INVESTIGATION OF PERFORMANCE METRICS IN REGRESSION ANALYSIS AND MACHINE LEARNING-BASED PREDICTION MODELS

Vagelis Plevris[1], German Solorzano[2], Nikolaos P. Bakas[3] and Mohamed El Amine Ben Seghier[4]

[1] Department of Civil and Architectural Engineering, Qatar University
P.O. Box: 2713, Doha, Qatar
e-mail: vplevris@qu.edu.qa

[2] Dept. of Civil Engineering and Energy Technology, OsloMet–Oslo Metropolitan University
Pilestredet 35, Oslo 0166, Norway
e-mail: germanso@oslomet.no

[3] Research and Development Department, RDC Informatics
Athens, Greece
e-mail: n.bakas@rdc.gr

[4] Faculty of Civil Engineering, Ton Duc Thang University
Ho Chi Minh City, Vietnam
e-mail: benseghier@tdtu.edu.vn

Key Words: *error metric, performance metric, evaluation, regression model, machine learning, neural network, prediction model.*

Performance metrics (Evaluation metrics or error metrics) are crucial components of regression analysis and machine learning-based prediction models. A performance metric can be defined as a logical and mathematical construct designed to measure how close the predicted outcome is to the actual result. A variety of performance metrics have been described and proposed in the literature. Knowledge about the metrics' properties needs to be systematized to simplify their design and use. In this work, we examine various regression related metrics (14 in total) for continuous variables, including the most widely used ones, such as the (root) mean squared error, the mean absolute error, the Pearson correlation coefficient, and the coefficient of determination, among many others. We provide their mathematical formulations, as well as a discussion on their use, their characteristics, advantages, disadvantages, and limitations, through theoretical analysis and a detailed numerical example. The 10 unitless metrics are further investigated through a numerical analysis with Monte Carlo Simulation based on (i) random guessing and (ii) the addition of random noise with various noise ratios to the predicted values. Some of the metrics show a poor or inconsistent performance, while others exhibit good performance as evaluation measures of the "goodness of fit". We highlight the importance of the usage of the right metrics to obtain good predictions in machine learning and regression models in general.