Using Artificial Intelligence to Automate the Quantity Takeoff Process

Ebrahim Karan, epk008@shsu.edu
Associate Professor, Department of Engineering Technology, Sam Houston State University, Huntsville, Texas, USA

Vehideh Karimi Mansoob, v.karimi@tamu.edu
Graduate Student, Department of Construction Science, Texas A&M University, College Station, Texas, USA

Ali Khodabandelu, khodaban@unlv.nevada.edu
Graduate Student, Department of Civil and Environmental Engineering and Construction of Science, University of Nevada, Las Vegas, USA

Sadegh Asgari, asgaris@merrimack.edu
Assistant Professor, Department of Civil Engineering, Merrimack College, North Andover, Massachusetts, USA

Atefeh Mohammadpour, mohammaa@pfw.edu
Associate Professor, School of Polytechnic, Purdue University Fort Wayne, Indiana, USA

Somayeh Asadi, sxa51@psu.edu
Associate Professor, Department of Architectural Engineering, Pennsylvania State University, State College, Pennsylvania, USA

Abstract
Quantity takeoff is the process of obtaining quantity measurements from construction plans and providing a list of materials needed to complete a project. Despite the recent and wide use of BIM tools, most construction plans and documents are still 2D drawings on electronic files or sheets of paper. Consequently, estimators and contractors need to measure items and then perform calculations from 2D drawings. This is why the quantity takeoff is the most time-consuming activity in creating the cost estimate and prone to error (like many other manual activities). It is fair to say that almost all new construction drawings are available (or created) as digital files and PDF is the most common file format for sharing the project plans and documents. A review of commercially available takeoff software for construction reveals that we are now able to accurately measure the quantities from digital files and even send them to spreadsheet programs, such as Microsoft Excel. These new software tools can greatly speed up the takeoff process and increase efficiencies. But they still need someone with a thorough understanding of architectural and construction terminologies and symbols to assign the measured quantities to the right cost items. Intelligent document processing is the next generation of automation and measuring quantities from a variety of PDF documents. In this study, we use Artificial Intelligence technologies such as natural language processing (NLP) and machine learning (ML) to identify and extract relevant quantities from a set of 2D construction drawings.

Keywords
Quantity Take-off, Cost Estimate, Artificial Intelligence

1 Introduction

The aim of cost estimating is to forecast, approximate, assess or calculate the probable cost of a project computed based on available information (Juszczyk 2017). Indeed, the cost estimating process has largely been a matter of importance to the success and failures of a construction project. The cost of a construction project needs to be estimated within a specified accuracy range, but the largest obstacles standing in front of a cost estimate, particularly in the early stage, are lack of preliminary information and larger uncertainties as a result of engineering solutions (Elmousalami 2020). Consequently, the critical process of cost estimation has become a major challenge for cost estimators. As such, to overcome this lack of detailed information, cost estimation techniques are used to approximate the cost within an acceptable accuracy range (Waty *et al.* 2018). Construction takeoff, as an essential aspect of the cost estimation process, undergoes many changes during the building process. In general, different phases of construction takeoff are highly dependent on the knowledge and experience of estimators and the use of historical company records that compiled cost information from previous completed projects (Famiyeh *et al.* 2017). Furthermore, all cost estimation software requires the use of associated and/or additional databases that contain the unit costs and other industry data needed to prepare an estimate (Lee *et al.* 2014). To determine the cost of each material item, the estimators may use their database of a third-party database (such as RSMeans in the US) or may seek a bid from a material supplier (Pratt 2018). Indeed, take-offs created via the use of computer tools still require human operators and interpreters no differently than take-offs produced manually. These computer tools can reduce the risk of miscalculation and save time and money, however, they do not eliminate the estimating skills and knowledge.

