This is a preprint draft. The published article can be found at: https://doi.org/10.1080/10286608.2022.2049257

Please cite this paper as:

Naser M.Z., Ross B. (2022). "An opinion piece on the dos and don'ts of artificial intelligence in civil engineering and charting a path from data-driven analysis to causal knowledge discovery." *Civil Engineering and Environmental Systems*. https://doi.org/10.1080/10286608.2022.2049257

An Opinion Piece on The Dos and Don'ts of Artificial Intelligence in Civil Engineering and Charting a Path from Data-driven Analysis to Causal Knowledge Discovery

M.Z. Naser[1,2], Brandon Ross[1]
[1]School of Civil and Environmental Engineering & Earth Sciences, Clemson University, Clemson, SC 29634, USA
[2]Artificial Intelligence Research Institute for Science and Engineering, Clemson University, Clemson, SC 29634, USA
E-mail: mznaser@clemson.edu, Website: www.mznaser.com, E-mail: bross2@clemson.edu

Synopsis

Artificial intelligence (AI) has been established as a universal *language* for solving science and engineering problems. Despite the rise of big data, the success of AI in parallel fields, and exciting works published at this frontier, some in the civil engineering community tie AI to a mystique stigma. And yet, there is also ever-growing inertia to embrace AI fully. The mystique of AI arises because 1) AI is not typically taught in a traditional civil engineering curriculum, 2) the majority of civil engineers remain appliers (as opposed to creators) of AI, and 3) commonly adopted AI algorithms leverage *blackbox* methods – the opposite to that commonly accepted in the civil engineering domain. We write this *opinion piece* with the aim of presenting a holistic look into the dos and don'ts of adopting AI into civil engineering.

Keywords: Artificial intelligence, Civil Engineering, Data-driven analysis, Causality, Knowledge discovery.

***Do** educate yourself on AI.*

Civil engineers are public servants with a duty for lifelong learning. Such a duty does not end by attaining a college degree but stretches beyond such a degree. A civil engineering curriculum doesn't and can't cover every aspect a civil engineer needs to learn, but rather covers the fundamentals that such an individual is likely to need during their tenure. While a civil engineering track of study rarely covers AI, AI-based solutions are being integrated into various civil engineering areas such as surveying, construction management, structural design, etc. A keen eye to the above can see a growing disconnect between our curriculum, academia, and industry. Given the rapid advancements in AI and the new opportunities it continues to bring, perhaps this disconnect is bound to disrupt our often calm, traditional, and steady discipline. Thus, the notion that AI is not an integrated component of our classical curriculum does not justify reticence to learn (or at the very least get familiar with) AI, especially now, where educational materials can be found on-demand and with ease.

This brings us to the following question, what is AI? Well, AI is a technology that enables machines to simulate human intelligence and perform tasks commonly undertaken by humans. AI has subfields emerge, specifically *machine learning (ML)* and *deep learning (DL)*. In ML, machines can automatically, and without explicit programming, learn from data to extract hidden patterns. DL falls under ML and is dedicated to creating a specialized form of multi-layered algorithms (or networks) that resemble the brain. The definitions of AI, ML, and DL do not specifically mention algorithmic details nor how these technologies practically relate to civil engineering. The same also infers that such technologies are primarily driven by data. Hence, a

This is a preprint draft. The published article can be found at: https://doi.org/10.1080/10286608.2022.2049257

Please cite this paper as:

Naser M.Z., Ross B. (2022). "An opinion piece on the dos and don'ts of artificial intelligence in civil engineering and charting a path from data-driven analysis to causal knowledge discovery." *Civil Engineering and Environmental Systems*. https://doi.org/10.1080/10286608.2022.2049257

question may arise. How can we infuse our prior knowledge of a system (say, civil engineering judgment) with AI technologies? Here is where *physics-informed AI* or *mechanistic AI* becomes effective and where we can enforce physical laws or ensure that a model satisfies such laws. Finally, we have a promising class of AI, referred to as *causal AI*. In causal AI, a model can be used to pinpoint the causes of a phenomenon, deliver insights, or simply be able to reason – all of which are features that other AI technologies fail to provide. From a civil engineering perspective, AI can be grouped under supervised learning (for regression and classification problems) and semi- and unsupervised learning (for clustering/labeling/control problems). Thus, to properly navigate the above technologies, we first need to identify a problem type and objective(s).

