

PREDICTION OF THE DELAY IN THE PORTFOLIO CONSTRUCTION USING NAÏVE BAYESIAN CLASSIFICATION ALGORITHMS

Kadhim Raheem ERZAIJ¹, Abbas M. BURHAN^{1,*}, Wadhah Amer HATEM², Rouwaida Hussein ALI³

¹ Civil Engineering Department, Construction Project Management, University of Baghdad, Baghdad, Iraq.

² Middle Technical University, Baquba Technical Institute, Baghdad, Iraq.

³ Civil Engineering Department, University of Baghdad, Baghdad, Iraq.

* corresponding author: abbasm.burhan@coeng.uobaghdad.edu.iq

Abstract

Projects suspensions are between the most insistent tasks confronted by the construction field accredited to the sector's difficulty and its essential delay risk foundations' interdependence. Machine learning provides a perfect group of techniques, which can attack those complex systems. The study aimed to recognize and progress a well-organized predictive data tool to examine and learn from delay sources depend on preceding data of construction projects by using decision trees and naïve Bayesian classification algorithms. An intensive review of available data has been conducted to explore the real reasons and causes of construction project delays. The results show that the postponement of delay of interim payments is at the forefront of delay factors caused by the employer's decision. Even the least one is to leave the job site caused by the contractor's second part of the contract, the repeated unjustified stopping of the work at the site, without permission or notice from the client's representatives. The developed model was applied to about 97 projects and used as a prediction model. The decision tree model shows higher accuracy in the prediction.

Keywords:

Delay;
Construction;
Decision tree;
Bayesian,
Portfolio.

1 Introduction

The project's success is the definitive aim of the numerous project stakeholders [1]; it is essential that the construction industry is time-sensitive and has several constraints. One of the most important is time as the increase in time will increase the cost. Projects suspensions are between the most insistent tasks confronted by the construction sector accredited to the sector's difficulty and its essential delay risk foundations' interdependence. Machine learning could provide a group of techniques that can attack those composite systems and solve complex problems. Lately, the construction sector has knowledgeable a quick-tempered development in the quantity of objective data produced daily and kept from the numerous disciplines through the project or the ability lifespan [2]. Such data facility generates a chance for valuable mining knowledge and expert solutions to the delay of the prevailing project.

Chan [3] checked the reasons for the delays by the client's sides, advisers, and contractors, and then inspected the influences disturbing efficiency. It explains the variances in insights of the relative meaning of influences among the three groups, revealing their practices, likely predispositions and absence of actual communicate. The influence of delays may involve an increase in time, cost, disputes, adjudication, litigation, and entire rejection. This study aims to develop a predictive data tool to examine and learn from delay sources. It depends on preceding data of construction projects by using decision trees and naïve Bayesian classification algorithms. A no inclusive list of ML methods useful in construction consisting of artificial neural network ANN [4-6] decision trees [7, 8], logistic

regression [9, 10], naïve Bayesian models [11, 12] and support vector machines [13-15]. However, the evaluation results indicated that the naïve Bayesian model better predicts the data set examined [16].

Developed a decision tree and naïve Bayesian classification algorithms model to facilitate precise project analysis of delay risk and prediction by objective data sources. The predictive presentations of two models were estimated through cross-validation tests, and the models were additionally equated using machine-learning-relevant performance indices. The evaluation results indicated that the naïve Bayesian model better predicts the data set examined.

2 Methodology

A comprehensive and intensive review of available data was conducted to explore the real reasons and causes of the delay and suspension of the work. The portfolio construction projects were implemented by the project management office funded and awarded by the state authority to cover the demand of services, education, security, and infrastructure. A study was conducted to classify and grow the machine learning using decision tree and naïve Bayesian classification algorithms models to facilitate precise project analysis of delay risk and prediction by objective data sources. The predictive presentations of two models were estimated through cross-validation tests, and the models were additionally equated using machine-learning-relevant performance indices. Fig. 1 shows the process of delay prediction adopted in this study, and Fig. 2 explains the analysis methodology and objectives of this study.

