Government Construction Project Budget Prediction Using Machine Learning

Wuttipong Kusonkhum, Korb Srinavin, Narong Leungbootnak, and Preenithi Aksorn
Department of Civil Engineering, Khon Kaen University, Khon Kaen, Thailand
Email: w.kusonkhum@gmail.com, korbsri@kku.ac.th, Narongl.fec@gmail.com, preenithi@kku.ac.th

Tanayut Chaitongrat
Faculty of Architecture, Urban Design and Creative Arts, Mahasarakham University, Mahasarakham, Thailand
Email: Tanayut.c@msu.ac.th

***Abstract*—The construction industry could not avoid the technology disruptive era. Therefore, the Thai government has created a new policy and directed all departments to implement big data technology. Big data technology includes Machine Learning (ML). The present study attempts to predict over-budget construction projects using an ML algorithm. Data were collected from the comptroller general's department of Thailand for over-budget project cases. Information about 692 projects completed in Thailand in 2019, covering all types of construction projects, was collected and analyzed. ML, an analytical technique for big data technology, was used as a tool in this study. In addition, k-Nearest Neighbors (KNN), an ML algorithm, was used to classify over-budget projects. The input data have four attributes: department of project, construction site location, type of project, and methods of procurement; the output is a yes/no decision on whether a project has been over budget. The dataset was preprocessed for analysis and modeled using the KNN function in Python 3. According to the test results, the KNN model achieves an accuracy (precision) of 0.86. Finally, the developed model has demonstrated that it can be used to predict the over-budget construction projects for the Thai government.**

***Index Terms*—government construction project, big data technology, machine learning, k-nearest neighbors, over-budget project**

I. Introduction

Machine Learning (ML) is a part of Artificial Intelligence (AI) that allows a system to learn from data rather than through basic programming. Moreover, it has become the most important topic within the AI field [1], [2]. Many studies have been conducted on how to use big data [3]-[6]. However, the use of ML algorithms and big data depends on the available data and objectives of an organization [1].

The Thai government also uses big data to improve the country. The Thai government has already mandated all of its organizations and departments to use emerging technologies, such as big data, in their works [7]. However, some issues, such as lack of infrastructure, quality of data, and tradition of data collection methods, have been encountered while using these emerging technologies [8].

Furthermore, project cost management is a task that can be improved by applying the se new technologies. At a global level, this dissertation has developed a unique methodology using ML methods to improve cost prediction at the early design stage of building construction. In addition, there is another study that has evaluated bidding price, which can directly affect the business of construction firms. Thus, to determine a bidding price for a new project to sign a contract is such a difficult and responsible decision. Most results of the study have been validated in real applications, which achieved good performance for training models [9]-[11].

Currently, the government procurement process is critical to Thailand's development since the government must budget and spend on various projects that would benefit the citizen's quality of life and the country's development [12]. The Thai government attempts to define the components of government construction procurement problems to eradicate corruption in the procurement process, especially in construction projects [13]. The estimated cost is a part of the procurement process, and ML is a technique that can be used for efficient monitoring [10], [14].

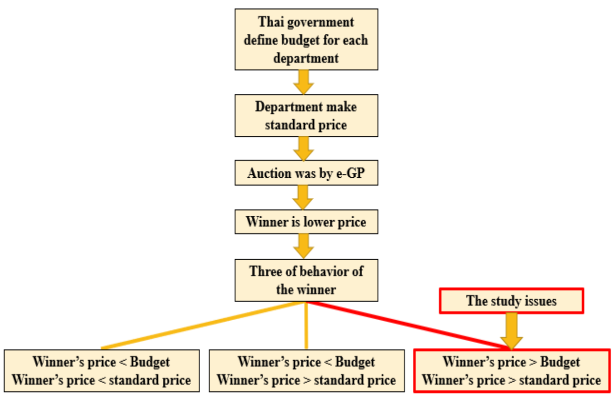

Figure 1. Thai procurement process.

In 2019, the comptroller general's department of Thailand found many construction projects with inflated

Manuscript received May 17, 2021; revised November 2, 2021.

doi: 10.12720/jait.13.1.29-35

budget prices. This situation has cost the Thai government's budget plan of approximately US$1.5 million per year, as shown in detail in Fig. 1. Accordingly, the current study has developed an ML-based prediction model to predict the projects with inflated budgets.