***Do** be curious and ask, how can AI help me?*

Engineers are curious by nature. We are trained to solve problems using logical and mathematical procedures. We also devise experiments or develop numerical models to examine or predict phenomena. In this pursuit, we generate data and document observations – many of which remain shelved away in reports and publications. A question then arises, did we fully utilize such data? Perhaps adopting a new method of investigation can identify hidden patterns that may have fallen between the cracks of idealizations and simplifying assumptions built into traditional methods (e.g., statistical methods, and especially when our data does not truly satisfy principles of statistics) (Rahmanpanah et al., 2020; Tran-Ngoc et al., 2020). The open literature continues to document scholarly works where predictions from AI methods outperform those obtained from building code provisions and statistical methods. Given the power of AI, shouldn't we use these new tools because they provide superior predictions and are built from data pools that are often more modern, diverse, and larger than those used in deriving codal provisions?[1] So, how can AI help a civil engineer? Well, many of our problems resemble a closed-loop system that is primed for automation. In such a system, a number of parameters are seen to associate with an outcome (or response). If we collect enough data on parameters and responses, then we could potentially better understand our problems and create/optimize novel solutions to overcome such problems[2]. The above is often defined as a data-driven approach to solving problems, i.e., arriving at answers by examining data. But, what about other problems that do not resemble a closed-loop system? Presumably those with an artistic or creative view. Current and future advancements in AI are leading to AI systems capable of maneuvering such problems. Let us keep an open mind for now. In a way, and on a philosophical level, learning AI can be an *asset* to our engineers.

***Do** ask if you need AI for a given problem*

Just because we have a new tool (in this case, AI) does not mean it is suited for all problems. Choosing to use AI is just like selecting other tools; the answer is based on the user's competency, preferences, past experience, and established professional practices. We tend to choose what we are familiar with and comfortable using. For example, and fundamentally speaking, we use codal

[1] An answer (or debate) to the above is likely to contain scientific, engineering, and philosophical arguments.
[2] An argument can be made for simply coding the relationship between parameters and response(s). However, coding requires knowledge of such a relationship (which is often resembled in a codal procedure or expression(s)). Thus, this argument can be countered by the superior performance of AI the associated use of lager data sets.

provisions because we trust such provisions, and we know that they are derived via collective and continuing efforts. We also use alternative tools when a previous approach fails to answer our inquiry or when we want to test the merit of the earlier approach[3]. As a profession, civil engineering has a rich history of appropriately applying simplifying assumptions to create elegant and efficient approaches to problems. Oftentimes these approaches are all that is needed. While AI may provide us with novel means to solve our problems, let us not be too hasty with this novelty and bypass classical civil engineering and statistical principles. Be intelligent in leveraging AI, and do not forget your roots.

***Do** remember the implication(s) of data-driven AI vs. causal AI.*

Data-driven approaches can come in handy since we often have plenty of data (thanks to the rise of big data and over a century of engineering experiments and publications) and access to high computing stations. This can result in the development of AI surrogates that can help us overcome bottlenecks associated with testing and simulations. There are a number of assumptions to data-driven approaches; most notably, 1) such approaches are tied to having comprehensive datasets, and 2) such data seem to repeat consistently. A skeptic may ask, do we have access or capability to create comprehensive datasets? And if so, how can we be sure of the quality and comprehensiveness of such datasets? Complete answers to these questions are beyond the scope of this piece; however, a rational argument for the sake of this discussion would be to *assume* that the datasets we have are comprehensive *enough*. This is a common assumption that we already make when choosing statistical methods and when analyzing our experiments (which traditionally comprise a smaller number of specimens). However, statistical methods are often constrained with assumptions (regarding the range of applicability of the methods themselves or type/quality/distribution of data, etc.), and we may not be able to arrive at suitable solutions[4].