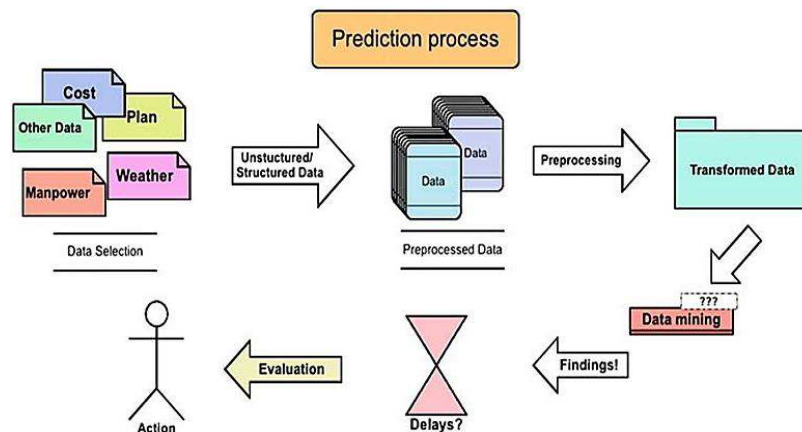


Fig. 1: Delay prediction process for this study.

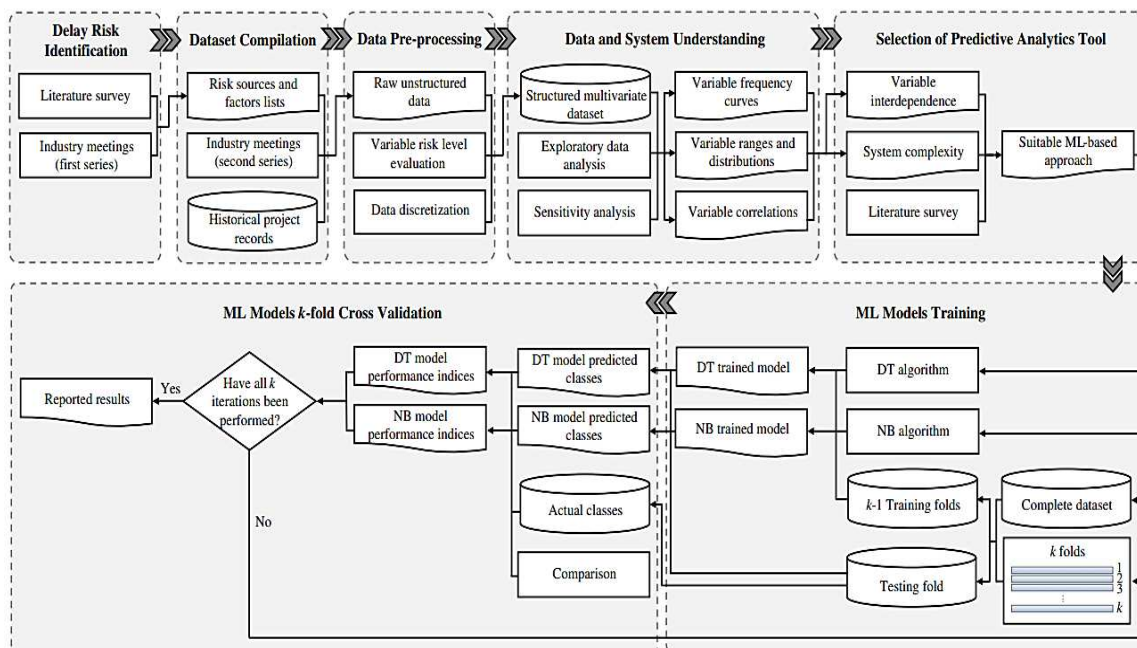


Fig. 2: Analysis methodology and objectives.

The study was over three years, considering the total number of projects is 364. Half of them are without credible documents of progress or tracking reports, where only 186 of them were well documented. The documented projects are extended over three categories of 137 new construction projects, 36 maintenance, and 13 rehabilitation projects. The data of 97 projects that own complete documentation has considered making the necessary diversity of projects type. To focus the survey process on reasons that will make real problems to identify the responsible stakeholder and make the required future arrangements and suitable practical remedies.

The delay reasons of the progress report showed large diversity in the delay of focused projects on the categories of work. The classification includes new construction rehabilitation, modification, and maintenance. There are neglect, non-integrity, lack of seriousness, and even corruption. In real such cases, push to postpone or stopping the work in the construction project out of control of the management office. After collecting data, a prediction model was used to predict the projects that will stop. Two algorithm tools used support vector machine and decision tree.