It is fair to say that almost all new construction drawings are available as digital files, and PDF is the most common file format for sharing the project plans and documents. In short, an estimator uses construction drawings as inputs and extracts material quantities. It is expected to have a list of all of the materials necessary to complete a project by taking off the quantities. Unit count, linear length, surface area, volume, and physical weights are the common units of measurement. Items such as doors, windows, or lights are listed by unit count. Linear length is used for materials such as lumber, pipe, and standard shape structural members. The surface area is used for surface and covering materials such as roof and wall finishing. Concrete as well as lumber (e.g. board feet) should be listed by their volume quantities, and the cost of steel and earthwork can also be estimated by weight. These quantities will be linked to cost values and combined with overhead, labour, subcontractor, and equipment costs to provide a final detailed estimate (Schaufelberger and Holm 2017).

Computers have become an inevitable part of cost estimating. Even in a manual quantity takeoff, cost estimators often utilize an excel spreadsheet, word processor, or other computer tools during the process. In contrast to a manual quantity takeoff, digital quantity take-offs are completed with the assistance of digital takeoff software. Taking off quantities has been facilitated by recent advances in computer-aided design (CAD) and building information modelling (BIM) programs (Abanda *et al.* 2017). A gap exists between the use of existing methods and the availability of an automated that measures quantities from a variety of 2D construction drawings. Research in artificial intelligence (AI) areas indicated that the recently developed machine learning and AI algorithms had potential application in developing a quantity takeoff and enhancing cost estimating services (Tijanić *et al.* 2019). The primary objective of this paper is to propose AI techniques for automation and measuring quantities from a variety of 2D construction documents. The ultimate goal is to accurately count and measure takeoff items for estimators lacking decades of field experience in the construction sector. To achieve this goal, this study explains the use of AI technologies such as natural language processing (NLP) and machine learning (ML) to identify and extract relevant quantities from a set of 2D construction drawings.

2 Review of Construction Takeoff Software

Quantity takeoff is a time-consuming, error-prone, repetitive, and sometimes tedious process if conducted manually. AI-powered algorithms have helped key players to tackle some of the challenges including cost and schedule (Mohammadpour *et al.* 2019). To overcome such difficulties, various software tools have been developed in recent years to automate the quantity takeoff to some degree. These tools feature some capabilities that help estimators to take off their desired quantities with further ease and higher accuracy. Takeoff software tools have consistently improved over time and now they offer a variety of cutting-edge features that can significantly raise the accuracy and speed of the quantity takeoff process. This section briefly introduces some of the most common and practical features of takeoff software tools, then compares nine of the most robust and well-known tools, and finally reveals some of the major gaps of the existing tools in achieving full automation. Some of the key features of the most common takeoff software tools are illustrated in Table 1.

Table 1. Key Features of Commercially Available Takeoff Software Tools

Feature	**Brief Explanation**
Object detection and count	Automatically detect a specific object on drawings and count the number of its appearances
1-D, 2-D, and 3-D measurement	Automatically measure the length, area, or volume of a line, 2-D, or 3-D shape specified by the user
Embedded items count	Automatically count the number of items that are inside or assembled to a predefined main item
Predefined configuration addition	Allow users to add and store items with customized configuration, embedded items, and assemblies
Collaborative environment	Allow various users to work on the same project through a cloud-based environment
Multi-layer plans	Allow users to define multiple layers and separately takeoff quantities on each layer
Multi-format import	Allow import of plans and drawings with different format types
Excel-compatible export	Automatically prepare Excel-exportable quantity takeoff reports

Although most of the developed takeoff software tools share the aforementioned features, there are some differences between them in terms of the convenience of the user interface, the level of allowable customization, the span of applicability, the compatibility with different platforms, etc. A comparison between nine of the most well-known takeoff software tools is provided in Table 2. Once a construction drawing is added to these tools, a list of all measurements and dimensions is generated and the estimator can apply prices to each item or object. Although digital quantity take-offs tools are the quickest method available in the market, the method is still time-consuming and relies heavily on the estimator’s experience.

