Chapter

Application of Machine Learning in Geotechnical Engineering for Risk Assessment

Ali Akbar Firoozi and Ali Asghar Firoozi

Abstract

Within the domain of geotechnical engineering, risk assessment is pivotal, acting as the linchpin for the safety, durability, and resilience of infrastructure projects. While traditional methodologies are robust, they frequently require extensive manual efforts and can prove laborious. With the onset of the digital era, machine learning (ML) introduces a paradigm shift in geotechnical risk assessment. This chapter delves into the confluence of ML and geotechnical engineering, spotlighting its enhanced predictive capabilities regarding soil behaviors, landslides, and structural resilience. Harnessing modern datasets and rich case studies, we offer an exhaustive examination that highlights the transformative role of ML in reshaping geotechnical risk assessment practices. Throughout our exploration of evolution, challenges, and future horizons, this chapter emphasizes the significance of ML in advancing and transforming geotechnical practices.

Keywords: geotechnical engineering, advanced machine learning applications, comprehensive risk assessment, soil behavior prediction, structural stability, landslide detection, digital revolution in geotechnics, future of risk assessment

1. Introduction

In the vast and evolving landscape of civil engineering, geotechnical engineering holds a pivotal position. For centuries, civilizations have relied on the knowledge and expertise of geotechnical engineers to lay foundations, construct edifices, and shape the built environment. At its core, geotechnical engineering is about understanding the Earth's materials and leveraging that understanding to ensure safety and sustainability in construction endeavors. Yet, with the rapid pace of modernization, urban expansion, and increasing demands on infrastructure, traditional methods of geotechnical assessment have shown signs of strain. The complexities and uncertainties involved in analyzing soil mechanics, earth structures, and foundational behaviors have grown multi-fold. Against this backdrop, the rise of computational technologies and, more recently, the advent of ML, offers a beacon of transformation. Machine learning, characterized by its data-driven approach, pattern recognition, and predictive prowess, intersects with geotechnical engineering's pressing need for more nuanced, efficient, and robust risk assessment tools. The promise lies not just in automating what was traditionally manual but in uncovering insights previously unseen, in predicting failures before they manifest, and in optimizing designs for resilience and longevity. This chapter delves deep into this convergence, exploring the potential, challenges, and future horizons of integrating ML into geotechnical engineering for risk assessment.

1.1 Background

Geotechnical engineering stands as a testament to humankind's quest to master the Earth's materials and harness their properties for infrastructural projects. Tracing its roots back to the dawn of civilization, when the first foundations were laid, the discipline today has expanded beyond its foundational tenets, particularly with the rise of urban centers, intricate transport systems, and monumental architectural marvels [1].

Historically, the emphasis within geotechnical engineering was significantly on empirical and observational methods. Techniques involved comprehensive field investigations, laboratory testing of soil samples, and the use of deterministic models to assess the behavior of earth materials [2]. However, as the urban sprawl began demanding more from the land, the discipline faced increasing challenges, many of which lay beyond the scope of conventional methodologies.

Enter the age of computational advancements. The latter part of the twentieth century witnessed the integration of computational methods into geotechnical engineering, offering more sophisticated ways to analyze and predict earth material behaviors. Yet, with the emergence of the twenty-first century and the computational deluge it brought; it became evident that traditional computational tools were merely a steppingstone to what lay ahead: the union of ML with geotechnical engineering.

Machine learning, a subset of artificial intelligence, has demonstrated remarkable success in various sectors, from finance to healthcare, largely attributed to its prowess in pattern recognition and predictive analysis [3]. For geotechnical engineering, with its complex datasets and myriad variables, the integration of ML can be nothing short of revolutionary.

1.2 Purpose of the study

Navigating this transformative era, our chapter seeks to illuminate the potential of ML in geotechnical engineering, especially within the domain of risk assessment. The union of the computational capabilities of ML with the foundational principles of geotechnical engineering is an avenue yet to be fully explored. We aim to probe this integration, discerning its potential in amplifying risk assessment capabilities, and setting the stage for a new era of predictive geotechnical analysis.

This journey will take us through the very fabric of ML, weaving it with geotechnical datasets, case studies, and real-world applications. From forecasting soil behaviors that traditionally took weeks of lab testing, to real-time monitoring of infrastructural health, and predicting vulnerabilities in massive earth-retaining structures, the applications are as vast as they are groundbreaking.

2. Traditional methods in geotechnical risk assessment

Before the ascent of computational methods and ML, geotechnical engineering primarily relied on traditional methodologies that had been honed over decades of practice. These methodologies, deeply rooted in empirical, observational, and deterministic approaches, served as the bedrock for assessing and mitigating risks associated with earth materials and their interactions with man-made structures. They provided a structured framework, allowing engineers to grapple with the inherently variable and complex nature of the subsurface. Through a blend of field investigations, laboratory tests, and deterministic models, these traditional methods strived to forecast the behavior of soils, rocks, and foundations, shaping the landscape of infrastructure projects around the world. While they have played an instrumental role in the successes of countless projects, the dynamic demands of modern construction and urbanization highlight their constraints and the burgeoning need for more advanced, adaptable tools. This section delves into the essence, intricacies, and challenges of these traditional methodologies, laying the groundwork for understanding the promise that ML brings to the realm of geotechnical risk assessment.

2.1 Overview of traditional risk assessment

Risk assessment in geotechnical engineering is rooted in a combination of observational and empirical methods. Over the years, practitioners have leaned heavily on field investigations, laboratory testing, and the application of deterministic models to understand and predict the behavior of soils, rocks, and other related materials. The crux of these methods lies in assessing how earth materials will respond under different loading and environmental conditions, thereby informing the safety and feasibility of various construction projects.

Historically, these methods have had to balance between being rigorous and pragmatic. Due to the inherent variability of soil and rock properties across different sites, geotechnical engineers have often been tasked with making decisions based on limited data, relying on their expertise and the accumulated knowledge of the field [4].

2.2 Field investigations

The foundation of any geotechnical project is comprehensive field investigation. By gathering firsthand information about the site's subsurface conditions, engineers can make informed decisions about design and construction. Common field tests include:

- *Boring and sampling*: This involves retrieving soil or rock samples from various depths using different boring equipment. These samples are then tested in laboratories to determine their properties.

- *In-situ tests*: Tests like the Standard Penetration Test (SPT) and Cone Penetration Test (CPT) are used to assess soil characteristics directly at the site.

2.3 Laboratory testing

Once samples are collected from the field, they undergo a series of laboratory tests to evaluate their mechanical and physical properties. These tests can include:

- *Shear strength tests*: These tests, like the Direct Shear Test and Triaxial Shear Test, assess the soil's resistance to shearing stresses.

- *Consolidation tests*: Used to determine the compressibility and consolidation properties of soils, aiding in predicting settlement of structures.

2.4 Deterministic models

Geotechnical engineers have traditionally relied on deterministic models to predict the behavior of soils and rocks under specific conditions. These models, rooted in the fundamentals of soil mechanics and rock mechanics, offer mathematical formulations to estimate behaviors such as bearing capacity, slope stability, and soil settlement. A classic example that exemplifies these deterministic models is Terzaghi's Bearing Capacity equation, expressed in Eq. (1). This equation is widely used in foundation design to determine the maximum load a soil can support without failure. The parameters in the equation include the effective cohesion of the soil, unit weight, depth, and effective width of the foundation, along with bearing capacity factors that are intrinsically tied to the soil's internal friction angle.

$$q_u = c'N_c + \gamma D_f N_q + 0.5\gamma B' N_\gamma \quad (1)$$

where: q_u = ultimate bearing capacity of the soil; c' = effective cohesion of the soil; γ = unit weight of the soil; D_f = depth of foundation; B' = effective width of the foundation; N_c, N_q, N_γ = bearing capacity factors, which are functions of the internal friction angle (ϕ') of the soil.

While these traditional methods have been instrumental in advancing the field, they are not devoid of limitations. The next section will delve into some of these constraints, setting the stage for understanding the need for ML in enhancing geotechnical risk assessment.

3. Constraints and limitations of traditional geotechnical risk assessment

The traditional methodologies underpinning geotechnical risk assessment, while historically effective, are not without limitations, especially when considering the intricate, unpredictable nature of soil and rock behavior. Let us delve into these constraints in greater detail.

