

# 1

## Overview of Machine Learning in Geomechanics

**Ioannis STEFANO**

*GeM, UMR 6183, CNRS, Ecole Centrale Nantes,  
Nantes Université, France*

### 1.1. What exactly is machine learning?

One of the best methods for making a puppy learn its name is to play the name game:

- 1) Take your puppy to a quiet place and gain its attention by calling its name in a happy and reassuring tone (e.g. “LoupI look at me!” or “LoupI come here!”).
- 2) As soon as you have gained its attention, reward it with a treat to encourage its good behavior.
- 3) Repeat this process several times until your puppy has learned to respond to its name.

---

For a color version of all the figures in this chapter, see [www.iste.co.uk/stefanou/machine1.zip](http://www.iste.co.uk/stefanou/machine1.zip).

*Machine Learning in Geomechanics 1*,  
coordinated by Ioannis STEFANO and Félix DARVE. © ISTE Ltd 2024.

In the above iterative process, the puppy learns from experience that reacting when you call its name is a good thing. By repeating this process, the puppy recognizes the pattern between similar, but not identical, vocal signals of its name (training data) and associates them with a specific behavior (task).

Many cognitive processes of humans involve a similar way for associating patterns with specific reactions; for instance, see how we learn to drive a car, speak a foreign language or even understand how a geomaterial behaves!

*Machine learning* (ML) is an anthropomorphic extension of the above process of “learning” to a computer program, that is, to the machine. By learning, the machine will progressively improve its performance in executing a specific task. According to Mitchell (1997), ML can be defined as follows:

*A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.*

The whole set of examples that the computer will use during learning is the *training set*, while each training example is called a training instance or sample. The whole process of learning is called *training*. The training set is usually distinguished into *training data*, *validation data* and *test data*. More details about those sets and their importance will be given in section 1.5 and in Chapters 2 and 5 of volume 1.

A popular example showing the power of ML was *AlphaGo*, a ML program, which beat the human world champion of Go (see also (Silver et al. 2016) and Chapter 4 of volume 2 on *reinforcement learning*). Go is a strategy board game for two players in which the aim is to cover more territory than the opponent (see Figure 1.1). Despite its relatively simple rules, Go is extremely complex. The *state-space complexity* of Go, that is, the number of legal game positions reachable from the initial position of the game (Allis 1994), is of the order of  $10^{170}$  (Tromp and Farnebäck 2007), which is much higher than that of chess, which

is of the order of  $10^{44}$  (Shannon 1950), and certainly higher than the number of atoms in the universe. This makes Go a great challenge for ML algorithms.



**Figure 1.1.** *Eighteen times Go world champion Lee Sedol playing with the machine (courtesy: Lee Jin-man/AP, The New Yorker, <https://www.newyorker.com/tech/annals-of-technology/alphago-lee-sedol-and-the-reassuring-future-of-humans-and-machines>)*

Applying the above-mentioned definition in an ML program for Go, the performance (P) will be measured by the ability of the machine to beat its opponent by gaining experience (E) by executing multiple times the task (T), which is playing the game with itself. Experience can be also obtained by processing existing games of other (human) players.

Another recent popular example of ML is the *ChatGPT* chatbot developed by OpenAI. ChatGPT was released in November 2022 and provoked a massive public reaction. Let us see what it answers when we ask it to describe itself and what it knows about ALERT Geomaterials.

**Ioannis:** *Describe yourself.*

**ChatGPT:** *I am ChatGPT, an AI language model developed by OpenAI. I have been trained on a vast amount of text data from the internet, books and other sources. My purpose is to assist*

*users in generating human-like text responses and providing information on a wide range of topics.*

*I don't have a physical form or personal experiences like humans do, but I can understand and generate text in multiple languages. I strive to provide helpful and accurate responses to the best of my abilities, drawing upon the knowledge I have acquired during my training.*

*I can engage in conversations, answer questions, offer explanations and assist with various tasks. However, please keep in mind that while I strive for accuracy, I may occasionally generate incorrect or incomplete information. It's always a good idea to verify important details from reliable sources.*

*Overall, I am here to assist and provide information to the best of my abilities. Feel free to ask me anything you'd like to know or engage in a conversation on any topic of your interest.*

**Ioannis:** *How do you work?*

**ChatGPT:** *As an AI language model, my functioning can be divided into two main phases: training and inference.*

*Training: During the training phase, large amounts of text data are used to train the model. In the case of GPT models like ChatGPT, this training involves a process called unsupervised learning. The model learns to predict what comes next in a given sequence of text, based on the patterns and relationships it discovers in the training data. By doing this over a vast dataset, the model gradually learns grammar, facts, reasoning abilities and the nuances of human language.*

*Inference: Once the training is complete, the model moves into the inference phase, where it can generate responses and interact with users. Here's a simplified overview of how it works during inference:*

*Input:* When you provide a prompt or a message, the model receives it as input. The input can be a question, a statement or any text you want to generate a response to.

*Tokenization:* The model breaks down the input text into smaller units called tokens. These tokens can represent words, subwords or characters. Tokenization helps the model understand the structure and meaning of the input.

*Processing:* The tokenized input is then passed through several layers of neural networks. These layers are designed to understand the relationships between the tokens and capture the context and meaning of the text.

*Generation:* Once the input is processed, the model generates a response. It predicts the most likely next tokens based on the context and the patterns it has learned during training. The generated tokens form the model's response.

*Iteration:* The process of tokenization, processing and generation can happen in an iterative manner. The generated response can be further used as input, creating a back-and-forth conversation or a continuous interaction.

*It's important to note that while ChatGPT can generate impressive responses, it may sometimes produce incorrect or nonsensical answers. It doesn't have real-world experiences or access to current information beyond its knowledge cutoff date. It's always a good idea to fact-check and verify important information from reliable sources.*

**Ioannis:** *What is ALERT Geomaterials in summary?*

**ChatGPT:** *In summary, ALERT Geomaterials is an international scientific association dedicated to the study of the mechanics and behavior of geomaterials, including soil, rock and other geological materials. It brings together researchers and professionals in the field of geomechanics and geotechnical engineering to promote collaboration, knowledge exchange*

*and scientific advancements. ALERT Geomaterials organizes workshops, conferences and training courses, providing a platform for researchers to share their findings and discuss topics related to soil and rock mechanics, slope stability, geotechnical engineering, experimental testing and numerical modeling. The association also publishes a journal focused on energy geotechnics and geoenvironmental engineering. Overall, ALERT Geomaterials plays a crucial role in fostering research and innovation in the field of geomaterials.*

Impressive! At this point, it becomes quite tempting to let ChatGPT finish this chapter or even both volumes! However, I assure you, we took the necessary time to entirely draft these volumes word by word. Read carefully ChatGPT's answers. Are all of them correct?

