THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學

國家鋼結構工程技術研究中心香港分中心
Chinese National Engineering Research Centre For Steel Construction (Hong Kong Branch)

Address: Hong Kong Polytechnic University, Phase 8, Hung Hom, Kowloon, Hong Kong.
Telephone: (852) 3400 8451 Email: cnerc.steel@polyu.edu.hk Website: https://www.polyu.edu.hk/cnerc-steel/

Project Title:

Investigation on the application of high strength steel H-sections in foundation engineering

Principal Investigator: **Dr Andy Y F Leung**

Project Team: **Dr Mark S C Hsu, Mr Vahid Asghari, Mr Dongya Han, Dr Horace M K Lo and Miss Ka Ming Yu**

Project Outline:

This project explores the issues associated with application of high strength steel sections in foundation engineering, such as the soil–structure interaction during and after construction, and the material variability and reliability considerations.

For example, in the steel production process, various substances are used to alloy with the steel, with compositions tailored to achieve different characteristics of the material. These substances may include manganese, silicon, phosphorus or sulphur in various forms. When added in different amounts and at different stages of the production, they can significant affect the physical and mechanical properties of the material such as the strength, ductility and weldability, etc. While our knowledge in metallurgy and material sciences have advanced, the complex relationships among the material composition, production process and various properties of the end product still involve much uncertainty. There may be 'hidden' correlations that are not yet well understood, but can be uncovered by harnessing large volume of production data through advanced data analytics such as machine learning techniques. This enables quantification of the rich experience in the past, for enhanced understanding of material properties and improved quality assurance for future production.

To facilitate better understanding of these issues and their implications to foundation engineering and underground engineering involving high performance steel, this study involves the following main objectives:

1) To investigate the application of machine learning or deep learning techniques such as deep neural networks (DNN) in the characterization of construction material properties, and to develop tailor-made models suitable for prediction of steel properties based on the components and manufacture process.

2) To examine the key issues associated with installation and operation of foundation elements with high strength steel sections. These include the soil-structure interactions considering the variability and uncertainty in various foundation and soil components.

Project progress

A user-friendly machine learning tool has been developed in 2019, calibrated and tested using datasets of both soil properties and steel properties. The former was used as a first trial as soils are known to exhibit significant uncertainty and variability arising from the natural sedimentation and formation processes. Currently, the tool adopts Deep Neural Networks (DNN) to obtain correlations among various index properties of the materials.

DNN are a member of deep learning family which enables the formulation of strong prediction of classifier models. These models are being used daily in image processing, voice recognition, and similar technologies. Compared to previous techniques such as artificial neural networks (ANN), DNN consists of many hidden layers between the input variables and the output variable. Their structures also involve different layer types such as dropout and/or convolutional layers. Consequently, the training and structuring of DNNs are complex and there is still no one-size-fits-all approach that is applicable to all conditions. For example, there are various options to optimize the design of DNN for specific applications, regarding aspects such as data pre-processing techniques, training curve analysis, structuring of neural network including number and types of layers and nodes, and other issues such as overfitting problems. Therefore, the first step of this study is to present a new DNN algorithm tailor-made to predict the properties of foundation materials including soils, steel and their associated variability, based on existing data of material composition and production process from various steel manufacturers in China.

Apart from the DNN model, the established framework includes guidelines for data pre-processing (data cleaning, detection of outliers and feature scaling), assessment of data adequacy by ‘learning curve analysis’, removal of outliers and fine-tuning of model hyper-parameters to identify and resolve overfitting problems.

