Machine Learning Applications

Hope, Hype, or Hindrance for Structural Engineering

By Henry V. Burton, S.E., Ph.D., and Michael Mieler, Ph.D.

Machine learning (ML) is a branch of artificial intelligence (AI) that uses algorithms to find patterns in data and make predictions about the future, essentially enabling computers to learn without being programmed explicitly beforehand. While AI and ML have been active fields of study and research since the 1950s, they have exploded in popularity over the past decade. This explosion is thanks to deep learning (DL), a type of ML that leverages big data and neural networks to tackle a diverse range of problems, from image recognition and fraud detection to customer support chatbots and language translation. *Figure 1* shows the relationship between AI, ML, and DL, while *Figure 2* shows the three major types of ML.

The recent success of ML applications in areas such as bioengineering, medicine, and advertising has been highly visible, creating a domino effect where others have begun to ask whether their respective fields of practice, including structural engineering, can be transformed or "revolutionized" by ML. Structural engineering researchers began to explore ML applications as early as the late 1980s. However, it is only within the last five years that the community of structural engineering researchers and practitioners has begun to seriously explore ways in which ML can improve the efficiency and/or accuracy of specific tasks or solve previously intractable problems. As with other fields, some have expressed legitimate concern that the potential benefits of ML to structural engineering are being overhyped and, in the worst case, exploited for marketing purposes.

This article identifies and reviews three areas of current and potential ML applications in structural engineering and discusses challenges and opportunities associated with each. The authors conclude with a general discussion of some of the challenges that must be addressed if ML is to be effectively used in structural engineering practice.

Improving Empirical Models

There is a long history of using statistical models based on experimental or field data to predict various types of parameters used in
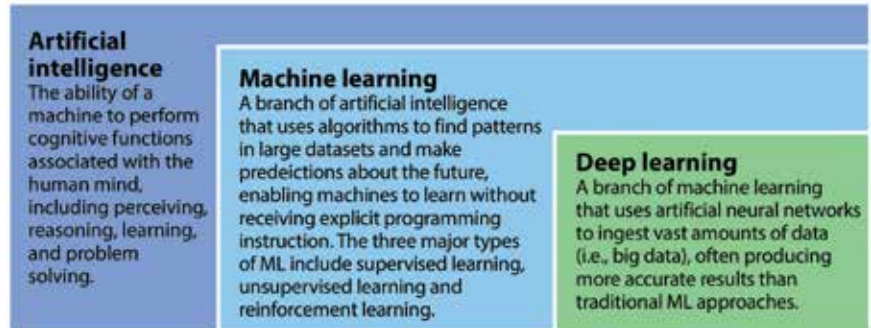

Figure 1. Overview of the relationship between AI, ML, and DL.

structural design, response analysis, and performance assessment. These types of models have been shown to be especially useful when physics-based models, which are often simplified for practical purposes, are unable to capture known physical mechanisms. For example, many of the analytical relationships provided in the ACI-318 *Building Code Requirements for Structural Concrete* are either fully empirical (i.e., based on regression using experimental data) or a hybrid (i.e., engineering equations with empirically derived parameters). Therefore, it is worthwhile to examine whether the predictive performance of these empirical relationships can be improved by using ML models and what, if any, tradeoffs are included in the latter.

For illustration purposes, examine Equation 18.10.6.2b of ACI-318-19, which is used to estimate the drift capacity (i.e., drift associated with a 20% peak strength loss) of special structural walls. The equation was developed by performing linear regression on a dataset of 164 physical experiments (Abdullah and Wallace, 2019) using the ratio of wall neutral axis depth-to-compression zone width, the ratio of wall length-to-compression zone width, wall shear stress ratio, and the configuration of the boundary zone reinforcement as the predictor (or independent) variables. Using the same dataset and predictors, a drift capacity model was developed using the Extreme Gradient Boosting (XGBoost) machine learning algorithm. In brief, XGBoost is a type of "tree-based" algorithm that works by repeatedly subdividing the dataset based on a set of criteria defined to maximize the accuracy of the resulting predictive model. More details of how this algorithm works can be found in Huang and Burton (2019).