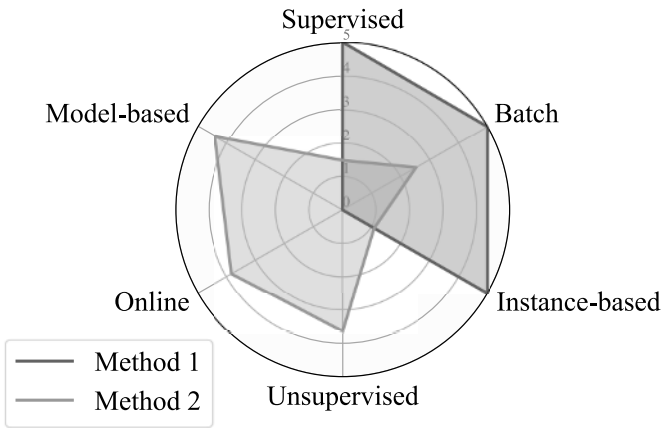
It is straightforward to see how the above definition for ML applies to ChatGPT and to many other applications of ML, such as speech, handwriting, face and object recognition, voice-to-text and vice versa, translation, text auto-correction and auto-completion, spam filtering, computer games, self-driving cars, medicine, forecasting, banking, security, marketing, control problems, engineering and, of course, (geo-)mechanics.

ML is an evolving field of knowledge and involves a plethora of methods and combinations of those. In the next section, we will try to categorize and classify them in groups.

Having described and defined what ML is, it is natural to ask what is *artificial intelligence* (AI)? ML and AI are closely related and the latter is considered to include the former. The exact definition of AI seems to be a bit foggy for the time being and depends on how we define the terms “artificial” and “intelligence”. In the author's opinion, it is easier to describe the characteristics of AI, rather than give a unique and exact definition of the term. Another example of a hard definition, in a totally different domain, is what is justice? It seems easier and more important to describe the characteristics of justice (e.g. equality for all) rather than give a precise definition of the term.

## 1.2. Classification of ML methods

There are numerous ML methods in the literature. Therefore, it is useful to classify them into different categories. Here, we follow the classification of (Géron 2019), who categorizes ML to *supervised learning* versus *unsupervised learning*, *batch learning* versus *online learning* and *instance-based learning* versus *model-based learning* methods. Of course, this is a rough classification and one method can combine different categories, as shown in Figure 1.2.



**Figure 1.2.** Classification of ML methods. Method 1 is a supervised, batch, instance-based method, while Method 2 has components belonging to different categories

### 1.2.1. Supervised versus unsupervised ML

In *supervised learning*, the training set includes the desired solutions/predictions, which are called *labels*.

For instance, imagine that your training set is hundreds of images with each one containing either a dog or a cat and that the training task is to distinguish the photos of dogs from those of cats. This is called a *classification* problem (see Chapter 4). If we give to the computer the information of which photos show dogs and which cats, or in other

words, if we label the training set, then the learning is called supervised. If the training is successful, then the computer will have learned to identify cats from dogs even in a new set of photos that it has not processed during its training.

On the contrary, if the training set is unlabeled, then the learning is called *unsupervised learning*. In this case, the computer will eventually understand the pattern of the two different animals shown in the photos and it will be able to distinguish dogs from cats in an unsupervised manner (see also *clustering* in Chapter 3 of volume 1). Of course, the machine will not have learned to call a cat, cat and a dog, dog, because we have not given this additional information, but it will have identified their differences and separated the data into two different classes.

Another ML problem is *regression* (see Chapters 2 and 5 of volume 1 and Chapters 3 and 5 of volume 2). In regression, the training set contains one or several numerical inputs, also called *features*, and the task is to predict one or several numerical outputs, also called *predictors*, that depend on the inputs. Consider as an example the prediction of the stress response of a geomaterial, which as we know depends on several input parameters, such as the applied strain, available information about the evolution of its microstructure (e.g. the position and the velocity of the grains of a sand obtained by a discrete element method [DEM] analysis), history and/or other features. As the training set contains both the output and the input, the training for predicting the stresses based on the above-mentioned features is supervised. However, the identification/extraction by the machine of a representation of the most important features of the microstructure that are related to the prediction of the stress response is unsupervised. Examples of unsupervised methods in ML are *feature extraction*, *anomaly detection* and *dimensionality reduction*, in which the aim is to reduce the size of the training set without losing important information with respect to a specific task or measure and *data compression* (among others).

In Table 1.1, we provide a list of important supervised and unsupervised methods in ML. However, not all ML methods can be categorized to supervised and unsupervised. A notable example is

*reinforcement learning* (RL), which does not require labeled data or a training set. For more details on RL, we refer to Chapter 4 of volume 2. Finally, when labels are not available for all the samples in the training set, we refer to *semisupervised learning*. Most semisupervised methods are a combination of supervised and unsupervised algorithms. An example of semisupervised learning is *active learning* (see Chapter 4 of volume 1 for more details).

ML method	Supervised	Unsupervised	Online	Presented in these volumes
Linear regression	✓			✓
Logistic regression	✓			✓
Polynomial regression	✓			✓
Lasso, Ridge	✓			✓
$k$ -Nearest neighbors	✓			✓
Support vector machines (SVM)	✓			✓
Decision trees	✓			
Random forests	✓			
Artificial neural networks (ANNs)	✓		✓	✓
Autoencoders		✓	✓	✓
Clustering		✓		✓
Principal component analysis (PCA)		✓	✓*	
Locally linear embedding (LLE)		✓		
Reinforcement learning (RL)			✓	✓

**Table 1.1.** Classification of some important methods in ML. The asterisk denotes that not all variations of the method fall into this category

### 1.2.2. Batch versus online ML

Another manner to classify ML algorithms is based on whether they can improve their predictions by providing them with new data that may become available after the first training.

