Application of Machine Learning Technology for Construction Site

Heejae Ahn[a], Dongmin Lee[b], Seongsoo Lee[c], Taehoon Kim[d], Hunhee Cho[e], and Kyung-In Kang[f]

[a,b,c,d,e,f]School of Civil, Environmental, and Architectural Engineering, Korea University, Republic of Korea
E-mail: ahn1948@korea.ac.kr, ldm1230@korea.ac.kr, ssleea@korea.ac.kr, kimth0930@korea.ac.kr, hhcho@korea.ac.kr, kikang@korea.ac.kr

Abstract –

Although several research studies on the application of machine learning to the construction field are actively underway, no study has yet been done on the areas where the application is primarily needed. Using the importance–performance analysis method, this paper identified the top five areas of construction sites where machine learning technology needs to be applied. Furthermore, it suggests application plans developed by using the Delphi method. The identified top five areas were unmanned tower crane, inspection of joint connections, prediction of construction safety accidents, operation of construction lift, layout of tower crane. This study is expected to facilitate the effective application of machine learning technology at construction sites in the future. Ultimately, the purpose of this study is to reduce waste of labor and the safety risks at construction sites through machine learning technologies.

Keywords –

Machine Learning; IPA method; Delphi method; Construction site

1 Introduction

In response to the reduced supply of functional labor, the increased demand for reduced construction times, and an increase in safety-related construction accidents, construction automation technologies are actively being developed. Artificial intelligence technology is a representative technology that can be used to realize construction automation. It was confirmed that the number of papers related to artificial intelligence in the field of construction in major journals increased from nine in 2000 to 50 in 2016 [1]. In particular, machine learning among artificial intelligence technologies is an effective way to respond to this trend. Much research related to machine learning is actively being carried out, such as construction accident prediction [2], analysis of construction equipment operation [3], and assumptions about the compressive strength of concrete [4]. However, it has not yet been properly identified which one of the various machine learning technologies is primarily needed to be applied for construction sites. Therefore, the purpose of this paper is to identify which elemental technologies are required on the construction site and to find those that need to be applied and developed first. Furthermore, the paper suggests approximate plans for the application of machine learning at construction sites. To achieve objective validity, experts in machine learning technology were consulted and a survey amongst field staff was conducted. First, a questionnaire was used to identify the areas of interest at construction sites that require urgent improvement. Alternative ways to improve those areas of interest using the proper machine learning technologies were suggested. The questionnaire surveyed the importance and satisfaction levels of specific elements of machine learning technologies that are considered to be applicable. In analyzing the results, the importance–performance analysis method (IPA) was used to identify the most urgent problems at construction sites.

To solve those problems, the particular problems that require improvement are presented with the appropriate machine learning skills using the Delphi method. Machine learning technologies such as artificial neural networks, support vector machines (SVMs), random forest, and genetic algorithms were selected as applicable methods to solve these problems.

The scope of the research included construction sites in and outside Korea, and includes the elements of construction work, such as construction safety, foundation work, and finishing work.

2 Literature review

2.1 Current study for application of machine learning in construction

As used in applications such as self-driving cars, drones and robots, artificial intelligence technology is an essential technology by which intelligent systems can overcome the performance limitations of traditional computing systems [5]. In particular, machine learning, one of the artificial intelligence technologies, enables the machine to directly evaluate a certain situation and take appropriate actions by collecting big data and learning from it on its own.

The application of machine learning technology in the construction field is already actively underway. As a result of analyzing articles published in major journals up to 2016, it has been established that the machine learning technologies used mostly in the construction field include artificial neural networks (50.8%), genetic algorithms (21.8%), composite models (16.3%), fuzzy theory (8.2%), and support vector machines (SVMs) (2.9%) [1].

Lee J et al. [6] suggested using machine learning technology in the analysis of reinforcement-bar images. In their study, a program for the analysis of reinforcement-bar images was separately realized using OTSU binarization through random forest and superpixel clustering. Park U et al. [7] constructed a selection model of retaining wall methods using a support vector machine, one of the machine learning technologies. As a result of its application in a real case, and with an accuracy rate of 93%, it has been proven that this particular machine learning technology can be used effectively in the construction field. Tixier AJ-P et al. [2] and Kim Y et al. [8] used machine learning technology not only to improve construction methods, but also to predict safety-related construction accidents. Kang I et al. [1] recently identified trends in the study of machine learning in the construction field by year, by category, by topic, and by technology. However, they failed to identify how certain elemental technologies were applied and the practical use of them.

