Geoscience Frontiers 7 (2016) 1–2

Contents lists available at ScienceDirect

China University of Geosciences (Beijing)

Geoscience Frontiers

journal homepage: www.elsevier.com/locate/gsf

Editorial

Progress of machine learning in geosciences: Preface

In the past two decades, artificial intelligence (AI) algorithms have proved to be promising tools for solving several tough scientific problems. As a broad subfield of AI, machine learning is concerned with algorithms and techniques that allow computers to "learn". The machine learning approach covers main domains such as data mining, difficult-to-program applications, and software applications. It is a collection of a variety of algorithms that can provide multivariate, nonlinear, nonparametric regression or classification. The remarkable simulation capabilities of the machine learning-based methods have resulted in their extensive applications in science and engineering. Recently, the machine learning techniques have found many applications in the geosciences and remote sensing. More specifically, these techniques are proved to be practical for cases where the system's deterministic model is computationally expensive or there is no deterministic model to solve the problem (Lary, 2010).

This special issue of *Geoscience Frontiers* aims to review the latest development of machine learning and its key applications in solving problems in geoscience and remote sensing domain. We assemble a set of scientific contributions that provide a window to a successful application of machine learning and its branches to challenging in the field.

The opening paper of this issue by Lary et al. (2016) presents a review of the machine learning applications in geosciences and remote sensing. The authors outline the unique features of some of the machine learning techniques with a specific attention to genetic programming paradigm. Furthermore, nonparametric regression and classification illustrative examples are presented to demonstrate the efficiency of machine learning. The problems investigated by the authors are characterizing airborne particulates and identifying dust sources.

Esmaili and Mohaghegh (2016) focus on full field reservoir modeling of shale assets using advanced data-driven analytics. They developed an artificial intelligence-based model that is conditioned to all available field measurements (e.g. production history, measured reservoir characterizations including geomechanical and geochemical properties) as well as measured hydraulic fracturing variables like slurry volume, proppant amount and sizes, injection rate etc. Such model has the potential to provide operators with an alternative to history-match, predict and assess reserves in oil and gas producing shale reservoirs. The integrated framework presented by the authors enables reservoir engineers to compare and contrast multiple scenarios and propose field development strategies.

Sparrow and Mercer (2016) investigate the predictability of US tornado outbreak seasons using ENSO and northern hemisphere geopotential height variability. Their study is focused on diagnosis of seasonal predictability of tornado outbreak frequency in the United States through 500-hPa, 1000-hPa, and ENSO interannual variability indices. To this aim, authors formulate a linear stepwise multivariate linear regression (SMLR) and a support vector regression (SVR) model using 16,500-hPa RPC score predictors, two 1000-hPa RPC score predictors, and the Niño 3.4 indices. They conclude that the nonlinear SVR technique reduces root mean square errors produced by the control SMLR technique by 28% and provided more consistent forecasts. The results obtained from this study would be useful to both forecasters and the general public in terms of preparing for the upcoming severe weather season.

A comprehended review of the AI applications in pile foundations is carried on by Shahin (2016). The author presents the salient features associated with the modeling development of these AI techniques. The paper also discusses the strength and limitations of the selected techniques compared to other available modeling approaches. This review paper is focused on behavior of pile foundations including bearing capacity prediction, settlement estimation, and modeling of load-settlement response. The author concludes that the AI techniques perform better than, or at least as good as, the most traditional methods.

Zhang and Goh (2016) present two machine learning methods, called multivariate adaptive regression splines (MARS) and back propagation neural network (BPNN) for assessing pile drivability. The main goal is to formulate maximum compressive stresses (MCS), maximum tensile stresses (MTS), and blow per foot (BPF) in terms of several influencing variables. The authors develop the models upon a database of more than 4000 piles. They conclude that BPNN and MARS models for the analyses of pile drivability provide good predictions. MARS is found to be more computationally efficient than BPNN as it builds flexible models.

