XLI CILAMCE
XLI Ibero-Latin American Congress on Computational Methods in Engineering
16-19 November, 2020 FOZ DO IGUAÇU PARANÁ - BRAZIL

Concrete compressive strength prediction with machine learning

Pedro B. A. Moreira[1], Victor M. Silva[1]

[1]*Student, Dept. of Engineering, University Veiga de Almeida*
Avenida das Américas, 22631-004, Rio de Janeiro, Brazil
pedrobermoreira@gmail.com
[2]*Assistant Professor, Dept. of Engineering, IBMEC/RJ*
TODO: Address, Zip-Code, State/Province, Country
TODO: somebody3@somewhere.com

Abstract. TODO: REDUCE TO 200 WORDS, CURRENTRLY IT HAS 270. Compressive strength is the main characteristic of concrete. The correct prediction of this parameter means cost and time reduction. This work built predictive models for 6 different ages of concrete samples (3, 7, 14, 28, 56, and 100 days). A set of data obtained in previous studies was used, a total of 1030 samples, with 9 variables: compressive strength, age, and 7 ingredients (water, cement, fine aggregate, coarse aggregate, fly ash, blast furnace slag, and superplasticizers). Another 6 variables were added to represent the proportions of the main ingredients in each sample (water/cement, fine aggregate/cement, coarse aggregate/cement, fine aggregate/coarse aggregate, water/coarse aggregate, and water/fine aggregate). The predictive models were developed in R language, using the caret package with the Parallel Random Forest algorithm and repeated cross-validation technique to optimize the parameters. The results were satisfactory and compatible with other studies using the same data set. The most important model, 28 days old, obtained RMSE of 4.717. The 3-day model obtained the best result, RMSE of 3.310. The worst result was the 56-day model, with RMSE of 5.939. The work showed that the compressive strength of concrete can be predicted. The choice of creating a model for each age, instead of using age as a predictor, allowed to get compatible results with the available data at each age. It was a promising alternative since good results were achieved by training with just one algorithm. This work facilitates exploration and new efforts to predict the compressive strength of concrete, it can be used as a baseline to predict with different algorithms or the combination of several.

Keywords: Concrete, Compressive Strength, Machine Learning, Prediction

1 Introduction

Compressive strength is the main characteristic of concrete, measured by tests of international standards that consist of the breaking of specimens. Measurement at 28 days is mandatory and represents the grade of the concrete. Knowing in advance what the result will be obtained for a given age, based on the proportions of its ingredients, is of great interest to concrete manufacturers, construction companies, and civil engineers.

The compressive strength is a nonlinear function of its ingredients and age, making it difficult to establish an analytical formula, although some formulas have already been proposed. Hasan [1] proposed a mathematical model to predict from the results of tests of 7 and 14 days, and Kabir [2] from 7 days. However, machine learning techniques can be used to model this characteristic from real sample data, using only the ingredients.

Many previous studies use the same dataset used by Yeh [3] to predict the compressive strength of concrete. Alshamiri [4] got good results with the regularized extreme learning machine (RELM) technique, and Hameed [5] got even better results with the Artificial Neural Networks and cross-validation technique. This set of samples is so well known that there are many pages on the internet of unpublished studies that use it and have good results, such as Abban [6], Raj [7], Modukuru [8] and Pierobon [9]. At the end of the work, the results found are compared to the works cited here.

Unlike previous studies with this dataset, this work does data preparation differently. The age of the concrete is the most unique feature that contributes to its compressive strength. For this reason, age is treated separately in the machine learning models, creating models for each age group.

2 Materials and methods

2.1 Materials and reproducibility

The methodology was carried out using RStudio software (RStudio Team [10]), an integrated virtual environment for code development in R language (R Core Team [11]). Throughout the process, all code executed was documented in the same order as its execution and pushed to the github repository (TODO: CREATE GITHUB REPO BIBLIO AND ADD HERE). The repository contains a longer version of this paper, including all the code, required packages and versions. In order to guarantee reproducibility, whenever there was code that uses probabilistic operations, a seed was defined before its execution, ensuring results consistency when running on another machine.

3 Results

The test RMSE for each model in ascending order of age was 3.31, 4.36, 4.62, 4.72, 5.94 and 5.85 respectively. The Table 1 presents the details and results of each model, including the naive one. The Fig. 1 compares the actual and predicted values for the final models.