Table 2. Comparison of Common Takeoff Software Tools

Program Name	**Platform**	**User Interface**	**Applicability**	**Striking Features**
STACK	Online on any device	Powerful and modern	General contractors, concrete contractors, home builders, interior finishers, landscape specialists, masons, roofers, and others	Highly collaborative environment, compatibility with other software tools
PlanSwift	Windows operating system	Convenient	General contractors, concrete, drywall, electrical, flooring, framing, decking contractors,	Compatible data import, easy customization

			HVAC, insulation, landscape, masonry, painting, and plumbing	
Bluebeam Revu	On-device installation with real-time collaboration capabilities	Intermediate-level	Almost any specialty from design to bidding	Powerful managerial features
Countfire	Online on any device	Modern and intuitive	electrical contractors and estimators	Robust counting feature, powerful customer service
On Center	Windows and cloud-based versions	Clean and easy to follow	Almost any specialty from biding to building	Highly collaborative environment, powerful customization
PrebuiltML	Windows operating system	Acceptable	concrete, roofing, siding, flooring, masonry, and paint contractors, as well as commercial dealers, framers, and builders	Access to various materials database and cost information
Square	Online on any device	Modern and intuitive	General contractors, plumbers, roofers, and many other specialists	Easy customization
eTakeoff Dimension	Windows operating system	Weak and dated	Almost any specialty and general contractor	Easy customization
Buildee	Online on any device	Modern, simple, and intuitive	Almost any specialty	Fast and direct data import
STACK	Online on any device	Powerful and modern	General contractors, concrete contractors, home builders, interior finishers, landscape specialists, masons, roofers, and others	Highly collaborative environment, compatibility with other software tools
PlanSwift	Windows operating system	Convenient	General contractors, concrete, drywall, electrical, flooring, framing, decking contractors, HVAC, insulation, landscape, masonry, painting, and plumbing	Compatible data import, easy customization
Bluebeam Revu	On-device installation with real-time collaboration capabilities	Intermediate-level	Almost any specialty from design to bidding	Powerful managerial features
Countfire	Online on any device	Modern and intuitive	electrical contractors and estimators	Robust counting feature, powerful customer service
On Center	Windows and cloud-based versions	Clean and easy to follow	Almost any specialty from biding to building	Highly collaborative environment, powerful customization
PrebuiltML	Only on Windows operating system	Acceptable	concrete, roofing, siding, flooring, masonry, and paint contractors, as well as commercial dealers, framers, and builders	Access to various materials database and cost information
Square	Online on any device	Modern and intuitive	General contractors, plumbers, roofers, and many other specialists	Easy customization
eTakeoff Dimension	Windows operating system	Weak and dated	Almost any specialty and general contractor	Easy customization
Buildee	Online on any device	Modern, simple, and intuitive	Almost any specialty	Fast and direct data import

These software tools can greatly speed up the takeoff process and increase efficiencies. Despite the considerable recent improvements in takeoff software tools and their eye-catching capabilities, there is still a long way towards full automation of the quantity takeoff process. The rest of this paper is centred around the use of AI to further raise the automation of the takeoff process.

3 Research Method –AI Techniques for Estimating Costs

Developing an intelligent agent requires capturing the knowledge of experts and professionals in their areas of expertise. This knowledge can be stored in the form of programmable algorithms and rules; such as IF-THEN constructs. In the present study, we model or represent the knowledge of skilled estimators in a way that an AI system can process. As shown in Figure 1, the AI techniques vary greatly as to their purposes and the information about a facility's construction requirements and associated activities. We consider 12 distinct groups of construction projects to find the common ground between AI techniques used in the quantity takeoff and cost estimating process. When an estimator is given a set of construction drawings, the estimator determines what quantities (or items/services) should be considered. Such items and services are commonly listed in the project solicitation. Also, standards like MasterFormat assists cost estimators in organizing information and searching for specific information in consistent locations. The first step for developing an intelligent cost estimator agent is to generate a structured representation of the task domains. The most recent edition of MasterFormat including 35 divisions is used in the study to list and organize the items and services.