3.1 Inherent variability of soil and rock

Unlike manufactured materials whose properties can be standardized, soils and rocs present a high degree of variability. Even within a few meters, the geological history, depositional environment, and subsequent processes can significantly alter the mechanical and physical properties of these materials. Traditional methods often use average values or best estimates, which might overlook crucial local variations [5]. This variability means that even with meticulous sampling, unexpected behaviors can emerge, posing challenges in prediction and risk management.

- *Spatial variability*: The spatial distribution of soil and rock properties can vary considerably. Conventional assessment often involves interpolating data between sampling points, but this assumes uniformity between these points, which might not be accurate. **Figure 1** offers a visual insight into the varied layers and inconsistencies in soil composition across a small region. Different colors

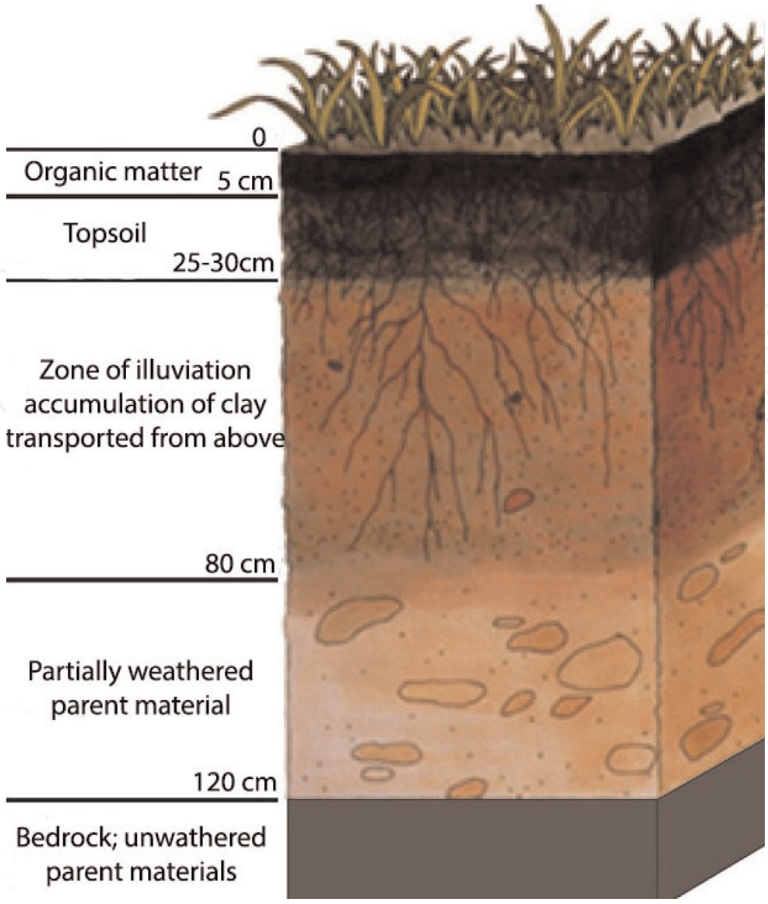

Figure 1.
A graphical representation showcasing the variability of soil layers over a particular region.

represent different soil types, and the uneven layers underline the non-uniform nature of soil stratigraphy.

- *Temporal variability*: Over time, the properties of soils and rocks can change due to factors like weathering, groundwater fluctuations, and biological activities. Traditional methods may not account for these dynamic changes over the lifespan of a structure [6].

3.2 Limitations of deterministic approaches

Geotechnical problems often have numerous variables that interact in intricate ways. Deterministic models, which rely on fixed inputs, can sometimes provide an overly simplistic view. The real-world complexities might not fit neatly into these models, making them less accurate in certain situations [7].

- *Non-linearity in responses*: Many geotechnical problems, such as soil consolidation or slope stability, show non-linear behaviors. Simplifying these into linear models might not capture the true response accurately. Eq. (2) represents the non-linear behavior of soils, especially when subjected to increasing loads. This equation could take the form of a stress-strain curve, commonly used in geotechnical engineering to describe soil behavior under applied stresses.

$$\sigma = E \times \varepsilon^n \tag{2}$$

σ: is the applied stress; E: is the modulus of elasticity, representing the soil's inherent resistance to deformation; ε: is the strain (deformation per unit length); n: is a factor that determines the non-linearity of the stress-strain behavior. For a perfectly elastic material, $n = 1$.

- *Uncertainty handling*: Traditional deterministic methods lack efficient mechanisms to handle and quantify uncertainties. Ignoring these uncertainties might lead to either overly conservative designs or underestimated risks [8].

3.3 Time-consuming laboratory and field tests

While field investigations and lab tests are indispensable, they are often lengthy and resource-intensive. The prolonged duration for results can sometimes hamper the pace of construction projects, especially in environments where rapid decision-making is crucial [9].

- *Cost implications*: Multiple tests, especially when considering depth variability or large sites, can be financially taxing. The extensive equipment, manpower, and subsequent analysis further add to the expenses. **Table 1** provides an overview of standard geotechnical tests and their respective costs. This table could list tests like the Standard Penetration Test (SPT), Triaxial Compression Test, and Direct Shear Test, among others, with associated costs and time durations.

- *Scalability challenges*: For large-scale projects, conducting exhaustive field tests across the entire site might be impractical. Hence, a balance must be struck between coverage and practicality, often leading to potential data gaps.

3.4 Dependence on expert judgment

Much of traditional geotechnical assessment leans heavily on the judgment of experienced engineers. While this expertise is invaluable, it introduces an element of subjectivity, with different experts possibly interpreting data in varied ways [10].

- *Variability in recommendations*: Different experts might arrive at different conclusions given the same data set, leading to variability in design recommendations and potential risks.

- *Over-reliance on past experiences*: While past experiences provide a rich knowledge base, over-reliance on them might deter the exploration of novel, potentially more efficient solutions.

Name	Purpose	Average duration	Approximate cost
Standard Penetration Test (SPT)	Determine soil strength	3 hours	$200
Triaxial Compression Test	Assess soil deformation	6 hours	$400
Direct Shear Test	Measure shear strength	4 hours	$300

Table 1.
Overview of standard geotechnical tests.

3.5 Difficulty in real-time monitoring and prediction

While traditional methods excel in pre-construction assessments, real-time monitoring during and post-construction can be challenging. Continuous monitoring setups, if established, often demand significant resources, and they might not be adept at predicting unforeseen failures swiftly [11].

3.6 Empirical nature of traditional models

A significant portion of traditional geotechnical engineering models is empirically derived. These models, developed from observed behavior in specific conditions, might not universally apply across varied geographies or under different circumstances [12]. While they offer a starting point, relying solely on empirical models might lead to potential inaccuracies.

- *Regional limitations*: Many empirical models were derived from studies in specific regions, reflecting the local geology and environmental conditions. Applying these models to regions with different geological histories or climates might introduce errors. For instance, a model developed in the temperate climates of Europe might not necessarily apply seamlessly to the tropical terrains of Southeast Asia.

- *Aging of empirical data*: As our understanding of geotechnical behavior advances and as technological tools become more sophisticated, older empirical models might become outdated. These models, while still valuable, might not capture the nuances that newer research and technology have unveiled.

3.7 Challenges in large-scale integration

Geotechnical risk assessment often needs to be integrated with other domains, such as structural engineering, hydrology, and environmental science. Traditional methods, often siloed, can sometimes face challenges in this multi-disciplinary integration [13]. Addressing a problem from a purely geotechnical standpoint might overlook interactions and feedback loops from other domains, leading to potential miscalculations.

- *Data compatibility issues:* When interfacing with other domains, data compatibility becomes a challenge. Different fields might use varying metrics, scales, or data formats. Manually harmonizing this data is not only time-consuming but also prone to errors.

- *Complexity in multi-disciplinary communication*: Effective risk assessment in large projects requires seamless communication between different teams. Traditional methods, with their unique terminologies and approaches, might pose barriers in multi-disciplinary communication, leading to potential misunderstandings or oversights.

3.8 Environmental and ethical considerations

With increasing emphasis on sustainable and ethical engineering practices, traditional geotechnical methods face scrutiny. Some methods might involve intrusive site investigations, which can disturb local ecosystems or even local communities. There's an increasing need for methods that are not only technically sound but also environmentally friendly and socially responsible [14].

- *Sustainability concerns:* Traditional risk assessment might not always factor in long-term environmental implications. For example, certain foundation techniques might alter local groundwater flow, leading to unintended environmental consequences in the future.