A timely alternative would be to explore AI models. Despite being blackboxes, AI models have been shown to perform well and can arrive at predictions with high accuracy that may far exceed that obtained from statistics and provisions from our codes and standards. However, we do not know how such models work, nor how they were able to arrive at such predictions (Rudin, 2019). So, this leaves us with a tool that seems to work well on our assumed-to-be comprehensive enough dataset, but we do not know how the tool works. *Would you trust such a tool?*[5] The point to take home is that we need to be cognizant of our methods of investigation and be able to critique the limits and merits of such methods. In some cases, data-driven blackbox AI may be the only available option, or it may be used to supplement traditional methods. In some instances, blackbox AI and real-world big data have been used to confirm results observed from limited field studies, laboratory tests, and engineering-based parametric models (Jonnalagadda et al., 2016). The dilemma of blackbox models can be overcome by physics-informed and causal AI (as these continue to flourish). While physics-informed AI adds constraints to the AI solution that mandate compliance with established physical laws and first principles, causal AI goes beyond that to get

[3] Let us acknowledge that other reasons may exist (i.e., need to publish novel scholarly works, etc.).
[4] Please refer to (Bzdok et al., 2018) for a detailed discussion on some of the limitations of statistical methods.
[5] The same can also be said to trusting an unknown version of an expression or an untested FE model.

to the cause of a phenomenon. While data-driven AI has helped us to solve some problems, physics-informed AI can further help us with more problems; casual AI is likely to help in many, if not all, problems. Why? Because in causal AI we are unlocking the true mechanisms that genuinely govern our phenomena. Remember, each approach has its merits, uses, and applications[6].

***Do** be careful of false causality and **do not** exaggerate the outcome of explainability methods.* Oftentimes poorly prepared datasets or those with loose/co-dependent/confounding features can lead to "artifacts" that may deceive AI models into thinking that such features are essential (Papadimitriou & Garcia-Molina, 2011). In such instances, the same features may also mislead the AI user into thinking that such features are important, especially users with limited experience (remember that at this point in time, civil engineers are likely to be appliers of AI). This may lead to poor decision making, which can come at a high cost in certain scenarios. Similarly, augmenting traditional AI models with explainability or interpretability[7] measures (such as Shapely values or partial dependence plots, etc.), while admirable, is still constrained to understanding the model's behavior across its dataset and possibly from a data-driven perspective that may or may not agree with our prior engineering knowledge, and most importantly may not convey the actual relationships used in attaining such predictions. Remember that explainability methods may not reveal hidden or mechanistic causality as proper causal AI models (Pearl, 2009). Still, augmenting blackbox models with such methods is still an improvement over that of pure blackbox models since explainability/interpretability methods may unveil or pinpoint new information that could form the foundation for new knowledge or verify existing knowledge (Naser, 2021; Tahmassebi, Motamedi, Alavi, & Gandomi, 2021).

---

[6] From an academic view, this piece does favor casual models. Our argument is simple, once we identify the driving mechanism(s) governing a phenomenon, then we know all the ins and outs of such a phenomenon and unlock valuable knowledge. However, causal models are in their early stages, and we remain aware of the usefulness and wide applications of data-driven and physics-informed approaches – many of which we have used in our publications. In other words, there is a room and place for each and all approaches.

[7] At this point in time, explainability and interpretability are often used interchangeably. Simply put, when a blackbox model is augmented with a measure(s) that allow us to peek through its internal mechanisms to understand how it works, such a model is referred to as explainable and/or interpretable model (note: other terms may include *eXplainable AI (XAI)*, *glassbox*, *whitebox* or *transparent AI* and a further discission/debate on scientific, engineering, and philosophical views to explainability and interpretability can be found elsewhere (Adadi & Berrada, 2018; Barredo Arrieta et al., 2020)). From an engineering view, we would like AI models that can articulate how a prediction is arrived at, and what are the mechanisms that the model uses to tie the inputs to such a prediction. Knowing so will allow us to figure out if such mechanisms agree with out physics principles and engineering judgement vs. arriving at a prediction via pure data association. A suitable AI model would have the properties of an equation. We know how an equation is derived, we know the assumptions used to derive such an equation and we also know how the input interacts without each other to arrive at a prediction. Such a model would be easy to breakdown and debug. If we can see how the model works, and understand why it works, then we can start to trust such model and perhaps draw new inferences with confidence. Note that the above argument on the resemblance of an AI model to an equation is also equivalent to that of a FE software.