Figure 3 (page 18) shows a plot of the observed (from the experimental data) versus predicted (by the model) drift capacity values for the linear regression (*Figure 3a*) and XGBoost (*Figure 3b*) models. The solid diagonal lines in the two plots represent the locations where the observed and predicted values are the same. Compared to the linear model, the XGBoost data points are more closely aligned along the diagonal, indicating superior predictive performance. A quantitative comparison of the two models can be obtained by computing the $D_{X\%}$ value, which is the percentage of data points

Figure 2. Three major types of ML.

with errors (the difference between observed and predicted) that are less than some predefined percentage (X) of the observed values. The $D_{10\%}$ value for the XGBoost model is 76%, which indicates that 76% of the errors in the predictions are within 10% of the observed values. In contrast, the $D_{10\%}$ value for the linear model is approximately 53%, which further indicates the improved accuracy provided by the XGBoost algorithm.

A defining feature of tree-based ML algorithms (including XGBoost) is that the resulting model cannot be expressed analytically. This creates obvious challenges with interpreting or interrogating the model and even implementing it in a building code or standard.

Therefore, one has to weigh the benefit of increased accuracy with the increased complexity of the XGBoost model. The latter can be addressed by developing a software application that implements the relevant ML model and uses visualizations (e.g., plots) to explain the relationship between the input and output variables.

The example presented in this section represents just one of several possibilities where the accuracy of existing empirical relationships for predicting structural response, capacity, or performance can be improved using ML models. However, as noted earlier, this often comes at a cost, which is frequently associated with the complexity of ML models. Other areas where ML algorithms are being examined as potential replacements for commonly used empirical models include component failure model classification, natural hazard damage assessment (e.g., predicting earthquake or flood damage to buildings), and structural health monitoring.

Surrogate Models

ML models could also potentially be used to reduce the time and effort associated with some computationally expensive structural analysis tasks. For example, performance-based design (PBD) often necessitates detailed nonlinear (geometric and material) analyses to understand structures' response to extreme loading. However, despite significant research advancements, PBD has not been widely adopted in practice. Even the 2nd generation performance-based earthquake engineering (PBEE) framework, which is considered by many to be the template for PBD, has not seen widespread adoption. This is partly because the state of structural engineering practice is such that most designers rely on linear analysis models

to estimate response demands. The use of nonlinear response history analysis as part of the design process is generally expensive, both computationally and in terms of labor, which often leads to unjustifiable design fees.

One strategy for reducing the computational expense and labor of nonlinear analyses is to use surrogate models to estimate response demands. Additional details on the procedure for developing surrogate models can be found in Moradi et al. 2018.

The process of creating and implementing surrogate models for structural response estimation is not without its challenges. First, the dataset used to train the ML model is generated using explicit

response analyses, which will be computationally expensive. However, the increased availability of computational and storage resources within the research community is likely to overcome this challenge. There are also limitations in terms of the lack of generalizability of surrogate models. ML models are generally good at interpolation (making predictions within the bounds of the dataset). Still, they do not perform as well at extrapolation (making predictions outside the bounds of the dataset). Therefore, the application limits of surrogate models should be carefully communicated to the users. Lastly, the challenges with interpretation discussed previously are also relevant to surrogate models.

