

Modelo em Aprendizagem de Máquina para Determinação da Capacidade de Carga em Fundações Diretas

Willian Lino Sousa

Graduado em Engenharia Civil, Universidade Federal do Ceará, Crateús, Brasil, willianlinosousa@alu.ufc.br

Marcio Avelino de Medeiros

Professor, Universidade Federal do Ceará, Crateús, Brasil, marcio.medeiros@crateus.ufc.br

Lívio Antônio Melo Freire

Cientista de dados, Petrobrás, Rio de Janeiro, Brasil, livio@crateus.ufc.br

RESUMO: O problema de estimar a capacidade de carga é essencial na engenharia civil, uma vez que o dimensionamento adequado das fundações é essencial para garantir a segurança e a estabilidade das estruturas. Estudos recentes vem empregando modelos em aprendizagem de máquina em problemas de fundações demonstrando significativa acurácia. O presente estudo buscou desenvolver um modelo para previsão da capacidade de carga em fundações diretas assentadas em diversos tipos de solos. Os resultados obtidos na revisão bibliográfica acerca do tema destacam a eficácia do uso de técnicas de machine learning em fundações diretas. Estudos mostraram que os modelos de aprendizado de máquina são capazes de fornecer estimativas precisas e confiáveis, superando as abordagens tradicionais baseadas em métodos empíricos. Além disso, a literatura também exploraram a influência de diferentes parâmetros geotécnicos e variáveis de entrada como: SPT, tipo de solo, cota de assentamento da sapata, presença do nível de água, entre outros, no desempenho dos modelos de machine learning. Eles mostraram que o uso de variáveis geotécnicas relevantes, como resistência ao cisalhamento do solo e características de compactação, pode melhorar significativamente a precisão dos dados de saída. O modelo criado foi calibrado com base na literatura já existente e foi determinado a partir de técnicas em Machine Learning, com o uso do algoritmo random forest, para previsão da capacidade de carga a partir da alimentação da máquina com os dados de entrada.

PALAVRAS-CHAVE: Aprendizagem de Máquina, Capacidade de Carga, Fundações Diretas.

ABSTRACT: The problem of estimating load capacity is essential in civil engineering since adequate dimensioning of foundations is essential to guarantee the safety and stability of structures. Recent studies have been using machine learning models in foundation problems, demonstrating significant accuracy. The present study sought to develop a model for predicting the load capacity of direct foundations laid in different types of soil. The results obtained in the accurate literature review on the topic highlight the effectiveness of using machine learning techniques in direct foundations. Studies have shown that machine learning models are capable of providing accurate and reliable estimates, outperforming traditional approaches based on empirical methods. Furthermore, the literature also explored the influence of different geotechnical parameters and input variables such as: SPT, soil type, footing installation level, presence of water level, among others, on the performance of machine learning models. They showed that the use of relevant geotechnical variables, such as soil shear strength and compaction characteristics, can significantly improve the accuracy of the output data. The model created was calibrated based on existing literature and was determined using Machine Learning techniques, using the random forest algorithm, to predict the load capacity based on feeding the machine with input data.

KEYWORDS: Machine Learning, Bearing Capacity, Direct Foudations.

1 INTRODUÇÃO

A Engenharia Geotécnica, especificamente na área de fundações, têm se atualizado cada vez mais na busca de modelos computacionais com maior acurácia em seus resultados. Ao longo dos últimos anos, com a disponibilidade maior de dados de campo, bem como o avanço computacional conquistado, o desenvolvimento de modelos em Machine Learning (aprendizagem de máquina) torna-se uma ferramenta poderosa na previsão de capacidade de carga de fundações.

Machine Learning é uma disciplina da área da Inteligência Artificial que se concentra no desenvolvimento de algoritmos e modelos que permitem aos sistemas de computador aprender e aprimorar seu desempenho em tarefas específicas, sem a necessidade de programação explícita (Alpaydin, 2010.)

As técnicas de machine learning têm sido utilizadas em várias áreas da engenharia civil, principalmente na geotecnia, como para susceptibilidade de deslizamento de terra (Marjanović et al., 2011; Tien Bui et al., 2016), para previsão de parâmetros geotécnicos (Ching & Phoon, 2019; Fikret Kurnaz & Kaya, 2018; Prayogo & Susanto, 2018; Puri et al., 2018) e detecção de anomalias em barragens (Belcher et al., 2016; Fisher et al., 2017, 2016).