This is a preprint draft. The published article can be found at: https://doi.org/10.1080/10286608.2022.2049257

Please cite this paper as:

Naser M.Z., Ross B. (2022). "An opinion piece on the dos and don'ts of artificial intelligence in civil engineering and charting a path from data-driven analysis to causal knowledge discovery." *Civil Engineering and Environmental Systems*. https://doi.org/10.1080/10286608.2022.2049257

***Do** remember that the devil is in the data (and in the details too).*

Most of us are familiar with the finite element (FE) method, wherein FE numerical models are developed to simulate physical phenomena (say, to predict the deflection of a loaded beam). A typical FE model primarily comprises the geometrical model that reflects the beam's geometry and possible movements and a material model that conveys the behavior of the material of such a beam. Thus, for the same geometry, a modeler can achieve different deflection responses if the material model is varied between concrete, steel, and timber. From this perspective, the FE model (as a model) does not recognize that we are trying to simulate a beam, let alone a specific type of a beam with particular material properties, but rather it takes our inputs and solves for deflections. The same is also true in the case of AI models. The solution is only as good as the data and the details upon which the AI model is based. It is naïve to assume an AI model's quality of predictions exceeds the quality of its source data or the logic used to build it[8]. Remember that there is *little room* to finetune the setting of a given algorithm to improve its prediction capability, and there is significant room to enhance data quality and a logical path used to build the AI model. Do go over some of the classical and recent works that provide rules of thumb to develop proper AI models for civil engineering problems (Behnood & Golafshani, 2020; Carpenter & Barthelemy, 1994; Degtyarev, 2020; Naser, 2021; Sun, Burton, & Huang, 2021; Tarawneh, Momani, & Alawadi, 2021).

***Do** master simple AI models before learning complex/exotic/causal models.*

AI platforms are readily available, including those with built-in packages such as *Scikit*, *R*, and *Python*. This may cause some users to lose track of many "simple", or "beginner" AI models. It is then advisable to start with small and manageable problems to visualize how a simple AI model performs before moving towards "complex", "exotic"[9] or "causal" models. This step-by-step learning approach is fundamental in civil engineers. New civil engineering students are first introduced to the fundamentals of statics and mechanics and then drawn to the theory of structures and structural design principles. During this journey, examples are first articulated and solved using simple problems (i.e., equilibrium of a point) before moving to more advanced problems (e.g., design of joints). Visualization and learning by example are two of the inherent characteristics of engineers. Use your training to your advantage and *start with simple models*. You may be, or dare say be, surprised at how many challenging problems can be solved using simple models. Remember your objective(s) and draw a path. If you are trying to understand a phenomenon, then aim for causality models. Or, if you are exploring a new problem domain or data set, then perhaps try a data-driven approach or physics-informed approach. There is a fine line to each, and all approaches are bound to intersect at some point. The choice of an AI model or strategy depends on your level of experience, data set, and objective(s).

[8] Please refer to (Lones, n.d.) For a good discussion on the role of data in AI modeling.

[9] Complex models are those that require more than a mere application of an algorithm (i.e., joining/ensembling of different algorithms into one algorithm) and exotic models are those that use one algorithm to optimize or finetune the parent algorithm.

*You **do not** always need complex or exotic models!*

Many of the existing AI models have been developed by our colleagues in the mathematics and computer science departments. Such models have also been extensively tested and validated across multiple problems and datasets. Due to the above, AI models (simple, as well as exotic) perform well across different fields and are likely to do well in civil engineering, too (even in their default settings (Naser et al., 2021)). Thus, remember to aim to satisfy the Law of Parsimony of *Occam's Razor*, which implies that simplicity is a goal in itself (Hildebrand, 1938). Start with simple AI models, and then apply complex ones as needed. Simple models are easy to develop and finetune. Such models are easier to debug and trust; bugs and trust are bound to rise at some point in time continue to hinder the integration of AI. While it is exciting to see works involving intricate and novel AI models, such models can also (and simply) be contrasted against simple models that often take a fraction of the time to build and interpret than more complex models.

*Ask yourself, **do** I need a causal model?*

In some instances, our goal or objective is not to discover new knowledge. Perhaps one would like to explore a dataset or simply develop a quick solution to a problem. In other instances, we may already know the inherent mechanism(s) behind a phenomenon or simply not be interested in arriving at the true cause-and-effect. Thus, a causal model is not a necessity for a successful AI analysis.