II. Background of the Study

The background explains the knowledge and theories that are related to project costs, procurement management, big data, and the k-Nearest Neighbors (KNN) classification method.

A. *Project Cost Management*

Project cost management involves four processes to complete a project with the approved budget: planning, estimating, budgeting, and control. First, Plan Cost Management defines how a project cost will be estimated, allocated, managed, monitored, and controlled. Second, Estimate Cost is a process of approximation development for monetary resources, which is necessary to complete a project. Third, Budget Determination is the process of aggregating the estimated costs of those individual activities or work packages to establish an authorized cost baseline. Fourth, Control Costs is a process concerned with monitoring the status of a project to update the project costs and manage changes to the cost baseline [14], [15]. After the above-stated processes, the procurement process can begin [8], [15].

B. *Project Procurement Management*

Project procurement management includes processes that are required to purchase or receive products, services, or results that are required from outside the project. Contracts, purchase orders, memoranda of agreements, and internal service level agreements must be administered as part of the procurement process. There are three processes for the procurement process: planning, conducting, and controlling [14].

Normally, the Thai government conducts the procurement process using a prequalification and lower price bidding approach [15], [16]. There are three types of procurement methods used by the government for projects, which consist of bidding, chosen, and specific [16].

C. *Big Data for Construction Management*

Big data have been used in some applications, such as Building Information Modeling (BIM), by managing the data of construction projects [17], [18].

ML and AI provide construction businesses with a competitive advantage in the market [19]. ML is a subfield of the AI field, and it is set up for computational tasks. ML learns automatically and directly from data of specific tasks [20]. Classification is an ML algorithm. The algorithm learns from data to classify new data. KNN is a simple classification method that has been used in the present study [21].

D. *ML in Construction*

In the construction industry, ML-based models have been widely developed and used in knowledge building. İt has become an educational trend for the new style of construction management systems. In addition, there is the implementation of managing construction model data using BIM and managing CAD data systematically to support the cooperation of stakeholders [2]. Furthermore, a support vector machine has been used in construction cost estimation. This machine would compare prediction performance with other methods [22].

However, the performance of an ML-based model depends on data quality and availability and the objective of construction model building. Otherwise, the developed model would not be efficient and could not be implemented. In Thailand, the overall operational ability in construction management is low, which affects the data quality or efficiency of projects in the organizations [23].

In Thailand, the government agency that is in charge of looking after and developing procurement management in constructions has developed a system called Electronic Government Procurement (e-GP). In addition, the system could collect annual procurement data of the entire country efficiently. Nevertheless, the system is not developed to fix any specific problems of the organization [13].

E. *Issue of Thai Government Procurement*

Currently, the construction industry has been developing cost management systems using emerging technologies [11]. ML methods are used to improve the cost management process in construction projects. In addition, the prediction has been used for the early design stage of building construction since cost estimates are important to project feasibility studies. It as well impacts on the final project's success. Cost estimates provide significant information that can be used in project measurement and assessment.

In addition, the Thai government is responsible for setting the standard cost pricing of each construction project [13]. The mistakes of the government's decision could affect the bidding prices from bidders. Moreover, the bidding price directly affects the construction project. Thus, to determine a bidding price for a new project to sign a contract is such a difficult and responsible decision [10]. This is a point of the study, to use and develop the old data to improve the process and the conceptual framework shown in Fig. 2.

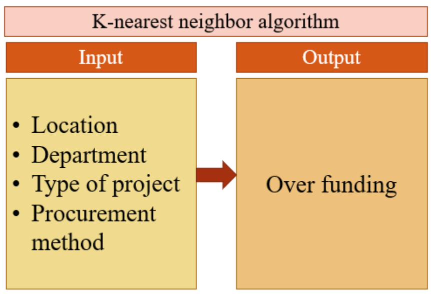

Figure 2. Conceptual framework.

F. *K-Nearest Neighbors*

KNN is a non-parametric method that has been used in classification and regression.

