TC Corner

TC304/309/210 Machine Learning Dialogue for Geotechnics 2019

1 Introduction

Digital transformation is occurring at a rapid pace in many industries. ISSMGE has taken leadership in hastening this transformation by establishing a new TC309 in Machine Learning and Big Data. TC309 has since organized two International Symposium on Machine Learning and Big Data in Geoscience (Norwegian Geotechnical Institute, Oslo, Norway, October 21 – 22, 2018; Tongji University, Shanghai, China, July 28 – 30, 2019). As a continuation of these activities, a machine learning (ML) dialogue for geotechnics was held in the National Taiwan University on 14 December, 2019, following the 7th International Symposium on Geotechnical Safety and Risk (ISGSR 2019) organized under the auspices of the Geotechnical Safety Network (GEOSNet) between December 11 and 13 2019 in Taipei. The machine learning dialogue was hosted by Prof. Jianye Ching (Chairman of ISSMGE TC304) at the Department of Civil Engineering, National Taiwan University, and was coordinated by Prof. Kok-Kwang Phoon (National Singapore University), Prof. Zi-Jun Cao (Wuhan University), and Prof. Yu Wang (City University of Hong Kong). This dialogue is supported by ISSMGE TC304, TC309 and TC210. Thirty-five experts from 13 countries and regions were invited to attend the dialogue (see a group photo in Figure 1).

The aim of the ML dialogue is to discuss the opportunities and challenges in developing and applying ML to geotechnical engineering research and practice. The focus is to discuss big ideas that can transform research and practice in completely new ways. Coordinators prepared reading materials to provide background information on ML and their current nascent applications in geotechnical engineering and pose key questions and desired outcomes to stimulate and structure the discussion. The reading materials, including a 2020 Georisk Spotlight paper on “The story of statistics in geotechnical engineering” (Phoon, 2020) were circulated to the participants 3 weeks before the dialogue. The participants were expected to actively engage in small group discussions during the dialogue. The organization and discussion rules of the ML dialogue were detailed in the proceedings of ISGSR 2019 (Phoon et al., 2019). The article summarizes the proceedings (including opening ceremony and presentation, group discussion, and all-group discussion) and key conclusions of the ML dialogue.

Figure 1. Participants of the machine learning dialogue at the Department of Civil Engineering, National Taiwan University, 14 December 2019

TC Corner (Con't)

TC304/309/210 Machine Learning Dialogue for Geotechnics 2019

2 Opening Ceremony and Presentation

At the opening ceremony, Prof. Jianye Ching warmly welcomed all the participants of the dialogue, and Dr. Zhongqiang Liu (Chairman of ISSMGE TC309, Norwegian Geotechnical Institute) reported outcomes from the pre-workshop panel discussion session that took place in the 2nd International Workshop on Machine Learning and Big Data in Geoscience organized by ISSMGE TC309 in Shanghai, July 28-30, 2019. Following the opening ceremony, Prof. Kok-Kwang Phoon gave a presentation to explain the purpose of the ML dialogue in the context of digital transformation. He called the digital transformation of geotechnical engineering as GEO 4.0 and highlighted that the new wicked problems (e.g., resilience and sustainability of both new and aging infrastructure) can only be solved by deploying digital technologies in the spirit of GEO 4.0. Prof. Kok-Kwang Phoon also proposed 7 "**E**"s in the GEO 4.0 research agenda covering **E**ssence (core asset is data), **E**conomic value (value of data to industry), **E**xchange (data sharing), **E**xtremes (dealing with outliers), **E**rrors (dealing with uncertainties), **E**xtrapolation (dealing with overfitting), and **E**xplanation (white box ML). Prof. Kok-Kwang Phoon concluded with three desired outcomes for the ML dialogue:

- What are the research questions? (be as specific as possible)
- What is our wish list? ("blue sky" ideas)
- What does TC304/TC309 need to do to lead the GEO 4.0 agenda?