***Do not** only concentrate on validating AI models and **do** think about model deployment.*

It is natural for an AI model to achieve high-performance metrics on selected, historical, or relatively "small" datasets (Baylor et al., 2017). Yet, this does not mean the model will perform as well when it is deployed in real-time, real-world conditions. In a way, reporting sky-high performance metrics in training or validation is impressive – and perhaps warranted and expected in scholarly articles. But one should be wary of misconstruing such metrics as a means for generalization or of justifying causal inferences. Remember, a model is trained and validated on a certain dataset[10], then deployed to the real world (where it is exposed to data likely to be of lesser quality and/or more variety than the initial set). During the training process, we hope to uncover why a certain phenomenon occurs, and during the deployment process, we hope to be able to capture the phenomenon on hand. While creating an AI surrogate to overcome a problem is possible via a data-driven approach, uncovering why a phenomenon occurs is a causal pursuit that may go beyond that of a data-driven approach.

***Do not** chase forced goodness.*

A common misconception associated with AI modeling is that they are only deemed fit if if they achieve high performance metrics (Corbett-Davies & Goel, 2018; Naser & Alavi, 2021). This exercise has initiated an unwarranted race towards overly tuned models to specific datasets[11] as well as inflated works with marginal merit. The reader is urged to remember that error metrics and performance fitness indicators are mathematical/logical constructs that describe the difference/spread between AI predictions and ground truths. While attaining high metrics may

[10] Unfortunately, the life cycle of most scholarly works ends here.

[11] This is unlike the race to develop "better", "new" or "improved" generalizing models.

imply that the model on hand captures the examined phenomenon *within* a dataset with an acceptable level of confidence, one needs to remember that not all metrics are created equally. Thus, the user needs to identify the most suitable metrics for the given problem/data and then also needs to understand the assumptions used in deriving such metrics to overcome their limitations – which will also help in better understanding model performance. False goodness occurs when inappropriate metrics are applied or when their outcome is misinterpreted (Arp et al., 2021). A false sense of causality arises once adequate models are assumed to "understand" or "grasp" or "map" the phenomenon on hand in a causal manner.

***Do** understand how performance metrics work and what they imply.*

Oftentimes, conventional metrics used in engineering and science problems (i.e., Pearson correlation coefficient (R), etc.) describe the overall performance of a particular model across a dataset. This may innocently hide poor performance in local regions within a dataset and notably near extreme ranges or outliers. Critically, the discussion on performance metrics tends to conceal the taken-for-granted assumptions within the metrics. These assumptions seem to be rarely or thoroughly checked (Arp et al., 2021). Similarly, different problem formulations tackled by AI may also require the adoption of specific metrics. For instance, the metrics used in regression and classification problems are different since both articulate distinct ways to evaluate performance (Botchkarev, 2019; Naser & Alavi, 2020).

***Do not** rely on one algorithm, nor one goodness metric.*

While, naturally, a user may be inclined to favor a specific AI model due to preference of familiarity, it is advisable to explore other algorithms. The famous "*No Free Lunch*" theory implies that some algorithms are not more sensible than others, but rather the averaged performance of all algorithms can be declared equivalent across all possible problems (Wolpert & Macready, 1997). From this perspective, an AI user is expected to attempt to explore more than one adequate algorithm (or algorithm architecture), or perhaps a series of adequate algorithms, in their analysis. Similarly, the above can also be extended towards selecting a combination of performance metrics, as opposed to a single metric, when examining the predictive capability of a given model. The critical point is to use multiple validation practices and to explore the auditability of new tools for AI analysis.

***Do** test, test, and further test.*

In general, AI modeling is relatively quick – especially when compared to traditional physical testing or FE simulations. This allows us to conduct further and affordable testing on our AI models – especially when most AI models are hard blackboxes that we do not know how they exactly work (remember that most civil engineers are appliers of AI, and not developers). As such, a user is encouraged to examine AI models with scrutiny. In addition to considering multiple algorithms and performance metrics, one may also opt to apply a combination of data transformation methods (i.e., linear, logarithmic, etc.) or validation methods (i.e., k-fold validation, train/validate/test splits, out of the bag testing, include confidence intervals, benchmark testing[12] , etc.) or ensembling

[12] Benchmarked databases and algorithms on key structural engineering problems have been recently published (see Naser et al. (2021)).

techniques. In addition, complimenting AI model testing with sensitivity and reliability studies can go a long way – especially in a design scenario where a certain level of reliability is needed (such as that often required in civil engineering building codes).