3 Results and discussion

3.1 Survey results

Most of the periodic reports in the developing countries are absent in most work except the termly of annually tracking reports where there are no regular daily or weekly reports. Such negligence for Daily or weekly reports is the background to the emergence and evolution of problems in the project. The systematic procedure tracking daily and weekly will contribute to defining the natural causes of delay issues. They include the most required information such as project financings, values, parties, locations, and milestones. The data collected are shown in Table 1 and Table 2, which shows the projects considered to examine the main causes of delay, halt of work, and postponed. Table 3 shows that the factor of delay of payments represents one to third of the whole reasons.

Table 1: The projects of two years duration.

No	Description of the project	Duration [Days]	Value of contract [\$]	Situation	Delay time
1	Construction of tourism hotel	730	1956351	Stopped	1680
2	Minaret center of religious studies	730	2770402	Stopped	1550
3	Construct general director building	730	472339	Stopped	900
4	Construct elevated water tanks	730	200000	Stopped	2520
5	Sewerage treatment station	730	9050000	Stopped	0
6	Construct of General director	730	497201	Stopped	0
7	Construct director of the assembly	730	488350	Stopped	0
8	Construct director of the assembly	730		Working	180
9	Construct director of the assembly	730	488350	Stopped	990
10	Construct of center of Municipal guards	660	132780	Taking over	300
11	Renew water supply network city	660	8500000	Ongoing working	0
12	Construct a primary school 8 classes	540	487779	Taking over	2100
13	Culture center	480	109366	Ongoing Working	990
14	Youth houses	540	475376	Stopped	570
15	Construct doweling of physicians	540	294072	Taking over	240

Table 2: The projects of one-year duration.

No	Name of project	Duration [Days]	Value of contract [\$]	Situation	Extension	Delay time
1	Veterinary clinic	365	142179	Ongoing working		540
2	Agriculture Society	365	116133	Stopped		840
3	Center of Breeders society	365	116133	stopped	180	930
4	Construct tourism building	365	132067	stopped		660
5	Building of Evaluation & Measurement	365	112593	stopped		990
6	Construct a 8 class School	365	458351	Ongoing		840
7	Construct a 6 class School	365	422610	Complete taking over		none
8	Construct 7 classes with utilities	365	475873	Compete taking over		570
9	New classes of intermediate school	365	240851	Ongoing		90
10	Construction new Kindergarten	365	206588	Ongoing	-	-
11	Added new classes to religious school	365	206588	Stopped		990
12	Modification & develop of health center	365	231489	Complete taking over	180	420
13	Managerial complex in	365	497500	Complete		900
14	Public park	365	417096	Complete taking over		300
15	Streets lighting system of district	365	128313	stopped		540
16	Construct a backup dam	365	50000	stopped		1110
17	Execution of paved roads and expand	365	499150	stopped	270	1170
18	Execution of paved roads of internal	365	479700	stopped		
19	Central workshop of transportation	365	441965	Stopped	180	930
20	Fuel station	365	320038	Stopped		870
21	Court building	365	304322	Stopped		
22	Security building	365	227741	Stopped		210
23	Construct Al-Hana society	365	225512	Stopped		480
24	Construct a house of guests 1	365	135806	Stopped		300
25	Construct a house of guests- 2	365	103553	Taking over		365
26	Water supply station	365	240000	Stopped	180	1020
27	Drinking water network division 2	365	164052	Stopped		
28	Water supply station 2	365	243121	Stopped		
29	Water tank 1000 m ³	365	111854	Stopped	180	1440
30	Eight class school	365	378844	Stopped		1050
31	Medical clinic	365	248185	Stopped		1110
32	Guest house	365	235155	Stopped		960
33	Municipal Guard Center	365	162855	Working		60
34	Water supply station	365	309546	Stopped		900
35	Army recruitment building	365	213109	Stopped		720
36	Court & prosecutor building	365	294344	Stopped		1020
37	Establishment of animal feed stores	365	140972	Taking over		-
38	Construct school (8 classes)	365	401321	Stopped		1180
39	Construct dwelling of the doctors	365	89494	Stopped		1530
40	Expansion existing clinic	365	93891	Taking over		-
41	Construct utilities in a clinic	365	77452	Taking over		-
42	Construct the municipal guards office	365	162855	Stopped		420
43	Street lighting	365	370453	Stopped		300
44	Construct water pump station	365	240000	Stopped	180	960
45	Post office building	365	118355	Taking over		-
46	Communication line of two cities	365	372950	Stopped		780
47	Court & prosecutor building	365	299333	Stopped	180	990