Each division can be further divided into some sections. For example, erosion control is a section in the earthwork division. The estimator goes through each section and finds the drawing that contains the relevant information. NLP is an AI technology that helps machines read construction drawings and find drawings that may contain the relevant information for the takeoff process. We use the RSMeans estimating database to list the keywords associated with each section and division. For example, "silt fence" or "erosion control mat" can be listed for estimating the cost of the erosion control section. The artificial agent can also learn from experience without explicit adding of the keywords. In this case, a combination of NLP and machine learning is used to add new keywords to the knowledge database. If an estimator sees a concrete washout area in the erosion and sedimentation control plan, the estimator includes the cost of this washout area in the erosion control section. Similarly, the artificial agent uses NLP to understand the legend in a drawing and uses machine learning to include that in the list of items or services for a given section (if there is a cost item in the database). Otherwise, the intelligent agent learns to ignore the item (e.g. boundary line or limits of construction).

The next step is the quantity takeoff. There are a variety of commercially available software tools suitable for taking off quantities and measuring areas and lengths. The estimator utilizes these tools or CAD programs to extract the size of the concrete slab, number of windows, or length of silt fence shown in the erosion control plan. Image processing is a technology to generate information (e.g. quantities) out of an image (or a digital construction drawing). Image processing that uses machine learning enables us to measure objects in the drawing and automate the PDF analysis process. Machine learning plays a major role here because the artificial agent should learn the style and shape of objects in the legend and then measure the size (or quantity) of objects with a similar style or shape in the drawing. In addition, the scale of the drawing should be used to convert the measurement to correct dimensions. Image processing is used for linear scales (also called graphical scales) and NLP is used for verbal scales or representative fraction scales.

Converting quantities into construction cost estimates requires considerable experience and knowledge of available resources, existing construction methods, and market conditions. In some cases, the method of construction is stated in the project specification (e.g. method of concrete delivery or placement). Even in those cases, the estimator has options to choose for performing the task. For example, an exterior painting specification may state a three-coat system for a job but the estimator can choose between spray painting and brushwork. In such cases, the available resources or the lowest price determine the unit price and/or assemblies. This situation applies to many labour-intensive jobs too, where available skilled workers result in a different unit price. Our analysis of the distinct groups of construction shows that there are a finite number of scenarios that an estimator has the options to choose for the method of construction.

Therefore, it is possible to write a specific set of IF-THEN rules to choose the unit price and/or assemblies. These deterministic expert systems have proven themselves in estimating and bidding for construction projects (Smith 2017). As shown in Figure 1, estimators tend to choose the most likely unit price when provided with some options. The project specifications and availability of resources limit those options. The most likely unit price is not necessarily the median value, but instead, market conditions will dictate where unit prices lie (e.g. lower price in an elastic market).