Likewise, machine learning technologies are actively being studied in various areas of the construction field, but those technologies that are primarily needed have not yet been scientifically prioritized. In this paper, the elemental machine learning technologies primarily required in the construction field are identified and application plans are proposed.

2.2 Artificial neural network

An artificial neural network is a structure that has been modeled to imitate interactions among the components of the brain of a living creature. The biological brain consists of neurons that are connected to each other through synapses, and it solves problems by changing the degree of connections. Similarly, an artificial neural network also learns about solutions for certain problems by adjusting the weight between artificial neurons. Thus, it is easy to solve complex problems, since solutions can be learned not only from linear functions but also from nonlinear functions based on input variables [9].

2.3 Support vector machine

The support vector machine (SVM) is considered as one of the most robust and accurate methods among data mining algorithms [10]. Based on structural risk minimization, SVM creates a separating hyperplane that can divide multiple data types into two classes. SVM is originally an algorithm for the classification of binary data. Therefore, to distinguish between multiple data types in more than two classes, either the one-against-all (OAA) or pairwise method must be used. Because of the advantages of the SVM algorithm in solving nonlinear problems, it can be used at a construction site that has many variables.

2.4 Random forest

Random forest is an ensemble technique that overcame the shortcomings of decision trees by using bagging and random space methods. Decision trees vary greatly depending on the learning data and therefore it is difficult to use them in generalization [11]. However, random forest generates and learns multiple decision trees from the data. This characteristic allows users to avoid overfitting in the classification and enables more proper generalization than decision trees.

2.5 Clustering

Clustering is a one of the unsupervised learning methods that classifies data by creating criteria and sorting data itself by classifying data without any base information. Clustering is usually used for processing image data and can be more useful when used with deep-learning technology than when used independently. The most famous and simplest example of clustering is K-means clustering. The technology is widely used in various fields, including classification of movies or TV scenes, and clustering related to users of social networking services [12].

2.6 Genetic algorithm

Genetic algorithm is one of the techniques for stochastic search or learning and optimization. The algorithm is based on Charles Darwin's theory of survival of the fittest and Mendel's laws of genetics. The former is the theory that creatures that are well-adapted to nature survive but those that are not are eliminated. The latter is the law that the genetic characteristics of offspring are inherited from their parents [13]. Based on these characteristics, the genetic algorithm can be especially useful for optimization of various fields.

3 Method

3.1 Questionnaire development

Prior to conducting the questionnaire survey, six major journals of construction were used to define relevant categories for the questionnaire. The validity of each category was verified by two experts from the construction field. In this questionnaire survey, each category addresses a part of the construction field where it is considered necessary that machine learning technology must be applied. The number of categories for the questionnaire determined through this process was set at 12. The detailed contents of the categories are listed below in Table 1.

Table 1. Detailed categories of questionnaire survey

Categories	Division
Operation of construction lift	A
Layout of tower crane in high-rise building construction	B
Assumption about compressive strength of concrete	C
Counting the numbers of reinforcemet bar	D
Inspection of joint connections	E
Unmanned tower crane	F
Determination the lifting order between the tower cranes	G
Selection of retaining wall methods	H
Prediction of safety-related construction accidents	I
Forecast of deep excavation wall movement	J
Inspection for correct installation of the forms	K
Determination the time of procurement	L

In this study, the survey was conducted to establish the level of importance and performance of each category on a five-point scale. The number of participants in the survey was 30. Most of the participants were working for construction companies then and had experience in working at construction sites. Their average experience at construction sites was 12 years. The positions of the participants were broadly from employees to managers. The experience at construction sites and their assigned roles in their companies of the participants are presented in Table 2.

Table 2. Information about the questionnaire survey participants

Experience in construction site		Main tasks in companies	
Years	Number of participants	Tasks	Number of participants
~1	1	Construction	17
1~5	6	Office work	4
5~10	7	Quality control	4
10~20	11	Design	3
20~	5	Others	2

3.2 IPA method for prioritizing elements that need improvement

The IPA method is an easily-applied technique to measure the importance and performance for users of each surveyed category. Before using the IPA method, a survey is usually conducted that can check the importance and performance on a five-point or a seven-point scale. The survey results are then analyzed with an appropriate method and displayed on a four-quadrant grid. On the four-quadrant grid, the vertical axis indicates importance, and the horizontal axis indicates performance [14]. Note that the second quadrant includes the categories with high importance and low performance, which indicates where the primary improvements need to be made. The IPA method is highly applicable not only to marketing and the business field, but also to machinery, education, and even the construction field.