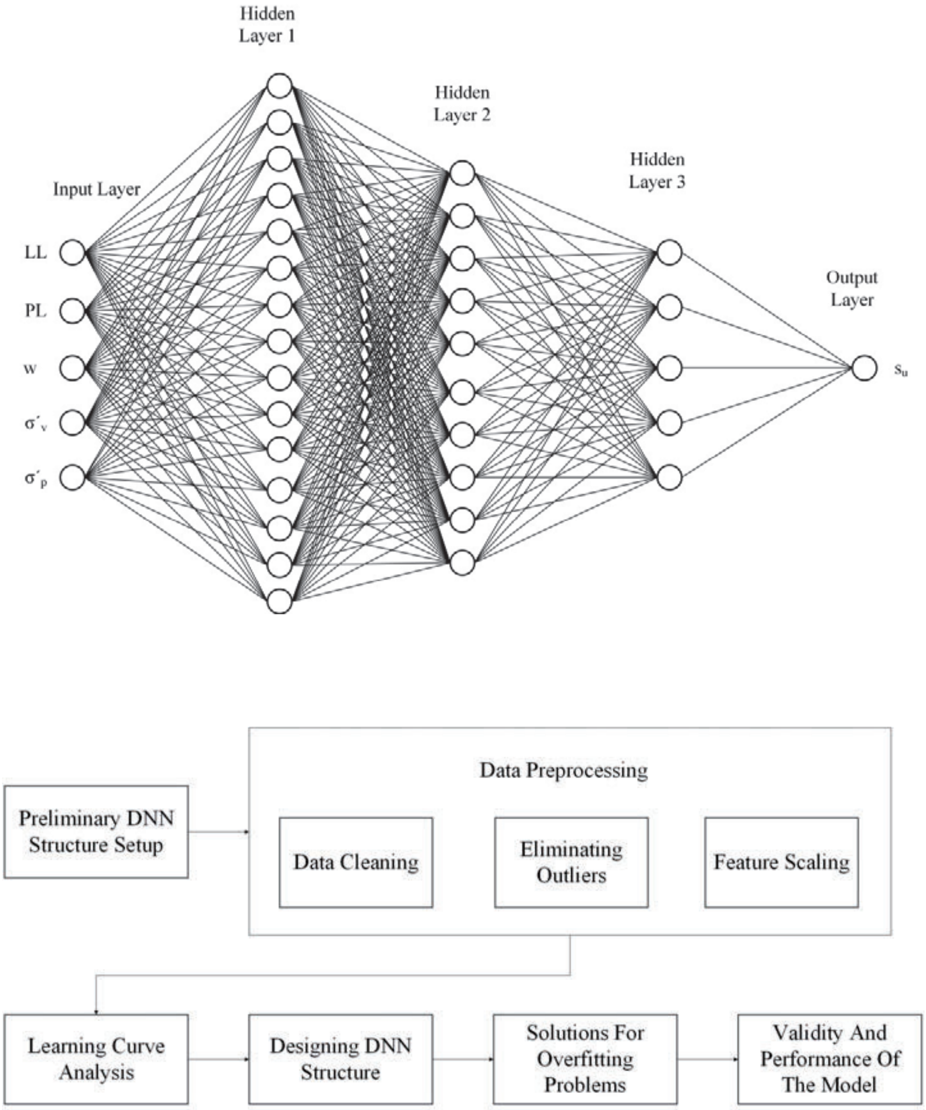
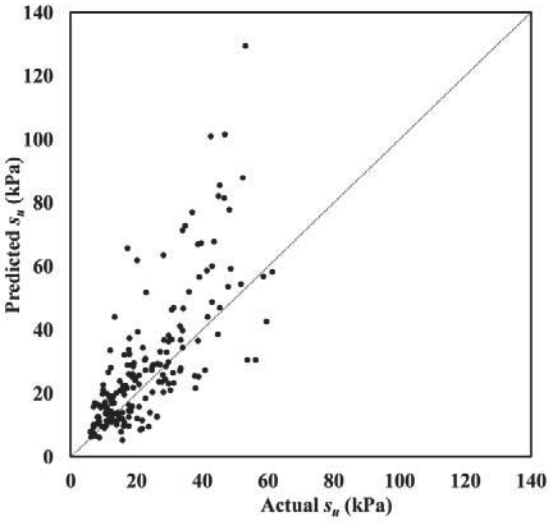
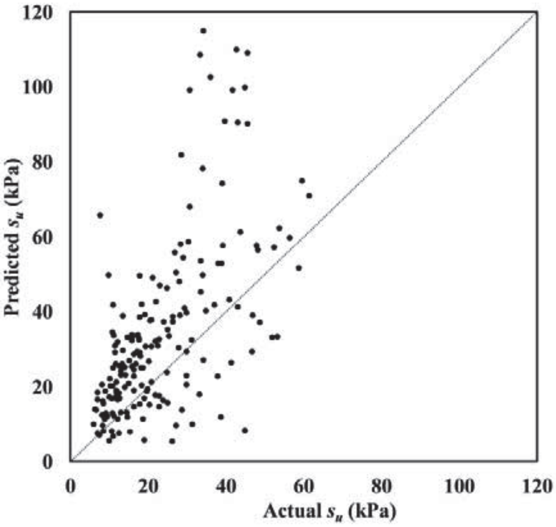