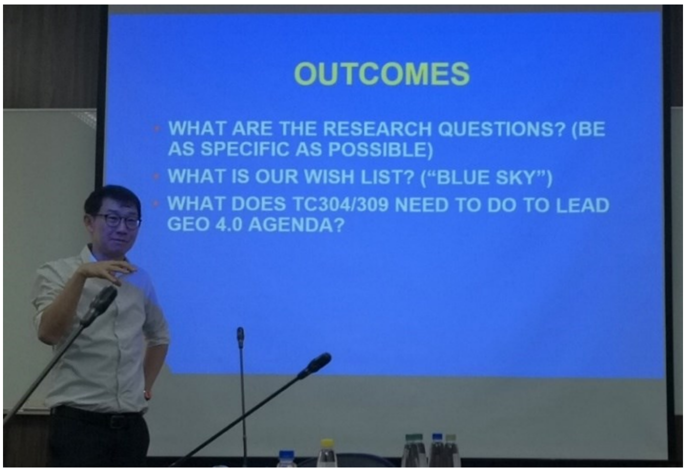

Figure 2. Opening presentation delivered by Prof. Kok-Kwang Phoon

Figure 3. Discussion by Group I

TC Corner (Con't)

TC304/309/210 Machine Learning Dialogue for Geotechnics 2019

3 Group Discussion

Following the open ceremony, ▫participants were divided into 5 small groups (see Table 1), with about 7 persons (including one facilitator and one reporter) in each group for individual group discussion. The identified facilitators guided their group discussions to achieve at least one desired outcomes. Each facilitator/reporter presented key ideas arising from the group discussion. The presentation was limited to two minutes.

Table 1. Organization of group discussion

Group ID	Facilitator	Reporter	Group members
Group I	Kok-Kwang Phoon	Jinhui Li	Jianye Ching, Wuzhang Luo, Wojciech Pula, Johan Spross, Xiaohui Tan
Group II	Yu Wang	Wenping Gong	Adeyemi Aladejare, Michele Calvello, Lulu Zhang, Bram Van Den Eijnden, Changhong Wang
Group III	Zhongqiang Liu	Dongming Zhang	Tom Charlton, Jinsong Huang, Jonathan Nuttall, Zhiyong Yang, Tengyuan Zhao
Group IV	Iason Papaioannou	Ivan Depina	Jing-Sen Cai, Zi-Jun Cao, Michael Hicks, Thi Minh Hue Le, Limin Zhang
Group V	Ikumasa Yoshida	Takayuki Shuku	Shinichi Akutagawa, Richard Bathurst, Hyun-Ki Kim, Andy Yat Fai Leung, Yu Otake, Te Xiao

The discussion in Group I (see Figure 3) started with sharing experiences on ML-related projects. During the discussion, three aspects were addressed, including data sharing, accessibility of national databases, and the value of a real flagship project for future discussions. Group 1 proposed a "data sandbox" that involves multiple industry players and researchers to "play" in the sandbox in terms of developing data-driven algorithms of significant value to industry while maintaining data confidentiality and respecting intellectual property rights. It is proposed that the researchers engage a single industry player to work on a win-win agreement to create a legally safe space for data sharing.

The discussion in Group II (see Figure 4) started with data. It was emphasized that data in ML shall not only be limited to conventional geotechnical data, but also include non-conventional data. One key point for ML algorithms is how to use or develop geotechnical ML algorithms to obtain valuable geotechnical information from these non-conventional data for decision making in engineering practice. Another important point raised during the discussion is the value-add or benefit from using ML to geotechnical engineering practice. It was pointed out that industry partnership is the key and concrete and high-profile successful examples of ML applications are urgently needed.

TC Corner (Con't)

TC304/309/210 Machine Learning Dialogue for Geotechnics 2019

The discussion in Group III (see Figure 5) mainly focused on practice related issues for ML research. Three major aspects were discussed, including application fields, circumstances for the use of ML, and ML methods. Value of engineering judgement was highlighted. It was pointed out that ML methods can be used with knowledge from multiple disciplines. The importance of physics-based domain knowledge was emphasized by adding physical constraints in data analysis and/or developing hybrid models considering both physics and data. The role of ML methods was considered to be supplementary to physics-based modeling.