Table 1. Final models

Model	mtry	CV	Repetitions	Naive RMSE (test)	Final RMSE (train)	Final RMSE (test)
3 days	6	30	10	9.303229	3.905196	3.310370
7 days	2	10	10	13.443646	4.475981	4.361987
14 days	13	30	10	7.593319	5.136687	4.620515
28 days	11	30	10	14.283824	5.847334	4.716698
56 days	8	30	10	12.702112	6.702565	5.939163
100 days	8	10	10	12.614652	6.381940	5.851088

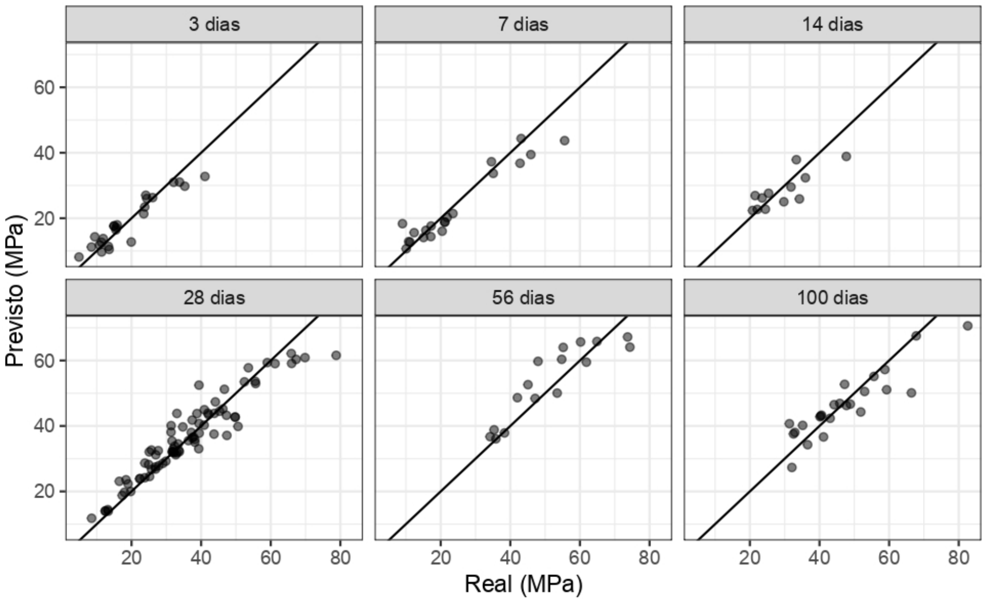

Figure 1. Actual vs Predicted values for each model

4 Discussion and conclusion

The models built present satisfactory results and prove that the compressive strength of concrete can be predicted relatively easily.The alternative adopted to create a model for each set of age proved to be a valid method, instead of using the age as a predictor along with the ingredients like all studies cited in the introduction with the same dataset. This stratification managed to obtain different results for each age group. Although the RMSE of these studies and the found here were close. The Table 2 shows the comparison between these studies and the 28 days model developed here.

Table 2. Comparison to other works with same dataset

Author	Year	Algorithm	RMSE
Pierobon [9]	2018	5 algorithms Ensemble	4.150
This work	2020	Parallel Random Forest	4.717 (28 day model)
Hameed [5]	2020	Artificial Neural Networks	4.736
Raj [7]	2018	Gradient Boosting Regressor	4.957
Modukuru [8]	2020	Random Forest Regressor	5.080
Alshamiri [4]	2020	Regularized Extreme Learning Machine	5.508
Abban [6]	2016	Support Vector Machines with Radial Basis Function Kernel	6.105

Following the line of reasoning of this work, it can be performed with different algorithms, the results found here used only one (Parallel Random Forest), other algorithms can present better results. Another option is to create an ensemble of various algorithms, just like Pierobon [9], but with the separation of age sets proposed here. In addition, it can be performed with a larger dataset, ideally with a similar number of samples in each age group and a more homogeneous distribution of compressive strength.

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] Hasan, Md e Kabir, A., 2011. Prediction of compressive strength of concrete from early age test result.
[2] Kabir, Ahsanul e Hasan, M. e. M. M., 2012. Predicting 28 days compressive strength of concrete from 7 days test result. pp. 18–22.
[3] Yeh, I.-C., 1998. Modeling of strength of high-performance concrete using artificial neural networks. *Cement and Concrete Research*, vol. 28, pp. 1797–1808.
[4] Alshamiri, Abobakr e Yuan, T.-F. e. K. a., 2020. Non-tuned machine learning approach for predicting the compressive strength of high-performance concrete. *Materials*, vol. 13, pp. 1023.
[5] Hameed, Mohammed e Khalid, M., 2020. *Prediction of Compressive Strength of High-Performance Concrete: Hybrid Artificial Intelligence Technique*, pp. 323–335.
[6] Abban, D., 2016. Concrete compresive strength.
[7] Raj, P., 2018. Predicting compressive strength of concrete.
[8] Modukuru, P., 2020. Concrete compressive strength prediction using machine learning.
[9] Pierobon, G., 2018. A comprehensive machine learning workflow with multiple modelling using caret and caretensemble in r.
[10] RStudio Team, 2020. *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.
[11] R Core Team, 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.