In *batch learning*, the machine has to be trained over a fixed training data set, without being able to add more data to the training set. Therefore, batch ML methods cannot improve their performance in a specific task by providing them with more data after the end of the training.

On the contrary, ML methods that support *online learning* allow to modify and increase the initial training set. Consequently, they are more flexible and suitable for large training sets, as they can be partially loaded into the memory of the computer and used whenever needed. Old data, over which the computer has been already trained, can also be erased to save space. Online learning is also very convenient when a constant flow of information exists, contrary to batch learning, in which the machine must be retrained over the whole data set. In other words, in batch learning, when new data become available the training set has to be updated and the training has to be repeated from scratch. This can be fine for some applications, but in many others it could have a very high computational and data storage cost.

Not all ML algorithms support online learning (see Table 1.1).

### **1.2.3. Instance-based versus model-based ML**

One more way to categorize ML methods is based on the way data is learned in order to make predictions for data outside the training set. Two ML categories can be distinguished: *instance-based* (also known as *memory-based* or *lazy*) learning and *model-based* learning (also known as *physics-based* in some applications).

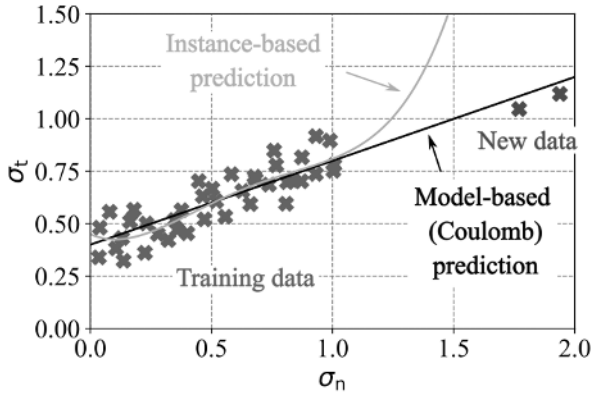
In instance-based learning, training data are simply interpolated in a high (usually) dimensional space. Then, new predictions are made based on how close or how similar new data are to those used for training. Similarity is measured on the basis of a distance measure, depending on the data and the problem at hand. In this sense, data are “memorized” by the machine, thus the term memory based. The quality of the predictions of data outside the training set is determined by how well new data are represented by the data of the training set. Instance-based approaches are straightforward to apply in any data set

and they do not require any particular knowledge about the structure of the data or other characteristics that they might have. The hope of the user of instance-based methods is the machine will eventually identify by itself the hidden patterns in the data and give correct predictions even for data outside the training set. Examples of instance-based methods are *artificial neural networks* (ANNs), *decision trees*, *random forests*, *k-nearest neighbors* and many *clustering* techniques, *locally linear embedding* (LLE) and *principal component analysis* (PCA).

Despite the versatility and the many advantages of instance-based ML methods, they have an important drawback. In physics and engineering, we know that data have to respect at least some fundamental principles, for instance, the conservation of mass, energy and linear and angular momentum. Therefore, instance-/memory-based predictions that do not respect these conservation laws are unacceptable and can be even dangerous for applications. Data in this case have to be processed and learned by the machine under a model that includes the laws of physics.

Model-based ML adopts a model with some model parameters, which are optimized during training in order to optimally represent the data of the training set. Then the model is used to make predictions for unseen input data. Of course, if the adopted model is poor, the predictions will be poor as well.

Both instance-based and model-based approaches can introduce bias: the former because of limited data over which they were trained and the latter because of the model chosen. To fix the ideas, an example of instance-based and model-based ML is given in Figure 1.3. The training set contains the shear stress at failure (*predictor*) for a given normal stress (*feature*) of a series of experimental tests of a frictional interface. An instance-based method could give very poor predictions for unseen data, while a model-based prediction will be as good as the model is for describing the data. In this simple example, a Coulomb model was adopted for the model-based approach.



**Figure 1.3.** Instance-based versus model-based ML for predicting failure (slip) of a frictional interface based on experimental data (fictitious)

### 1.3. ML and geomechanics

The applications of ML in science are nowadays numerous and increasing. The same holds for applications of ML in geomechanics. It is out of the scope of this chapter to provide a comprehensive literature review, but it is worth mentioning some research directions in *constitutive modeling*, *geotechnics*, *geophysics* and *image correlation* that can be a starting point for the interested reader.

Geomaterials are among the most complex materials to study and model. The main reason is that their mechanical response is governed by multiphysics couplings at multiple spatial and temporal scales, which result in a macroscopic strongly nonlinear and dissipative response.

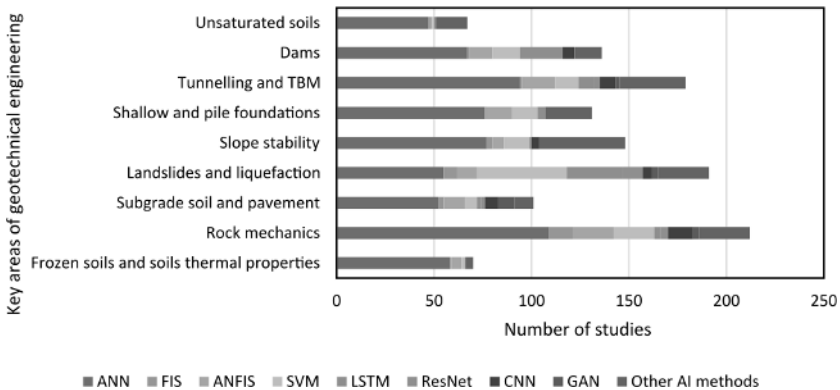
Empirical constitutive models are often used to describe the mechanical behavior of geomaterials. Classification and regression methods can be used for choosing the appropriate constitutive model that best fits the experimental data (see (Marques et al. 2022; Morand and Helm 2019; Guo et al. 2021; Schulte et al. 2023; Pouragha et al. 2020; Zhang et al. 2021a) and references therein). Another challenging task is to capture the macroscopic behavior of geomaterials based on

the behavior of their microstructure and its evolution due to loading. Multiscale approaches can be employed for this purpose, requiring the solution of a boundary value problem (BVP) of an elementary cell of the microstructure at each point and time increment of the macroscopic analysis. Then, based on homogenization, the intrinsic constitutive behavior of the microstructure can be upscaled to the macroscopic level. For this purpose, mixed numerical schemes, such as the  $FE^2$  method (Feyel 2003; Lloberas Valls et al. 2019; Eijnden et al. 2017) and  $FEM \times DEM$  (Nitka et al. 2011; Nguyen et al. 2014) are often employed, depending on the nature of the microstructure. However, the computational cost of these methods is extremely high for real-scale applications, if not prohibitive. ML and ANN are one of the most promising ways to speed-up this multiscale process.