The discussion in Group IV (see Figure 6) started with a review of different ML methods and their classifications (e.g., supervised and unsupervised methods). Thereafter, each group member discussed the types of ML that they have applied in geotechnical engineering problems. Then, the group discussion focused on identifying a number of challenges in the further application and promotion of ML in geotechnical engineering, such as those in data sharing, enhancement of the predictability of ML approaches through infusing the physics of the problem, etc. The importance of education of potential users, e.g., through short courses addressed to the industry, was also highlighted.

Figure 4. Discussion by Group II

Figure 5. Discussion by Group III

Figure 6. Discussion by Group IV

TC Corner (Con't)

TC304/309/210 Machine Learning Dialogue for Geotechnics 2019

The discussion in the Group V (see Figure 7) centered on "Our Wish List". The wish list achieved in the discussion included database on infrequent geotechnical events (e.g., slope failure), access to large amount of data, more investment in site investigation in construction projects, open web database, user-friendly supercomputer and quantum computer, better collaboration between industry and academics to design monitoring programs for instrumented structures and analysis of data, a new journal devoted to ML in geotechnical engineering, collaboration with other ISSMGE TCs, reducing fatalities/injuries from geotechnical failures by using monitoring, fast data processing and ML.

Figure 7. Discussion by Group V

4 All-Group Discussion

The all-group discussion session was open to all participants to express their views and reactions. The discussions covered how to demonstrate value of data in geotechnical designs, possible venues for publication of ML-related geotechnical studies, how to incorporate physics-based domain knowledge into geotechnical data analytics, and the ML wish list. Finally, Prof. Kok-Kwang Phoon summarized the discussions and closed the dialogue with following concluding points:

- Industry is a necessary partner for ML-based studies in geotechnical engineering. There are two compelling reasons. One, ML must bring transformative value to industry, not universities. This value is ultimately linked to the end users that the industry serves. Two, real data are generated and owned by projects and they are necessary ingredients for ML development (not simulated data).
- A "data sandbox" is needed to enable and accelerate data sharing between industry and research. Theoretical and pure methodology development of ML is not seen to be the immediate priority.
- Future discussions on this topic, including organization of future workshops, should be tied to a specific flagship project with the industry. This "AlphaGeo" project (cf. Google's "AlphaGo" project) is ideally game changing, rather than improving some aspects of existing practice in an incremental way.

TC Corner (Con't)

TC304/309/210 Machine Learning Dialogue for Geotechnics 2019

Acknowledgements

The dialogue was supported by ISSMGE TC304 (Chair, Prof Jianye Ching), TC309 (Chair, Dr Zhongqiang Liu), and TC210 (Chair, Prof Limin Zhang). It could not be held successfully without the enthusiastic participation and contributions from all the invited experts (refer to Table 1). The coordinators would like to thank Prof. Jianye Ching for hosting the dialogue at the National Taiwan University and all facilitators and reporters for their efforts in coordinating the preparation for the dialogue and guiding group discussions. Special thanks go to Prof. Jinhui Li (Harbin Institute of Technology, Shenzhen), Prof. Yu Wang (City University of Hong Kong), Prof. Dongming Zhang (Tongji University), Dr. Iason Papaioannou (Technical University of Munich), and Dr. Takayuki Shuku (Okayama University) for contributing a summary of their group discussions to this article. The assistance provided by Mr. Qin-Xuan Deng (PhD Candidate, Wuhan University) in preparing for this ML dialogue, including taking the photographs included in this article, is gratefully acknowledged.

References

Phoon,K.K. (2020). The story of statistics in geotechnical engineering. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 14(1), 3-25.

Phoon,K.K., Cao, Z.J., Wang, Y. (2019). ISSMGE TC304/309/210 Machine Learning Dialogue for Geotechnics 2019. Proceedings of the 7th International Symposium on Geotechnical Safety and Risk (ISGSR 2019), Taipei, Taiwan, December, 11–13, 2019.