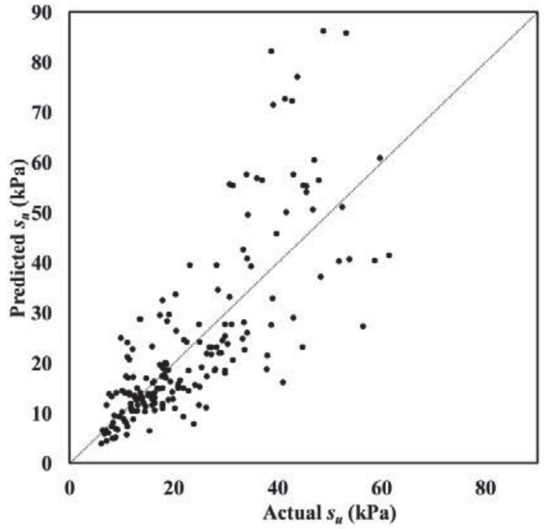
Figure: Example DNN structure (top) and key steps in establishing model for correlation of material properties (bottom)



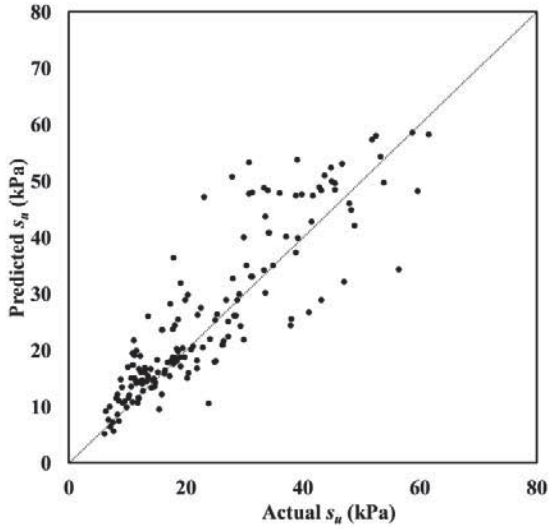
(a) Hansbo forumula



(b) Ching and Phoon formula (i)



(c) Ching and Phoon formula (ii)



(d) DNN approach in this study

Figure: Improved DNN algorithm (d) for prediction of material properties in comparison with previous formulas in the literature (a)-(c)

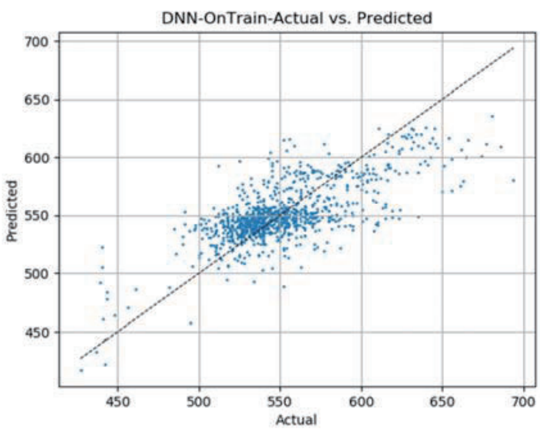
As a first step, the DNN approach is implemented to investigate the correlation of foundation soil properties with a database of approximately 1000 entries of soil strength and other index properties. The above figure shows that the approach is capable of producing reasonable predictions of material properties considering the significant uncertainties in soils. It also outperforms previously published formulas which were developed either based on theoretical considerations, empirical relationships or combinations of both.