Recent works have shown that ANN can successfully encapsulate several aspects of the constitutive behavior of the underlying microstructure and provide the necessary information to the macroscopic scale with reduced calculation cost (see (Ghaboussi et al. 1991; Lefik and Schrefler 2003; Mozaffar et al. 2019; Liu and Wu 2019; Huang et al. 2020; Vlassis and Sun 2021; Zhou et al. 2021; Wang et al. 2019; Rocha et al. 2021; Becker et al. 2022; Wu et al. 2023; Sun et al. 2022; Zhang et al. 2023; Peng et al. 2021) to mention few). Going a step further, ANN can be designed in such a way to respect, by construction, the laws of physics (Karniadakis et al. 2021; Raissi et al. 2019), symmetries (Heider et al. 2020) and thermodynamics (Masi et al. 2021; Masi and Stefanou 2022, 2023; Hernandez et al. 2021) (see also Chapter 3 of volume 2). Once trained, these approaches can tremendously speed up the solution of difficult multiscale problems, guarantee the respect of the thermodynamic restrictions in their predictions and enable the extraction of the hidden state variables of the material. The latter can, in turn, shed light on the importance and the role of specific micromechanisms to the overall macroscopic behavior of complex (geo-)materials. The above collection of ML approaches is enriched by the so-called “data-driven” methods, which present an alternative formulation, whereby optimal material states are sought within a dataset that most closely satisfies

momentum and energy conservation principles (Karapiperis et al. 2020; Kirchdoerfer and Ortiz 2016, 2018; Karapiperis et al. 2021) (see also Chapter 1 of volume 2).

Moving to applications of ML in geotechnics, according to the recent review of (Baghbani et al. 2022) (see also (Zhang et al. 2021b)), more than 1,200 articles can be found in the literature starting from the early 1990s. According to the same source, a net burst of production of scientific articles is observed after 2017. Note that more than half of these works use ANN. Several areas of geotechnical engineering are covered (see Figure 1.4), such as frozen soils and soil thermal properties, rock mechanics, subgrade soils and pavements, landslides (see (Tehrani et al. 2022) for a recent review), liquefaction, slope stability, shallow foundations, piles, tunneling and tunnel boring machines, dams and unsaturated soils, among others. In 2018, the increasing interest of the geotechnical community in ML led (Zhang and Liu 2022) to the creation of a new technical committee (TC) in the International Society for Soil Mechanics and Geotechnical Engineering (ISSMGE), entitled TC309 “Machine Learning and Big Data” (<https://www.issmge.org/committees/technical-committees/impact-on-society/machine-learning>, see also TC304).



**Figure 1.4.** Number of published articles using ML for geotechnical applications (slightly modified from Baghbani et al. (2022))

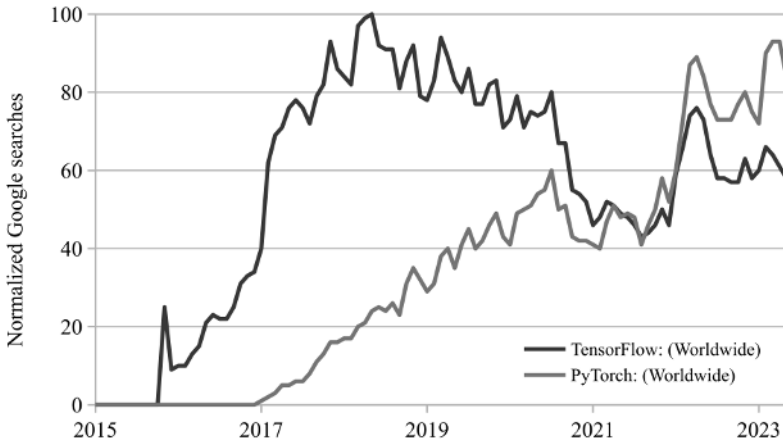
ML also has promising applications in geophysics. For instance, it finds applications in geophysical exploration, reservoir engineering and drilling (Sircar et al. 2021 and references therein). Moreover, it was used for creating synthetic accelerograms based on numerical simulations or databases of real earthquake signals (Gatti and Clouteau 2020; Lehmann et al. 2022). ML was used as well in an attempt for earthquake prediction – the holy grail in seismology – showing that earthquake-like events in the laboratory could be predicted by identification by the machine of seismic precursor patterns (Johnson et al. 2021; Beroza et al. 2021; Laurenti et al. 2022; Ridzwan and Yusoff 2023). RL was also used for controlling earthquake-like events (Papachristos and Stefanou 2021) (see Chapter 4 of volume 2) that more exact mathematical theories show are controllable (Stefanou 2019; Gutiérrez-Oribio et al. 2022).

Given the striking advances of ML in self-driving cars, and object and face recognition (Balaban 2015), it is natural to expect ML methods to also find numerous applications in image processing in experimental geomechanics and geotechnics. *Convolutional neural networks* (CNNs) (see Chapter 5 of volume 2 for more details) are the basic ingredient of most ML techniques used in image and video processing, without forgetting more traditional compression techniques that are also considered as ML (e.g. PCA). Boukhtache et al. (2021) presents a review of digital image correlation (DIC) with deep learning. We also refer to some previous studies (Boukhtache et al. 2023; Cheng et al. 2023a; Duan et al. 2023) for some recent developments. Accuracy is one of the main issues for those methods in order to outperform the current state of the art (see ALERT Doctoral School 2022 (Ando et al. 2022)). Focusing on granular materials, (Cevallos et al. 2023) propose a CNN to automatically distinguish properly segmented digital grains with up to 90% accuracy, while (Cheng et al. 2023b) present an ML-based strategy to estimate the contact force chains of uniformly sized spherical granular materials using particle kinematics and inter-particle contact evolution data measured by X-ray micro-tomography.