O aprendizado de máquina já vem sendo implantado na geotécnica em diversos setores e finalidades, como por exemplo o aplicativo Kroki, que calcula a capacidade de carga de acordo com a teoria e metodologia de Aoki-Velloso (1975), Décourt- Quaresma (1978), e Teixeira (1996).

O Kroki utiliza técnicas de aprendizado de máquina para automatizar parte desse processo. Através de um conjunto de dados de estacas previamente analisadas e informações geotécnicas, o aplicativo é capaz de aprender padrões e relações entre os dados. Com base nesse aprendizado, ele pode fazer estimativas e prever a capacidade de carga de novas estacas com base em suas características.

O trabalho de Oliveira, Yuri. et al. (2020), mostra como funciona a modelagem da ML no campo da geotecnia, no qual, apresentou uma máquina que fez estimativa de coeficiente de mola de estacas, usando como validação provas de cargas estáticas. Esse artigo utilizou para modelar sua máquina o modelo de aprendizagem Random Forest Regressor, no qual se baseia na construção de várias árvores de decisão.

Uma árvore de decisão é um modelo de aprendizado de máquina supervisionado que é frequentemente usado para resolver problemas de classificação e regressão. É uma representação gráfica de um conjunto de regras de decisão que são aplicadas sequencialmente para chegar a uma decisão final.

O objetivo do trabalho é a criação de um modelo em aprendizagem de máquina, usando dados de entrada já conhecidos que foram obtidos em boletins de sondagem, como os parâmetros de um solo, cota do nível de água, cota de assentamento da sapata, ensaios SPT, para calibrar um modelo em Machine Learning, para a previsão da capacidade de carga em fundações diretas assentadas em diversos tipos de solo.

2 METODOLOGIA

O estudo foi desenvolvido a partir da criação de um banco de dados sobre os parâmetros do solo, onde esse banco de dados foi utilizado para o treinamento de aprendizagem de máquina, que foi obtido os valores de capacidade de carga, como mostrado na Figura 1.

Para esse estudo foi criado um banco de dados sobre os parâmetros do solo obtidos em sondagens que foram retiradas de nuvens de armazenamento das licitações executadas de diversas prefeituras espalhadas pelo Brasil. Nessas sondagens vinham o valor SPT, tipo de solo, cota, nível de água e para complementar os dados foram calculados ângulo de atrito, peso específico do solo e peso da camada de solo, para todas as sondagens. Para o cálculo do ângulo de atrito foi utilizada a metodologia de Texeira (1996) como mostra a equação 1 e para coesão foi utilizada a metodologia de Godoy (1996) como mostra a equação 2.

$$\phi = 2\sqrt{20 \cdot N_{SPT} + 15} \quad (1)$$

$$C = 10 \cdot N_{spt} \quad (2)$$

2.1 Modelo em *Machine Learning*

Nesse trabalho foi utilizado o algoritmo random forest pois são dados tabulares e poucos dados, esse algoritmo é adequado para essa configuração de dados. Para as sondagens foi usado um total de noventa e

cinco boletins de sondagens. Posteriormente, os dados foram particionados em dois conjuntos: um destinado ao treinamento e seleção do modelo de aprendizado de máquina, e outro para o teste e avaliação da capacidade de generalização do modelo. Essa separação foi realizada de maneira que setenta e cinco por cento das sondagens foram utilizadas para treinamento e vinte e cinco por cento para testes.

Para o funcionamento do modelo construído, é necessário solicitar os dados de entrada, que são tipo de solo, SPT, cota e peso da camada de solo acima da sapata, os demais parâmetros são previstos pelo modelo.