48	Building of dwelling of police border	365	275057	Working	180	690
49	Construct underground water tanks 300 m ³	365	151577	working	-	510
50	Guard Fence around the border office	365	191172	Finish	-	1920
51	Construct underground water tanks 300 m ³	365	151577	Stopped	-	1050
52	Construct of sport team center	365	179934	Finished/Taking over		-
53	The lighting of enteral streets of the old city	365	147980	Finished/Taking over		1140
54	Building of projects management office	365	398907	stopped		
55	Fence of city graves	365	196353	stopped		780
56	Center of municipality guard building	365	168375	Working		60
57	Service bureau building	365	42173	Taking over		450
58	Post office building	365	117339	stopped		1200
59	Communication complex	365	450418	stopped		1020
60	Services and fence of high school	365	80795	Taking over		1050
61	Construct 6 classes school	365	425978	Working	180	750
62	Construction of a drinking water well	365	556990	Stopped		1560
63	Construct a training field	365	429218	Stopped	180	1590
64	Water supply network	365	220000	Stopped		840
65	Municipality guards camp	365	168375	Working		60
66	Management of residential affairs	365	133485	stopped	300	1440
67	Lighting of internal roads	365	492490	Stopped		1110
68	Construct a well for drinking water	365	103957	Stopped		2100
69	Establishment of a desert quarantine	365	91960	Stopped		900
70	Building agricultural reservation	365	139956	Stopped		
71	Drilling three agricultural wells	365	1110317	Stopped		2100
72	Construct the office of tourism assistance	365	8000		180	-
73	Established of Religious school	365	467323	Stopped		120
74	Construct celebration Square	365	220000	Stopped		1410
75	Water supply station	365	481190	Working		480
76	Post office	365	117331	Taking over	180	-
77	Popular assistance headquarters	365	133485	Stopped	180	1020
78	Headquarters / General Administration	365	133776	Stopped	300	1140
79	Primary school - Six classes	365	425978	Stopped	180	1470
80	Center of woman training	365	92027	Taking over		900
81	Elevated tank	365	228877	Taking over	180	330
82	Concrete pavements for highway	365	4689250			1170

Table 3: The real multi-reasons of causes of delay.

Causes	Main of in charge	Percentage
Delay of payments	Employer	34 %
Difference of opinion	Contractor	17 %
Stop work	Employer	11 %
Inability to implement	Contractor	10 %
Unregistered reasons	Both parties	9 %
Possession of site	Contractor	9 %
Termination	Client or mutual	3 %
Lack of seriousness	Contractor	3 %
Leave the job site	Contractor	2 %
Several simple reasons	Contractor	2 %

The most factor that relates to the delay in the construction industry is the delay of reimbursable the contractor's. It is caused mainly by the employer to be the most common problem in delaying projects in developing countries. The interim payment of main contractors will cover debits of suppliers, procurement parts, or even nominated subcontractors, will pay after a specified day of submitting the payments, which are prepared according to schedule and execute programs on site. But the most of the time, the employer in developing countries, especially when the contracts with local partners, try to ignoring and disregard such times and delay the payments of contractors due to matters that belong to him like funding or liquidity availability, therefore, as a consequence the contractor will halt the works because of liquidity. The least one is to leave the job site caused by the contractor, the repeated unjustified stopping of the work, without permission or notice from the client's representatives.

3.2 Machine learning results

The Navis Base and Decision tree algorithm were used to predict the stopping time of the projects. At the same time, The RAPID MINER program was used to implement these algorithms and the results as shown in Fig. 3. The data was loaded, and by trial and error, the splitting ratio was conducted as 75:25 to give the maximum accuracy.

Machine learning, naïve Bayes classifiers are a private of simple "probabilistic classifiers" depend on relating Bayes' formula with robust (naïve) independence norms among the features. They are amongst the greenest Bayesian network models. [17]. The results were the accuracy about 77 % and the confusion matrix as shown in Figures 4 and 5. In comparison, the using of the decision tree the results as shown in Fig. 6. The decision tree shows an accuracy of about 89 %, which is much better than Naïve Bayes due to the type of data string.