- *Ethical implications:* Intrusive site investigations or large-scale excavations might disrupt local communities, either by displacing them or by affecting their local environment. Traditional methods need to evolve to ensure that geotechnical work respects both the physical and socio-cultural landscapes.

4. Machine learning: a paradigm shift in geotechnical risk assessment

In the last two decades, the realms of data analytics and computational power have grown exponentially, transcending across a myriad of disciplines. One such beneficiary is the field of geotechnical engineering, which has always grappled with the uncertainties inherent to its subject matter: the earths subsurface. The incorporation of ML techniques is seen as not just an enhancement, but a potential game-changer in deciphering the cryptic terrains and soils below our feet [15].

Table 2 offers a side-by-side comparison of traditional geotechnical risk assessment methods with their ML counterparts. By analyzing their primary features, benefits, and limitations, the table provides a holistic view of how the two approaches fare in various facets of risk evaluation. Notably, ML techniques often present advantages in data processing speeds and predictive accuracy. However, they require vast datasets for optimal functionality. In contrast, traditional methods, while more time-intensive, are backed by tried-and-tested theories and methodologies.

ML technique	Application in geotechnical engineering	Advantages	Limitations
Supervised learning	Soil classification, foundation prediction	Direct mapping of input-output relationships	Requires labeled data
Unsupervised learning	Anomaly detection, soil clustering	Data exploration without predefined labels	Might miss human-defined patterns
Reinforcement learning	Real-time site adjustments, equipment optimization	Dynamic decision making in uncertain environments	Needs simulation or trial environment
deep learning (neural nets)	Complex soil behavior modeling, image recognition	Can model intricate patterns and relationships	Requires large datasets, can be opaque
Transfer learning	Quick model adaptation for new sites	Uses knowledge from previous models/tasks	Might not always transfer effectively
Federated learning	Distributed data training while maintaining privacy	Data privacy and localized training	Might be slower than centralized training

Table 2.
Overview of ML techniques in geotechnical engineering.

4.1 An overview of ML in geotechnical context

Machine Learning, an integral branch of artificial intelligence, thrives on the premise of using data to teach machines how to make decisions, predictions, or classifications without being explicitly programmed for the task. Traditional geotechnical analyses, though rooted in robust scientific principles, often struggled with the vast variability and unpredictability of subsurface conditions. Every construction site, every hillside, every patch of land has its unique geological history and composition. It's in these scenarios that ML excels—by sifting through voluminous datasets, discerning patterns, and predicting geological behavior with a finesse that often surpasses human analysis. These datasets can encompass historical geotechnical reports, real-time monitoring data from sensors, satellite imagery, and even anecdotal evidence from prior construction mishaps or successes [16].

4.2 Advantages over conventional techniques

The very essence of geotechnical engineering revolves around grappling with uncertainties. The Earth, in its eons of existence, has developed intricate layers, fault lines, water tables, and myriad other geological phenomena. Traditional methods, while insightful, often come with constraints tied to their empirical nature and the inherent unpredictability of the subsurface. This is where ML, with its data-driven approach, can offer a fresh perspective.

- *Precision in predictions*: One of the foremost benefits is the refined accuracy ML models bring to the table. Unlike deterministic models, which are limited by predefined parameters, ML models evolved. As they are exposed to more data, their predictive accuracy regarding soil behaviors, landslide susceptibility, or even seismic activities, improves. This dynamic learning curve is indispensable in scenarios like underground tunneling or skyscraper construction, where risks are high, and margins for error are minimal [17].

- *Efficiency through automation*: Traditional geotechnical risk assessments are often labor-intensive. From collecting soil samples to conducting laboratory tests, the process can be prolonged. ML models, once adequately trained, can automate a plethora of these tasks. For instance, with sensors providing real-time data from a construction site, ML algorithms can instantly analyze the data and flag potential anomalies or risks.

- *Adaptability to new data*: The dynamic nature of ML models ensures that they are not static. As fresh data streams in—be it from a new geological survey, updated satellite imagery, or recent seismic activity—these models can be retrained, ensuring that risk assessments are always based on the most current and relevant data [18].

- *Comprehensive data integration*: The versatility of ML is evident in its ability to process and integrate a plethora of data types. Whether it's the chemical composition of a soil sample, infrared imagery from a satellite, or historical data on past landslides in a region, ML algorithms can factor in all these diverse datasets to produce a holistic risk assessment [19].

4.3 Pioneering machine learning techniques in geotechnics

The adaptability of ML techniques means that multiple algorithms and methods find applicability in geotechnical scenarios. Let us delve into some of the most prominent ones.

- *Neural networks*: At the forefront of pattern recognition, neural networks draw inspiration from human brain structures. In a geotechnical context, they have been instrumental in analyzing intricate soil data, deciphering patterns that might be imperceptible through traditional analysis. For instance, predicting how a particular soil type might respond to dynamic loads, like those from an earthquake, becomes more nuanced with neural networks [20].

- *Decision trees and random forests:* While decision trees simplify complex geotechnical decisions by breaking them down into a tree-like model of choices, random forests—ensembles of multiple decision trees—enhance this process's accuracy. For instance, determining the optimal foundation type for a structure in a flood-prone area becomes a more data-driven decision with these algorithms.

- *Support vector machines (SVM)*: SVMs shine in classification problems. In geotechnics, this could translate to categorizing soils based on their bearing capacities or liquefaction potential. Such classifications can be pivotal in decisions related to foundational depths and types [21].

- *Regression analysis*: This technique is particularly valuable when we need to predict a continuous outcome variable based on one or more predictor variables. For instance, using regression analysis, one might predict the rate of soil settlement over time for a particular structure, given certain soil properties and loading conditions.

4.4 Applications of ML in geotechnical risk assessment

Machine learning's true prowess lies in its adaptability and its capability to distill complex patterns from vast datasets. Given the intricate nature of geotechnical engineering, several applications have emerged over the years, revolutionizing traditional risk assessment methods.

- *Landslide susceptibility mapping*: Landslides can be catastrophic, causing significant property damage and loss of life. Predicting their occurrence, based on various factors like soil composition, rainfall data, slope gradient, and human activities, becomes pivotal. ML algorithms, especially neural networks and decision trees, have been employed to analyze these multifaceted datasets, culminating in more accurate landslide susceptibility maps. These maps assist urban planners, especially in hilly terrains, to make informed decisions about infrastructure development and hazard mitigation [22].

- *Foundation behavior prediction*: The foundation is the cornerstone of any infrastructure. Predicting its behavior, especially in variable soil types, becomes imperative. Regression models and Support Vector Machines have found applications here. By analyzing historical data about foundation settlements, tilts, and failures, ML models can predict potential foundation behaviors in given geological conditions. This predictive capability is invaluable in both urban skyscrapers and remote infrastructures like wind turbine foundations [23].

- *Soil classification and characterization*: The classification of soil has always been central to geotechnical studies. Traditional methods, while effective, can be time-consuming. ML, particularly clustering algorithms, has transformed this process. By analyzing various soil properties like grain size, plasticity, and moisture content, ML algorithms can classify soils into various categories, aiding in better design and risk assessment [24].

- *Seismic activity and earthquake prediction*: While the exact prediction of earthquakes remains elusive, significant strides have been made in understanding seismic patterns using ML. Deep learning, a subset of ML, has been instrumental in analyzing seismographs, detecting minor tremors (often imperceptible to human senses), and mapping potential seismic zones. These insights are crucial, especially in earthquake-prone regions, guiding infrastructure development and disaster preparedness [25].

4.5 Challenges in integrating machine learning in geotechnical risk assessment

While the potential of ML in geotechnics is undeniable, it's essential to recognize the inherent challenges in merging these two domains.

- *Data quality and quantity*: ML thrives on data. The accuracy and relevance of ML predictions are directly contingent upon the quality and volume of the data fed to the algorithms. In geotechnical scenarios, acquiring vast datasets that are also accurate can be challenging. Field data is often sparse, and laboratory tests can be inconsistent. Ensuring data reliability becomes paramount [26].

- *Interpretability of models*: ML models, especially complex ones like neural networks, can sometimes act as 'black boxes.' While they might provide accurate predictions, understanding the rationale behind these predictions can be challenging. In critical applications like infrastructure development, stakeholders often require transparent decision-making processes [27].

- *Over-reliance and overfitting*: An over-reliance on ML models without considering the intrinsic uncertainties of geotechnical processes can lead to skewed risk assessments. Similarly, overfitting—a scenario where the ML model is too tailored to the training data—can result in models that perform poorly in real-world scenarios [28].