Do not *assume your model is unbiassed.*

In real life, data may be expensive to obtain and may be limited in size and comprehensiveness (Belkin et al., 2019). Thus, compiled datasets may not always be considered fair or inclusive. This is particularly problematic in social sciences, with datasets taken from human populations that do (or often don't) include data from minority groups. The same can also be a problem in engineering datasets that are thought to be complete. For example, datasets on urban water distribution in relation to resident income or race or wildfire-trigged evacuation in rural areas may not reflect the entire population and may leave minoritized groups out. One must not forget that historical data can be helpful to guide preliminary designs, yet a user should be aware that historical data may, or may not, be representative of recent and current conditions. Do remember that many models are biased since they are the result of *human operators* and *historical trends*[13].

Do *share your data and details of your models.*

To better facilitate the integration of AI into civil engineering, it is advisable for AI users to share full details (including data and model settings) of their works whenever possible. Noting the availability of free and comprehensive data repositories (i.e., Mendeley, GitHub), such data sharing can be easily appended. The same is also true for creating AI-based applications or software that can be shared with others. This practice will help further verify existing models, extend working models, and expedite adopting AI by providing the community with new data, case studies, and pre-built models[14]. Perhaps in the near future, we could arrive at standardized procedures for creating and validating AI models in our domain. To get there, we need a series of community initiatives and collaborative efforts. A simple and doable initiative at this point in time is to share our data and models.

Do *integrate AI into your future works.*

AI methods are expected to continue to advance and become more prominent in the years to come. It is then advisable to start to incorporate AI into our curriculum, research, and practice. For example, short modules can be added to existing courses, and practitioners can attend seminars and workshops on AI development and deployment. In addition, efforts may lead to special issues or edited volumes/collections with AI themes to be sponsored by leading journals, etc. Perhaps a smaller component of AI can be added to ongoing research projects/proposals. The boarder civil engineering industry homes immense and underutilized data. A line of communication should be established between academia and industry partners. Remember, the industry needs our graduates, and academics need to understand our industry's needs.

[13] We would like to point out that while this section is relatively shorter than other sections, the potential broader impacts print of *bias* and *trust* with regard to AI in civil engineers is large and requires a series of dedicated investigations.

[14] The same can also be true for FE models. Admittedly, such models may be much larger in size and may require specialized software to operate as opposed to lines of codes of an AI model.

Acknowledgment

We would like to thank the Editor, Prof. Paul Jowitt, and anonymous reviewers for their kind look and constructive comments and for lending us the platform of the Journal of Civil Engineering and Environmental Systems to showcase this opinion piece.

Data Availability

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of Interest

The authors declare no conflict of interest.

References

Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*. https://doi.org/10.1109/ACCESS.2018.2870052

Arp, D., Quiring, E., Pendlebury, F., Warnecke, A., Pierazzi, F., Wressnegger, C., Cavallaro, L., & Rieck, K. (2021). *Dos and Don'ts of Machine Learning in Computer Security*.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*. https://doi.org/10.1016/j.inffus.2019.12.012

Baylor, D., Breck, E., Cheng, H. T., Fiedel, N., Foo, C. Y., Haque, Z., Haykal, S., Ispir, M., Jain, V., Koc, L., Koo, C. Y., Lew, L., Mewald, C., Modi, A. N., Polyzotis, N., Ramesh, S., Roy, S., Whang, S. E., Wicke, M., … Zinkevich, M. (2017). TFX: A TensorFlow-based production-scale machine learning platform. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/3097983.3098021

Behnood, A., & Golafshani, E. M. (2020). Machine learning study of the mechanical properties of concretes containing waste foundry sand. *Construction and Building Materials*. https://doi.org/10.1016/j.conbuildmat.2020.118152

Belkin, M., Hsu, D., Ma, S., & Mandal, S. (2019). Reconciling modern machine-learning practice and the classical bias–variance trade-off. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1903070116

Botchkarev, A. (2019). A new typology design of performance metrics to measure errors in machine learning regression algorithms. *Interdisciplinary Journal of Information, Knowledge, and Management*. https://doi.org/10.28945/4184

Bzdok, D., Altman, N., & Krzywinski, M. (2018). Points of Significance: Statistics versus machine learning. In *Nature Methods*. https://doi.org/10.1038/nmeth.4642

Carpenter, W. C., & Barthelemy, J. (1994). Common Misconceptions about Neural Networks as Approximators. *Journal of Computing in Civil Engineering*. https://doi.org/10.1061/(asce)0887-3801(1994)8:3(345)

Corbett-Davies, S., & Goel, S. (2018). The measure and mismeasure of fairness: A critical review of fair machine learning. In *arXiv*.