In addition to potential applications in PBD, surrogate models can be used in regional assessments of natural hazard impacts. More specifically, the earthquake engineering research community has begun to explore the use of surrogate models as part of the workflow for regional seismic risk assessment utilizing the FEMA P-58 PBEE procedure. At this scale, the inventory size is typically on the order of tens or hundreds of thousands of buildings, such that explicit nonlinear response simulation becomes unfeasible. The Seismic Performance Prediction Program (SP3) (**www.hbrisk.com**), a commercial online tool that is used for PBEE assessments, has developed a "structural response prediction engine" that essentially uses the surrogate model concept to enable users to bypass explicit nonlinear response history analyses. However, the details of the adopted statistical or ML methods have not been made publicly available.
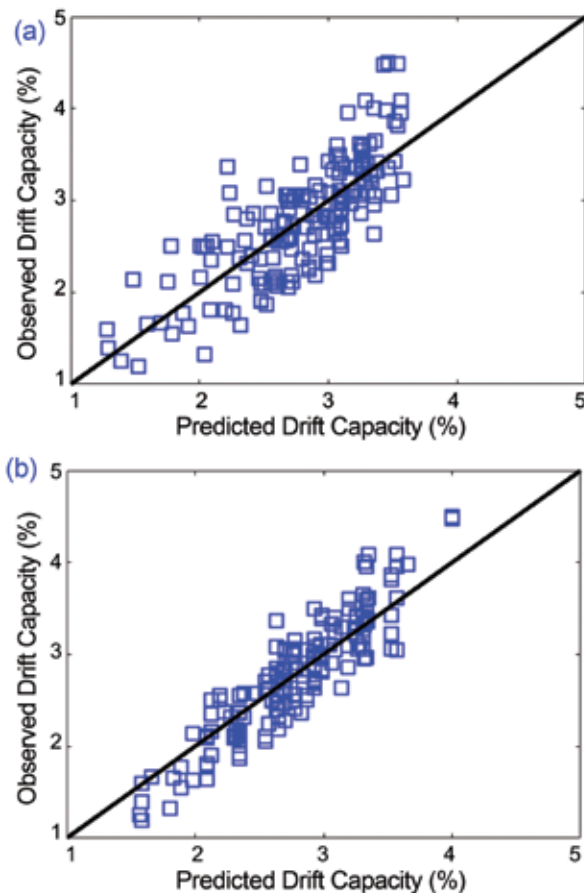
Information Extraction

ML models for automating information extraction from images, video, and written text have found widespread applications in several fields, including engineering, medicine, and different branches of science. A reasonable argument can be made that, among the various ML technologies, those that automate information extraction from different media sources hold the greatest promise in structural engineering applications. This is partially supported by the rapid growth in the popularity of this area of ML-related structural engineering research within the past three years. Additionally, there is anecdotal evidence that, unlike the previously discussed areas of potential application, these information extraction technologies have started to make their way into practice.

Computer vision (CV) is a subcategory of AI that gives computers the ability to extract meaningful information from images and videos. DL algorithms are essential to the functionality of modern CV techniques. For almost a decade, the research community has been exploring the use of CV techniques to classify structural system and component types, detect defects or damage

Figures 3a and 3b. Observed versus predicted drift capacity values; a) linear regression, b) XG Boost model.

in structures (e.g., cracks in concrete, loosened bolts), automate the development of as-built models using images, and identify seismic deficiencies. Through informal inquiries, the authors also found that structural engineering practitioners are using CV to detect corrosion in offshore structures, monitor the structural integrity of bridges during construction, and identify the presence of soft stories in buildings.

Natural language processing (NLP) is a branch of AI focused on empowering computers to extract useful information from written text. NLP technologies have found widespread applications in advertising, litigation tasks, and medicine. Despite being one of the least explored ML technologies in structural engineering, researchers have begun to investigate the use of NLP to rapidly assess the damage to buildings caused by natural hazard events using text generated by field inspections. Additionally, practitioners have begun to explore NLP's use to catalog building information by automating text extraction from structural drawings.

Challenges

For ML applications to become an effective part of structural engineering practice, several challenges must be addressed. The first involves sufficient access to diverse and high-quality data. Most, if not all, of the previously described structural engineering applications have utilized relatively small, homogeneous datasets, which implies that the resulting models cannot be generalized. The development of genuinely representative datasets requires that open access repositories with rigorous quality control be instituted. The research community has begun to take the necessary steps towards achieving this vision by establishing platforms such as DesignSafe, Structural ImageNet, and the DataCenterHub (**http://datacenterhub.org**).

The second challenge involves the "black box" nature of some ML algorithms, giving rise to interpretation and quality control issues. However, it is important to distinguish between those algorithms that are truly black boxes (e.g., deep neural networks whose predictions often cannot be explained by ML experts) and those that are somewhat interpretable (e.g., XGBoost) but merely unfamiliar to the structural engineering community. It is reasonable to suggest that ML not be used by individuals who are completely unfamiliar with the algorithms. This is especially true for structural design and analysis tasks where safety and reliability are a primary concern. In the long term, the inclusion of basic data science courses in university civil/structural engineering curricula could increase the number of structural engineering practitioners that have a working knowledge of some commonly used ML algorithms.