In this paper, each category of the IPA method consisted of an area that needs improvement through the introduction of machine learning technology. These categories were selected through consultation with experts from the construction field and experts in the field of artificial intelligence technology. The reason for using the IPA method was that it was impossible to deal with everything that needed improvement and the categories had to be prioritized by using an appropriate method.

4 Results and Discussion

4.1 Analysis of IPA results

The results of the questionnaire survey of each category are outlined in Table 3. To easily identify the priority of categories that need improvement, the contents of Table 3 are displayed on the four-quadrant grid as shown in Figure 1.

Table 3. Average value of importance-performance by category

Category	Importance	Performance	Category	Importance	Performance
A	3.800	2.400	G	3.833	2.567
B	4.267	2.767	H	4.200	2.900
C	4.033	2.733	I	4.167	2.600
D	3.767	2.533	J	4.067	2.900
E	4.167	2.600	K	4.033	2.667
F	3.800	2.000	L	3.700	2.400

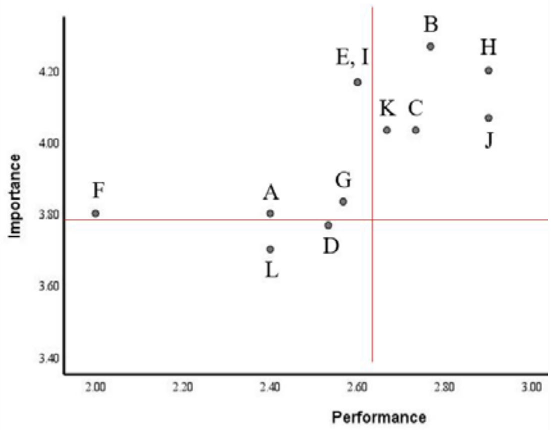

Fig 1. The values of importance-performance displayed on the four-quadrant grid

As shown in Figure 1, categories A, E, F, G, and I are located on the second quadrant. It means that these are the primary categories that need to be improved. However, the importance/performance values of the categories do not correspond exactly to Figure 1. To identify the top five categories easily, their importance/performance values were displayed as shown in Table 4. This study suggests application plans of machine learning technologies for construction sites only for these top five categories.

Table 4. Top five of categories that have the highest importance/performance values

Categories	Importance/performance value
Unmanned tower crane	1.9000
Inspection of joint connections	1.6026
Prediction of construction safety accidents	1.6026
Operation of construction lift	1.5833
Layout of tower crane in high-rise building construction	1.5422

4.2 Application plans using the Delphi method

To develop application plans for machine learning technology for the top five categories, two consultations with a group of four experts in machine learning technology were conducted. The following are proposed application plans using the Delphi method.

4.2.1 Unmanned tower crane

The tower crane is a very important piece of lifting equipment that can make or break construction work, especially in high-rise building construction [15]. It means expenses and the duration of the construction work can vary depending on the ability of the tower crane operator. Moreover, operating the tower crane is very dangerous work for the tower crane operator. Therefore, the commercialization of unmanned cranes is essential for construction sites. There are several elemental technologies for the development of unmanned cranes.

First, voice recognition technology is needed. When a person gives orders through a radio, the unmanned crane should recognize the orders and act on them. To recognize voice, the sound waves should be learned. Because of the difference in the waveforms of sounds among people for any one command, several similar waveforms should be learned as shown in Figure 2.

The waveforms are saved as coordinates. Later, when a person gives an order through the radio, the unmanned crane can find similar coordinates and act on the order by itself. In this procedure, classification technology such as clustering can be applied. This technology can increase the accuracy of voice recognition on noisy construction sites.

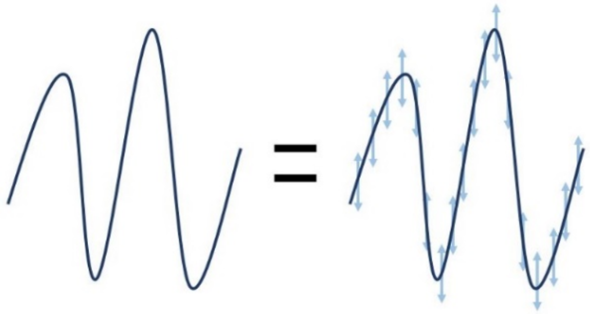

Fig 2. Original waveform and similar waveforms of voice

Second, location recognition technology is needed. When there are several cranes, the movements and distances between the cranes should be perceived by each unmanned tower crane, as the tower cranes may collide with each other while moving. This problem can be solved using global positioning system technology.