KNN has been successfully used for classification in many applications. This method measures the distance between items in a test set and each item in a training set; note that the K training set items are the KNN. After that, the test set items are classified according to the class of the most common among the KNN, and each neighbor is allowed to vote. (When there is a tie, all training set items that are not farther away than the Kth nearest neighbors are included in the voting process, so the total number of voters exceeds K.) A few researchers have considered the best way to measure the distance. Their approaches have included using global and local metrics; however, these metrics are problem-specific. So far, Euclidean distance has been the most widely used metric, where the distance between two points x_r and x_s is given by the square root of the sum (possibly weighted) of the squared distances over each co-ordinate [24]-[27].

G. ML for Thai Government Procurement

Above all, the Thai government procurement process has been successful in collecting and storing data in an electronic system [13], and it also has a procurement process management principle for construction projects [9]. Furthermore, the system could support emerging technologies, which tend to develop and build knowledge on the basis of prior data to solve existing problems and find benefits for organizations [21].

In addition, data collection from government research studies has revealed some unusual data. Thus, the Thai government agency in charge of procurement wants to develop and identify data anomalies to determine whether the collected data are sufficient to use for ML [21], [23].

III. Research Methodology

The authors requested permission from the comptroller general's department to collect data from the Government Procurement electronic system (e-GP). The required data are the information about all over-budget government projects in Thailand.

Following the collection of the necessary data, the data were divided into two parts: 80% of the data were used for model training, and 20% of the data were used for model validation. The training data were analyzed using KNN and verified by a confusion matrix [28], [29].

A. KNN Model Development

The collected data need to be prepared in a CSV file to ensure that there is no missing value and unknown category. A computer program is necessary to perform KNN analysis. The computer program was written in Python language and ran on Anaconda software. Table I shows some codes.

B. Verifying the KNN Model

The KNN model was verified for its accuracy, precision, and recall by constructing a confusion matrix and using the following equations (1-3) [29].

$$\text{Accuracy} = (TP-TN)/TP-TN-FP-FN) \quad (1)$$

$$\text{Precision} = TP/(TP-FP) \quad (2)$$

$$\text{Recall} = TP/(TP-FN) \quad (3)$$

where TP = true positive, TN = true negative, FP = false positive, and FN = false negative.

TABLE I. Example Code of the KNN Model

Input	import pandas as pd import seaborn as sns import matplotlib.pyplot as plt import numpy as np
Input	from sklearn.compose import ColumnTransformer, make_column_transformer from sklearn.preprocessing import OneHotEncoder, StandardScaler from sklearn.impute import SimpleImputer from sklearn.neighbors import KNeighborsClassifier from sklearn.pipeline import Pipeline, make_pipeline
Input	col_trans = make_column_transformer((OneHotEncoder(sparse=False, handle_unknown='ignore'), ['DEPT_NAME','Site_Loction', 'REAL_METHOD_ID','Type of project']),) col_trans

Accuracy is the ratio of the total number of correct classifications to the total number of predicted classifications, as calculated by a model. Precision is defined as the ability to obtain consistent results from repeated measurements. Precise values differ from one another because of random error, a type of observational error in information retrieval. In addition, Recall is the percentage of relevant documents that are successfully retrieved [30], [31].

C. Data Collection

Table II shows that there are five attributes in input data collected from the traditional system of the Thai government. Each parameter has a different factor, depending on its definition. The location attribute has 72 factors, including 72 cities in Thailand, where projects are being executed. The location attribute is one of the factors influencing construction cost estimation [28]. The department name attribute has 48 factors, including 48 departments handling construction projects. A project owner is one of the stakeholders involved in cost management [14], [15]. The type of project attribute has three factors, including building roads and irrigation projects [16]. This attribute is an important factor that influences cost estimate and project budgeting [14]. The procurement method attribute has three factors, including the chosen bidding method and specific method. The procurement method is different for each country, but the goal is the same, i.e., to eradicate corruption [13]. The procurement method is a significant factor influencing a contractor's cost [13], [16].

The case of projects attribute is the consideration of whether a project is under budget or over budget according to the Thai government policy. It has two factors.

IV. Result

All the attributes were collected, including projects' departments, construction site location, type of project, and methods of procurement. The most preferable procurement method for the Thai government is the specific method, which accounts for 86.85% of all procurement methods, as shown in Table III.