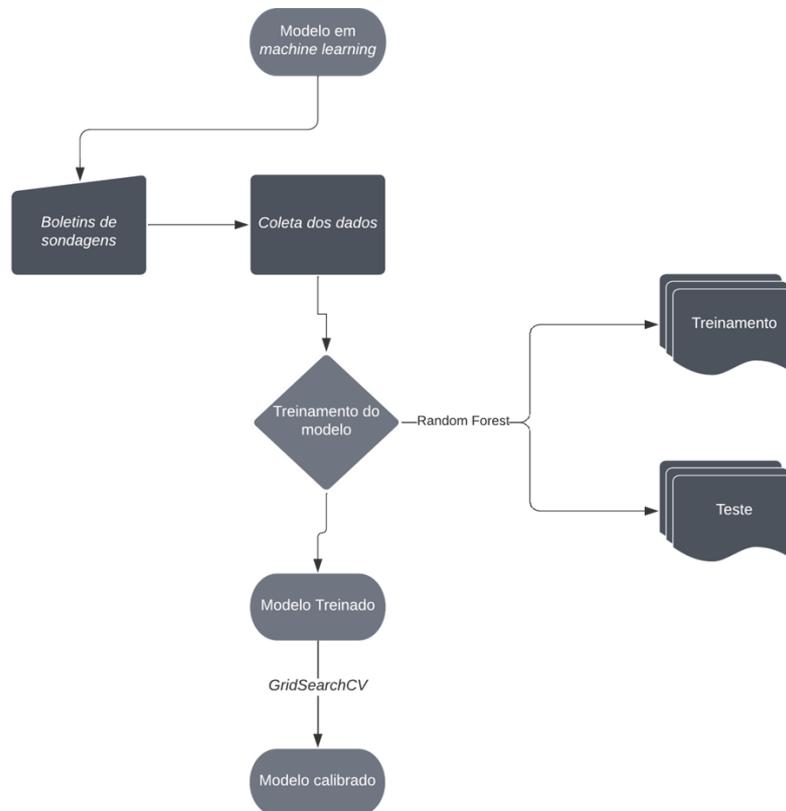


Figura 1. Fluxograma dos passos do trabalho

2.2 Treinamento do modelo

O valor de saída do modelo em machine learning foi a capacidade de carga de sapata com geometria previamente definida, que para esse modelo são sapatas quadradas ($L=B$) como mostrado na Figura 2, dentre as dimensões possíveis: 1,00 x 1,00 m, 1,50 x 1,50 m, 2,00 x 2,00 m, 2,50 x 2,50 m e 3,00 x 3,00 m.

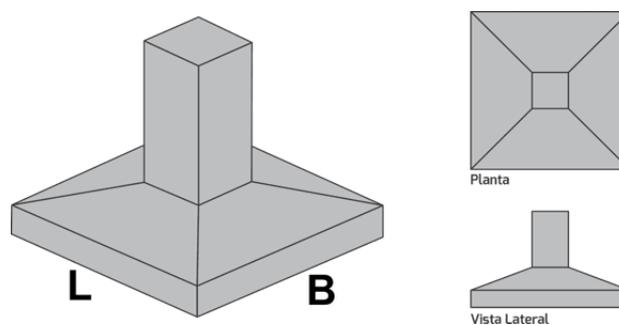


Figura 2. Esquema da sapata utilizada no modelo

2.3 Validação Cruzada utilizando K -folds

Na validação cruzada utilizando k -folds, o conjunto de dados é dividido em k seguimentos. Dessa maneira, um classificador é treinado usando $k-1$ partes da amostra e um valor de erro é calculado testando o

classificador no segmento restante – repetindo o procedimento até que todas as amostras sejam usadas como teste. Por fim, a estimativa k-cv do erro é o valor médio dos erros cometidos em cada iteração (RODRIGUEZ; PEREZ; LOZANO; 2009). A Figura 3 ilustra esse processo.

ITERAÇÃO 1	TESTE	TREINO	TREINO	TREINO	TREINO
ITERAÇÃO 2	TREINO	TESTE	TREINO	TREINO	TREINO
ITERAÇÃO 3	TREINO	TREINO	TESTE	TREINO	TREINO
ITERAÇÃO 4	TREINO	TREINO	TREINO	TESTE	TREINO
ITERAÇÃO 5	TREINO	TREINO	TREINO	TREINO	TESTE

Figura 3. Validação Cruzada para 5 folds

3 RESULTADOS E DISCUSSÃO

Após o treinamento do modelo, na etapa pré-processamentos dos dados foi feito a plotagem dos gráficos valor predito pelo modelo vs. valor real, onde foi comparado o valor predito com o valor real e foi gerado uma reta aproximada para descrever a linearidade dos resultados que visa traçar uma reta onde consegue explicar de forma mais exata a distribuição dos dados por meio de uma reta, para todas as geometrias definidas. Na Figura 4 é possível observar os resultados do modelo em *machine learning* obtido para a capacidade de caga das sapatas quadradas.