Third, material identification technology is needed to enable the tower crane to stop exactly where hoisting of the material is needed. In this case, using sensors to identify material is a more appropriate method to use than machine learning.

Proper combination of these three technologies could facilitate the development of the unmanned crane.

4.2.2 Inspection of joint connections

Construction work on joint connections requires special care and inspections on these areas are not easy. In particular, if a building consists of reinforced concrete, it is also necessary to ensure that the steel is distributed properly, as intended. Since the inspections are carried out by inspectors, there is a risk of incorrect inspections and a waste of labor. Therefore, if a technology can be developed to take a photograph of the joint connections and automatically compare it to a drawing, these problems can be solved.

For such comparisons, image classification technology is needed. Among the machine learning technologies, random forests and superpixel clustering are appropriate methods for image classification. Before image classification can be done, a technology to capture internal images of reinforced concrete, such as x-ray or MRI, is needed. Then, many images of reinforced concrete need to be learned to determine which pixel corresponds to which area. These data values make up many decision trees of the random forest. Superpixel clustering is used for separating pixels in different areas. Combined with the random forest, it can be used to correct a pixel that is incorrectly classified. The overall procedure of image classification is shown in Figure 3.

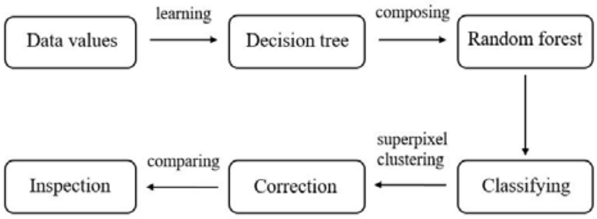

Fig 3. Overall procedure of image classification for inspection

Through this application plan, autonomous inspection of the joint connections can be done. This application plan was based on the work of Lee J et al [6].

4.2.3 Prediction of safety-related construction accidents

Safety-related construction accidents should be prevented by taking into consideration the varying and dynamic natures of construction sites [8]. Therefore, notifying workers in a timely manner of potential accidents that are most likely to happen would be very helpful to prevent such safety-related construction accidents. There are many variables at construction sites, such as types of work in progress, weather, and location. Proper prediction of safety-related construction accidents can be achieved by compiling several variables.

To predict safety-related construction accidents, the artificial neural network is useful, as is shown in Figure 4.

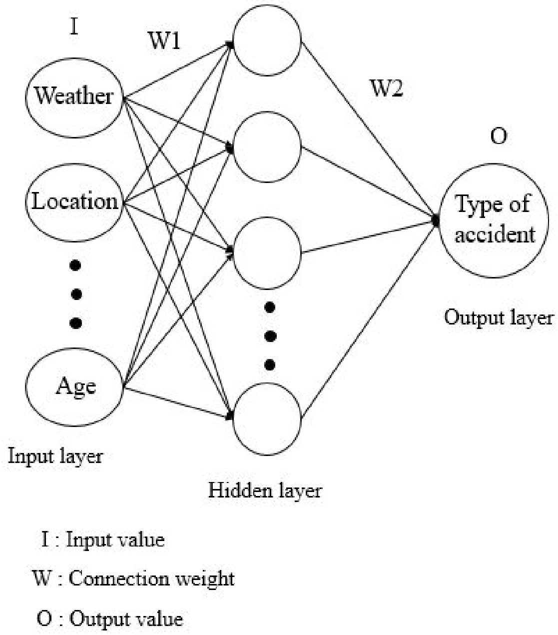

Fig 4. A structure of the artificial neural network

The variables that can affect construction safety are input values. Output values are the types of accidents, such as falling, crashing, and being hit by machines or objects. First of all, to build a model of an artificial neural network, large volumes of data from real cases of safety-related construction accidents should be learned. Then, using proper learning methods, such as error back-propagation, the connection weight for each node is determined. If input values are entered later, the connection weights are multiplied to determine the appropriate output value that is the most probable safety-related construction accident. In this way, the workers can be notified daily of the most probable safety-related construction accidents to happen. Detailed work on this topic was done by Kim Y et al [8].