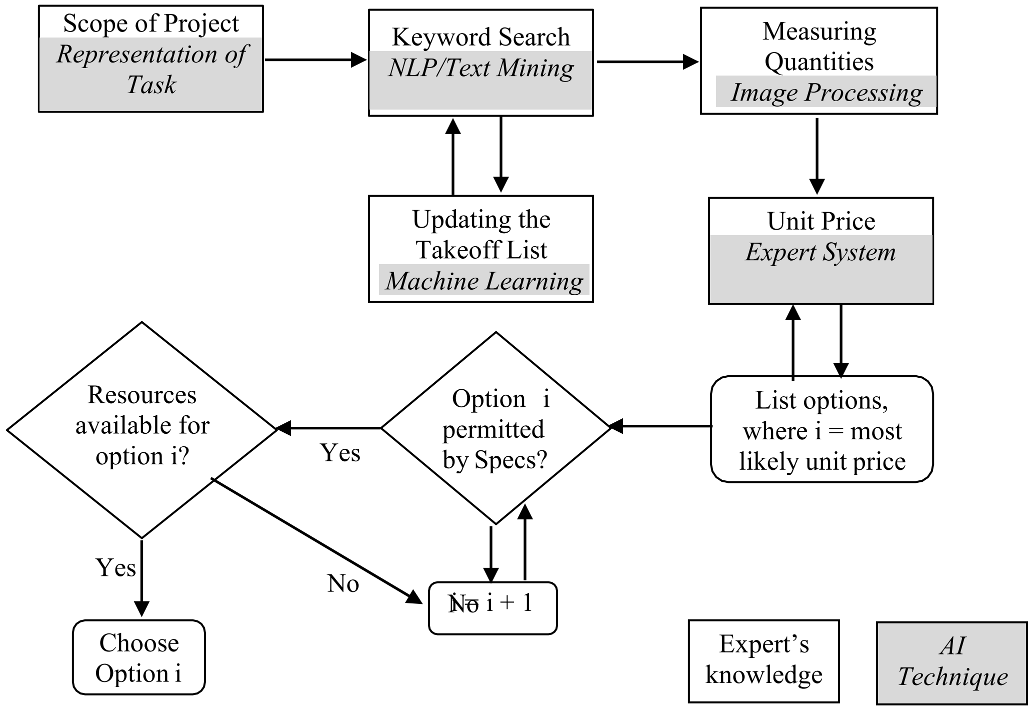

Figure 1. AI techniques used in the takeoff process

The quantity takeoff process enables us to calculate the cost of materials, labours, and equipment and as a result, provide the foundation for calculating the indirect costs. The project cost is calculated by adding project overheads, mark-up cost or profit, contingency, and allowance for construction risks. The AI techniques used in the takeoff process are described in the following subsections using simple examples. Table 3 lists the potential applications of AI techniques in the quantity takeoff and cost estimating process of 12 distinct groups of construction projects.

Table 3. Role of AI Techniques in Automating Construction Cost Estimation

Groups of Construction	NLP/Text Mining	Machine Learning	Image Processing	Expert System
Demolish	High	Moderate	Moderate	Moderate
Concrete	High	High	High	Low
Masonry	High	High	High	High
Metals	High	Moderate	High	Low
Wood	Moderate	High	High	Low
Thermal and Moisture Protection	High	Moderate	High	Low
Openings	High	Moderate	Moderate	Moderate
Finishes	High	High	High	High
Earthwork	High	High	High	High
Plumbing	Moderate	High	High	High
Electrical	Moderate	High	Moderate	High
Mechanical (HVAC)	High	High	Moderate	High

3.1 NLP/Text Mining

Text Mining (TM), also referred to as text data mining is the process of deriving high-quality information from textual data. TM involves automatically extracting information from different written resources to discover new, previously unknown information (Williams and Gong 2014). TM uses a variety of algorithms and methodologies including NLP for structuring and processing the input text. NLP is mainly concerned with how computers can be used to understand, process, and analyse natural language text or speech. Because NLP aims for human-like performance, it is considered an AI discipline (Al Qady et al. 2010). The application of NLP technology has significantly increased in healthcare, risk management, insurance, customer service, and advertising over the past decade.

A 3,000 square foot (around 280 m^2) residential building in Texas is used as the basis for our case study. The construction plan set contains 30 pages of drawings and specifications. To estimate the cost of walls, we train the agent to index the keyword “wall” and generate a tag cloud to find the most frequently used words in the drawings. We extracted more than 156 “wall” keywords in the project drawings were extracted and listed the drawings with at least ten results in Table 4 (e.g. site plan with two results or window and door schedule with two results are not listed). With the aid of a concordance tool, we could identify the preceding and following contexts in the concordance of the keyword “wall”. NLP also is used to clean regular expressions and filter out most of the unwanted texts. Note that only the preceding and following words with high frequency are listed in Table 4.

