Constructability of districts: capabilities of productivity and logistics big data for machine learning prediction

Dimosthenis Kifokeris[1, *], Christian Koch[1] and Yiannis Xenidis[2]
Chalmers University of Technology
Aristotle University of Thessaloniki
e-mail: dimkif@chalmers.se

Abstract

Big data, reflecting both qualitative information and quantitative material, can be used within the construction management processes of complex and large-scale building activities, such as the development of whole districts in urban areas. Such big data is probably largely focused on transport routes, productivity and site logistics portfolios. However, despite the capabilities offered by construction informatics, such data has scarcely been utilized systematically and in its full capacity for descriptive and predictive purposes. Such a systematic data utilization process can be framed through the lens of the novel construction management concept of district constructability, namely the extension of constructability into the collective level of entire districts. Constructability is here understood as the optimal use of construction knowledge and experience in planning, design, procurement, and field operations, to achieve the project objectives of time, cost and quality, and omit the gap between the as-designed and as-built project states. District constructability moves from individual projects to an overall metric for the facilitation of construction knowledge and experience implementation when undertaking large-scale construction activities (e.g. the erection of numerous buildings) for the development of entire districts; thus, it can be realized, among others, through the achievement of optimal construction productivity rates and smooth logistics operations. To combine all the aforementioned, and simultaneously fully and meaningfully exploit the capabilities that construction productivity and logistics big data may present for the assessment of district constructability, data mining can be utilized, namely the set of processes that computationally discover and "comprehend" patterns in datasets. More particularly, machine learning, here defined as the exploration of algorithms that enable computing systems to "learn" and make data-driven predictions by building a model from a sample dataset and without being explicitly programmed, can be at the methodological forefront of fully exploiting all data found in transport routes, buffer facilities, productivity rates and logistics portfolios. In this paper, the capabilities of the information structures found in the data for developing machine learning models predicting the district constructability in new large-scale urbanization activities, are examined.

Keywords: District constructability, productivity, logistics, big data, machine learning.

1. Introduction

Construction management is the research and application field that, apart from deeper systematic understanding, aim at providing the methodologies and tools implemented for the management of construction projects from their initiation until their delivery, so that their objectives of time, cost and quality are optimized (Knutson et al., 2008). As with individual projects, construction management as activity is integral in large-scale construction activities, as they are set out in long-scope urban development strategy plans for whole districts, or even towns, cities, and their metropolitan areas (see, for example, the plan of the Planning and Building Committee, 2014, for the town of Gothenburg, Sweden); such activities can include the erection of buildings (e.g. residential, offices etc.), as well as stand-alone or supporting infrastructural projects (Göteborgs Stad, 2019). Key to successful construction management, also when building districts, is the collection, understanding, and processing of relevant big data (Bilal et al., 2016; Chen & Lu, 2018); such big data can include quantitative and qualitative productivity-related indicators, such as productivity rates (Kitchin, 2014), as well as elements related to construction logistics and supply chains portfolios (Yigitcanlar et al., 2008). Within construction informatics – namely, the interdisciplinary applied field related to construction, information systems and computer science and studying the issues related to the design, processing, representation, implementation, communication and use of construction-specific information in humans and software (Turk, 2006) – methodologies and tools are explored for such meaningful utilization of big data for construction management (Turk, 2007), including data mining and machine learning (ML) (Turk, 2007; Bilal et al., 2016; Chen & Lu 2018).