TABLE II. Input of Data

Attributes	**Factor**
Location	72
Department name	48
Type of project	3
Procurement method	3
Case of project	2

TABLE III. Method of Procurement

Method of procurement	Amount (projects)	%
Bidding	48	6.94
Chosen	43	6.21
Specific	601	86.85

Most Thai government construction projects are road and building construction (Table IV) and Table V shows that under-budget projects account for 79.34%, whereas over-budget projects account for 20.66% of all data. The data were prepared in a data frame. In addition, the data should be preprocessed before being used with the ML algorithm, as shown in Table VI and Table VII.

TABLE IV. Type of Project

Project type	Amount (projects)	%
Building	299	34.97
Road	281	35.84
Irrigation	112	14.45

TABLE V. Over-/Under-budget Projects

Case of projects	Amount (projects)	%
Under budget	549	79.34
Over budget	143	20.66

TABLE VI. Example of Data Frame

DEPT NAME	Site Location	REAL METHOD ID	Type of project	Funding (Million bath)	Mid Std. (Million bath)	Win Bid (Million bath)	Dif.Price (Million bath)	Y
D4	S1	B3	T1	30.00	28.70	30.00	1.29	No
D1	S1	B2	T3	12.43	12.01	12.08	65.00	No
D2	S1	B1	T1	13.32	12.28	13.29	1.01	No

The data were split into two groups: training and test sets. The test set accounts for 20% of all data. The data were split into two groups: training and test sets. The test set accounts for 20% of all data. The KNN model achieves an accuracy of 0.8561, as can be seen in Table VIII.

TABLE VII. Data of Site Location

Input	df[['Site_Location']]	
Output	0	S1
	1	S1
	2	S1
	….	….
	690	S1
	691	S1
692 rows × 1 columns		

TABLE VIII. Accuracy of Model Testing Function

Input	pipe.fit(X_train, y_train) pipe.score(X_test, y_test)
Output	0.8561151079136691

The confusion matrix (Fig. 3) was used to calculate the model's classification accuracy. The matrix showed that the model correctly predicted 119 of 139 cases (Table IX), outperforming another study that studied a student's graduation year [32] and the number of true negatives and positives is shown in Table X.

		Actual class	
		P	N
Predicted class	P	TP	FP
	N	FN	TN

Figure 3. Confusion matrix.

TABLE IX. Results from the Confusion Matrix

Input	metrics.confusion_matrix(y_test, predicted)
Output	array([[107, 4], [25, 3]], dtype=int64)

Similarly, the model's precision can also be calculated using the confusion matrix. The precision of the model is 0.8592 and can be seen in Table XI. The recall is the percentage of relevant instances that were retrieved [31]. A perfect recall score of 1.0 means that all relevant documents were retrieved by a search [33]. The recall value of the model is 0.9914, as shown in Table XI.

According to test results, the model achieved an accuracy of 86%. The accuracy is similar compared with that of other studies in Table XII. However, the accuracy is high because of the relation of the input data, the normal behavior of contractor companies, and the abnormal work of government officers [34]. In the input data, the important attribute that improves the model's prediction accuracy is the method of procurement attribute [13], [16]. The Thai government has three procurement methods, and one of these methods was designed for small-scale projects, such as small buildings and reinforcement of concrete roads. The specific method was designed by the government for preferential contractors, whom the government could guarantee their capacity to complete projects in a short time. The benefit of this method is that constructed facilities would be provided to people in a shorter time than the bidding method.

The precision can be divided into two cases (i.e., Yes and No cases), as shown in Fig. 4. For the first case, the Yes case, the model achieves a precision of 56%. The Yes case means that the model predicts the winner's price is higher than both the standard price and budget during the bidding process. The accuracy of the Yes case decreases as the training data approach a "No" case. This is the main reason for the Yes case's low accuracy.