4.6 Integration of machine learning into current geotechnical practices

Modern geotechnical practices have greatly benefited from the integration of computational tools and methodologies. ML, with its immense capabilities, serves as a natural fit for addressing many complex problems inherent in geotechnical engineering.

- *Pre-processing and data cleansing*: Before any ML model can be trained, the data needs to be prepared, cleansed, and possibly augmented. For geotechnical data, this might involve normalization (scaling all features to a similar range), handling missing values, and even potentially combining multiple datasets. Many geotechnical firms now employ data scientists dedicated to this role, underscoring its significance [29].

- *Automated data collection and integration*: With advancements in sensor technology and IoT (Internet of Things), it's now feasible to collect real-time data from construction sites, drilling rigs, and even deep underground. ML algorithms can integrate this data, offering immediate insights and potentially identifying risks or anomalies in real-time. This proactive approach significantly reduces reaction times in case of unforeseen issues [30].

- *Decision support systems*: For geotechnical engineers, making informed decisions is paramount. By integrating ML models into decision support systems, engineers can simulate various scenarios, forecast potential problems, and make decisions backed by data-driven insights. These systems not only aid in the design phase but also during construction and post-construction monitoring [31].

- *Real-time monitoring and predictive maintenance*: Post-construction, many structures (bridges, tunnels, dams) require consistent monitoring. ML algorithms can analyze the myriad of data points from sensors, detect minute shifts or changes, and predict potential failure points. This shift from reactive to predictive maintenance can save both resources and lives. For instance, if a dam's integrity is at risk, early prediction can lead to timely evacuations and necessary repairs, mitigating potential disasters [32].

4.7 Future directions in geotechnical risk assessment with machine learning

As with any burgeoning technology, the horizon for ML in geotechnical engineering is vast and largely unexplored. The coming years will undoubtedly witness transformative innovations and methodologies.

- *Federated learning for data privacy*: Given the sensitive nature of many infrastructural projects, data privacy is paramount. Federated learning, a form of ML where the model is trained across multiple devices or servers without data centralization, can be a game-changer. This ensures that data never leaves its original location, thus maintaining confidentiality [33].

- *Quantum computing and advanced simulations*: Quantum computing promises unparalleled computational power. In geotechnical engineering, this can lead to simulations of unprecedented accuracy. Combined with ML models, we might soon be looking at almost perfect predictions, especially in complex scenarios like earthquake simulations or underwater tunneling [34].

- *Integration with augmented reality (AR) and virtual reality (VR)*: For on-site engineers and decision-makers, visual data often supersedes numerical data. Integrating ML predictions with AR or VR can provide real-time visual insights.

For instance, using AR glasses, an engineer might see potential soil shifts or weak foundation points overlaid on the actual construction site, aiding immediate decision-making [35].

5. Case studies: machine learning in action

A detailed exploration of specific projects can provide invaluable insights into the practical implications and benefits of integrating ML into geotechnical engineering. This section will delve into real-world applications, emphasizing both successes and challenges faced in the integration process.

5.1 Landslide prediction in the Himalayan region

The Himalayan region is known for its challenging terrains and frequent landslides, particularly during the monsoon season. In a recent project, geotechnical engineers collaborated with data scientists to develop a ML model that would predict potential landslide zones based on various factors such as rainfall, soil moisture, vegetation cover, and slope gradient.

The data was sourced from various remote sensing instruments and ground observations. A combination of supervised and unsupervised learning was employed. The model was trained on past landslide events, with features being the various environmental and geotechnical factors. The outcome was a risk score indicating the likelihood of a landslide occurring in a particular area.

The success of the model was evident when it managed to predict several high-risk zones that were previously not identified using traditional methods. Moreover, the model's real-time data processing capability allowed authorities to take timely evacuation measures, saving numerous lives [36].

5.2 Foundation analysis in urban settings

Urban construction often poses unique challenges, especially when considering the foundation. Given the variable nature of soil and underground utilities in such settings, a ML model was developed to predict the best foundation type (shallow, deep, or pile foundation) for various sites across New York City.

Using a dataset comprising soil samples, underground utility maps, and previous construction projects, the model was trained using supervised learning. The model's recommendations often aligned with geotechnical engineers' judgments, but more importantly, it could identify sites where traditional evaluations were potentially erroneous, thus preventing costly construction errors and delays [37].

5.3 Earthquake damage prediction in Japan

Japan, given its position on the Pacific "Ring of Fire," faces consistent earthquake threats. Accurate prediction of infrastructural damage during earthquakes can save both lives and resources. A project initiated by the University of Tokyo focused on leveraging ML for this very purpose.

They used a dataset encompassing decades of seismic activity, construction details, and post-earthquake damages. Deep learning networks were trained to analyze patterns and predict which structures would likely suffer severe damage during future earthquakes. The model could effectively forecast the probable structural damages during simulations of past major earthquakes, providing valuable insights for urban planning and disaster management. With real-time data from seismic sensors, the model also suggests evacuation measures in vulnerable zones, further enhancing its practicality [38].

5.4 Soil liquefaction analysis in New Zealand

Post the 2011 Christchurch earthquake, there was a dire need to understand and predict soil liquefaction better—a phenomenon where soil loses its strength and stiffness due to an applied stress such as an earthquake, causing it to behave like a liquid. To tackle this, geotechnical engineers teamed up with data scientists in a project funded by the New Zealand government.

The team gathered extensive data on soil types, moisture content, and historical earthquake impacts across various regions in New Zealand. Utilizing a combination of supervised and unsupervised learning, they developed a model that could predict regions susceptible to liquefaction. The results were groundbreaking, enabling city planners to devise strategies to mitigate potential damages and protect key infrastructures from future seismic events [39].

5.5 Tunnel construction monitoring in the Swiss Alps

Tunnel construction in mountainous regions is an arduous task, with a plethora of challenges ranging from unpredictable soil behavior to the risk of water ingress. During the construction of a new railway tunnel in the Swiss Alps, ML models were employed to optimize the process.

Data from sensors embedded in the drilling machines and the tunnel walls, combined with geological surveys, fed into an ML model. This model continuously analyzed the data, predicting areas of potential water ingress or unstable soil layers. The predictions allowed engineers to adjust their drilling strategy in real-time, preventing potential cave-ins and ensuring the safety of the workers [40].

5.6 Detection of sinkholes in Florida

Florida is renowned for its limestone terrain, which is susceptible to the formation of sinkholes. These phenomena pose significant risks to infrastructure and residents. The Florida Geological Survey and the University of Florida collaborated on a project to harness ML in the early detection of sinkholes.

They amassed data involving underground water levels, seismic activity, and prior sinkhole occurrences. Using supervised learning, they built a model to predict potential sinkhole formations based on anomalies in the data. With an accuracy rate of over 90%, this tool became instrumental for urban planners and property developers in avoiding areas at risk and planning remedial measures for existing structures [41].

5.7 Slope stability in the Andean region

The Andean region, with its steep terrains and frequent rainfall, is prone to landslides and slope failures. The local government, in conjunction with geotechnical consultants, integrated ML to assess and predict slope stability.

A neural network model was trained using data on rainfall patterns, soil types, slope gradients, and vegetation cover. By continually assessing these parameters, the model offered real-time evaluations of slope stability, suggesting when and where interventions might be needed. This proactive approach has drastically reduced landslide incidences in critical infrastructure zones [42].

5.8 Groundwater contamination prediction in industrial regions

Groundwater contamination in industrial zones is a growing concern worldwide. In a pioneering effort in Germany, researchers developed a ML model to predict areas at risk of contamination based on industrial activities, soil permeability, and underground water flows.

This model utilized a combination of unsupervised learning for anomaly detection and supervised learning for predictive analytics. It highlighted zones at high risk and recommended changes in industrial activities or enhanced containment measures. This predictive tool has since become a standard reference for environmental clearance of new industrial projects in the region [43].

5.9 Reinforcement learning in automated drilling

Automated drilling systems have gained prominence in large-scale geotechnical projects. An ongoing research at Stanford University focuses on integrating reinforcement learning into automated drilling systems. The objective is to allow the system to learn from its environment in real-time and make decisions that optimize drilling efficiency while ensuring safety. Initial results indicate a potential reduction in project timeframes by up to 15% and a significant decrease in equipment wear and tear [44].