Particularly, data mining is the set of processes used to discover and comprehend patterns in datasets (Bilal et al., 2016; Tan et al., 2018). ML is used for state-of-the-art data mining (Bilal et al., 2016; Witten et al., 2017), and is generally defined as the exploration of algorithms that enable computing systems to “learn”, i.e. develop new algorithms linking data, and make data-driven predictions by building models from sample datasets, without being explicitly programmed (Witten et al., 2017); complementarily, it can be said that ML systems are computer systems that automatically improve through experience (Jordan & Mitchell, 2015; see also Sarkar et al., 2013; and Portugal et al., 2018). ML is frequently classified in three types: supervised, unsupervised, and hybrid. Supervised ML utilizes algorithms that are trained and validated using labeled datasets, in a context where it is assumed that the reasoning of the application domain is known (Tan et al., 2018). The task of the respective algorithm is to learn the way they should act based on real training data, validate such gained knowledge, and then apply it on new instances for predictive purposes (Portugal et al., 2018). Exemplary algorithms used in supervised ML include decision trees, decision forests, logistic regression, support vector machines, kernel methods, and Bayesian classifiers (Portugal et al., 2018). Unsupervised ML deals with unlabeled datasets having hidden patterns (Witten et al., 2017), and can be understood as “the analysis of unlabeled data under assumptions about structural properties of the data” (Jordan & Mitchell, 2015). In unsupervised ML, algorithms do not operate on a training set; the respective systems are rather presented with some data about a domain and have to develop relational models from that data “on their own”, by running internal procedures (Portugal et al., 2018). Exemplary algorithms used in ML include vector quantization clustering, and generative adversarial networks (Hastie et al., 2009). Hybrid ML mixes approaches, including semi-supervised and reinforcement learning (Jordan & Mitchell 2015; Portugal et al., 2018); it is increasingly preferred in current research efforts (Portugal et al., 2018; Amasyali & El-Gohary, 2019). Even more recently, deep learning, building on the foundations of ML, has received much attention; its systems utilize gradient-based optimization algorithms to adjust parameters throughout multilayered networks, based on errors at their outputs (Jordan & Mitchell, 2015).

As noted earlier for the case of large-scale construction activities, the big data to be used within construction management can be both qualitative (e.g. lessons-learned databases), and quantitative (e.g. cost and time overheads) (Chen & Lu, 2018; Yeung et al., 2018). However, and despite the capabilities offered by construction informatics, it has scarcely been utilized systematically and in its full capacity for descriptive and predictive purposes (Bilal et al., 2016); it is mostly used either in a simple informatory manner (Bilal et al., 2016), or for more narrow applications (Bilal et al. 2016; Tixier et al., 2016). Empirical knowledge is still the main driver of state-of-art construction management, even when aided by cutting-edge methodologies and tools utilized within construction informatics, and especially

ML and its aforementioned variants (Bilal et al. 2016; Tixier et al., 2016). Construction managers still prefer to mainly use hands-on experience; but while tacit experience is essential, a more holistic data utilization could enhance the managers' decision-making and action-taking (Kumar & Reinartz, 2018).

Such a ML-aided holistic data utilization for construction management, can prove essential in the messy environment of large-scale construction activities, like the development of entire districts. This can be even more crucial in cases of rapid urbanization; such an intensive activity and its associated complex processes, especially within densely populated areas, may result in several construction management issues, like productivity- and logistics-related ones (e.g. delayed deliveries, complicated supply chain coordination, and low on-site productivity) (Dubois et al., 2017). But to reach such a meaningful and systematic utilization with practical and useful results, the relative contextual framework must be devised, the associated types and orders of datasets must be identified, and the suitable technical aspects (e.g. the ML algorithms) of the model leading to the realization of the framework must be investigated, tested and verified. The aim of this paper is to devise a conceptual framework, for the exploitation of big data to build ML models acting as decision-making and action-taking helpers for construction managers operating within large-scale urban activities (culminated in the case of whole district construction). In the second section of the paper, the introduction of the concept of district constructability will act as the contextualization of the framework. In the third section, forms of big data generated in district development, and especially the ones related to productivity and construction logistics, will be investigated in terms of their capabilities and suitability for use within ML models appraising district constructability. In the fourth section, early considerations for the realization of the conceptual framework will be showcased. Following will be the conclusions and recommendations for future work.