Table 4. Result of the text analysis for the keyword wall (a case study)

Drawing	Count	Frequent Context
Material & Room Finish Schedule	9	Blocking Wall (2)/ Adjacent Wall (3)
Entry Level Plan	23	Exterior Wall (4)/ Interior Wall (3)/ Wall (6)
Lower Level Plan	22	Exterior Wall (5)/ Interior Wall (4)/ Retaining Wall (2)
Wall Sections	9	Siding Wall (5)/ Retaining Wall (2)
Interior Elevations	9	WD Wall (2)/Wall Sconce (2)/ Wall Paper (4)
Interior Details	11	Wall Framing (2)/ Finish Wall (2)/ Wall Partition (2)

3.2 Image Processing

Digital image processing (DIP), a subfield of digital signal processing, is the application of a computer to process and manipulate digital images using efficient algorithms (Gonzalez and Woods 2018). The NLP technology leads the intelligent agent to focus on the drawing with the most frequent cases (e.g. lower level plan for "wall" in the case study). The most common measurements made by an image processing system are the number of similar objects (count), perimeter (length), and area. Continuing with the wall example, 32 items are obtained from the Lower Level Plan with the pixel as the basic unit. The processed plan (or image) is calibrated to establish the relationship between each pixel and the size of the real object. The digital image is resized using the drawing scale of 1/4" = 1'-0" in this example. The result of the image processing is shown in Table 5 for the Lower Level Plan. The wall items are specified by the NLP/Text Mining. We can train the agent with machine learning techniques to ignore the Count and Area measurements.

Table 5. Result of the image processing for wall measurement (a case study)

Wall Type	Count	Length (m)	Area (m^2)
Exterior 2x6 Wall with Brick	0	0	0
Exterior 2x6 Wall with Hardie Siding	4	51	18.1
Interior 2x4 Wall	14	46	15.3
Interior 2x6 Wall	1	6	0.8
2x4 Partial Height Wall	13	43	13.2
2x4 Wall Below	0	0	0

3.3 Machine Learning

ML is a branch of artificial intelligence, the study of computer programs that automatically learn from experience and improve their performance for some tasks (Rafiei and Adeli 2018). In the wall example, a machine learning technique called dimensionality reduction is used to remove the count and area quantities from a data set. In general, it is not possible to directly measure the quantity of vertical objects such as walls from 2D top views such as floor plans. Another machine learning method called artificial neural network is used to convert the extracted measurements to desired quantities. In the wall example, a hidden layer for the height of the walls is added to calculate the surface area.

3.4 Expert System

An expert system (ES) is a computer program that is designed to imitate the decision-making ability of human experts by reasoning through a knowledge base, represented mainly as if-then rules rather than through conventional procedural code, using an inference engine (Kulkarni *et al.* 2017). In the wall example, the framing of the wall is specified and the artificial agent needs to process the framing in board footage. This can be learned by the list of materials defined for the wall framing.

4 Findings and Conclusions

Estimators regularly spend many hours working on quantity takeoff to estimate the cost of the projects. This study focused on the use of AI technologies such as NLP, ML, and ES to identify and extract quantities from a set of 2D construction drawings. AI techniques vary depends on their applications and purposes for a project's requirements and associated activities. We consider 12 distinct groups of construction projects to find the common ground between AI techniques used in the quantity takeoff and cost estimating process. An AI-based system, with access to drawings of the buildings, can develop quantity takeoff based on the knowledge it gains from the drawings and data in the database. To address that, we modeled the knowledge of skilled estimators in a way that an AI system can process. To ensure the cost estimating process fits the AI strategy, the common takeoff software tools were reviewed and the role of AI techniques in automating construction and cost estimation was studied. The construction drawings of a residential building were selected as a case study to implement AI techniques including NLP to extract the number of walls in the drawings for quantity takeoff and DIP and ML to estimate the wall areas as well as the number of walls (vertical objects) from floor plans (2D top view). This knowledge was stored in the form of programmable algorithms and rules; such as IF-THEN constructs.