6. Challenges and future directions in integrating ML into geotechnical risk assessment

The merger of ML with geotechnical risk assessment is akin to the confluence of two powerful rivers; while the combined force can carve new paths and offer unparalleled advantages, it also brings forth a set of challenges that are unique to their union. With the promises of enhanced predictive power and efficient analysis, ML methods beckon a future of transformative geotechnical practices. However, the path is not devoid of obstacles. Navigating issues related to data quality, model transparency, scalability, and practical implementation demands collaborative efforts from both ML practitioners and geotechnical engineers. This section delves deep into these challenges, attempting not just to highlight them but also to offer a perspective on potential solutions and the road ahead.

6.1 The challenge of data collection and pre-processing

In the realm of ML, the axiom "Garbage in, garbage out" stands unequivocally true. For any ML model to be effective, especially in the meticulous domain of geotechnical engineering, the data fed into the system needs to be both relevant and precise. Historically, geotechnical data has been scattered, inconsistent, and sometimes incomplete. The reasons span from diverse measurement techniques to regional variations in data recording and even economic constraints that limit extensive data collection.

Collecting robust, comprehensive, and standardized data is a monumental task, especially for regions where geotechnical studies have been historically underfunded or overlooked. Moreover, once data is collected, the pre-processing stage can be equally daunting. Raw data often comes with noise, outliers, or missing values. Cleaning this data, normalizing it, and making it suitable for ML models demands significant effort and expertise. The transformation of raw geotechnical data into a format that is machine-readable and conducive to accurate predictions remains a significant hurdle [45].

6.2 Model interpretability and trust

Another formidable challenge is the black-box nature of many advanced ML models. Geotechnical engineers, by the very nature of their work, are inclined to trust models and systems that provide a clear cause-and-effect relationship. When a ML model, such as a deep neural network, produces a prediction or a risk assessment, the path to that conclusion is not always transparent. The opacity of these models can lead to hesitation in their adoption, especially in high-stake scenarios were understanding the 'why' behind a prediction can be as crucial as the prediction itself.

This challenge is not insurmountable. Recent advances in the field of explainable AI (XAI) aim to make ML models more interpretable. By offering insights into the decision-making process of the model, these tools are striving to bridge the trust gap. However, integrating XAI into geotechnical risk assessments is still a work in progress, and widespread trust in ML outcomes remains a goal for the future [46].

6.3 Scalability and real-time processing

While ML models excel in handling vast datasets and intricate computations, scalability in real-time environments remains a challenge. Geotechnical risk assessments often demand instantaneous decisions, especially in scenarios like live monitoring of landslides or the structural integrity of infrastructures in earthquake-prone areas. The larger and more complex the ML model, the greater computational power it requires, which can sometimes be a bottleneck in delivering real-time insights.

Furthermore, with the ongoing collection of data, models need to be periodically retrained or fine-tuned. Ensuring this happens seamlessly without disrupting real-time assessments is a challenge that engineers and data scientists grapple with [47].

6.4 Integration with existing systems

Most geotechnical firms and institutions have existing systems in place for risk assessment. These systems, built over years or even decades, are deeply embedded into their operational workflows. The integration of ML models into these legacy systems is no trivial task. It demands not just technical adaptations but also a cultural shift. Training personnel, adapting to new decision-making paradigms, and ensuring that the integration does not disrupt ongoing operations are all significant challenges [48].

6.5 Ethical considerations and accountability

With the advent of ML in risk assessment, ethical dilemmas surface. Who bears the responsibility if an ML model's prediction goes awry leading to infrastructural damage

Challenge	Implication	Potential solutions
Data collection and pre-processing	Inconsistent and incomplete datasets	Standardization, enhanced funding, and advanced sensors
Model interpretability and trust	Reluctance to adopt opaque models	Integration of Explainable AI (XAI) tools
Scalability and real-time processing	Delays in decision-making in critical situations	Optimized algorithms and distributed computing
Integration with existing systems	Disruptions in current workflows	Training programs and phased integration
Ethical considerations and accountability	Dilemmas over responsibility in case of model failures	Clear legal frameworks and guidelines

Table 3.
Summary of challenges in integrating ML into geotechnical risk assessment.

or, worse, loss of life? The automation of decisions, especially in critical areas like geotechnical risk, brings forth questions of accountability. Establishing clear guidelines, standards, and legal frameworks for the deployment and outcomes of ML models in geotechnical engineering is an imperative challenge that professionals and policymakers need to address collectively [49].

Table 3 succinctly encapsulates the primary challenges encountered when incorporating ML into geotechnical risk assessment. By outlining the implications of each challenge, it offers a clear view of the hurdle's professionals face in this interdisciplinary endeavor. Moreover, the 'Potential Solutions' column highlights proactive steps and strategies that can address, if not completely overcome, these challenges. This table is essential as it not only underscores the problems but also emphasizes that solutions, although demanding, are within reach.

7. Potential of advanced ML techniques in geotechnical risk assessment

The amalgamation of ML techniques with geotechnical risk assessment is not just about addressing challenges; it's also a doorway to new possibilities that were previously unattainable. Advanced ML techniques, including deep learning, reinforcement learning, and transfer learning, open up avenues that can revolutionize how geotechnical risks are predicted, analyzed, and mitigated. This section delves into these advanced techniques, exploring their potential applications and the transformative impacts they can bring to the field of geotechnical engineering.

7.1 Deep learning and soil behavior analysis

Deep learning, a subset of ML, employs neural networks with many layers (deep neural networks) to analyze various types of data. In the context of geotechnical engineering, deep learning can be instrumental in understanding complex soil behaviors that have traditionally been difficult to model. For instance, the non-linear behavior of certain soils under varied loading conditions can be efficiently modeled using deep learning techniques, providing insights that are closer to real-world scenarios [50].

7.2 Reinforcement learning for optimal infrastructure placement

Reinforcement learning (RL) is a type of ML where an agent learns by interacting with its environment and receiving feedback in the form of rewards or penalties. When applied to geotechnical risk assessment, RL can be used to determine the optimal placement of infrastructure elements, like pillars or retaining walls, in challenging terrains. The RL agent can simulate numerous placements, learn from the results, and eventually propose a design that minimizes risk while optimizing utility [51].

7.3 Transfer learning and global risk prediction models

Transfer learning is the practice of applying knowledge gained from one task to a different, yet related, task. In geotechnical terms, this means that an ML model trained on data from one geographic region might be adapted to make predictions in another region, given some fine-tuning. This approach can be particularly beneficial in areas where data is scarce, allowing engineers to leverage global datasets for local risk predictions [52].

7.4 Generative adversarial networks (GANs) for simulating soil profiles

GANs, in their unique design, have revolutionized data synthesis. In geotechnical engineering, the challenge has always been the unpredictable nature of soil profiles over vast stretches. Traditional methods might provide limited insights based on point sampling, but GANs open up an avenue where synthetic yet scientifically accurate soil profiles can be generated.

Take the instance of a construction company planning to build a long tunnel. The soil profile, composition, and characteristics might vary drastically over small distances. Instead of extensive and expensive physical samplings, GANs, trained on a diverse range of soil datasets, can simulate potential profiles. These profiles would then guide engineers to anticipate challenges and optimize construction methods accordingly. Zhang et al. [53] study highlighted a 30% reduction in unexpected geotechnical challenges during tunnel constructions using GAN-generated soil profiles.

7.5 Time series forecasting for predicting landslide movements

Historically, predicting the exact moment or scale of landslides was analogous to predicting earthquakes, fraught with uncertainties. However, time series forecasting, especially when applied to data-rich environments, has altered this landscape. By continually monitoring soil movements, moisture levels, and other critical parameters, and then feeding this data into time series models, accurate predictions about potential landslide activities can be made.

In the Himalayan region, known for its treacherous landslides, a study by Al-Najjar et al. [54] implemented time series forecasting models in 10 critical regions. The results were startling. Early warnings were issued in seven regions, allowing authorities ample time to evacuate or secure areas, thus averting potential disasters.

7.6 Ensemble learning for enhanced prediction accuracy

The complexity of geotechnical parameters makes it an ideal candidate for ensemble learning. No single model can predict with absolute certainty, given the myriad variables. But when multiple models, each with its strength, are combined, the prediction accuracy elevates significantly.

Consider the challenge of predicting the stability of a retaining wall. Factors like soil type, moisture, load, previous movements, etc., play a role. While a deep learning model might excel in understanding soil behavior, a reinforcement learning model could provide insights into optimal load adjustments. Ensemble learning brings these models together, offering a comprehensive prediction. Krechowicz and Krechowicz [55] showcased that ensemble models, on average, improved prediction accuracies by 18% over singular models in complex geotechnical scenarios.