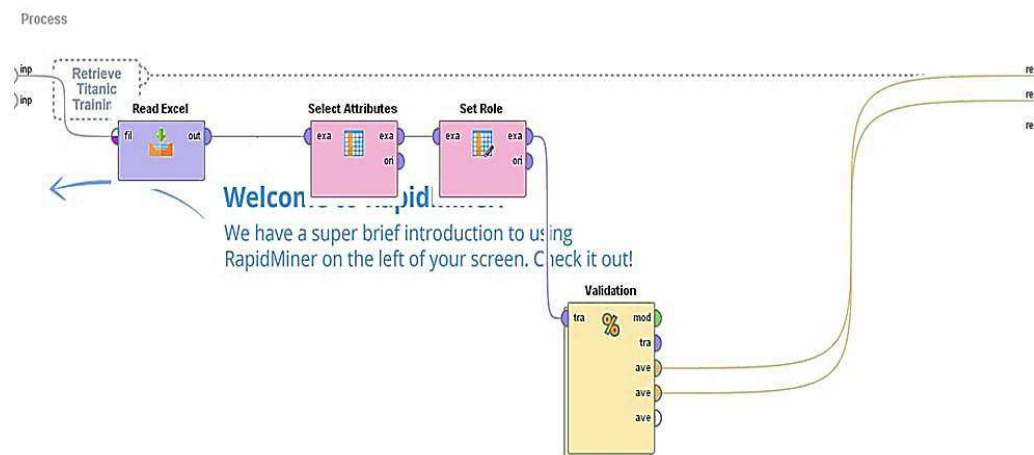


Fig. 3: Load and set the splitting data in.

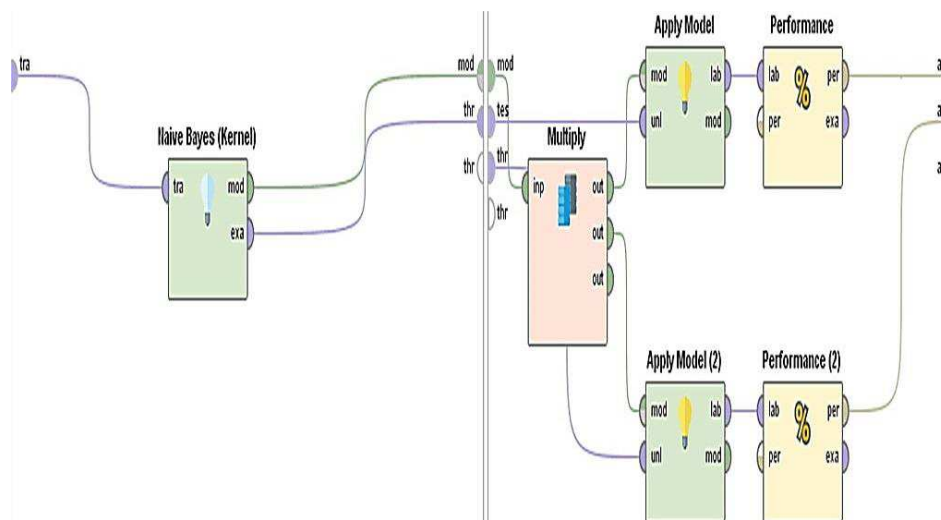


Fig. 4: Naïve Bayes in the program.

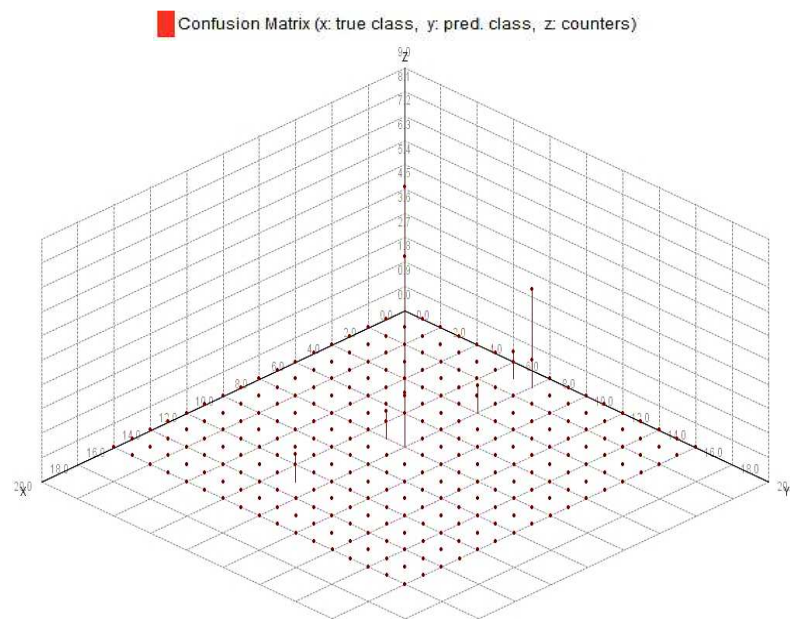


Fig. 5: Confusion matrix.