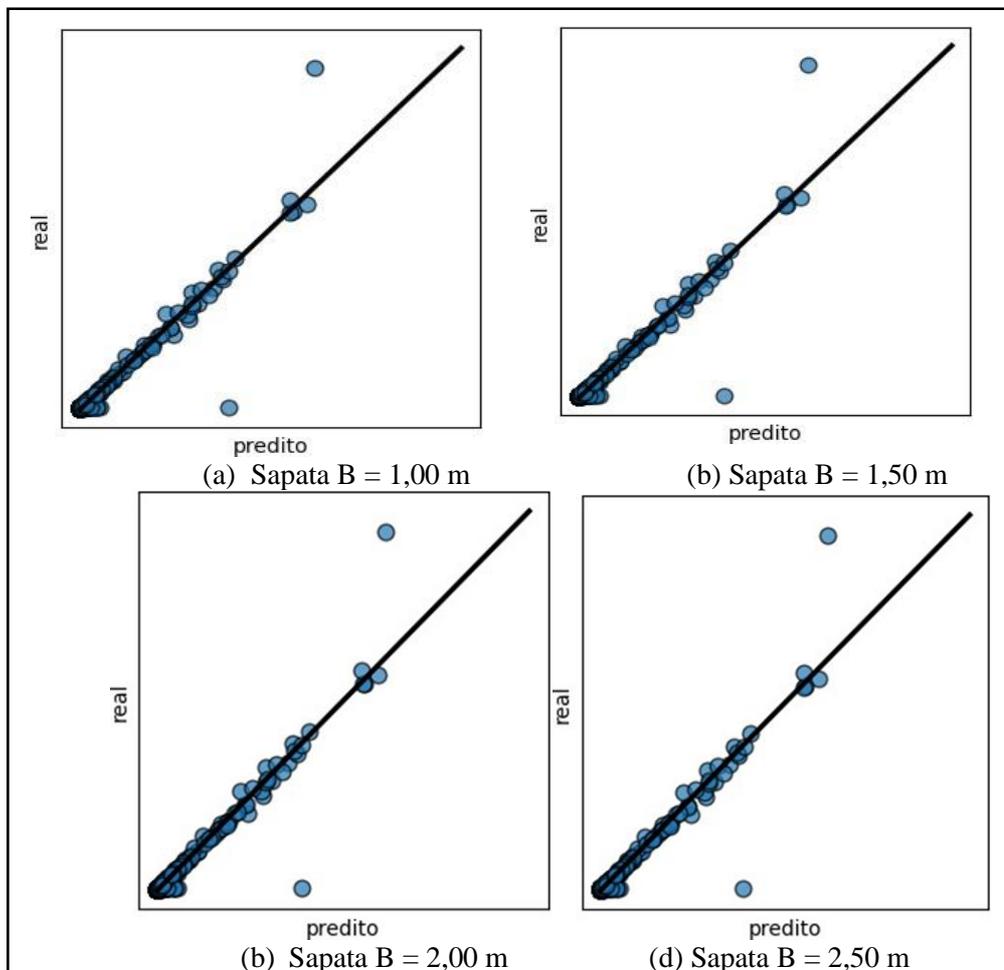


Figura 4. Relação dos valores obtidos para as sapatas de 1,0 a 2,50 m de base

Já na Figura 5 é apresentado o modelo relacional obtido para sapata de maior dimensão. Bem como observado na Figura 4, para a sapata de maior dimensão também pode-se verificar poucos pontos *outliers*, o que indica uma boa correlação obtida pelo modelo em aprendizagem de máquina.

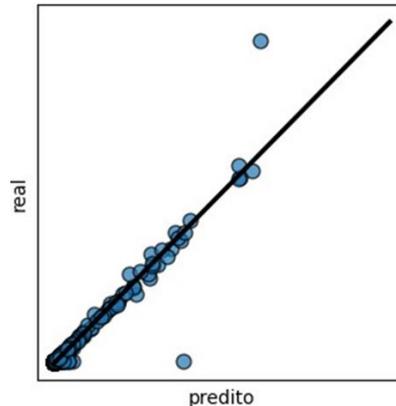


Figura 5. Correlação obtida para a sapata com B = 3,00 m

Os resultados demonstrados avaliam a performance do modelo nas amostras de teste, onde dados conhecidos são usados para prever a capacidade de carga em fundações diretas. No entanto, para atribuir maior confiança às previsões, foi utilizada a biblioteca GridSearchCV, para uma otimização das previsões também na fase de treinamento, utilizando $k = 5$ folds.

Foi utilizada, ainda, a função GridSearchCV para otimização dos valores previstos pelo modelo, onde foram aplicados hiperparâmetros e escolhido o qual obteve o melhor resultado. Esses resultados são apresentados na Tabela 1.