2. Context: district constructability

Constructability is "the optimum use of construction knowledge and experience in planning, design, procurement, and field operations to achieve overall project objectives" (Construction Industry Institute, 1986). It is a crucial aspect of optimal construction management, and it encompasses buildability ("the extent to which the design of a building facilitates ease of construction, subject to the overall requirements for the completed building" (Construction Industry Research and Information Association, 1983) as its design- and early construction-related aspect. Constructability, also including early contractor involvement, is implemented through the whole initiation, execution, and delivery project lifecycle phases to optimize the project's performance objectives of time, cost, and quality (Construction Industry Institute, 1986), as well as client satisfaction (Poon et al., 1999). Such an implementation is achieved with constructability programs, namely "the application of a disciplined, systematic optimization of construction-related aspects of a project during the planning, design, procurement, construction, test, and start-up phases by knowledgeable, experienced construction personnel who are part of a project team" (Construction Management Committee of the American Association of Civil Engineers, 1991). For the realization of constructability programs, several methodological and application frameworks have been integrated with constructability, such as – indicatively – planning and operations performance evaluation, hybrid value engineering, knowledge management, cost/benefit analysis, total quality management, object-oriented analysis, total building performance, regression analysis (Kifokeris & Xenidis, 2017), and technical project risk analysis (Kifokeris & Xenidis, 2019). Furthermore, numerous related cognitive, mathematical, programming, and software methodologies and tools have been developed to appraise and/or assess constructability in terms of quantitative and qualitative project features' assessment, schedule-cost-quality management and decision-making, program review, information feedback, and knowledge management and dissemination (Kifokeris & Xenidis, 2017), including, among others, diverse ML models (Skibniewski et al., 1997; Ugwu et al., 2005; Le et al., 2018; Kifokeris & Xenidis, 2019).

Among others, important constructability aspects are a holistic view on logistics (including, but not limited to, supply chain integration, on-site resources flow management, and close cooperation of the related actors), and the optimization of the productivity of the whole project lifecycle, and especially during on-site operations (Kifokeris & Xenidis, 2017). Even for large-scale construction activities, such as the development of entire urban districts, constructability of individual projects (e.g. high-rise buildings) can be realized, among others, through the achievement of optimal construction productivity rates and smooth logistics operations (Kifokeris, 2018). In conjunction with that, the overall performance of construction activities in the district level can be contextualized accordingly and appraised in terms of optimized productivity and smooth logistics operations – as reflected in the relative big data generated in each case – which centrally include quantitative and qualitative productivity-related indicators (such as productivity rates (Kitchin, 2014)) and elements related to construction logistics and supply chains portfolios (Yigitcanlar et al., 2008). Considering the two aforementioned points and by exploiting (a) the direct connection of constructability to the overall project objectives rather than narrow applications, (b) its affiliation with construction knowledge and experience implementation, and (c) the capabilities of construction informatics (and especially ML) in extracting and processing productivity- and logistics-related data, a novel predicting ML system aiming at holistically enhancing the decision-making, action-taking and knowledge communication of construction managers affiliated with the urban development of entire districts, can be formulated.

To capitalize on the points made above and create a contextualization for the previously mentioned predicting system, we hereby propose the concept of district constructability. District constructability extends constructability from individual projects to an overall, collective metric for the facilitation of construction knowledge and experience implementation when undertaking large-scale construction activities (e.g. the erection of numerous buildings) for the development of entire districts, thus acting as a qualitative performance indicator for urban development. Central factors in the appraisal of district constructability are qualitative and quantitative indices and metrics connected to on-site construction productivity and construction logistics operations on the district level. Therefore, in the abovementioned potential predicting system, district constructability can provide the context of its conceptualization and realization, since in the core of this system there can be a model for the prediction of the way construction productivity rates and logistics and supply chain issues in the district level can affect the associated district constructability.

3. Big data for district constructability appraisal

Following the contextualization of the previous section, forms of big data generated in district development, and especially the ones related to productivity and construction logistics, will be investigated in terms of their capabilities and suitability for use within ML models appraising district constructability. As the basis for this investigation, the data found in the productivity report "Produktivitetsläget i svenskt byggande 2014" [Productivity status in Swedish construction 2014] (Koch & Lundholm, 2018) – based on the work by Josephson (2013) – was used. The aforementioned report adopts the metrics of cost (SEK, Swedish crowns) and work hours per square meter of total gross area, for the measurement of productivity for different building types. In addition, logistics problems are identified and qualitatively assessed on a five-point Likert scale. It is assumed that project output depends on the relevant conditions (such as the performance of the project organization) as input, and then the production process takes place and "causes" costs and working time (namely, productivity), as well as logistics issues, as the output. Significant stakeholders are identified for each project, such as the clients, the contractors, and the suppliers.