## 1.4. Libraries for ML

Today, many libraries exist for ML (see [https://en.wikipedia.org/wiki/Machine\\_learning#Software](https://en.wikipedia.org/wiki/Machine_learning#Software) for an updated list). Most of the existing libraries provide a Python interface.

Some general purpose libraries for ML are Numpy (Harris et al. 2020), Pandas (McKinney 2010), SciPy (Virtanen et al. 2020) and Scikit-learn (Pedregosa et al. 2011). For ANNs, TensorFlow (Abadi et al. 2015) and PyTorch (Paszke et al. 2019) are equally popular at present (see Figure 1.5). Both offer parallelization and GPU support for training large ANN and handling large collections of data. The examples presented in these volumes make use of PyTorch (see Chapter 5 of volume 1 for more details). As always, the best library is the library that we know the best, provided that it is open source and it allows us to do what we want!



**Figure 1.5.** Normalized Google searches over time for TensorFlow and PyTorch ANN libraries

## 1.5. Bias in ML and limitations

The success of ML algorithms in performing complicated and complex tasks makes them convenient tools for many applications.

However, we have to be aware of their limitations. In other words, we have to use the right screwdriver for the right screw! In section 1.2.3 (see Figure 1.3), we showed how the choice of the ML method and the quality of the data can lead to biased predictions, even in a very simple problem.

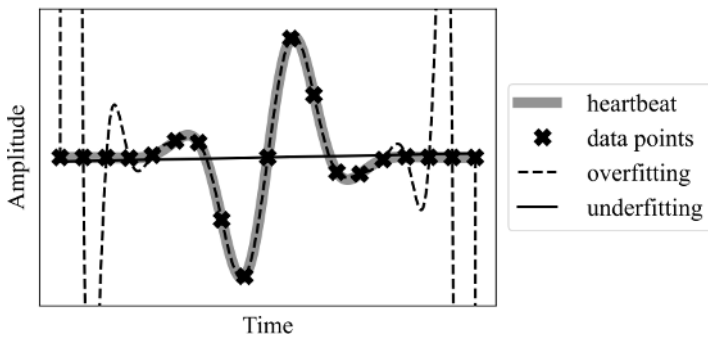
Applications of ML algorithms have demonstrated gender (Kelion 2019), racial (Dressel and Farid 2018; Perkwitz 2021), hiring (Votto et al. 2021) and other biases (Varsha 2023). ML methods are algorithms which should be used with care and knowledge of the underlying limitations. This is not different from the application of other methods. For instance, we know how wrong finite element predictions can be when the appropriate finite elements are not used, when convergence analyses are not performed or when we want to model a softening material without regularizing the underlying mathematical problem. The inappropriate use of specific methods has traditionally led to spectacular failures with uncountable casualties and economic loss. ML methods will not be an exception, unless we understand them better and use them with caution.

As far as it concerns ML and geomechanics, luckily, we have at our disposal established and undeniable principles that have to be respected in any application. Conservation principles, the laws of thermodynamics and other physics at various spatiotemporal scales should be incorporated into the ML algorithms in order to assure adequate and safe predictions. Physics-based approaches in ML (see Chapters 1, 2 and 3 of volume 2) gain more and more attention from the scientific community and can become the natural environment for marrying the established know-how of decades of research in geomechanics with ML. This could give fresh ideas and an opportunity to push further the current state of the art in our fields. The incorporation of physics in ML could eventually inspire new ML methods in other domains too, which today suffer from inevitable bias (see social sciences where data are always limited).

Another limitation of most of the available ML methods today is their greediness for data. We do not need thousand of photos to teach a child what a car is! With ML though we need tons of data! In many

applications, there is an abundance of data (see, for instance, the data that are produced, but not saved, at each increment at each Gauss point during a nonlinear finite element analysis in a geomechanics problem). In some other applications though, data may not be enough (see, for example, X-rays of thousand of specimens; Tengattini et al. 2020).

Noise in the data, overfitting and underfitting are some other points that we have to pay attention to. Data with a lot of noise can make the learning process hard and render the predictions unreliable and of poor generalization (see also Chapters 1–3 of volume 2). Data preparation to assure good quality is of paramount importance then.



**Figure 1.6.** Data points of a synthetic heartbeat, interpolated by two polynomial functions. The high-degree polynomial exactly fits the data points, but overfits the signal. The low-degree polynomial poorly fits the data points and underfits the signal. Both regressions are not acceptable, because they either predict a superhuman heartbeat or a dead person!

The choice of a ML method allowing to fit data in a very high dimensional space can lead to overfitting. In this case, the training data can be very well represented, but predictions for unseen data can be far off. The opposite happens with underfitting, where the space for fitting is too low for identifying and reproducing the inherent patterns in the data. A classical example of overfitting and underfitting in polynomial regression is shown in Figure 1.6. Thankfully, there are ways for estimating in practice how well an ML model performs. This is mainly achieved by splitting the training set into training, validation and test

data. *Regularization* methods can also help in avoiding this problem. For more details, we refer to Chapters 2, 4 and 5 of volume 1.

## 1.6. What to expect from these volumes?

These volumes aim to explain what ML is, what its main methods are and how they can be used for solving problems in (geo-)mechanics. Most of the chapters were written to provide a pedagogical introduction to the most important methods in ML and the fundamental notions behind them.

It is not possible to cover all the available ML methods in the existing literature and, without any doubt, many important methods were inevitably left out. For instance, we will not discuss genetic algorithms, PCA and related methods, particle swarm optimization, fuzzy logic algorithms, ML methods based on control theory and many others. We hope, however, to have provided a good selection of ML methods for an introductory course.

We expect the material presented in these volumes to help the readers to:

- demystify and understand what ML is;
- access the fundamental notions of the most important ML methods;
- use ML in simple examples, get aware of pitfalls and understand the need for physics-based ML methods for solving problems in (geo-)mechanics.

These volumes are addressed to undergraduate and graduate level. The minimum requirements for accessing them are:

- knowledge of Python programming language<sup>1</sup>;
- basic concepts in mathematics (calculus and elements of differential calculus and numerical analysis);

---

<sup>1</sup> For learning Python, we propose the book *Python Crash Course: A Hands-On, Project-Based Introduction to Programming* by Eric Matthes and/or many excellent tutorials that can be found on the internet.