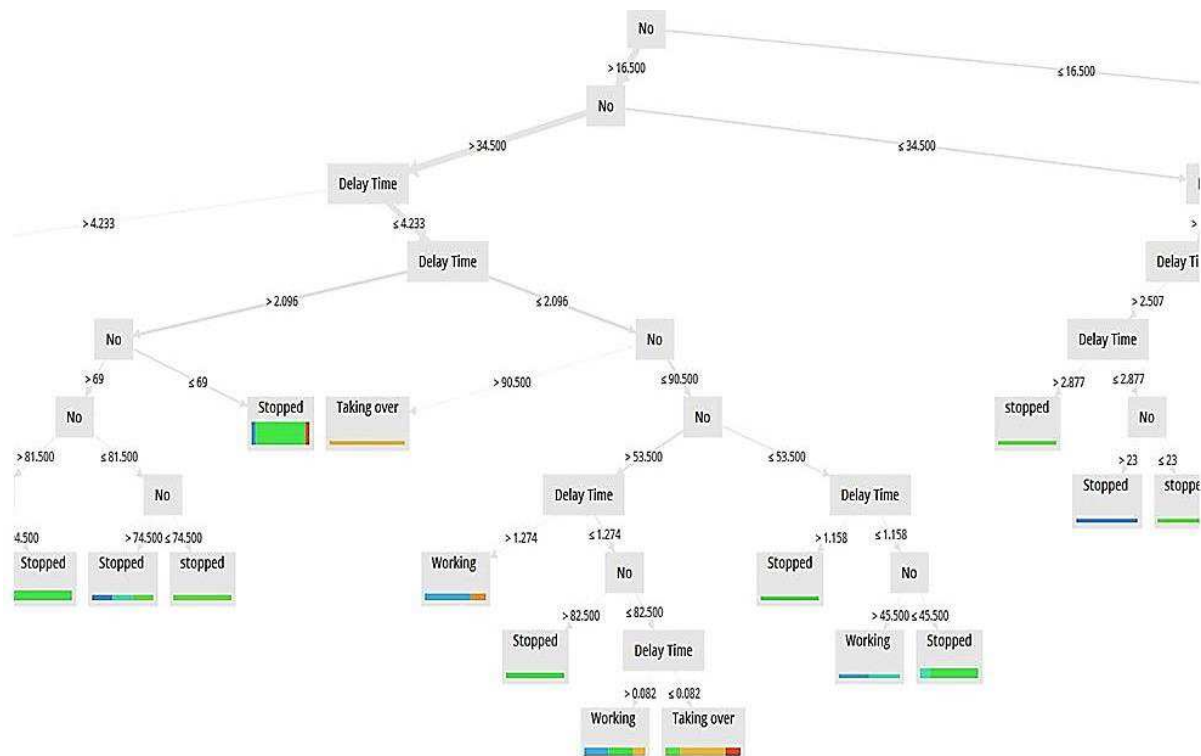


Fig. 6: Decision tree.

The results show that the decision tree is the better algorithm in predicting, and these results agree with Yaseen and others (2020) [21], Naji and others (2018) [20].

4 Conclusion

The current revision donated to this attempt by classifying and relating ML algorithms to mature two construction project delay, predictive models, depending on decision tree and naïve Bayesian classification algorithms. The developed model could contribute to decreasing the time of delay and consequently could reduce cost as well. The most factor that relates to the delay in the construction industry is a delay of payment of contractor's compensations caused by the employer. The model was applied to about 97 projects and used a prediction model, where the decision tree showed higher accuracy in the prediction.