The data in "Produktivitetsläget i svenskt byggande 2014" was collected through telephone interviews supported by questionnaires. In this way, answers are based on the respondent's own perspective (Koch & Lundholm, 2018). Construction projects encompassed in the survey are primarily premises. These include daycare centres, schools, office buildings, administrative buildings, sports and recreation facilities, hospitals and elderly care centres, church buildings, nursing homes, stores, industrial properties, and group-built family houses. In each individual project, the client's project manager and the contractor's site manager answered the relevant questionnaires, during the period of October to November 2014. The questions covered project aspects, such as the relevant processes, organization, costs, time, work progress, and team performance, and their number was limited. In the respective questionnaires, the clients received 23 questions and the site managers 21. The surveys were sent to 1000 individuals, with 580 valid responses (58% overall answering rate). The survey was answered by 324 contractor representatives (72% answering rate), and 256 clients (62% answering rate).

The relevant statistics are interesting in revealing the most central existing issues regarding productivity, construction logistics and supply chain management in the sites investigated; the issues of on-site congestion, transportation challenges and storage bottlenecks, are experienced by site managers at around 40% of the studied projects (Koch & Lundholm, 2018). Congestion is thus recurrent to an extent, yet an exception compared to 60% of the sites not reporting it. Some of the districts corresponding to site managers reporting congestion were situated in Stockholm and Malmö; however, and surprisingly enough, congestion was also evident in much less populated towns. In total, the central productivity and logistics issues concerned ten districts ("kvartärer") or and four single or multi-project areas having similar development needs (e.g. brownfields).

In the district development cases, the productivity rates and the logistics and supply chain issues were calculated and appraised, respectively, in the ways mentioned above. Logistics and supply chain issues primarily pertained to the good cooperation of the project group regarding on-site supply chain tasks, disturbances in the relative flows, on-site congestion, keeping of the delivery timetable, difficulties in material and equipment transportation and storage due to on-site narrow spaces, limitations in the construction production and logistics preparation, the extent of available staffing for the construction works (recognized as a risk source for constructability in Kifokeris & Xenidis, 2018, and therefore potentially extensible to the case of district constructability), and the informed selection of the material and equipment suppliers (Koch & Lundholm, 2018).

The data found in the above also needs to be further complemented with more data found in logistics portfolios and productivity studies of construction projects, and can then be utilized as a basis for a translation into independent variables that suit predictive ML systems, e.g. ones using support vector machines and/or support vector regression for classification or regression, respectively, through supervised ML. The productivity rates can be translated into continuous numerical variables, having as benchmarks median productivity rate values in relation to e.g. the size, number and type of the individual projects constructed during the whole district development activities. In addition, the logistics and supply chain management issues appraised through the Likert scales can be translated into multinomial categorical variables, or processed into binomial variables 1 or 0, for "yes" or "no" for binary classification problems.

To properly train the respective ML models, district constructability should also be translated into the dependent variable. Evaluation of district constructability is at present mostly built on experiential knowledge supplemented by some modeling, according to interviews conducted by the authors. However, in the relevant literature, there has not been yet, to the best of the authors' knowledge, a meaningful representation of constructability as a continuous variable – apart from some early conceptual attempts like the one of Yu & Skibniewski, 1999. Therefore, it may be sensible to base the representation of district constructability on efforts treating constructability itself as a discrete variable (e.g. on the binomial constructability variable in Kifokeris & Xenidis, 2019). As district constructability is, both in itself and as an object of machine learning modelling, a hereby introduced new concept, it may be difficult to properly define a continuous domain for its numerical values, along with all the associated thresholds and benchmarks. Therefore, a multinomial representation (e.g. via a three- or five-point Likert scale, or via whole-number percentages) of district constructability could be more informative and in line with the current related research trends.