Tabela 1. Hiperparâmetros testados	
Hiperparâmetro	Faixa
“N_estimators”	[200, 300, 400, 500]
“Max_features”	[‘sqrt’, ‘log2’]
“Max_depth”	[5, 10, 15, none]
“criterion”	[‘square_error’]
“Random_state”	[18]

Nessa etapa o GridSearchCV foi executado, obteve-se os melhores hiperparâmetros para uma melhor previsão dos valores de capacidade de carga: {'criterion': 'squared_error', 'max_depth': 10, 'max_features': 'sqrt', 'n_estimators': 400, 'random_state': 18, cv=GroupKFold (n splits=5)}.

Na etapa de pós-processamento dos dados com o uso do GridSearchCV foi possível fazer a plotagem dos gráficos valor predito pelo modelo X valor real, onde foi comparado o valor predito com o valor real e foi gerado uma reta aproximada para descrever a linearidade dos resultados que visa traçar uma reta onde consegue explicar de forma mais exata a distribuição dos dados por meio de uma reta, para todas as geometrias definidas. Para os poucos valores que ficaram fora da reta, foram valores que o modelo não conseguiu aprender na fase de treinamento. Nas Figura 6 são apresentados os resultados comparativos do valor obtido pelo modelo vs. o valor considerado como real, após a aplicação da função.

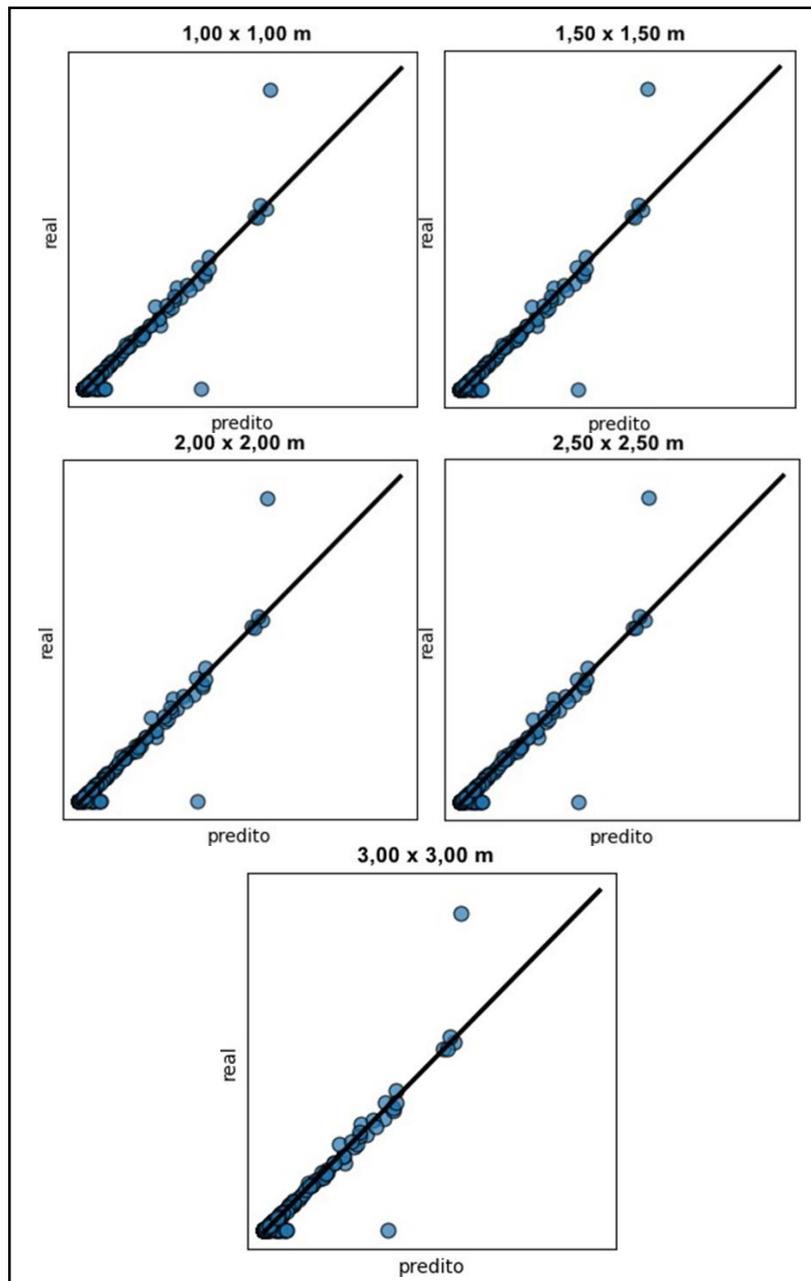


Figura 6. Valor predito vs. valor real do modelo obtido após a função GridSearchCV

Por meio da comparação dos valores preditos pelo modelo e o valor real, é possível afirmar que o modelo se ajustou com razoável precisão. Isso fica evidente ao observar a alta taxa de acertos e baixos desvios, o que significa que não houve uma grande diferença entre o valor real e o valor predito das amostras.