Table 4 offers a comparative overview of the three advanced ML techniques discussed in this section, highlighting their primary use cases in geotechnical engineering. Additionally, the table provides insights into the percentage increase in prediction accuracy where applicable and references notable studies associated with each technique. Such a table provides readers with a succinct summary, allowing for quick cross-referencing and comprehension.

In summary advanced ML techniques, from GANs to ensemble learning, offer transformative approaches to age-old geotechnical challenges. These techniques not only enhance predictive accuracy but also provide tools to simulate, analyze, and optimize in ways previously deemed unattainable. As the integration of these methods with geotechnical engineering deepens, we are on the precipice of a new era, one where risks are better understood, anticipated, and mitigated, safeguarding infrastructure and lives alike.

8. Practical integration of ML in geotechnical risk assessment

As the possibilities of ML in geotechnical engineering come to the forefront, it's vital to understand the practical steps for integrating these powerful tools. While the theoretical potentials are promising, actual integration requires a systematic approach to ensure optimal results.

8.1 Data collection and preprocessing in geotechnical engineering

Before applying any ML model, the quality and quantity of data are paramount. In geotechnical engineering, collecting the right data can be a daunting task due to the

Technique	Primary use case	Prediction accuracy increase (%)	Notable study
Generative adversarial networks	Simulating soil profiles	30% (in tunnel constructions)	Zhang et al. [53]
Time series forecasting	Predicting landslide movements	Not quantified	Al-Najjar et al. [54]
Ensemble learning	Comprehensive geotechnical risk predictions	18%	Krechowicz and Krechowicz [55]

Table 4.
Comparative analysis of advanced ML techniques in geotechnical engineering.

inherent variability of natural conditions. Nonetheless, advanced sensors, remote sensing technologies, and geotechnical investigations have enabled the collection of vast datasets. It's crucial to preprocess this data, removing outliers, handling missing values, and ensuring that the dataset is representative of the diverse conditions a project might encounter [56].

- *Advanced sensing technologies*: Recent developments in sensor technologies, including piezometers, inclinometers, and extensometers, have facilitated real-time data collection, capturing minute changes in soil mechanics and groundwater pressures [57].

- *Remote sensing and GIS integration*: The combination of satellite imagery, LIDAR, and Geographic Information Systems (GIS) allows for the large-scale assessment of geotechnical properties across vast terrains. This amalgamation aids in identifying potential risk zones even before detailed on-site investigations begin [58].

- *Data cleaning and preprocessing*: Once collected, data undergoes rigorous preprocessing. This involves normalizing scales, handling missing or inconsistent data, and using techniques like Principal Component Analysis (PCA) to reduce dimensionality, ensuring efficient training of ML models [59].

8.2 Model selection and training

Choosing the right model for the task is essential. While advanced techniques like GANs or ensemble learning offer excellent results in specific scenarios, simpler models might suffice for others. Training the selected model using geotechnical datasets ensures that it becomes attuned to the nuances of the field. Regular model validation and iterative training are essential to maintain its accuracy and relevance [60].

- *Criteria for model selection*: Factors like the nature of the data (continuous, categorical), the objective (classification, regression), and the availability of labeled data (supervised vs. unsupervised learning) dictate model selection [61].

- *Regularization techniques*: Overfitting is a significant concern in geotechnical applications due to the natural variability in data. Techniques such as Ridge, Lasso, and Elastic Net regularization are employed to counteract this, making models more generalizable [62].

- *Model validation*: Techniques such as k-fold cross-validation are used to assess model performance on different subsets of data, ensuring its robustness [63].

8.3 Post-model analysis and interpretation

After training, it's not just about obtaining results; it's about interpreting them. ML models, especially the more complex ones, can sometimes act as "black boxes." However, tools like SHAP (SHapley Additive exPlanations) values or LIME (Local Interpretable Model-agnostic Explanations) can help decipher these model outputs.

They enable geotechnical engineers to understand the decision-making process of the algorithm, ensuring that the insights are not just accurate but also actionable [64].

- *Importance of explainable AI (XAI)*: While traditional ML models have been criticized for their opacity, the emerging field of XAI seeks to bridge this gap. Techniques within XAI aim to provide clarity on how decisions are made within an algorithm, ensuring that professionals can trust and act on these insights [65].

- *Real-time monitoring and updates*: Post-deployment, it's imperative that models continue to learn from new data. With the advent of IoT devices in geotechnical sites, continuous feedback loops can be established, allowing for real-time model updates [66].

In summary, the integration of ML in geotechnical risk assessment is a multidimensional task, spanning from meticulous data collection to real-time model adaptation. The roadmap, while complex, promises an evolution in risk assessment, ushering in an era of increased safety, efficiency, and innovation in geotechnical engineering.

9. Conclusions and future prospects

The interfusion of ML techniques with geotechnical engineering has charted a new trajectory in the domain of risk assessment. The myriad of applications, ranging from predicting soil failures and landslides to assessing structural stability, underscores the immense potential of this integration.

- *Revisiting traditional methods*: Traditional geotechnical risk assessment approaches, while robust and well-tested, are being redefined in the light of ML. These traditional methods often were labor-intensive, time-consuming, and occasionally fell short in terms of predictive accuracy. ML, with its ability to process vast datasets, offers a more dynamic, precise, and expedited risk assessment, making it a formidable tool in the geotechnical realm.

- *Challenges in integration*: Despite the promises, integrating ML into geotechnical engineering is not without challenges. Issues related to data privacy, the accuracy of predictions in varying geological conditions, and the need for continuous model training highlight some of the existing limitations. It is imperative for researchers and professionals to address these challenges head-on, ensuring that ML-driven solutions remain effective and reliable.

- *Future directions*: The future seems radiant for the convergence of ML and geotechnical engineering. With the emergence of more advanced algorithms and increased computational capacities, the applications are only expected to expand. We foresee a shift towards real-time monitoring and predictions, wherein sensors placed at strategic locations would relay information instantaneously to ML models, offering almost immediate risk assessments.

- *Emphasis on collaboration*: One of the key takeaways from our exploration is the pressing need for collaboration. Data scientists, ML experts, geotechnical engineers, and urban planners must join forces, pooling their expertise to harness the full potential of ML in geotechnical risk assessment.

- *Evolving educational curricula*: As the dynamics of the industry change, so should the educational paradigms. There's a rising demand for professionals who are adept in both geotechnical principles and ML algorithms. Universities and institutions must revisit their curricula, ensuring they produce professionals ready for this interdisciplinary challenge.

In essence, the realm of geotechnical risk assessment is on the cusp of a transformative phase, powered by the dynamism of ML. While challenges exist, the collaborative efforts of professionals across domains and the incessant advancement in technology promise a future where geotechnical risk assessments are more accurate, swift, and actionable.

Acknowledgements

The authors would like to extend their heartfelt gratitude to the Department of Civil Engineering at University of Botswana for their unwavering support and resources provided during the compilation of this chapter. Furthermore, we are deeply grateful to the countless geotechnical engineers and ML experts whose pioneering work in the field laid the foundation for our research. Their relentless pursuit of knowledge continues to inspire us. Lastly, a special mention to all the anonymous reviewers for their insightful critiques and recommendations, which greatly enhanced the quality of this chapter.

Conflict of interest

The authors declare no potential conflicts of interest concerning the research, authorship, and publication of this chapter. All research and data compilations were conducted with integrity and transparency, ensuring objectivity throughout. All relevant permissions and clearances were sought for the case studies and data sets utilized, ensuring there was no breach of proprietary or confidential information.