To sum up these insights, in Table 1 (see next page) there is an exposition of the exemplary independent variables that can be potentially derived from big data found in studies on productivity, logistics, and supply chain management, as well as the translation of district constructability into a dependent variable. Such a list of variables can be enriched with yet more relative elements generated within similar research and practical studies, like, for example, production flow inventories, descriptions of construction site spatial and schedule clashes, number of reworks, material quantity problems, optimal vehicle rounds, district disturbances, and the existence and proper function of buffer facilities for vehicles and goods.

Table 1: Exemplary independent and dependent variables for district constructability appraisal, as derived from Koch & Lundholm (2018)

Independent variables	**Type**	**Example of value**
Productivity rate	Continuous	0,1 (%)
Level of project group cooperation for on-site supply chain tasks	Discrete multinomial	{1,…,5}
Flow disturbance	Discrete multinomial	{1,…,5}
On-site congestion	Discrete multinomial	{1,…,5}
Keeping of delivery timetable	Discrete multinomial	{1,…,5}
Difficulties in material and equipment transportation and storage	Discrete multinomial	{1,…,5}
Limited construction production and logistics processes preparation	Discrete multinomial	{1,…,5}
Enough workforce for optimal undertaking of construction tasks	Discrete multinomial	{1,…,5}
Informed selection of material and equipment suppliers	Discrete multinomial	{1,…,5}
Dependent variable	**Type**	**Example of value**
District constructability	Discrete multinomial	{1,…,5}

In addition, indicators and metrics of the infrastructure in and around the district (e.g. access routes, points of entry, traffic diversion and/or emergency roads) should be considered, for a more holistic representation of real situations in district development.

4. Modelling aspects of the conceptual framework

After the relative contextualization with the introduction of district constructability, and the investigation of the capabilities of forms of productivity- and logistics-related big data generated in district development for use within a ML model appraising district constructability, the early conceptual steps and considerations for the actual formulation of such a model are given in Figure 1 (see next page).

What is showcased in Figure 1 can be furtherly explained in the following:

Step 1. Data collection. For a large number of building projects that are part of the ongoing entire district development, the data suppliers will provide quantitative and qualititative construction productivity and logistics data (e.g. site productivity rates, production flow inventories, descriptions of construction site spatial and schedule clashes, material quantity problems, material and equipment transport routes and bottlenecks, indicators about buffer facilities). In addition, they will supply a qualitative district constructability labelling of the respective districts (e.g. using a five-point Likert scale). The interpretation of the qualitative labels as levels of district constructability achievement, can be obtained through interviews along with the data providers.

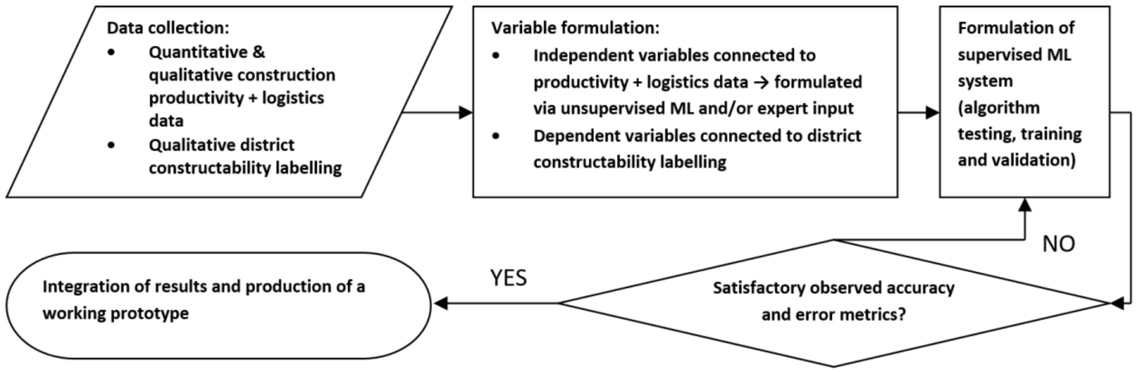

Figure 1: Explanatory simplified flowchart describing the conceptual framework

Step 2. Variable formulation. Independent variables: Depending on the form of the construction productivity and logistics data, meaningful independent variables (e.g. "Number of reworks") measured through the values of the collected data, will be produced either through unsupervised ML techniques (e.g. vector quantization, linguistic clustering), or qualitative techniques relying on expert input (e.g. brainstorming). Dependent variables: These will be the district constructability achievement levels, and can act, for example, as multinomial discrete classification variables.