Patel and Chatterjee (2016) present a computer vision-based rock-type classification algorithm for fast and reliable identification without human intervention. They develop a laboratory scale vision-based model using probabilistic neural network (PNN) where they use color histogram features as input. A total nine features are used as input for the PNN classification model. The authors validate the model using a test data set and conclude that their proposed vision-based model can perform satisfactorily for classifying limestone rock-types. Also, they prove that their

Peer-review under responsibility of China University of Geosciences (Beijing).

http://dx.doi.org/10.1016/j.gsf.2015.10.006

proposed method performs substantially better than three other classification algorithms.

The study done by Viswanathan and Samui (2016) aims at determination of rock depth using three machine learning techniques, namely Gaussian Process Regression (GPR), Least Square Support Vector Machine (LSSVM) and Extreme Learning Machine (ELM). The authors use Latitude (Lx) and Longitude (Ly) to formulate the rock depth. They conclude that the developed ELM, GPR and LSSVM produce spatial variability of rock depth at Chennai. It showed that the used machine learning methods are robust for prediction of rock depth.

Khan et al. (2016) develop prediction models for residual strength of clay based on a new method, called functional networks (FN). The authors perform a comparative study between FN, SVM and artificial neural network (ANN) based on statistical parameters like correlation coefficient (R), Nash–Sutcliff coefficient of efficiency (E), absolute average error (AAE), maximum average error (MAE) and root mean square error (RMSE). It is found that FN has a better prediction performance than ANN. They perform a sensitivity analysis to ascertain the importance of various inputs in the prediction of the residual strength of clay. Besides, the authors provide a prediction equation that can be used by the practicing geotechnical engineers to calculate the residual friction angle value if the index properties of the soil are available.

Peak ground acceleration (PGA) is a well-known engineering parameter of an earthquake for seismic structural analysis and risk assessment. Gandomi et al. (2016) propose a new model to predict PGA utilizing a new method coupling ANN and simulated annealing (SA), called SA-ANN. They formulate PGA in terms of earthquake source to site distance, earthquake magnitude, average shear-wave velocity, faulting mechanisms, and focal depth. The authors utilize a huge database of strong ground-motion recordings of 36 earthquakes in Iran. Their proposed model is verified for a part of the database beyond the training data domain. They compare the SA-ANN model with the simple ANN, as well as 10 other well-known models in the field. The results show that SA-ANN is superior to the single ANN and other existing attenuation approaches. Finally, they extract an explicit formula that can be easily used in a spreadsheet or hand calculations to predict PGA, especially in Iran's tectonic regions.

Kashani et al. (2016) propose imperialistic competitive algorithm (ICA) for locating the critical failure surface and computing the factor of safety (FOS) in a slope stability analysis. The FOS relating to each trial slip surface is calculated using a simplified algorithm of the Morgenstern-Price method. The authors use four benchmark test problems to explore the performance of the algorithm. It is concluded that ICA can provide reliable, accurate and efficient solutions for locating the critical failure surface and relating FOS. Moreover, they prove that the ICA algorithm is the most proficient algorithm among the other existing algorithms because of smaller FOS with low standard deviation.

In the closing paper of this special issue, Alavi and Sadrossadat (2016) propose new nonlinear prediction models for the ultimate bearing capacity of shallow foundations resting on non-fractured rock masses. The authors utilize a novel evolutionary computational approach, called linear genetic programming. They use a comprehensive set of rock socket, centrifuge rock socket, plate load and large-scaled footing load test results to develop the models. The results indicate that the models proposed by the authors accurately characterize the bearing capacity of shallow foundations. Moreover, the derived models reach a notably better prediction performance than the traditional equations. The authors provide transparent programs that can be used for further analysis of the bearing capacity, as well as optimization purposes.