O erro absoluto (valor predito – valor real) e as métricas sobre o erro absoluto foi calculado. Os valores de capacidade de carga preditos pelo modelo foi comparado a um valor de capacidade de carga obtido a partir do método teórico de Terzaghi. As métricas são representadas em unidades de MPa na Tabela 2.

Tabela 2. Métricas calculadas para os resultados de capacidade de carga obtidas

Geometria	1,0 x 1,0 m	1,5 x 1,5 m	2,0 x 2,0 m	2,5 x 2,5 m	3,0 x 3,0 m
Média	4,24	4,26	4,28	4,31	4,33
Mediana	0,58	0,55	0,56	0,56	0,56
Desvio padrão	12,29	12,35	12,41	12,47	12,54

É possível notar que as precisões dos modelos são significativamente exatos para previsão de capacidade de carga com baixo valor da média dos erros, como também um valor pequeno de desvio padrão para as amostras analisadas, com poucas exceções. De fato, destaca-se uma diminuição das métricas de erro para valores de capacidade de carga acima de 1 MPa, o que comprova um modelo em machine learning bem calibrado e com resultados precisos.

4 CONCLUSÕES

Esse trabalho criou um modelo em machine learning para previsão de capacidade de carga em fundações diretas com geometrias definidas, que são de 1 a 3 metros. A partir do modelo treinado obtido, pode-se chegar às seguintes conclusões:

- i. O banco de dados foi criado com parâmetros do solo bem diversificado, que contemplam diversos tipos de solo, para uma melhor calibragem do modelo, pois um banco de dados robusto significa um treinamento mais eficiente.
- ii. A calibragem do modelo em machine learning para estimativa da capacidade de carga obteve resultados satisfatórios e precisos, pois os valores preditos pelo modelo foram similares aos valores reais em todas as geometrias de sapatas previstas.
- iii. As métricas de erro usadas para validação do modelo em machine learning teve resultados significativos, e foram importantes para declarar o modelo preciso.

REFERÊNCIAS BIBLIOGRÁFICAS

- Alpaydin, E. 2010. *Introduction to machine learning*. MIT press Cambridge.
- Associação Brasileira de Normas Técnicas (2023). *NBR 6122: Projetos e Execução de Fundações*, Rio de Janeiro, p. 8.
- Belcher, W., Camp, T., & Krzhizhanovskaya, V. V. (2016). Detecting Erosion Events in Earth Dam and Levee Passive Seismic Data with Clustering. *Proceedings - 2015 IEEE 14th International Conference on Machine Learning and Applications, ICMLA 2015*, 903–910.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Ching, J., & Phoon, K.-K. (2019). Constructing Site-Specific Multivariate Probability Distribution Model Using Bayesian Machine Learning. *Journal of Engineering Mechanics*, 145(1), 04018126.
- Cintra, J. C. A., Aoki, N., Albiero, J. H. (2011). *Fundações diretas - projeto geotécnico*. São Paulo : Oficina de Textos, 2011.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- Marjanović, M., Kovačević, M., Bajat, B., & Voženílek, V. (2011). Landslide susceptibility assessment using SVM machine learning algorithm. *Engineering Geology*, 123(3), 225–234.
- Oliveira, Y., Duboc, G., Aquino, F., Guimarães, M., Gonçalves, A., Oliveira, P. (2020). *Machine Learning aplicado na estimativa do coeficiente de mola de estacas com validação por meio de provas de carga estáticas*.
- Pinto, L. A. G., Da Silva, J. M.. Aprendizagem de máquina na determinação de parâmetros de resistência de filitos brandos. *Research, Society and Development*, v. 12, n. 1, p. e19012139693-e19012139693, 2023.
- Rodriguez, J. D.; Perez, A.; Lozano, J. A. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 32, n. 3, p. 569–575, 2009.
- Terzaghi, K., Peck, R. B., & Mesri, G. (1996). *Soil Mechanics in Engineering Practice*. Wiley.