TABLE X. NUMBER OF TRUE NEGATIVES AND POSITIVES PREDICTED

Input	predicted=pipe.predict(X_test) predicted
Output	array(['No', 'Yes', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'Yes', 'No', 'Yes', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'Yes', 'No', 'Yes', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'No', 'Yes', 'No', 'Yes', 'No', 'No', 'No', 'No', 'No', 'No'], dtype=object)

TABLE XI. ACCURACY OF MODEL TESTING FUNCTION

Input	print(metrics.classification_report(y_test, predicted))				
Output		precision	recall	f1-score	support
	No	0.86	0.99	0.92	117
	Yes	0.75	0.14	0.23	22
	Accuracy			0.86	139
	Macro avg	0.80	0.56	0.58	139
	Weighted avg	0.84	0.86	0.81	139

Therefore, government officers, as one of the factors influencing the procurement process, can be involved in several activities during the specific method as long as their activities are not being inspected and tracked [34]. However, there is a flaw in the procurement regulation that contributes to the government's poor image. The flaw is that this procurement method is suitable for only small projects (the project cost must be less than US$16,800). Furthermore, the government agency is the only one who can choose contractors directly without any complicated procedure. Although there is a department dedicated to project auditing, it cannot cover every single project. In the other study show that the government agency is one of important factor regarding corruption [35].

TABLE XII. ACCURACY OF KNN MODELS

No.	Field of study	Accuracy (%)
1	Fingerprinting big data: the case of KNN graph construction [36].	78.9
2	KNN-based knowledge-sharing model for severe change order disputes in construction [37].	93.82
3	KNN model-based approach in classification [38].	90.41

Accordingly, this suggests that more training data are needed to improve the model's performance, and that the more data algorithm the learns, the more accurate it is at predicting the Yes case [16], [31].

For the second case, the No case, the model achieves an accuracy of 56%. The No case means that the model predicts the winner's price is higher than the standard price but lower than the budget during the bidding process.

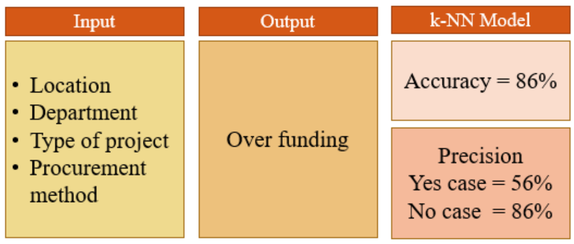

Figure 4. Table of precision.

V. CONCLUSION

The current study collected data from the traditional system of the Thai government. The data have four attributes: department name, site location, method of procurement, and type of project. The authors have developed an ML-based model for predicting over-budget projects. The accuracy of the developed model, which was developed with the KNN algorithm, was 0.86.

The present study has demonstrated that some of the archaic data from the Thai government are sufficiently efficient for developing new technologies, such as ML. Although ML is a component of big data for analyzing large amounts of data, it has a progress goal.

Finally, the present study can prove that even a small amount of data from the Thai government can be efficiently used by big data and ML techniques.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

W. Kusonkhum and K. Srinavin conducted and analyzed the research; W. Kusonkhum analyzed the data;

K. Srinavin, T. Chaitongrat, P. Aksorn, and N. Leungbootnak are supervised the research; all authors had approved the final version.

ACKNOWLEDGMENT

We would like to sincerely thank the Graduate School, Khon Kaen University for their funding support. In addition, we would like to thank the Comptroller General's Department for allowing their data to be used in this study. Without any of them, our research would not be accomplished.