4.2.4 Operation of construction lift

As more high-rise buildings are being built, the types and volume of materials that need to be lifted are increasing. Thus, the lifting plan of construction lifts is becoming increasingly important to increase efficiency of the construction work [16]. However, the decision-maker for the lifting plans faces an uncertain situation that can result in inaccurate prediction and definition of the various variables in dynamic construction sites [17]. Therefore, development of smart construction lifts that can operate themselves efficiently by referring to a schedule of construction and current situation, is needed.

The smart construction lift needs a combination of several functions. When calling the lift from a specific floor, the current loaded weight and the weight to be loaded should be considered. If the weight to be loaded will result in the total loaded weight exceeding the capacity of the lift, the lift should not answer the call. Furthermore, the lift should recognize the construction schedule by itself. If a task on a critical path is scheduled, the lift must always be waiting to execute the task. These functions of the construction lift can be realized using logical algorithms.

In addition, the ability of the lift to self-zone in context by itself is needed. The construction lift is similar to an elevator, but the number of floors it runs varies as the construction work progresses. To realize the function of self-zoning, it is more efficient to reflect on the latest data than on data of the remote past. Therefore, a recurrent neural network (RNN) that can handle time series data, is useful. RNN can further reflect the latest situation, by placing a higher connection weight on the latest time series data and forgetting the data of the remote past. In this case, the input values are information about the operation of the lift such as the number, time of driving, starting floor, and arrival floor. By learning large volumes of data about this information, the construction lifts can decide about running zones by themselves. The result of zoning the construction lift is shown as the output value. A combination of this machine learning technology and logical algorithms can realize the development of the smart construction lift.

4.2.5 Layout of tower crane in high-rise building construction

The importance of the tower crane was mentioned in Section 4.2.1. Not only the ability of the tower crane operator, but also the model, location, and number of tower cranes can have a significant influence on construction expenses and time. In particular, for high-rise building construction, the layout of tower cranes should be done by systematic calculation, not by human intuition.

In this category, the genetic algorithm is the proper method for optimization. To use the genetic algorithm, the function and the output variable of it must be determined. Using an advanced type of tower crane and a large number of tower cranes can shorten construction duration, but it can increase the daily rental fee of the tower cranes. Based on this principle, the output variable of the function was set to the total rental fee. It can be calculated from multiplying the daily rental fee and the rental duration for each tower crane. The input variables are the number, location, and model of the tower crane. A diagram of the optimization is shown in Figure 5. Next, the construction site should be displayed on a three-dimensional coordinate system. Each part of the tower crane should be displayed as well, as shown in Figure 6.

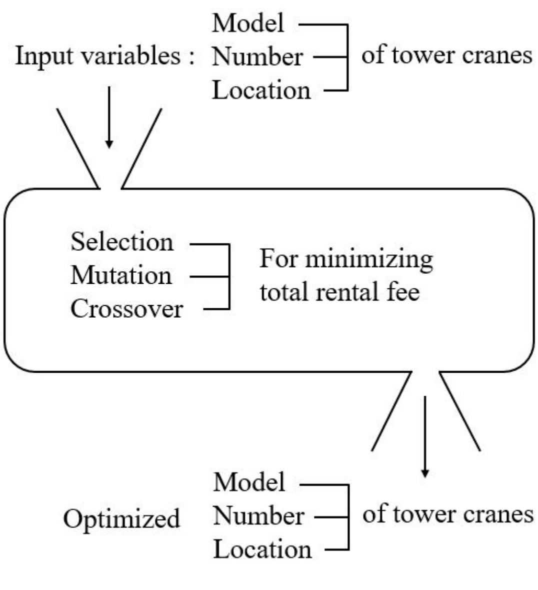

Fig 5. A diagram of the optimization

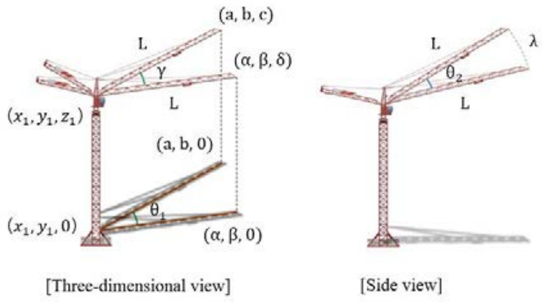

Fig 6. Images of tower crane displayed on three-dimensional coordinate systems

In Figure 6, (α, β, 0) is the location of the stock yard, (a, b, 0) is the location of the place where the material must be put down, and L is the length of a jib boom that varies depending on the model of the tower crane. Using appropriate functions, the distance that the tower crane should move can be calculated. Then, dividing the total distance by the operating speed of the tower crane, the rental duration can be calculated. The daily rental fee of a tower crane depends on the model of the tower crane.