Degtyarev, V. (2020). *Elastic buckling and post-buckling strength of CFS decks in bending: FE simulation results*. *2*. https://doi.org/10.17632/3KM4B3SFCW.2

Hildebrand, C. D. W. (1938). Ockham, Studies and Selections By Stephen Chak Tornay. LaSalle, Ill.: Open Court Publishing Company, 1938. viii, 207 pages. $1.75. *Church History*. https://doi.org/10.2307/3160457

Jonnalagadda, S., Ross, B. E., & Khademi, A. (2016). A modelling approach for evaluating the effects of design variables on bridge condition ratings. *Journal of Structural Integrity and Maintenance*, *1*(4), 167–176. https://doi.org/10.1080/24705314.2016.1240523

Lones, M. A. (n.d.). *How to avoid machine learning pitfalls: a guide for academic researchers*. Retrieved December 10, 2021, from http://www.macs.hw.ac.uk/

Naser, M., & Alavi, A. (2020). Insights into Performance Fitness and Error Metrics for Machine Learning. *Under Review*.

Naser, M. Z. (2021). An engineer's guide to eXplainable Artificial Intelligence and Interpretable Machine Learning: Navigating causality, forced goodness, and the false perception of inference. *Automation in Construction*, *129*, 103821. https://doi.org/10.1016/J.AUTCON.2021.103821

Naser, M. Z., Amir, ·, & Alavi, H. (2021). Error Metrics and Performance Fitness Indicators for Artificial Intelligence and Machine Learning in Engineering and Sciences. *Architecture, Structures and Construction 2021*, *1*, 1–19. https://doi.org/10.1007/S44150-021-00015-8

Naser, M. Z., Kodur, V., Thai, H.-T., Hawileh, R., Abdalla, J., & Degtyarev, V. V. (2021). StructuresNet and FireNet: Benchmarking databases and machine learning algorithms in structural and fire engineering domains. *Journal of Building Engineering*, 102977. https://doi.org/10.1016/J.JOBE.2021.102977

Papadimitriou, P., & Garcia-Molina, H. (2011). Data leakage detection. *IEEE Transactions on Knowledge and Data Engineering*. https://doi.org/10.1109/TKDE.2010.100

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*. https://doi.org/10.1214/09-SS057

Rahmanpanah, H., Mouloodi, S., Burvill, C., Gohari, S., & Davies, H. M. S. (2020). Prediction of load-displacement curve in a complex structure using artificial neural networks: A study on a long bone. *International Journal of Engineering Science*. https://doi.org/10.1016/j.ijengsci.2020.103319

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*. https://doi.org/10.1038/s42256-019-0048-x

Sun, H., Burton, H. V., & Huang, H. (2021). Machine learning applications for building structural design and performance assessment: State-of-the-art review. *Journal of Building Engineering*, *33*, 101816. https://doi.org/10.1016/j.jobe.2020.101816

Tahmassebi, A., Motamedi, M., Alavi, A. H., & Gandomi, A. H. (2021). An explainable prediction framework for engineering problems: case studies in reinforced concrete members modeling. *Engineering Computations*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/EC-02-2021-0096

Tarawneh, A., Momani, Y., & Alawadi, R. (2021). Leveraging artificial intelligence for more accurate and reliable predictions of anchors shear breakout capacity in thin concrete members. *Structures*. https://doi.org/10.1016/j.istruc.2021.03.074

Tran-Ngoc, H., Khatir, S., Le-Xuan, T., De Roeck, G., Bui-Tien, T., & Abdel Wahab, M. (2020). A novel machine-learning based on the global search techniques using vectorized data for damage detection in structures. *International Journal of Engineering Science*. https://doi.org/10.1016/j.ijengsci.2020.103376

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*. https://doi.org/10.1109/4235.585893