REFERENCES

[1] J. Hurwitz and D. Kirsch, *Machine Learning for Dummies*, IBM Limited Edition, vol. 75, 2018.
[2] R. Eadie, M. Browne, H. Odeyinka, C. McKeown, and S. McNiff, "BIM implementation throughout the UK construction project lifecycle: An analysis," *Automation in Construction*, vol. 36, pp. 145-151, 2013.
[3] F. Insights. (2015). Betting on big data: How the right culture, strategy and investments can help you leapfrog the com-petition. [Online]. Avaialble: www.forbes.com/forbesinsights
[4] S. Bai, M. Li, R. Kong, S. Han, H. Li, and L. Qin, "Data mining approach to construction productivity prediction for cutter suction dredgers," *Automation in Construction*, vol. 105, p. 102833, 2019.
[5] H. Naganathan, W. O. Chong, and X. Chen, "Building Energy Modeling (BEM) using clustering algorithms and semi-supervised machine learning approaches," *Automation in Construction*, vol. 72, pp. 187-194, 2016.
[6] C. Q. Poh, C. U. Ubeynarayana, and Y. M. Goh, "Safety leading indicators for construction sites: A machine learning approach," *Automation in Construction*, vol. 93, pp. 375-386, 2018.
[7] Royal Thai Government Gazette. Thailand's 20-Year National Strategy. [Online]. Available: http://www.ratchakitcha.soc.go.th/DATA/PDF/2561/A/082/T_0001.PDF
[8] Digital Government Development Agency (DGA). Debate: How Can Big Data Make Government's Services Smarter. [Online]. Available: http://www.egovernmentforum.com/assets/files/How%20Can%20Big%20Data%20Make%20Governments%20Serivices%20Smarter_Dr.Sak_20180830.pdf
[9] M. Y. Cheng, H. S. Peng, Y. W. Wu, and T. L. Chen, "Estimate at completion for construction projects using evolutionary support vector machine inference model," *Automation in Construction*, vol. 19, no. 5, pp. 619-629, 2010.
[10] S. Petruseva, P. Sherrod, V. Z. Pancovska, and A. Petrovski, "Predicting bidding price in construction using support vector machine," *TEM Journal*, vol. 5, no. 2, p. 143, 2016.
[11] S. Z. Doğan, *Using Machine Learning Techniques for Early Cost Prediction of Structural Systems of Buildings*, ProQuest Dissertations Publishing, 2005.
[12] A. Dzuke and M. J. Naude, "Procurement challenges in the Zimbabwean public sector: A preliminary study," *Journal of Transport and Supply Chain Management*, vol. 9, no. 1, pp. 1-9, 2015.
[13] N. Leungbootnak, T. Chaithongrat, and P. Aksorn, "An exploratory factor analysis of government construction pro-curement problems," in *Proc. the 4th International Conference on Engineering, Applied Sciences and Technology*, 2018, vol. 192, article number 02057.
[14] R. Burke, *Project Management: Planning and Control Techniques*, New Jersey, USA: John Wiley & Sons, 2013, p. 26.
[15] B. M. Jervis and P. Levin, *Construction Law, Principles and Practice*, McGraw-Hill College, 1988
[16] *The Government Procurement and Supplies Management Act B.E. 2560*, The Comptroller General's Department, Bangkok, 2017.
[17] Y. Jiao, *et al.*, "An augmented MapReduce framework for building information modeling applications," in *Proc. IEEE 18th International Conference on Computer Supported Cooperative Work in Design*, May 2014, pp. 283-288.
[18] J. R. Lin, Z. Z. Hu, J. P. Zhang, and F. Q. Yu, "A natural-language-based approach to intelligent data retrieval and representation for cloud BIM," *Computer-Aided Civil and Infrastructure Engineering*, vol. 31, no. 1, pp. 18-33, 2016.
[19] I. H. Witten and E. Frank, "Data mining: Practical machine learning tools and techniques with Java implementations," *Acm Sigmod Record*, vol. 31, no. 1, pp. 76-77, 2002.
[20] M. Bilal, *et al.*, "Big Data in the construction industry: A review of present status, opportunities, and future trends," *Advanced Engineering Informatics*, vol. 30, no. 3, pp. 500-521, 2016.
[21] S. Cost and S Salzberg, "A weighted nearest neighbor algorithm for learning with symbolic features," *Machine Learning*, vol. 10, no. 1, pp. 57-78, 1993.
[22] T. Mahfouz, "Unstructured construction document classification model through Support Vector Machine (SVM)," in *Proc. International Workshop on Computing in Civil Engineering*, 2011, pp. 126-133.
[23] C. Koomanachai, "Documentation associated with construction management: Case study of an especially large-scale and privately owned construction project," Chulalongkorn University, 2007.
[24] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.
[25] Z. Zhang, "Too much covariates in a multivariable model may cause the problem of overfitting," *Journal of Thoracic Disease*, vol. 6, no. 9, p. E196, 2014.
[26] B. Lantz, *Machine Learning with R*, Packt Publishing Ltd., 2013.
[27] H. R. Bernard and H. R. Bernard, *Social Research Methods: Qualitative and Quantitative Approaches*, Sage, 2013.
[28] K. K. Chitkara, *Construction Project Management*, Tata McGraw-Hill Education, 1998.
[29] A. Gondia, A. Siam, W. El-Dakhakhni, and A. H. Nassar, "Machine learning algorithms for construction projects delay risk prediction," *Journal of Construction Engineering and Management*, vol. 146, no. 1, p. 04019085, 2020.
[30] J. Hernández-Torruco, J. Canul-Reich, J. Frausto-Solis, and J. J. Méndez-Castillo, "Towards a predictive model for Guil-lain-Barré syndrome," in *Proc. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, August 2015, pp. 7234-7237.
[31] K. Allen, M. M. Berry, F. U. Luehrs Jr, and J. W. Perry, "Machine literature searching VIII. Operational criteria for designing information retrieval systems," *American Documentation (pre-1986)*, vol. 6, no. 2, p. 93, 1955.
[32] Z. Zhang, "Introduction to machine learning: k-nearest neighbors," *Annals of Translational Medicine*, vol. 4, no. 11. 2016.
[33] I. Ahmad, H. Sulistiani, and H. Saputra, "The application of fuzzy k-nearest neighbour methods for a student graduation rate," *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 1, no. 1, pp. 47-52, 2018.
[34] D. L. Olson and D. Delen, *Advanced Data Mining Techniques*, Springer Science & Business Media, 2008.
[35] T. Chaitongrat, N. Leungbootnak, W. Kusonkhum, W. Deewong, S. Liwthaisong, and K. Srinavin, "Measurement model of good governance in government procurement," *IOP Conference Series: Materials Science and Engineering*, vol. 639, no. 1, p. 012024, October 2019.
[36] R. Guerraoui, A. M. Kermarrec, O. Ruas, and F. Taïani, "Fingerprinting big data: The case of KNN graph construction," in *Proc. IEEE 35th International Conference on Data Engineering*, April 2019, pp. 1738-1741.
[37] J. H. Chen, "KNN based knowledge-sharing model for severe change order disputes in construction," *Automation in Construction*, vol. 17, no. 6, pp. 773-779, 2008.
[38] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach in classification," in *Proc. OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"*, November 2003, pp. 986-996.