The genetic algorithm automatically calculates the minimized total rental cost by changing the number, location, and model of the tower crane. In this way, the number, locations, and models of the tower cranes that will minimize the total cost can be obtained.

5 Conclusion

In this paper, the parts of the construction field that primarily need improvement were identified from the results of a questionnaire survey and the IPA method. Furthermore, using the Delphi method, improvement plans with machine learning technologies were suggested.

The results of the IPA method showed that the areas in the construction field that need urgent improvement are varied, and include unmanned tower cranes, inspection of construction, safety-related construction accidents, operation of construction lifts, and layout of tower cranes. It means that even though various research studies are actively underway to address problems in the construction field, many areas still exist that need further improvement.

This study will be used for the effective application of machine learning at construction sites by prioritization categories. Ultimately, it is expected to contribute to the construction field by facilitating more economical and safer construction sites through minimizing unnecessary waste of labor at construction sites. However, there are still additional areas to be examined in future studies. Above all, the proposed application plans of machine learning technologies for construction sites in this paper were too sketchy. Detailed plans for application and materialization of models for each category have not yet been proposed. Therefore, in a future study, the specific model for each of the top five categories that need urgent improvements should be developed using machine learning technologies. Furthermore, the effectiveness of the specific models should be properly verified.

Acknowledgement

This research was supported by a grant (18AUDP-B106327-04) from the Architecture & Urban Development Research Program funded by the Ministry of Land, Infrastructure and Transport of the Korean Government.

References

[1] Kang I. Moon J. and Park J. Recent research trends of artificial intelligent machine learning in architectural field. *Journal of the Architectural Institute of Korea Structure & Construction*, 33(4):63–68, 2017.

[2] Tixier AJ-P. Hallowell MR. Rajagopalan B. and Bowman D. Application of machine learning to construction injury prediction. *Automation in Construction*, 69:102–114, 2016.

[3] Cheng C-F. Rashidi A. Davenport MA. and Anderson DV. Activity analysis of construction equipment using audio signals and support vector machines. *Automation in Construction*, 81:240–253, 2017.

[4] Lee S. and Lee H. A study on estimation method of concrete compressive strength based on machine learning algorithm considering mixture factor. *Journal of the Korea Institute of Building Construction*, 17(1):152–153, 2017.

[5] Kim B. Artificial intelligence and trend analysis of national policy recommendations. *ICT Policy*, 23(1):74–93, 2016.

[6] Lee J. Nam S. and Kim S. Automatic Analysis of Reinforcement-bar Images Using Machine Learning Techniques. *Korea Computer Congress 2017*, 21-23, 2017.

[7] Park U. and Kim J. A study on the selection model of retaining wall methods using support vector machines. *Korean Journal of Construction Engineering and Management*, 7(2):118–126, 2006.

[8] Kim Y. Yoo W. and Shin Y. Application of artificial neural networks to prediction of construction safety accidents. *Journal of the Korean Society of Hazard Mitigation*, 17(1):7–14, 2017.

[9] Kim K. and Kim K. A scheme on detecting unknown-attacks using ant clustering algorithm and artificial neural network. In *Proceedings of the Korea Institutes of Information Security and Cryptology Conference*, pages 63–65, 2015.

[10] Wu X. Kumar V. *The top ten algorithms in data mining*, Taylor & Francis Group, 2009.

[11] Hoonki90. ML&DM/Ensemble Learning. On-line: http://gentlej90.tistory.com/37?category=729784, Accessed: 24/01/2018.

[12] Song H. Lee S. and Kwak H. Clustering method for similar user with mixed data in SNS. *Journal of the Korea Society of Computer and Information*, 20(11):25–30, 2015.

[13] Goldberg D. *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley, 1989.

[14] John A. and John C. Importance-performance analysis. *Journal of Marketing*, 41(1):77–79, 1977.

[15] Kim Y. Cho J. Kim M. Cho K. and Hyun C. A productivity analysis of tower crane installation progress based on simulation technique. *Journal of the Korea Institute of Building Construction*, 10(2):33–40, 2010.

[16] Shin Y. Construction lift planning system in tall building construction. *Journal of the Architectural Institute of Korea Structure & Construction*, 27(11):227–235, 2011.

[17] Lee J. Jeon Y. and Park C. Models for predicting hoisting times of tower crane in the high-rise building construction. *Korea Journal of Construction Engineering and Management*, 472–475, 2004.