It is important to note that AI is based on past decisions, activities, or performances and needs to train data. Data acquisition and storage is one of the main challenges related to AI combined with restrictions on data sharing and data ownership that can cause legal challenges for the company. As a result, the larger construction companies have more resources and data to benefit more from AI applications. Researchers respond to this challenge by developing ML and NLP datasets that represent or include construction drawings. Meanwhile, AI applications in cost estimating can create positive momentum and it makes sense to start with small use cases and built upon them for more extensive use of AI for quantity takeoff and cost estimation in construction projects. The study proposed AI techniques appropriate for taking off quantities in a 2D construction drawing and explained their role in the takeoff process.

5 Acknowledgement

The fourth author gratefully acknowledges the support of the 2019 Zampell Family Faculty Fellowship for conducting his part in this study. Any opinions, findings, conclusions, and recommendations expressed in this manuscript are those of the authors and do not reflect the views of the Zampell Family.

6 References

Abanda, F. H., Kamsu-Foguem, B., and Tah, J. (2017). "BIM–New rules of measurement ontology for construction cost estimation." *International Journal of Engineering Science and Technology*, 20(2), 443-459.

Al Qady, M., Kandil, A. (2010). "Concept relation extraction from construction documents using natural language processing." *Journal of construction engineering and management*, 136(3), 294-302.

Elmousalami, H. H. (2020). "Artificial intelligence and parametric construction cost estimate modeling: state-of-the-art review." *Journal of Construction Engineering and Management* 146(1), 03119008.

Famiyeh, S., Amoatey, C. T., Adaku, E., and Agbenohevi, C. S. (2017). "Major causes of construction time and cost overruns." *Journal of Engineering, Design and Technology*, 15(2), 181-198.

Gonzalez, R. C., and Woods, R. E. (2018). "Digital Image Processing, Hoboken." NJ: Pearson.

Juszczyk, M. (2017). "The challenges of nonparametric cost estimation of construction works with the use of artificial intelligence tools." *Journal of Procedia engineering*, 196(2017), 415-422.

Kulkarni, P., Londhe, S., and Deo, M. (2017). "Artificial neural networks for construction management: a review." *Journal of Soft Computing in Civil Engineering*, 1(2), 70-88.

Lee, S.-K., Kim, K.-R., and Yu, J.-H. (2014). "BIM and ontology-based approach for building cost estimation." *Journal of Automation in construction*, 41, 96-105.

Mohammadpour, A., Karan, E., & Asadi, S. (2019). "Artificial intelligence techniques to support design and construction." ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction (Vol. 36, pp. 1282-1289). IAARC Publications.

Pratt, D. (2018). *Fundamentals of construction estimating*, Cengage Learning, Boston, MA USA.

Rafiei, M. H., and Adeli, H. (2018). "Novel machine-learning model for estimating construction costs considering economic variables and indexes." *Journal of construction engineering management*, 144(12), 04018106.

Schaufelberger, J. E., and Holm, L. (2017). *Management of construction projects: a constructor's perspective*, Taylor & Francis, New York, NY USA.

Smith, A. J. (2017). *Estimating, Tendering and Bidding for Construction Work*, Macmillan International Higher Education, New York, NY.

Tijanić, K., Car-Pušić, D., Šperac, M., and Applications (2019). "Cost estimation in road construction using artificial neural network." *Journal of Neural Computing*, 32, 9343-9355.

Waty, M., Alisjahbana, S. W., Gondokusumo, O., Sulistio, H., Hasyim, C., Setiawan, M. I., Harmanto, D., and Ahmar, A. S. (2018). "Modelling of waste material costs on road construction projects." *International Journal of Engineering & Technology*, 7(2), 474-477.

Williams, T. P., and Gong, J. (2014). "Predicting construction cost overruns using text mining, numerical data and ensemble classifiers." *Journal of Automation in Construction*, 43, 23-29.