– nice problems in mind that could combine ML and geomechanics!

## 1.7. Acknowledgments

The author would like to acknowledge the support of the European Research Council (ERC) under the European Union's Horizon2020 research and innovation program (Grant agreement no. 757848 CoQuake).

## 1.8. References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. et al. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems [Online]. Available at: <https://www.tensorflow.org/>.
- Allis, L.V. (1994). Searching for solutions in games and artificial intelligence. Doctoral Thesis, Maastricht University.
- Ando, E., Marks, B., Hurley, R., Dijkstra, J.A. (2022). Advanced experimental geomechanics. Lecture Notes, The Alliance of Laboratories in Europe for Education, Research and Technology [Online]. Available at: [https://www.alertgeomaterials.eu/data/school/2022/2022\\_ALERT\\_school.pdf](https://www.alertgeomaterials.eu/data/school/2022/2022_ALERT_school.pdf).
- Baghbani, A., Choudhury, T., Costa, S., Reiner, J. (2022). Application of artificial intelligence in geotechnical engineering: A state-of-the-art review. *Earth-Science Reviews* 228, 103991 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0012825222000757>.
- Balaban, S. (2015). Deep learning and face recognition: The state of the art. *Biometric and Surveillance Technology for Human and Activity Identification XII*, Baltimore, MD, 94570B [Online]. Available at: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2181526>.
- Becker, S., Devijver, E., Molinier, R., Jakse, N. (2022). Unsupervised topological learning approach of crystal nucleation. *Scientific Reports* 12(1), 3195 [Online]. Available at: <https://www.nature.com/articles/s41598-022-06963-5>.

- Beroza, G.C., Segou, M., Mostafa Mousavi, S. (2021). Machine learning and earthquake forecasting—next steps. *Nature Communications* 12(1), 4761 [Online]. Available at: <https://www.nature.com/articles/s41467-021-24952-6>.
- Boukhtache, S., Abdelouahab, K., Berry, F., Blaysat, B., Grédiac, M., Sur, F. (2021). When deep learning meets digital image correlation. *Optics and Lasers in Engineering*, 136, 106308 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0143816620306588>.
- Boukhtache, S., Abdelouahab, K., Bahou, A., Berry, F., Blaysat, B., Grédiac, M., Sur, F. (2023). A lightweight convolutional neural network as an alternative to DIC to measure in-plane displacement fields. *Optics and Lasers in Engineering*, 161, 107367 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0143816622004183>.
- Cevallos, S.B., Jerves, A.X., Mital, U., Medina, D.A., Quinteros, V.S., Mulas, M., Torgersrud, N. (2023). Towards a more accurate characterization of granular media 2.0: Involving AI in the process. *Computers and Geotechnics*, 160, 105510 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0266352X23002677>.
- Cheng, X., Zhou, S., Xing, T., Zhu, Y., Ma, S. (2023a). Solving digital image correlation with neural networks constrained by strain-displacement relations. *Optics Express*, 31(3), 3865 [Online]. Available at: <https://opg.optica.org/abstract.cfm?URI=oe-31-3-3865>.
- Cheng, Z., Wang, J., Xiong, W. (2023b). A machine learning-based strategy for experimentally estimating force chains of granular materials using X-ray micro-tomography. *Géotechnique*, 1–13 [Online]. Available at: <https://www.icevirtuallibrary.com/doi/10.1680/jgeot.21.00281>.
- Dressel, J. and Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, 4(1), eaao5580 [Online]. Available at: <https://www.science.org/doi/10.1126/sciadv.aao5580>.

- Duan, X., Xu, H., Dong, R., Lin, F., Huang, J. (2023). Digital image correlation based on convolutional neural networks. *Optics and Lasers in Engineering* 160, 107234 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0143816622002871>.
- Eijnden, A.P.V.D., Bésuelle, P., Collin, F., Chambon, R., Desrues, J. (2017). Modeling the strain localization around an underground gallery with a hydro-mechanical double scale model; effect of anisotropy. *Computers and Geotechnics*, 85, 384–400. doi: 10.1016/j.compgeo.2016.08.006.
- Feyel, F. (2003). A multilevel finite element method (FE<sup>2</sup>) to describe the response of highly non-linear structures using generalized continua. *Computer Methods in Applied Mechanics and Engineering* 192(28–30), 3233–3244.
- Gatti, F. and Clouteau, D. (2020). Towards blending physics-based numerical simulations and seismic databases using generative adversarial network. *Computer Methods in Applied Mechanics and Engineering* 372, 113421 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S004578252030606X>.
- Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Sebastopol, CA.
- Ghaboussi, J., Garrett, J.H., Wu, X. (1991). Knowledge based modeling of material behavior with neural networks. *Journal of Engineering Mechanics*, 117(1), 132–153.
- Guo, Z., Bai, R., Lei, Z., Jiang, H., Liu, D., Zou, J., Yan, C. (2021). CPINet: Parameter identification of path-dependent constitutive model with automatic denoising based on CNN-LSTM. *European Journal of Mechanics – A/Solids*, 90, 104327 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S099775382100098X>.
- Gutiérrez-Oribio, D., Tzortzopoulos, G., Stefanou, I., Plestan, F. (2022). Earthquake control: An emerging application for robust control. Theory and experimental tests. *IEEE Transactions on Control Systems Technology*, 31, 1747–1761.