Wuttipong Kusonkhum is a Doctoral degree student in the Department of Civil Engineering, Faculty of Engineering, Khon Kaen University, Thailand. He obtained his B.Eng. in Civil Engineering from Mahasarakham University and M.Eng. in Civil Engineering from Khon Kaen University. His research interests involve project management and construction industry development.

Korb Srinavin is an Associate Professor in the Department of Civil Engineering, Faculty of Engineering, Khon Kaen University, Thailand. He obtained his B.Eng. in Agricultural and Civil Engineering and M.Eng. in Structural Engineering and Construction from Asian Institute of Technology. He gained his Ph.D. in Construction Engineering and Management from Griffith University Australia. His research interests involve project management and construction industry development.

Narong Leungbootnak is an Associate Professor in the Department of Civil Engineering, Faculty of Engineering, Khon Kaen University, Thailand. He obtained his B.Eng. in Civil Engineering and M.Eng. in Structural Engineering from Khon Kaen University. He gained his Ph.D. in Construction Engineering and Management from Asian Institute of Technology. His research interests involve project management and construction industry development.

Tanayut Chaitongrat, Ph.D., is an Assistant Professor in the Faculty of Architecture, Urban Design and Creative Arts, Mahasarakham University, Thailand. He earned B.S.Tech.Ed. in Civil Engineering from king Mongkut Institute of Technology North Bangkok, then M.Eng. and PhD in Civil Engineering from Khon Kaen University. Dr. Chaitongrat is a head of a Construction & Project management Center and has been working as an Editor-in-Chief of Journal of Architecture, Design and Construction. He is also specialized in Public Procurement in Construction, Construction Project Management, Planning Management, Monitoring and Controlling System, and Evaluation & Grading Project.

Preenithi Aksorn is an Assistant Professor in the Department of Civil Engineering, Faculty of Engineering, Khon Kaen University, Thailand. He obtained his B.Eng. from Khon Kaen University and M.Eng. from Thammasat University. He obtained his PhD in Construction Engineering and Infrastructure Management from Asian Institute of Technology. His area of research is within the field of sustainability of local infrastructure project.