References

- [1] FRIMPONG, Y. - OLUWOYEB, J. - CRAWFORD, L.: Causes of delay and cost overruns in construction of groundwater projects in a developing country. *International Journal of Project Management*, Vol. 21, Iss. 5, 2003, pp. 321-326.
- [2] BILAL, M. - OYEDELE, L. - QADIR, J. - MUNIR, K. - AJAYI, S. - AKINADE, O. - OWOLABI, H. - ALAKA, H. - PASHA, M.: Big Data in the construction industry: A review of present status, opportunities, and future trend. *Advanced Engineering Informatic*, Vol. 30, Iss. 3, 2016, pp. 500-521.
- [3] CHAN, D. - KUMARASWAMY, M.: Compressing construction durations: lessons learned from Hong Kong building projects. *International Journal of Project Management*, Vol. 20, Iss.1, 2002, pp. 22-35.
- [4] ELAZOUNI, A. M.: Classifying construction contractors using unsupervised-learning neural networks. *Journal of Construction Engineering and Management*, Vol. 132, Iss.12, 2006, pp. 1242-1253.
- [5] CHAO, L. C. - CHIEN, C. F.: Estimating project S-curves using polynomial function and neural networks. *Journal of Construction Engineering and Management*, Vol.135, Iss.3, 2009, pp. 169-177.
- [6] HERAVI, G. - ESLAMDOOST, E.: Applying artificial neural networks for measuring and predicting construction-labor productivity. *Journal of Construction Engineering and Management*, Vol. 141, Iss. 10, 2015, 04015032.
- [7] DESAI, V. S. - JOSHI, S.: Application of decision tree technique to analyze construction project data. *International Conference on Information Systems, Technology and Management ICISTM 2010: Information Systems, Technology and Management*, pp. 304-313.
- [8] CHI, S. - SUK, S. - KANG, Y. - MULVA, S.: Development of a data mining-based analysis framework for multi-attribute construction project information. *Advanced Engineering Informatics*, Vol. 26, Iss. 3, 2012, pp. 574-581.
- [9] CHEUNG, S. O. - YIU, T. W. - CHAN, H. W.: Exploring the potential for predicting project dispute resolution satisfaction using logistic regression. *Journal of Construction Engineering and Management*, Vol. 136, Iss. 5, 2010, pp. 508-517.
- [10] ALZHRANI, J. I. - EMSLEY, M. W.: The impact of contractors' attributes on construction project success: A post construction evaluation. *International Journal of Project Management*, Vol. 31, Iss. 2, 2013, pp. 313-322.
- [11] JIANG, X. - MAHADEVAN, S.: Bayesian probabilistic inference for nonparametric damage detection of structures. *Journal of Engineering Mechanics*, Vol. 134, Iss. 10, 2008, pp. 820-831.
- [12] GONG, J. - CALDAS, C. H. - GORDON, C.: Learning and classifying actions of construction workers and equipment using Bag-of-VideoFeature-Words and Bayesian network models. *Advanced Engineering Informatics*, Vol. 25, Iss. 4, 2011, pp. 771-782.
- [13] CHENG, M. Y. - WU, Y. W.: Evolutionary support vector machine inference system for construction management. *Automation in Construction*, Vol. 18, Iss. 5, 2009, pp. 597-604.
- [14] LAM, K. C. - LAM, M. C. K. - WANG, D.: Efficacy of using support vector machine in a contractor prequalification decision model. *Journal of Computing in Civil Engineering*, Vol. 24, Iss. 3, 2010, pp. 273-280.
- [15] GONDIA, A. - SIAM, A. - EL-DAKHAKHNI, W. - NASSAR, A. H.: Machine Learning Algorithms for Construction Projects Delay Risk Prediction. *Journal of Construction Engineering and Management*, Vol. 146, Iss. 1, 2019, 04019085.
- [16] Mc CALLUM, A.: Graphical Models, Lecture2: Bayesian Network Representation. Retrieved 22 October 2019.
- [17] GHAZVINI, K. - ZANDIEH, M. - VAFAMEHR, M.: Exploring KPIs Utilization Effects on Decision Making for the Architectural Design Process in Industrial Buildings. *Civil and Environmental Engineering*, Vol. 16, Iss. 1, 2020, pp. 198-209.
- [18] ŠESTAKOVÁ, J. - IŽVOLT, L. - MEČÁR, M.: Degradation - Prediction Models of the Railway Track Quality. *Civil and Environmental Engineering*, Vol. 15, Iss. 2, 2019, pp. 115-124.
- [19] NAJI, H. I. - AL-ZUBAIDI, E. A. - ALI, R. H.: Risk Response Failure in Construction Projects, 2018.
- [20] YASEEN, Z. M., et al.: Prediction of risk delay in construction projects using a hybrid artificial intelligence model. *Sustainability*, 2020, 12.4: 1514.