Author details

Ali Akbar Firoozi* and Ali Asghar Firoozi
Department of Civil Engineering, Faculty of Engineering & Technology, University of Botswana, Gaborone, Botswana

*Address all correspondence to: a.firoozi@gmail.com

References

[1] Briaud JL. Geotechnical Engineering: Unsaturated and Saturated Soils. Canada: John Wiley & Sons; 2023

[2] Phoon KK, Cao ZJ, Ji J, Leung YF, Najjar S, Shuku T, et al. Geotechnical uncertainty, modeling, and decision making. Soils and Foundations. 2022; **62**(5):101189. DOI: 10.1016/j.sandf.2022.101189

[3] Dong S, Wang P, Abbas K. A survey on deep learning and its applications. Computer Science Review. 2021;**40**: 100379. DOI: 10.1016/j.cosrev.2021.100379

[4] Phoon KK, Ching J, Wang Y. Managing risk in geotechnical engineering–from data to digitalization. In: Proc., 7th Int. Symp. on Geotechnical Safety and Risk (ISGSR 2019). 2019. pp. 13-34

[5] Olaiz AH, Zapata CE, Soltanpour Y. A Bayesian forecast framework for climatic parameters in geotechnical modeling. In: Geo-Risk. 2023. pp. 88-97

[6] Uzielli M, Lacasse S, Nadim F, Phoon KK. Soil variability analysis for geotechnical practice. Characterization and Engineering Properties of Natural Soils. 2006;**3**:1653-1752

[7] Phoon KK, Kulhawy FH. Characterization of geotechnical variability. Canadian Geotechnical Journal. 1999;**36**(4):612-624. DOI: 10.1139/t99-038

[8] Wang L, Wu C, Tang L, Zhang W, Lacasse S, Liu H, et al. Efficient reliability analysis of earth dam slope stability using extreme gradient boosting method. Acta Geotechnica. 2020;**15**: 3135-3150. DOI: 10.1007/s11440-020-00962-4

[9] Atkinson J. An Introduction to the Mechanics of Soils and Foundations: Through Critical State Soil Mechanics. United States: McGraw-Hill Book Company (UK) Ltd.; 1993

[10] Mayne PW, Coop MR, Springman SM, Huang AB, Zornberg JG. Geomaterial behavior and testing. In: Proceedings of the 17th International Conference on Soil Mechanics and Geotechnical Engineering, Egypt. Vol. 1, 2, 3 and 4. IOS Press; 2009. pp. 2777-2872. DOI: 10.3233/978-1-60750-031-5-2777

[11] Brandl H. Energy foundations and other thermo-active ground structures. Géotechnique. 2006;**56**(2):81-122. DOI: 10.1680/geot.2006.56.2.81

[12] Madhusudhan BR, Boominathan A, Banerjee S. Effect of specimen size on the dynamic properties of river sand and rubber tire shreds from cyclic triaxial and cyclic simple shear tests. In geotechnical characterization and modelling. In: Proceedings of IGC. Springer Singapore; 2020, 2018. pp. 453-465. DOI: 10.1007/978-981-15-6086-6_37

[13] Ferrario MF, Bonadeo L, Brunamonte F, Livio F, Martinelli E, Michetti AM, et al. Late quaternary environmental evolution of the Como urban area (northern Italy): A multidisciplinary tool for risk management and urban planning. Engineering Geology. 2015;**193**:384-401. DOI: 10.1016/j.enggeo.2015.05.013

[14] Terzaghi K, Peck RB, Mesri G. Soil Mechanics in Engineering Practice. Canada: John Wiley & Sons; 1996

[15] Zhang W, Li H, Li Y, Liu H, Chen Y, Ding X. Application of deep learning

algorithms in geotechnical engineering: A short critical review. Artificial Intelligence Review. 2021;**54**(8):5633-5673. DOI: 10.1007/s10462-021-09967-1

[16] Abbaszadeh Shahri A, Kheiri A, Hamzeh A. Subsurface topographic modeling using geospatial and data driven algorithm. ISPRS International Journal of Geo-Information. 2021;**10**(5): 341. DOI: 10.3390/ijgi10050341

[17] Tabarsa A, Latifi N, Osouli A, Bagheri Y. Unconfined compressive strength prediction of soils stabilized using artificial neural networks and support vector machines. Frontiers of Structural and Civil Engineering. 2021; **15**:520-536. DOI: 10.1007/s11709-021-0689-9

[18] Phoon KK, Zhang W. Future of machine learning in geotechnics. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards. 2023;**17**(1):7-22. DOI: 10.1080/17499518.2022.2087884

[19] Dodangeh E, Choubin B, Eigdir AN, Nabipour N, Panahi M, Shamshirband S, et al. Integrated machine learning methods with resampling algorithms for flood susceptibility prediction. Science of the Total Environment. 2020;**705**: 135983. DOI: 10.1016/j.scitotenv.2019.135983

[20] Deng X, Liu P, Liu X, Wang R, Zhang Y, He J, et al. Geospatial big data: New paradigm of remote sensing applications. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2019;**12**(10): 3841-3851. DOI: 10.1109/JSTARS.2019.2944952

[21] Zhou J, Huang S, Wang M, Qiu Y. Performance evaluation of hybrid GA–SVM and GWO–SVM models to predict earthquake-induced liquefaction potential of soil: A multi-dataset investigation. Engineering with Computers. 2022;**38**(5):4197-4215. DOI: 10.1007/s00366-021-01418-3

[22] Merghadi A, Yunus AP, Dou J, Whiteley J, ThaiPham B, Bui DT, et al. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. Earth-Science Reviews. 2020;**207**: 103225. DOI: 10.1016/j.earscirev.2020.103225

[23] Goodell JW, Kumar S, Lim WM, Pattnaik D. Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. Journal of Behavioral and Experimental Finance. 2021;**32**:100577. DOI: 10.1016/j.jbef.2021.100577

[24] Vij A, Vijendra S, Jain A, Bajaj S, Bassi A, Sharma A. IoT and machine learning approaches for automation of farm irrigation system. Procedia Computer Science. 2020;**167**:1250-1257. DOI: 10.1016/j.procs.2020.03.440

[25] Szakács A. Precursor-based earthquake prediction research: Proposal for a paradigm-shifting strategy. Frontiers in Earth Science. 2021;**8**: 548398. DOI: 10.3389/feart.2020.548398

[26] Chen J, Vissinga M, Shen Y, Hu S, Beal E, Newlin J. Machine learning–based digital integration of geotechnical and ultrahigh–frequency geophysical data for offshore site characterizations. Journal of Geotechnical and Geoenvironmental Engineering. 2021; **147**(12):04021160. DOI: 10.1061/(ASCE)GT.1943-5606.0002702

[27] Bhattacharya S, Demirci HE, Nikitas G, Prakhya GKV, Lombardi D, Alexander NA, et al. Chapter 11 - Physical modeling of interaction

problems in geotechnical engineering. Modeling in Geotechnical Engineering. Academic Press; 2021. pp. 205-256. DOI: 10.1016/B978-0-12-821205-9.00017-4. ISBN 9780128212059

[28] Pan Q, Qu X, Liu L, Dias D. A sequential sparse polynomial chaos expansion using Bayesian regression for geotechnical reliability estimations. International Journal for Numerical and Analytical Methods in Geomechanics. 2020;**44**(6):874-889. DOI: 10.1002/nag.3044

[29] Kim HJ, Mawuntu KBA, Park TW, Kim HS, Park JY, Jeong YS. Spatial autocorrelation incorporated machine learning model for geotechnical subsurface modeling. Applied Sciences. 2023;**13**(7):4497. DOI: 10.3390/app13074497

[30] Carri A. Innovative application of iot technologies to improve geotechnical monitoring tools and early warning performances. In: Critical Thinking in the Sustainable Rehabilitation and Risk Management of the Built Environment: CRIT-RE-BUILT. Proceedings of the International Conference; November 7-9, 2019, Iași, Romania. Switzerland: Springer Nature; 2020. p. 142

[31] Mahdi IM, Ebid AM, Khallaf R. Decision support system for optimum soft clay improvement technique for highway construction projects. Ain Shams Engineering Journal. 2020;**11**(1):213-223. DOI: 10.1016/j.asej.2019.08.007

[32] Hallaji SM, Fang Y, Winfrey BK. Predictive maintenance of pumps in civil infrastructure: State-of-the-art, challenges and future directions. Automation in Construction. 2022;**134**:104049. DOI: 10.1016/j.autcon.2021.104049

[33] Yang Y, Lu Y, Mei G. A federated learning based approach for predicting landslide displacement considering data security. Future Generation Computer Systems. 2023;**149**:184-199. DOI: 10.1016/j.future.2023.07.021

[34] Criekemans D. Chapter 2 'Geotechnical Ensembles': How new technologies change geopolitical factors and contexts in economy, energy and security. Geopolitics and International Relations. Leiden, The Netherlands: Brill | Nijhoff; 2021. DOI: 10.1163/9789004432086_004

[35] Rodríguez Piedrabuena A. Feasibility study of using augmented reality in geotechnical site inspection [Bachelor's Thesis]. Spain: Universitat Politècnica de Catalunya; 2021. Available from: http://hdl.handle.net/2117/358184