- Harris, C.R., Millman, K.J., Walt, S.J.V. D., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J. et al. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. doi: 10.1038/s41586-020-2649-2.
- Heider, Y., Wang, K., Sun, W. (2020). SO(3)-invariance of informed-graph-based deep neural network for anisotropic elastoplastic materials. *Computer Methods in Applied Mechanics and Engineering* 363, 112875.
- Hernandez, Q., Badías, A., González, D., Chinesta, F., Cueto, E. (2021). Deep learning of thermodynamics-aware reduced-order models from data. *Computer Methods in Applied Mechanics and Engineering*, 379, 113763 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0045782521000992>.
- Huang, D.Z., Xu, K., Farhat, C., Darve, E. (2020). Learning constitutive relations from indirect observations using deep neural networks. *Journal of Computational Physics*, 416, 109491.
- Johnson, P.A., Rouet-Leduc, B., Pyrak-Nolte, L.J., Beroza, G.C., Marone, C.J., Hulbert, C., Howard, A., Singer, P., Gordeev, D., Karaflos, D. et al. (2021). Laboratory earthquake forecasting: A machine learning competition. *Proceedings of the National Academy of Sciences*, 118(5), e2011362118 [Online]. Available at: <https://pnas.org/doi/full/10.1073/pnas.2011362118>.
- Karapiperis, K., Stainier, L., Ortiz, M. and Andrade, J. (2020). Data-driven multiscale modeling in mechanics. *Journal of the Mechanics and Physics of Solids*, 147, 104239. doi: 10.1016/j.jmps.2020.104239.
- Karapiperis, K., Ortiz, M., Andrade, J. (2021). Data-driven nonlocal mechanics: Discovering the internal length scales of materials. *Computer Methods in Applied Mechanics and Engineering*, 386, 114039. doi: 10.1016/j.cma.2021.114039.
- Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, 3(6), 422–440 [Online]. Available at: <https://www.nature.com/articles/s42254-021-00314-5>.

- Kelion, L. (2019). Apple's 'sexist' credit card investigated by US regulator. *BBC* [Online]. Available at: <https://www.bbc.com/news/business-50365609>.
- Kirchdoerfer, T. and Ortiz, M. (2016). Data-driven computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 304, 81–101. doi: 10.1016/j.cma.2016.02.001.
- Kirchdoerfer, T. and Ortiz, M. (2018). Data-driven computing in dynamics. *International Journal for Numerical Methods in Engineering*, 113(11), 1697–1710.
- Laurenti, L., Tinti, E., Galasso, F., Franco, L., Marone, C. (2022). Deep learning for laboratory earthquake prediction and autoregressive forecasting of fault zone stress. *Earth and Planetary Science Letters*, 598, 117825 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0012821X22004617>.
- Lefik, M. and Schrefler, B.A. (2003). Artificial neural network as an incremental non-linear constitutive model for a finite element code. *Computer Methods in Applied Mechanics and Engineering*, 192(28–30), 3265–3283.
- Lehmann, F., Gatti, F., Bertin, M., Clouteau, D. (2022). Machine learning opportunities to conduct high-fidelity earthquake simulations in multi-scale heterogeneous geology. *Frontiers in Earth Science*, 10, 1029160 [Online]. Available at: <https://www.frontiersin.org/articles/10.3389/feart.2022.1029160/full>.
- Liu, Z. and Wu, C.T. (2019). Exploring the 3D architectures of deep material network in data-driven multiscale mechanics. *Journal of the Mechanics and Physics of Solids*, 127, 20–46. arXiv: 1901.04832.
- Lloberas Valls, O., Raschi Schaw, M., Huespe, A.E., Oliver Olivella, X. (2019). Reduced finite element square techniques (RFE2): Towards industrial multiscale FE software. In *COMPLAS 2019: XV International Conference on Computational Plasticity: Fundamentals and Applications*. International Centre for Numerical Methods in Engineering (CIMNE).

- Marques, A., Pereira, A., Ribeiro, B., Prates, P.A. (2022). On the identification of material constitutive model parameters using machine learning algorithms. *Key Engineering Materials*, 926, 2146–2153 [Online]. Available at: <https://www.scientific.net/KEM.926.2146>.
- Masi, F. and Stefanou, I. (2022). Multiscale modeling of inelastic materials with thermodynamics-based artificial neural networks (TANN). *Computer Methods in Applied Mechanics and Engineering*, 398, 115190 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0045782522003450>.
- Masi, F. and Stefanou, I. (2023). Evolution TANN and the identification of internal variables and evolution equations in solid mechanics. *Journal of the Mechanics and Physics of Solids*, 105245 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0022509623000492>.
- Masi, F., Stefanou, I., Vannucci, P., Maffi-Berthier, V. (2021). Thermodynamics-based artificial neural networks for constitutive modeling. *Journal of the Mechanics and Physics of Solids* 147, 104277 [Online]. Available at: <http://arxiv.org/abs/2005.12183>.
- McKinney, W. (2010). Data structures for statistical computing in Python. In *Proceedings of the 9th Python in Science Conference*, Austin, TX.
- Mitchell, T.M. (1997). *Machine Learning*. McGraw-Hill, New York.
- Morand, L. and Helm, D. (2019). A mixture of experts approach to handle ambiguities in parameter identification problems in material modeling. *Computational Materials Science*, 167, 85–91 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0927025619302058>.
- Mozaffar, M., Bostanabad, R., Chen, W., Ehmann, K., Cao, J., Bessa, M.A. (2019). Deep learning predicts path-dependent plasticity. *Proceedings of the National Academy of Sciences* 116(52), 26414–26420.

- Nguyen, T.K., Combe, G., Caillerie, D., Desrues, J. (2014). FEM  $\times$  DEM modelling of cohesive granular materials: Numerical homogenisation and multi-scale simulations. *Acta Geophysica* 62(5), 1109–1126 [Online]. Available at: <http://link.springer.com/10.2478/s11600-014-0228-3>.
- Nitka, M., Combe, G., Dascalu, C., Desrues, J. (2011). Two-scale modeling of granular materials: A DEM-FEM approach. *Granular Matter*, 13(3), 277–281.
- Papachristos, E. and Stefanou, I. (2021). Controlling earthquake-like instabilities using artificial intelligence. *arXiv* [Online]. Available at: <http://arxiv.org/abs/2104.13180>.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L. et al. (2019). PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*. Curran Associates, Inc, Vancouver [Online]. Available at: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Peng, G.C.Y., Alber, M., Buganza Tepole, A., Cannon, W.R., De, S., Dura-Bernal, S., Garikipati, K., Karniadakis, G., Lytton, W.W., Perdikaris, P. et al. (2021). Multiscale modeling meets machine learning: What can we learn? *Archives of Computational Methods in Engineering*, 28(3), 1017–1037 [Online]. Available at: <https://link.springer.com/10.1007/s11831-020-09405-5>.
- Perkowitz, S. (2021). The bias in the machine: Facial recognition technology and racial disparities. *MIT Case Studies in Social and Ethical Responsibilities of Computing*, February [Online]. Available at: <https://mit-serc.pubpub.org/pub/bias-in-machine>.