The third challenge, which is a side effect of its current popularity, is that AI is sometimes exalted as a "cure-all" for problems across different domains. As a marketing tool, there is a tendency to lead with the predictive accuracy of various models. A hypothetical example of such a claim is that "our model can predict earthquake damage to buildings with 95% accuracy." Important nuances such as the differences between the model development and application context and uncertainties in the prediction outcomes are often lost in such declarations. It is therefore critical that developers of structural engineering ML technologies communicate openly and honestly with users about limitations and potential pitfalls. This is especially important when the application context has implications for the safety of populations on a broad scale.

The fourth challenge centers on determining whether ML is suitable for a particular structural engineering task. First and foremost, it is important to examine whether there are tangible benefits to its application (e.g., increased productivity or improved prediction accuracy relative to existing models). In those cases where there are actual benefits, potential users also need to weigh them against tradeoffs, such as reduced interpretability. Another important consideration is whether the dataset used to train the ML model is within the domain of the potential application. As noted earlier, many ML models are good at interpolation but less capable of extrapolating beyond the training dataset, which brings into question their ability to generalize. Lastly, the high predictive accuracy of some ML algorithms (e.g., deep learning) within the context of training and testing is such that the uncertainty associated with their predictions on "unseen" datasets is often overlooked. The centrality of the role that structural engineers play in ensuring the safety and functionality of the built environment is such that uncertainty in design and performance outcomes cannot be ignored.

In summary, the authors' opinion is that there are some promising areas where ML can provide meaningful benefits to the practice of structural engineering. However, effective implementation requires that the community of practitioners and researchers come together to address some of the challenges that have been identified in this article. In the meantime, both researchers and practitioners should proceed with "cautious exploration" of ML applications in structural engineering while being mindful of the profession's commitment to the public's safety and well-being.▪

Dr. Henry V. Burton is an Associate Professor and the Englekirk Presidential Chair in Structural Engineering in the Department of Civil and Environmental Engineering at the University of California, Los Angeles. (hvburton@ucla.edu)

Dr. Michael Mieler is a Senior Risk and Resilience Engineer in Arup's San Francisco office. (michael.mieler@arup.com)

References

Abdullah, S., and Wallace, J. (2019). Drift capacity of reinforced concrete structural walls with special boundary elements, *ACI Structural Journal* 116(1), 183-194.

Cook, D, Wade, K, Haselton, C, Baker, J, and DeBock, D. (2018). *A structural response prediction engine to support advanced seismic risk assessment*, Proceedings of the 11th National Conference in Earthquake Engineering, Los Angeles, California, USA.

FEMA. (2012). *Seismic Performance Assessment of Buildings*, FEMA-P58, Applied Technology Council, Redwood City, CA.

Huang, H., and Burton, H. V. (2019). Classification of in-plane failure modes for reinforced concrete frames with infills using machine learning. *Journal of Building Engineering*, 25, 100767.

Mangalathu, S., Jang, H., Hwang, S. H., and Jeon, J. S. (2020). Data-driven machine-learning-based seismic failure mode identification of reinforced concrete shear walls. *Engineering Structures*, 208, 110331.

Mangalathu, S., Sun, H., Nweke, C. C., Yi, Z., and Burton, H. V. (2020). Classifying earthquake damage to buildings using machine learning. *Earthquake Spectra*, 36(1), 183-208.

Mangalathu, S., and Burton, H. V. (2019). Deep learning-based classification of earthquake-impacted buildings using textual damage descriptions. *International Journal of Disaster Risk Reduction*, 36, 101111.

McKinsey & Company (2018). *An executive's guide to AI*. **www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai#**

Moradi, S., Burton, H. V., and Kumar, I. (2018). Parameterized fragility functions for controlled rocking steel braced frames. *Engineering Structures*, 176, 254-264.

Samuel, AL (1959). "Some studies in machine learning using the game of checkers." *IBM Journal of Research and Development* 3(3): 535–554.

Sun, H., Burton, H., and Wallace, J. (2019). Reconstructing seismic response demands across multiple tall buildings using kernel-based machine learning methods. *Structural Control and Health Monitoring*, 26(7), 1-26.