Step 3. System formulation. A ML system trained and validated with the collected data, in the way it is expressed through the independent and dependent variables previously defined, will be formulated. A possibility is to choose a multinomial classification supervised ML scheme to be trained and validated, as it can be derived from the data form and amount, and the variables' type and number. This choice can be specified as a result of multiple experiments conducted within a suitable platform (like the Waikato Environment for Knowledge Analysis – WEKA) (Witten et al., 2017), with numerous algorithms, such as variations of the support vector machines and the random forest algorithms. Given the pre-study of the previous section on the possible representation of the exemplary variables, such a multinomial classification supervised ML scheme can operate with algorithms like naive Bayes classifiers, decision trees, random forests, k-nearest neighbors, support vector machines (SVM), and types of artificial neural networks, as they are considered suitable for such classification problems (Witten et al., 2017).

Error and observed accuracy metrics can be used during each training and validation iteration of the respective algorithms, to determine the actual correctness of the algorithmic results. Among such error metrics are included the Cohen's kappa, mean absolute error, true positive rate, false positive rate, precision, recall, F-measure (Witten et al., 2017), and the Matthews correlation coefficient (Chicco, 2017). When the observed accuracy rates after the training and validation of the respective algorithm are sufficiently high and/or satisfactory, and at the same time the aforementioned metrics are within thresholds characterized as good or optimal (Witten et al., 2017), there can generally be little to no room for further optimization for the respective modelling with the current dataset t has been presented with (Witten et al., 2017).

Auxiliary mathematical, methodological and software tools may be utilized to various extents within Steps 1-3, such as (a) non-negative matrix factorization for data normalization and pre-processing (Steps 1-2), (b) multi-input Analytical Hierarchy Process (AHP), for variable labelling (Step 2), (b) the "kernel trick", to aid in the non-linear function of certain supervised ML algorithms, if elected (Step 3), (d) n-fold cross-validation, for the simultaneous training and validation of certain supervised ML algorithms, if elected (Step 3), (e) the WEKA platform (Step 3), (f) Surprise Scikit (Steps 2-3), and (g) the programming language Python (Steps 2-3).

Step 4. Integration of results and production of a working prototype. The ML system can be integrated as a working prototype within construction management plans in the district level, for the verification of its predicting results – namely, the appraisal of the level of district constructability during the development of new districts, given the values of the productivity- and logistics-related metrics utilized as independent variables, as those values are generated in the course of the district development. This may take place through suitable programming routines and/or graphical user interfaces (such as PyQt, featured in the Anaconda platform).

Such a novel methodological framework and subsequent modelling can not only break ground with the proposition and appraisal of a new concept connected to urban development in the district level, but can furtherly strengthen the placement of ML within construction informatics, for the benefit of construction managers and related disciplines.

5. Conclusions

The urban development of entire districts represents a wide array of interconnected construction activities through multiple individual building and infrastructural projects; it can also generate big data primarily culminated in metrics such as productivity rates and identified issues related to on-site logistics and supply chain management. Apart from their individual exploitation for informatory reasons, these points of data can also be used as a means to measure yet more holistic metrics, which can provide construction managers (and other key stakeholders) with a higher-level overview of the district development process, thus helping in more well-founded decision-making and action-taking.