- Pouragha, M., Eghbalian, M., Wan, R. (2020). A note on applicability of artificial intelligence to constitutive modeling of geomaterials. *Journal of Machine Learning for Modeling and Computing*, 1(2), 157–170 [Online]. Available at: <http://www.dl.begellhouse.com/journals/558048804a15188a,2cbcbe11139f18e5,478af42571e4a7ae.html>.
- Raissi, M., Perdikaris, P., Karniadakis, G.E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. doi: 10.1016/j.jcp.2018.10.045.
- Ridzwan, N.S.M. and Yusoff, S.H.M. (2023). Machine learning for earthquake prediction: A review (2017–2021). *Earth Science Informatics*, 16(2), 1133–1149 [Online]. Available at: <https://link.springer.com/10.1007/s12145-023-00991-z>.
- Rocha, I., Kerfriden, P., Van Der Meer, F. (2021). On-the-fly construction of surrogate constitutive models for concurrent multiscale mechanical analysis through probabilistic machine learning. *Journal of Computational Physics: X*, 9, 100083 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S2590055220300354>.
- Schulte, R., Karca, C., Ostwald, R., Menzel, A. (2023). Machine learning-assisted parameter identification for constitutive models based on concatenated loading path sequences. *European Journal of Mechanics – A/Solids*, 98, 104854 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0997753822002844>.
- Shannon, C.E. (1950). XXII. Programming a computer for playing chess. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 41(314), 256–275 [Online]. Available at: <http://www.tandfonline.com/doi/abs/10.1080/14786445008521796>.

- Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M. et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484–489 [Online]. Available at: <https://www.nature.com/articles/nature16961>.
- Sircar, A., Yadav, K., Rayavarapu, K., Bist, N., Oza, H. (2021). Application of machine learning and artificial intelligence in oil and gas industry. *Petroleum Research*, 6(4), 379–391 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S2096249521000429>.
- Stefanou, I. (2019). Controlling anthropogenic and natural seismicity: Insights from active stabilization of the Spring–Slider model. *Journal of Geophysical Research: Solid Earth*, 124(8), 8786–8802 [Online]. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1029/2019JB017847>.
- Sun, X., Bahmani, B., Vlassis, N.N., Sun, W., Xu, Y. (2022). Data-driven discovery of interpretable causal relations for deep learning material laws with uncertainty propagation. *Granular Matter*, 24, 1. doi: 10.1007/s10035-021-01137-y.
- Tehrani, F.S., Calvello, M., Liu, Z., Zhang, L., Lacasse, S. (2022). Machine learning and landslide studies: Recent advances and applications. *Natural Hazards*, 114(2), 1197–1245 [Online]. Available at: <https://link.springer.com/10.1007/s11069-022-05423-7>.
- Tengattini, A., Lenoir, N., Andò, E., Giroud, B., Atkins, D., Beaucour, J., Viggiani, G. (2020). NeXT-Grenoble, the neutron and X-ray tomograph in Grenoble. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 968, 163939 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0168900220304198>.

- Tromp, J. and Farneback, G. (2007). Combinatorics of Go. In *Computers and Games*, Hutchison, D., Kanade, T., Kittler, J. et al. (eds). Springer, Berlin, Heidelberg [Online]. Available at: [http://link.springer.com/10.1007/978-3-540-75538-8\\_8](http://link.springer.com/10.1007/978-3-540-75538-8_8).
- Varsha, P.S. (2023). How can we manage biases in artificial intelligence systems – A systematic literature review. *International Journal of Information Management Data Insights*, 3(1), 100165 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S2667096823000125>.
- Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J. et al. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17, 261–272.
- Vlassis, N.N. and Sun, W. (2021). Component-based machine learning paradigm for discovering rate-dependent and pressure-sensitive level-set plasticity models. *Journal of Applied Mechanics*, 1–13 [Online]. Available at: <https://asmedigitalcollection.asme.org/appliedmechanics/article/doi/10.1115/1.4052684/1121215/Component-based-machine-learning-paradigm-for>.
- Votto, A.M., Valecha, R., Najafirad, P., Rao, H.R. (2021). Artificial intelligence in tactical human resource management: A systematic literature review. *International Journal of Information Management Data Insights*, 1(2), 100047 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S2667096821000409>.
- Wang, K., Sun, W., Du, Q. (2019). A cooperative game for automated learning of elasto-plasticity knowledge graphs and models with AI-guided experimentation. *Computational Mechanics*, 64(2), 467–499 [Online]. Available at: <http://link.springer.com/10.1007/s00466-019-01723-1>.

- Wu, M., Xia, Z., Wang, J. (2023). Constitutive modelling of idealised granular materials using machine learning method. *Journal of Rock Mechanics and Geotechnical Engineering*, 15(4), 1038–1051 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S1674775522001688>.
- Zhang, W. and Liu, Z. (2022). Editorial for machine learning in geotechnics. *Acta Geotechnica*, 17(4), 1017–1017 [Online]. Available at: <https://link.springer.com/10.1007/s11440-022-01563-z>.
- Zhang, P., Yin, Z.-Y., Jin, Y.-F. (2021a). State-of-the-art review of machine learning applications in constitutive modeling of soils. *Archives of Computational Methods in Engineering*, 28(5), 3661–3686 [Online]. Available at: <https://link.springer.com/10.1007/s11831-020-09524-z>.
- Zhang, W., Li, H., Li, Y., Liu, H., Chen, Y., Ding, X. (2021b). Application of deep learning algorithms in geotechnical engineering: A short critical review. *Artificial Intelligence Review*, 54(8), 5633–5673 [Online]. Available at: <https://link.springer.com/10.1007/s10462-021-09967-1>.
- Zhang, N., Zhou, A., Jin, Y.-F., Yin, Z.-Y., Shen, S.-L. (2023). An enhanced deep learning method for accurate and robust modelling of soil stress–strain response. *Acta Geotechnica* [Online]. Available at: <https://link.springer.com/10.1007/s11440-023-01813-8>.
- Zhou, X.-H., Han, J., Xiao, H. (2021). Learning nonlocal constitutive models with neural networks. *Computer Methods in Applied Mechanics and Engineering* 384, 113927 [Online]. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0045782521002644>.