	brnn	989,1	848,4	713,9	16,7%
All	rf	863,1	857,2	642,4	13,9%
	knn	1005,4	715,6	640,6	23,3%
	brnn	1021,3	823,4	742,8	18,6%

Table 5. RF with new datasets' predicted values and absolute mean error for case study

	Pile number 120	121	122	Absolute mean error	R^2
Tx	909.3	887.3	654.4	16.8%	0.763
Dq	877.0	834.8	659.0	12.9%	0.758
Mey	855.5	860.2	625.2	14.4%	0.760
Av	902.8	759.3	603.0	18.2%	0.743
All	897.2	877.4	626.7	17.1%	0.755

alternative to the problem of predicting the load capacity of piles. This conclusion is supported by other studies found in the literature.

The case study performed also reinforces the ability of ML models to be used for the given problem. The average errors achieved by these techniques were lower than those of the semi-empirical methods when exposed to new examples.

In addition, the analysis of the RF model trained with the dataset containing soil type information was not very conclusive. Despite a slight decrease in the mean error for the case study examples, R^2 did not present major changes.

An important limitation of this work is the size of the dataset. It is well known that ML models perform only as well as the quality and size of the dataset they use for testing and training. Therefore, the authors hope that the dataset used in this work (available via Zenodo at `https://zenodo.org/record/6600964#.Ypa-PqjMKUk`) will be expanded and used in future works.

The authors also suggest for future works to explore other combinations of inputs to the problem and also other types of data, such as CPT survey results.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] ABNT. Projeto e execução de fundações. In *Assiciação Brasileira de Normas Técnicas*. ABNT, Rio de Janeiro, Brazil, 2010.

[2] L. Décourt and A. Quaresma. Capacidade de carga de estacas a partir de valores de spt. In *Congresso Brasileiro de Mecânica dos Solos e Engenharia de Fundações*, pp. 45–53. Associação Brasileira de Mecânica dos Solos e Engenharia Geotécnica, 1978.

[3] A. H. Teixeira. Projeto e execução de fundações. In *Anais do SEFE*, volume 1, pp. 377, 1996.

[4] N. Aoki and D. Velloso. An approximate method to estimate the bearing capacity of piles. In *Proc., 5th Pan-American Conf. of Soil Mechanics and Foundation Engineering*, pp. 367–376, Buenos Aires, Argentina. International Society of Soil Mechanics and Geotechnical Engineering, 1975.

[5] G. Meyerhof. Bearing capacity and settlements of pile foundations. *Journal of Geotechnical Engineering*, vol. 102, pp. 197–228, 1976.

[6] M. Ferentinou and M. Sakellariou. Computational intelligence tools for the prediction of slope performance. *Computers and Geotechnics*, vol. 34, n. 5, pp. 362–384, 2007.

[7] L. O. Carvalho and D. B. Ribeiro. Soil classification system from cone penetration test data applying distance-based machine learning algorithms. *Soils and Rocks*, vol. 42, n. 2, pp. 167–178, 2019.

[8] P. Samui. Support vector machine applied to settlement of shallow foundations on cohesionless soils. *Computers and Geotechnics*, vol. 35, n. 3, pp. 419–427, 2008.

[9] J. C. A. Cintra and N. Aoki. *Fundações por estacas: projeto geotécnico*. Myron C. Clark, São Paulo, Brasil, 2010.

[10] K. Terzaghi. *Theoretical Soil Mechanics*. John Wiley and Sons, New York, 1943.

[11] E. Alpaydin. *Introduction to machine learning*. MIT press, 2020.

[12] B. Kuo, J. Yang, T. Sheu, and S. Yang. Kernel-based knn and gaussian classifiers for hyperspectral image classification. In *2008 International Geoscience and Remote Sensing Symposium*, pp. 1006–1008. IEEE, 2008.

[13] T. K. Ho. Random decision forests. In *Proceedings of* $3^r d$ *International Conference on Document Analysis and Recognition*, volume 1, pp. 278–282, 1995.

[14] O. Moselhi, T. Hegazy, and P. Fazio. Potential applications of neural networks in construction. *Canadian Journal of Civil Engineering*, vol. 19, n. 3, pp. 521–529, 1992.

[15] V. N. Vapnik. An overview of statistical learning theory. *IEEE transactions on neural networks*, vol. 10, n. 5, pp. 988–999, 1999.

[16] B. Lobo. *Método de previsão de capacidade de carga de estacas: aplicação dos conceitos de energia do ensaio SPT*. PhD thesis, Universidade Federal do Rio Grande do Sul, Escola de Engenharia, Porto Alegre, Brasil, 2005.

[17] A. Vianna. Análise de provas de carga estática em estacas pré-moldadas cravadas na cidade de curitiba e região metripolitana. M.S. thesis, Universidade de São Paulo, São Carlos, Brasil, 2000.

[18] O. Santos Jr. Previsão do comportamento carga-recalque de estacas pré-moldadas de concreto. M.S. thesis, Universidade de São Paulo, São Carlos, Brasil, 1988.

[19] T. S. O. Morais, C. H. C. Tsuha, L. A. Bandeira Neto, and R. M. Singh. Effects of seasonal variations on the thermal response of energy piles in an unsaturated brazilian tropical soil. *Energy and Buildings*, vol. 216, pp. 1 – 19, 2020.