We hope that the new findings and perspectives from the papers presented in this special issue will provide new insights and guidelines for further applications of the machine learning techniques in geosciences and remote sensing. We would like to express our sincere thanks to Prof. M. Santosh, Co Editor-in-Chief of Geoscience Frontiers for his support in making this special issue a reality. We are also grateful to Dr. Lily Wang, Editorial Assistant of Geoscience Frontiers for her dedicated assistance during the review process of manuscripts. We would like to thank all the authors for providing interesting articles and patronage for the timely submission and revision. We are also thankful to the reviewers for providing fast and constructive reviews for the submitted papers, including papers that did not reach the final stage of publication.

References

Alavi, A.H., Sadrossadat, E., 2016. New design equations for estimation of ultimate bearing capacity of shallow foundations resting on rock masses. Geoscience Frontiers 7 (1), 91–99. http://dx.doi.org/10.1016/j.gsf.2014.12.005.

Esmaili, S., Mohaghegh, S.D., 2016. Full field reservoir modeling of shale assets using advanced data-driven analytics. Geoscience Frontiers 7 (1), 11–20. http://dx.doi.org/10.1016/j.gsf.2014.12.006.

Gandomi, M., Soltanpour, M., Zolfaghari, M.R., Gandomi, A.H., 2016. Prediction of peak ground acceleration of Iran's tectonic regions using a hybrid soft computing technique. Geoscience Frontiers 7 (1), 75–82. http://dx.doi.org/10.1016/j.gsf.2014.10.004.

Kashani, A.R., Gandomi, A.H., Mousavi, M., 2016. Imperialistic Competitive Algorithm: a metaheuristic algorithm for locating the critical slip surface in 2-dimensional soil slopes. Geoscience Frontiers 7 (1), 83–89. http://dx.doi.org/10.1016/j.gsf.2014.11.005.

Khan, S.Z., Suman, S., Pavani, M., Das, S.K., 2016. Prediction of the residual strength of clay using functional networks. Geoscience Frontiers 7 (1), 67–74. http://dx.doi.org/10.1016/j.gsf.2014.12.008.

Lary, D.J., 2010. Artificial intelligence in geoscience and remote sensing. In: Imperatore, P., Riccio, D. (Eds.), Geoscience and Remote Sensing, New Achievements. IN-TECH, Vukovar, Croatia, pp. 1–24.

Lary, D.J., Alavi, A.H., Gandomi, A.H., Walker, A., 2016. Machine learning in geosciences and remote sensing. Geoscience Frontiers 7 (1), 3–10. http://dx.doi.org/10.1016/j.gsf.2015.07.003.

Patel, A.K., Chatterjee, S., 2016. Computer vision-based limestone rock-type classification using probabilistic neural network. Geoscience Frontiers 7 (1), 53–60. http://dx.doi.org/10.1016/j.gsf.2014.10.005.

Shahin, M., 2016. State-of-the-art review of artificial intelligence applications in pile foundations. Geoscience Frontiers 7 (1), 33–44. http://dx.doi.org/10.1016/j.gsf.2014.10.002.

Sparrow, K., Mercer, A., 2016. Predictability of U.S. tornado outbreak seasons using ENSO and northern hemisphere geopotential height variability. Geoscience Frontiers 7 (1), 21–31. http://dx.doi.org/10.1016/j.gsf.2015.07.007.

Viswanathan, R., Samui, P., 2016. Determination of rock depth using artificial intelligence techniques. Geoscience Frontiers 7 (1), 61–66. http://dx.doi.org/10.1016/j.gsf.2015.04.002.

Zhang, W., Goh, A.T.C., 2016. Multivariate adaptive regression splines and neural network models for prediction of pile drivability. Geoscience Frontiers 7 (1), 45–52. http://dx.doi.org/10.1016/j.gsf.2014.10.003.

Amir H. Alavi*
Department of Civil and Environmental Engineering, Michigan State University, USA

Amir H. Gandomi
BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI 48824, USA

David J. Lary
Hanson Center for Space Science, University of Texas at Dallas, Richardson, TX 75080, USA

* Corresponding author.
E-mail address: alavi@msu.edu (A.H. Alavi)

Available online 10 November 2015