Models utilizing ML algorithms and miscellaneous methodological, mathematical and programming tools, can provide the framework for such a meaningful understanding, processing and exploitation of the aforementioned big data for descriptive and predictive purposes, as long as there are suitable concepts to contextualize this implementation in the construction and district development sector. Such a concept can be found in the hereby introduced notion of district constructability, which is an extension of the construction management concept of constructability from the level of individual projects to an overall, collective metric for the facilitation of construction knowledge and experience implementation when undertaking large-scale construction activities (e.g. the erection of numerous buildings) for the development of entire districts; thus, district constructability can acting as a qualitative performance indicator for urban development.

Qualitative and quantitative indices and metrics connected to on-site construction productivity and construction logistics operations on the district level can be used for the appraisal of district constructability. So, for the culmination of a relative ML model, independent variables can be generated from such metrics, and then computationally correlated with a dependent variable representing district constructability. Such independent variables can include, among others, the level of project group cooperation regarding on-site supply chain tasks, flow disturbances, on-site congestion, keeping of the delivery timetable, material and equipment transportation and storage difficulties, limitations in the construction production and logistics preparation, construction staff availability, material and equipment suppliers selection, optimal vehicle rounds, the function of buffer facilities, the level of disturbance due to the on-site construction and transportation activities, production flow inventories, descriptions of on-site spatial and schedule clashes, number of reworks, and material quantity problems.

A recommendation for further research is the actual realization of Steps 1 and 2 of the presented framework through access in the relative databases given by the interested stakeholders. Concurrently, the study of and experimentation with ML algorithms on a suitable platform should be undertaken, to prepare for the realization of Step 3. ML is a field that can present a wealth of opportunities for the development of solutions within the construction sector and construction management specifically, and even more in high/profile activities such as the development of entire urban districts.

References

Amasyali, K., & El-Gohary, N. (2019). Predicting Energy Consumption of Office Buildings: A Hybrid Machine Learning-Based Approach. In: Mutis, I., & Hartmann, T. (Eds.). Advances in Informatics and Computing in Civil and Construction Engineering: Proceedings of the 35th CIB W78 2018 Conference: IT in Design, Construction, and Management. (pp. 695-700). Cham: Springer Nature.

Bilal, M., Oyedele, L. O., Qadir, J., Munir, K., Ajayi, S. O., Akinade, O. O., Owolabi, H. A., Alaka, H. A., & Pasha, M. (2016). Big Data in the construction industry: A review of present status, opportunities, and future trends. Advanced Engineering Informatics, 30(3), 500-521.

Chen, X., & Lu, W. S. (2018). Scenarios for Applying Big Data in Boosting Construction: A Review. In: Chau, K. W., Chan, I. Y. S., Lu, W., & Webster, C. (Eds.). Proceedings of the 21st International Symposium on Advancement of Construction Management and Real Estate (pp. 1299-1306). Singapore: Springer Nature.

Chicco, D. (2017). Ten quick tips for machine learning in computational biology. BioData Mining, 10(35), 1-17.

Construction Industry Institute (CII) (1986). Constructability: A primer. Publication 3-1. Austin: University of Texas at Austin.

Construction Industry Research and Information Association (CIRIA) (1983). Buildability: An assessment. London: CIRIA Publications.

Construction Management Committee of the American Association of Civil Engineers (ASCE) (1991). Constructability and Constructability programs: White Paper. Journal of Construction Engineering and Management, 67–89.

Dubois, A. Hulthén, K, & Sundquist, V. (2017). Organising logistics and transport activities in construction. The International Journal of Logistics Management, 30(2), 620-640.

Göteborgs Stad (2019). Stadsutveckling Göteborg. Retrieved from https://stadsutveckling.goteborg.se/tidslinje/

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data mining, Inference, and Prediction. New York: Springer.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

Josephson, P. E. (2013). Produktivitetsläget i svenskt byggande 2013, nybyggnad flerbostadshus och kontor. Stockholm: Sveriges Byggindustrier.

Kifokeris, D. (2018). Risk-based Constructability Assessment in Civil Engineering Projects. (Doctoral dissertation, Aristotle University of Thessaloniki, Thessaloniki, Greece). Retrieved from http://ikee.lib.auth.gr/record/301230/files/GRI-2018-23059.pdf

Kifokeris, D., & Xenidis, Y. (2017). Constructability: outline of past, present, and future research. Journal of Construction Engineering and Management, 143(8), 04017035-1 – 04017035-13.

