Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 125 (2018) 509–517

Procedia Computer Science

www.elsevier.com/locate/procedia

6th International Conference on Smart Computing and Communications, ICSCC 2017, 7-8 December 2017, Kurukshetra, India

Prediction of Geotechnical Parameters Using Machine Learning Techniques

Nitish Puri*[a], Harsh Deep Prasad[a], Ashwani Jain[a]

[a]*Department of Civil Engineering, NIT Kurukshetra, Kurukshetra-136119, India*

Abstract

In the present study, relationships between in-place density using SPT N-value, compression index (C_c) using liquid limit (LL) and void ratio (e), and cohesion (c) and angle of internal friction (ϕ) using SPT N-value have been established using machine learning techniques. Geotechnical data up to a depth of 50 m from 1053 borehole locations covering almost every district in the state of Haryana have been considered to develop models and statistical correlations. A general trend has been recorded in the observed data and accordingly, the outliers have been excluded. Several models have been developed to establish functional correlations. These correlations have been ranked on the basis their coefficient of determination (R^2) value and mean absolute error (MAE). Subsequently, the model with the highest R^2 value and minimum mean absolute error has been considered for the development of correlations. Analysis has also been carried out for all the developed models to assess their individual performance. For this purpose, all the developed models have been evaluated by fitting a straight line between observed and modelled values, and in all the cases, a good value of R^2 has been observed. The R^2 values obtained for all the models range from 0.798 to 0.988. On comparison, it has been observed that the values of geotechnical parameters obtained are in close agreement with the existing work.

Peer-review under responsibility of the scientific committee of the 6th International Conference on Smart Computing and Communications.

Keywords: Geotechnical Properties; Machine Learning Techniques; Correlations; Haryana

* Corresponding author
E-mail address: nitishpuri.ce.89@gmail.com

1877-0509 © 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 6th International Conference on Smart Computing and Communications
10.1016/j.procs.2017.12.066

1. Introduction

In Geotechnical Engineering, empirical correlations are frequently used to evaluate various engineering properties of soils. Correlations are generally derived with the help of statistical methods using data from extensive laboratory or field testing. Linear Regression (LR) Analysis, Artificial Neural Network (ANN), Support Vector Machine (SVM), Random Forest (RF) and M5 model trees (M5P) are some of the types of machine learning techniques. These techniques learn from data cases presented to them to capture the functional relationship among the data even if the fundamental relationships are unknown or the physical meaning is tough to explain. This contrasts with most traditional empirical and statistical methods, which need prior information about the nature of the relationships among the data. ML is thus well suited to model the complex performance of most Geotechnical Engineering materials, which, by their very nature, exhibit extreme erraticism. This modeling capability, as well as the ability to learn from experience, has given ML techniques superiority over most traditional modeling approaches since there is no need for making assumptions about what could be the primary rules that govern the problem in hand. These techniques are being widely used to solve various civil engineering problems[1-10].

Geotechnical parameters like in-place density, compression index (C_c), coefficient of consolidation (C_v), strength characteristics (c, ϕ) are extensively used for the design of earthen dams, embankments, pavements, landfill liners and foundation of various Civil Engineering structures. Most of these parameters are determined in the laboratory and some are estimated on the field. Their calculation requires a specific laboratory equipment, an experienced geotechnical engineer with a team of skilled technicians. Thus, determination of these parameters is costly and time consuming. Also, soil is a highly erratic material as its performance is based on the processes due to which it is formed. Hence, correlations developed for one region may not be applicable for the other. This ascertains the need to develop region-based correlations to predict geotechnical properties.

In the present study, engineering parameters like in-place density, compression index (C_c), strength characteristics, namely cohesion (c) and angle of internal friction (ϕ) have been correlated with soil parameters determined in laboratory and in field. For this purpose, machine learning techniques like Linear Regression (LR) Analysis, Artificial Neural Network (ANN), Support Vector Machine (SVM), Random Forest (RF) and M5 Tree (M5P) have been used. Geotechnical data have been collected from various government and private organizations across Haryana and optimized for development of more accurate models. The results indicate that developed models are very accurate and provide a viable tool to site engineers and consultants for predicting missing data, and for cross checking the observed values.

2. Study Area, Data Collection and Methodology

Haryana is a non-coastal state in North India with its capital at Chandigarh. It is a moderate sized state having an area of 44,212 km^2, which is 40 times the area of Delhi. It ranks 19th in terms of area in the country. It is surrounded by the states of Uttarakhand, Himachal Pradesh and Shiwalik hills on the North, Uttar Pradesh on the East, Punjab on the West and Delhi, Rajasthan and Aravali hills on the South. It lies between 27°39' to 30°35' N latitude and 74°28' and 77°36' E longitude. The country's capital Delhi is surrounded by Haryana from three sides, forming the northern, western and southern borders of Delhi. Consequently, a large area of Haryana is included in the National Capital Region (NCR) for the purposes of planning for development. Haryana is a leading state in the country on both the industrial and agricultural front. The state has invested in the development of world class infrastructure facilities such as special economic zones (SEZs), Kundli-Manesar-Palwal (KMP) global corridor and Delhi-Mumbai industrial corridor (DMIC) [11].

Geotechnical data collected from Public Works Department (PWD), Delhi Metro Rail Corporation (DMRC), Northern Railways (NR), Haryana Urban Development Authority (HUDA), Nuclear Power Corporation of India Limited (NPCIL), Rail Vikas Nigam Limited (RVNL) and several geotechnical consultants have been used in the study. The developed geotechnical database has information for 1053 distinct locations in the State of Haryana covering almost each district up to a depth of 50 m.

The observed values of geotechnical properties for 1053 borehole locations have been considered for development of various models and statistical correlations. Sorting of relevant data has been carried out by observing a recurring trend and thus deleting the outliers from the data sets. The models were then ranked based on