[36] Riaz MT, Basharat M, Brunetti MT. Assessing the effectiveness of alternative landslide partitioning in machine learning methods for landslide prediction in the complex Himalayan terrain. Progress in Physical Geography: Earth and Environment. 2023;**47**(3):315-347. DOI: 10.1177/03091333221113660

[37] Seyedzadeh S, Rahimian FP, Oliver S, Rodriguez S, Glesk I. Machine learning modelling for predicting non-domestic buildings energy performance: A model to support deep energy retrofit decision-making. Applied Energy. 2020;**279**:115908. DOI: 10.1016/j.apenergy.2020.115908

[38] Velasco Herrera VM, Rossello EA, Orgeira MJ, Arioni L, Soon W, Velasco G, et al. Long-term forecasting of strong earthquakes in North America, South America, Japan, southern China and northern India with machine learning. Frontiers in Earth Science. 2022;**10**:905792. DOI: 10.3389/feart.2022.905792

[39] Quigley MC, Saunders W, Massey C, Van Dissen R, Villamor P, Jack H, et al. The utility of earth science information in post-earthquake land-use decision-making: The 2010–2011 Canterbury earthquake sequence in Aotearoa New Zealand. Natural Hazards and Earth System Sciences Discussions. 2020;**2020**: 1-35. DOI: 10.5194/nhess-20-3361-2020

[40] Song Q, Wu Y, Xin X, Yang L, Yang M, Chen H, et al. Real-time tunnel crack analysis system via deep learning. IEEE Access. 2019;**7**: 64186-64197. DOI: 10.1109/ACCESS.2019.2916330

[41] Kariminejad N, Mondini A, Hosseinalizadeh A et al. Detection of sinkholes and landslides in a semi-arid environment using deep-learning methods. UAV images, and Topographical Derivatives. 15 May 2023. PREPRINT (Version 1). DOI: 10.21203/rs.3.rs-2847897/v1. Available from: Research Square

[42] Bravo-López E, Fernández Del Castillo T, Sellers C, Delgado-García J. Landslide susceptibility mapping of landslides with artificial neural networks: Multi-approach analysis of backpropagation algorithm applying the neuralnet package in Cuenca. Ecuador. Remote Sensing. 2022;**14**(14):3495. DOI: 10.3390/rs14143495

[43] Elzain HE, Chung SY, Venkatramanan S, Selvam S, Ahemd HA, Seo YK, et al. Novel machine learning algorithms to predict the groundwater vulnerability index to nitrate pollution at two levels of modeling. Chemosphere. 2023;**314**: 137671. DOI: 10.1016/j.chemosphere.2022.137671

[44] Sircar A, Yadav K, Rayavarapu K, Bist N, Oza H. Application of machine learning and artificial intelligence in oil and gas industry. Petroleum Research. 2021;**6**(4):379-391. DOI: 10.1016/j.ptlrs.2021.05.009

[45] Zhang W, Gu X, Tang L, Yin Y, Liu D, Zhang Y. Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: Comprehensive review and future challenge. Gondwana Research. 2022;**109**:1-17. DOI: 10.1016/j.gr.2022.03.015

[46] Lee J, Azamfar M, Singh J, Siahpour S. Integration of digital twin and deep learning in cyber-physical systems: Towards smart manufacturing. IET Collaborative Intelligent Manufacturing. 2020;**2**(1):34-36. DOI: 10.1049/iet-cim.2020.0009

[47] Lu X, Xu Y, Tian Y, Cetiner B, Taciroglu E. A deep learning approach to rapid regional post-event seismic damage assessment using time-frequency distributions of ground motions. Earthquake Engineering & Structural Dynamics. 2021;**50**(6): 1612-1627. DOI: 10.1002/eqe.3415

[48] Kim HS, Sun CG, Lee MG, Cho HI. Multivariate geotechnical zonation of seismic site effects with clustering-blended model for a city area, South Korea. Engineering Geology. 2021;**294**: 106365. DOI: 10.1016/j.enggeo.2021.106365

[49] Künzler M, Huggel C, Ramírez JM. A risk analysis for floods and lahars: Case study in the Cordillera Central of Colombia. Natural Hazards. Oct 2012;**64**: 767-796

[50] Yaseen ZM. An insight into machine learning models era in simulating soil, water bodies and adsorption heavy metals: Review, challenges and solutions. Chemosphere. 2021;**277**:130126. DOI: 10.1016/j.chemosphere.2021.130126

[51] Zhang P, Yin ZY, Jin YF. Machine learning-based modelling of soil properties for geotechnical design: Review, tool development and comparison. Archives of Computational Methods in Engineering. 2022;**29**(2): 1229-1245. DOI: 10.1007/s11831-021-09615-5

[52] Jena R, Shanableh A, Al-Ruzouq R, Pradhan B, Gibril MBA, Ghorbanzadeh O, et al. An integration of deep learning and transfer learning for earthquake-risk assessment in the Eurasian region. Remote Sensing. 2023; **15**(15):3759. DOI: 10.3390/rs15153759

[53] Zhang C, Liang M, Song X, Liu L, Wang H, Li W, et al. Generative adversarial network for geological prediction based on TBM operational data. Mechanical Systems and Signal Processing. 2022;**162**:108035. DOI: 10.1016/j.ymssp.2021.108035

[54] Fang H, Shao Y, Xie C, Tian B, Shen C, Zhu Y, et al. A new approach to spatial landslide susceptibility prediction in karst mining areas based on explainable Artificial Intelligence. Sustainability. 2023;**8, 15**(4):3094. DOI: 10.3390/su15043094

[55] Krechowicz M, Krechowicz A. Risk assessment in energy infrastructure installations by horizontal directional drilling using machine learning. Energies. 2021;**14**(2):289. DOI: 10.3390/en14020289

[56] Mitelman A, Yang B, Urlainis A, Elmo D. Coupling geotechnical numerical analysis with machine learning for observational method projects. Geosciences. 2023;**13**(7):196. DOI: 10.3390/geosciences13070196

[57] Barzegar M, Blanks S, Sainsbury BA, Timms W. MEMS technology and applications in geotechnical monitoring: A review. Measurement Science and Technology. 2022;**33**(5):052001. DOI: 10.1088/1361-6501/ac4f00

[58] Jena R, Pradhan B, Beydoun G, Al-Amri A, Sofyan H. Seismic hazard and risk assessment: A review of state-of-the-art traditional and GIS models. Arabian Journal of Geosciences. 2020;**13**: 1-21. DOI: 10.1007/s12517-019-5012-x

[59] Kim HS, Ji Y. Three-dimensional geotechnical-layer mapping in Seoul using borehole database and deep neural network-based model. Engineering Geology. 2022;**297**:106489. DOI: 10.1016/j.enggeo.2021.106489

[60] Guan QZ, Yang ZX, Guo N, Hu Z. Finite element geotechnical analysis incorporating deep learning-based soil model. Computers and Geotechnics. 2023;**154**:105120. DOI: 10.1016/j.compgeo.2022.105120

[61] Xie J, Huang J, Zeng C, Jiang SH, Podlich N. Systematic literature review on data-driven models for predictive maintenance of railway track: Implications in geotechnical engineering. Geosciences. 2020;**10**(11):425. DOI: 10.3390/geosciences10110425

[62] Pei T. Integrating Geotechnical Domain Knowledge into Machine Learning for Slope Stability Predictions. [PhD Thesis]. USA: Penn State University; 2023

[63] Mali N, Dutt V, Uday KV. Determining the geotechnical slope failure factors via ensemble and individual machine learning techniques: A case study in Mandi, India. Frontiers in Earth Science. 2021;**9**. DOI: 10.3389/feart.2021.701837

[64] Xie J, Huang J, Zeng C, Huang S, Burton GJ. A generic framework for geotechnical subsurface modeling with

machine learning. Journal of Rock Mechanics and Geotechnical Engineering. 2022;**14**(5):1366-1379. DOI: 10.1016/j.jrmge.2022.08.001

[65] Zhang W, Pradhan B, Stuyts B, Xu C. Application of artificial intelligence in geotechnical and geohazard investigations. Geological Journal. 2023;**58**(6):2187-2194. DOI: 10.1002/gj.4779

[66] Phoon KK, Zhang LM, Cao ZJ. Special issue on "machine learning and AI in geotechnics". Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards. 2023;**17**(1):1-6. DOI: 10.1080/17499518.2023.2185938