Kifokeris, D., & Xenidis, Y. (2018). Application of linguistic clustering to define sources of risk in technical projects. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 4(1), 04017031 – 1-04017031-13.

Kifokeris, D., & Xenidis, Y. (2019). Risk source-based constructability appraisal using supervised machine learning. Automation in Construction, 104, 341-359.

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. GeoJournal, 79(1).

Knutson, K., Schexnayder, C. J., Fiori, C. M., & Mayo, R. (2008). Construction Management Fundamentals (2nd ed). New York: McGraw-Hill Education.

Koch, C., & Lundholm, M. (2018). Produktivitetsläget i svenskt byggande 2014. Lokaler, Grupphus och Anläggning. Stockholm: Sveriges Byggindustrier.

Kumar, V., & Reinartz, W. (2018). Data Mining. In: Customer Relationship Management. Berlin, Heidelberg: Springer Texts in Business and Economics.

Le, T., Le, C., Jeong, H. D., Gilbert, S. B., & Chukharev-Hudilainen, E. (2018). Requirement Text Detection from Contract Packages to Support Project Definition Determination. In: Advances in Informatics and Computing in Civil and Construction Engineering: Proceedings of the 35th CIB W78 2018 Conference: IT in Design, Construction, and Management. (p. 569-576). Cham: Springer Nature.

Planning and Building Committee (2014). Development Strategy Göteborg 2035. Gothenburg: Göteborgs Stad.

Poon, J., Potts, K., & Cooper, P. (1999). Development of a new best practice model for building projects. In: Hughes, W. (Ed.). Proc., 15th Annual ARCOM Conference (pp. 705-714). Liverpool: Association of Researchers in Construction Management Publications.

Portugal, I., Alencar, P., & Cowan, D. (2018). The Use of Machine Learning Algorithms in Recommender Systems: A Systematic Review. Expert Systems With Applications, 97(1), 205-227.

Sarkar, D., Bali R., & Sharma, T. (2013). Practical Machine Learning with Python. New York: Apress.

Skibniewski, M., Arciszewski, T., & Lueprasert, K. (1997). Constructability analysis: machine learning approach. Journal of Computing in Civil Engineering, 11(1), 8-16.

Tan, P. N., Steinbach, M., Karpatne, A., & Kumar, V. (2018). Introduction to data mining (2nd ed.). Boston: Pearson.

Tixier, A. J. P., Hallowell, M. R., Rajagopalan, B., & Bowman, D. (2016). Application of machine learning to construction injury prediction. Automation in Construction, 69, 102-114.

Turk, Ž. (2006). Construction informatics: Definition and ontology. Advanced Engineering Informatics, 20(2), 187-199.

Turk, Ž. (2007). Construction Informatics in European Research: Topics and Agendas. Journal of Computing in Civil Engineering, 21(3), 211-219.

Ugwu, O. O., Anumba, C. J., & Thorpe, A. (2005). Ontological foundations for agent support in constructability assessment of steel structures – a case study. Automation in Construction, 14(1), 99-114.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). Data mining: practical machine learning tools and techniques (4th ed). Cambridge, Massachusetts: Morgan Kaufmann, Elsevier.

Yeung, C. L., Wang, W. M., Cheung, C. F., Tsui, E., Setchi, R., & Lee, R. W. B. (2018). Computational narrative mapping for the acquisition and representation of lessons learned knowledge. Engineering Applications of Artificial Intelligence, 71, 190-209. https://doi.org/10.1016/j.engappai.2018.02.011

Yigitcanlar, T., O'Connor, K., & Westerman, C. (2008). The making of knowledge cities: Melbourne's knowledge-based urban development experience. Cities, 25(2), 63-72.

Yu, W., & Skibniewski, M. J. (1999). Quantitative constructability analysis with a neuro-fuzzy knowledge-based multi-criterion decision support system. Automation in Construction, 8(5), 553-565.