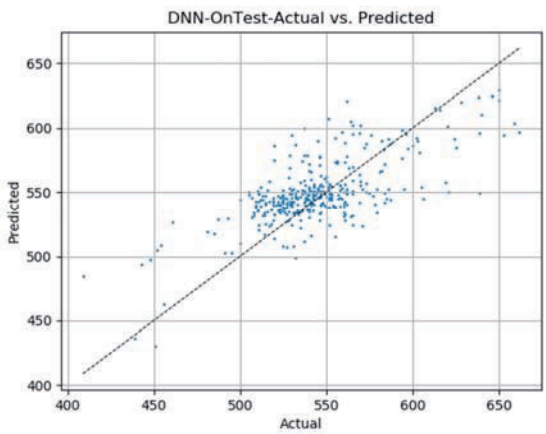
A trial database of steel properties was then established containing 2000 entries of yield strength and various contributing factors including compositions of elements such as carbon, silicon, Manganese, Phosphorus and Sulphur, and the identities of the manufacturers and adopted procedure standards in the manufacturing process.

Apart from the DNN algorithm, a number of other machine learning and deep learning approaches have been compiled and implemented on this dataset, to tailor for the specific features of steel properties arising from the manufacture processes. Some of the algorithms that have been adopted include:

- DNN
- ExtraTrees
- Support vector machine
- Ridge and Lasso Regression, etc.

with some of the results shown below.





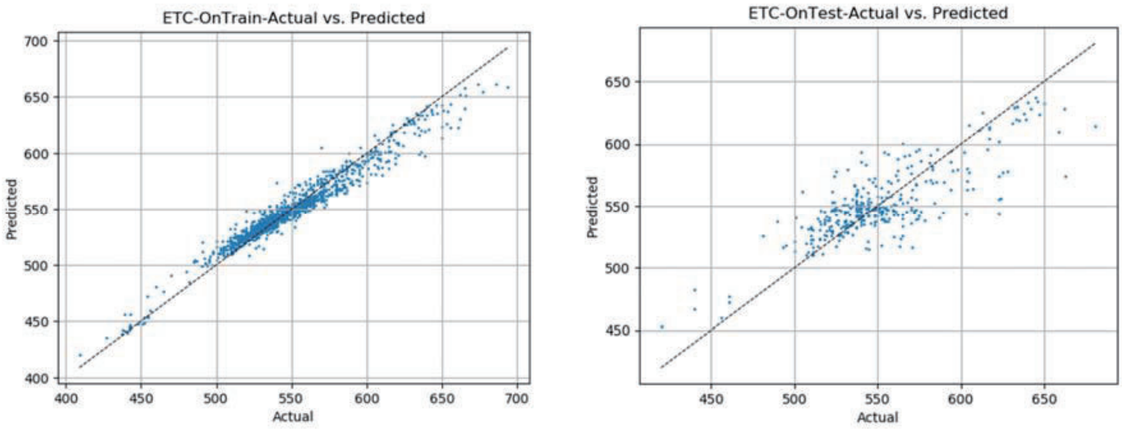


Figure: Predicted versus actual yield strength by DNN (top) and ExtraTrees (bottom)

With the limited dataset in the trial database, ExtraTrees appear to be more suitable for prediction of steel strength from the manufacture process. As machine learning methods generally improve with the number of data points and the number of attributes considered (information on raw material, mineral composition and other details of the manufacturing process), the performance is expected to improve further with an expanded database, which is currently being compiled.

Project deliverables

A Python code has been compiled for DNN analyses of engineering metrics, and uploaded for public access (https://github.com/vd1371/Developing-a-Deep-Neural-Network-based-Framework-for-Complex-Correlations-in-Engineering-Metrics), while the details of underlying methodology are discussed in the publications below. The code is a user-friendly tool that enables determination of complex correlations among engineering variables, and is particularly suitable for components of foundation engineering including soils and steel elements.

The publications arising from this project include:

[1] Asghari, V., Leung, Y.F. and Hsu, S. C. (2020). Deep neural network based framework for complex correlations in engineering metrics. *Advanced Engineering Informatics* (accepted for publication).

[2] Asghari, V., Leung, Y.F. and Hsu, S. C. (2019). Deep neural networks for prediction of undrained shear strength of clays. *Proceedings of the 7th International Symposium on Geotechnical Safety and Risk*, 554-559.

Future work and project significance

The user-friendly Python code can be readily applied for prediction of physical and mechanical properties such as strength, ductility and corrosion resistance of steel. As a key feature of machine learning, the program can be further calibrated/trained and refined as the database grows in the future with increasing availability of relevant information. The database will be expanded with tens of thousands of steel mill test data from various factories in China. A well-established machine learning algorithm will facilitate detection of anomalies in the production data and eventually benefit the quality assurance of steel production in China.