Analyzing the Factors Affecting the Soil Investigation Cost

Hafeth I. Naji
Department of Civil Engineering, University of Diyala, Baqubah, Iraq

Abstract: Geotechnical site investigation with appropriate planning and management consider significant to gain adequate and correct site information for designing a foundation of the building in a timely manner and with minimum cost for the needed effort. The important of this study is to analyze the factors effecting on soil investigations and its impact on the cost of the projects as there is strong relationship between the soil investigation cost and the cost of the projects and more importantly. The methodology of the study divided in two part, questionnaire and the use the techniques of data mining including fuzzy decision tree. The soil investigation consider very important in the construction projects have direct impact on the cost of the project as lack of topographic and geological maps of the region lead to misleading of what type of soil we dealing with hence thus will lead to inappropriate selection of the toll and equipment. The fuzzy decision tree show the higher accuracy and the variable importance flow show that the personal relationship has direct impact on cost.

Key words: Soil, soil investigation, cost, fuzzy logic, decision tree, KNIME

INTRODUCTION

The main part of the preliminary design work is site investigation or soil survey as consider very important in structure to insure the gathering of information with respect to the arrangement of strata and the level of ground water andalso to gather the samples for identification and testing. Additionally, a site investigation is often essential to measure the safety of an current structure or to investigate a case where failure has appeared (Simons *et al.*, 2002).

It is significant that before the program of field investigation is established, a preliminary site description depending on published literature and clients and information by the consultants be prepared. The site description should involve climate and ground temperature at commissioning of the building and at the end of its service life.

The goal of site exploration is to obtain information about the following: the extent and sequence of every soil and rock stratum in the area likely to be influenced by the proposed works, stratum nature and engineering soil properties and rock which may influence the design and method of construction of proposed structure and its foundation, the ground water location and possible corrosive consequence of soil and water on foundation materials (Ogallo, 2015). Like the thumb rule, there are no obligatory rules on either numbers or borings depths. It will depend on the judgment of engineering and site discovery. In common and to estimate the concept of boring layout planning (the depth specifically) (Al-Hashemi, 2016).

The important of this study is to analyze the factors effecting on soil investigations and its impact on the cost of the projects as there is strong relationship between the soil investigation cost and the cost of the projects and more importantly, the factors that effect on the soil investigation as the consider the success factors of the investigation hence success of the projects.

Literature review: Geotechnical site investigation with appropriate planning and management consider significant to gain adequate and correct site information for designing a foundation of the building in a timely manner and with minimum cost for the needed effort. The power and detail of the program of geotechnical site investigation in permafrost is composite. It be influenced by on: proposed building design criteria of the, previous knowledge of general site conditions and existing buildings performance, availability of equipment for drilling, time of the year the work needs to be done that may control the geotechnical site investigation technique and costs (Anonymous, 2010).

The factors associated with the ground have often been the source of contractual claims with important time and cost overruns on equally large and small construction projects. About 80 and 85% failures of the building and its damages, according to European statistics are related to unexpected and unfavorable ground conditions. Without satisfactory site investigation, clients are always exposed to the risk of costly delays, redesign and late delivery of the project arising out of unforeseen ground conditions. (Littlejohn, 1994). In 1992, a report was gathered by the government displayed that the final total cost comparing

with tender values for accomplished highway contracts sat show increased about 28%. Additional analysis of 17 of these contracts indicated 44% of the cost increase was related to earthworks and unexpected ground conditions (Clayton *et al.*, 2000). Likewise, the Department of the Environment, Transport and the Regions (DETR) annual report, emphasized that the seven biggest road projects were some £516 million over budget, due largely to unexpected ground conditions (Anonymous, 1998). This associated to an over-spend which accounted for a massive 63% increase in projected expenditure (Jones, 1998).

Hytiris made study to show the importance of the site investigation and its impact on the cost, the results of his study showed that suitable site investigations for low-rise buildings should in fact be a minimum of 0.42% of the cost of a project.

Therefore, this study consider important as reveal the factors effect on the soil investigation and hence effecting on the cost of the projects, the factors were analysis using a combination of the two techniques, fuzzy logic with decision tree using KNIME program.

The KNIME program show good performance, (Naji and Ali, 2017) used fuzzy decision tree to analysis the risk, the accuracy of the technique was about 91 and that indicate good accuracy. Al-Zubaidi *et al.* (2006) state that by using decision tree in WEKA program to analyze the risk in the construction projects, the technique shows very good accuracy in classifying the risks and the accuracy was about 92.7. Based on the previous the decision tree show very god technique in the analysis. One of the important techniques of analysis is fuzzy logic the originator of fuzzy logic is Lotfi Zadeh. Significant advancement was made by him in the stabilization of fuzzy logic as a scientific discipline. Fuzzy logic not a unique system of knowledge instead is a variety of methodologies suggesting logical consideration of knowledge that imperfectly and vaguely.

The definition a Fuzzy Logic System (FLS) is the nonlinear mapping of data set that consider an input to a scalar the data of output. A FLS contain four fundamental parts: fuzzifier, rules, inference engine and defuzzifier (Fig. 1).

Gradient boosting is technique used in a machine learning for both problems of classification and regression. It forms the model in a fashion of stage-wise like former boosting methods do and it simp lifies them by permitting optimization of an arbitrary differenti able loss function.

Similar to other boosting methods, gradient boosting goatherd weak "learners" into a one strong learner in fashion of an iterative minimizing.
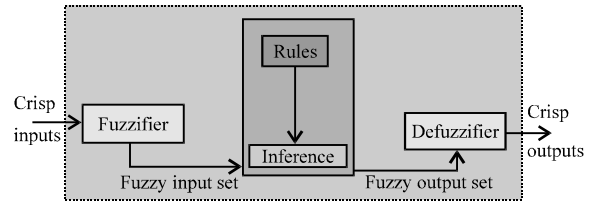

Fig. 1: A fuzzy logic system

MATERIALS AND METHODS

The methodology of the study divided in two part, questionnaire and the use the techniques of data mining. The questionnaire was distributed to the owners, the contractor and other parties involved in the project, 30 projects with 30 responder were taken. The questionnaire includes the factors effects on the soil investigation cost in construction projects, these factors were gathered from literature review and then presented on the responder in order to gather their opinion, five measurements were used which are too low, low, medium, high and very high shown in Table 1, the factors of the project are shown in table and the projects were taken in the periods 2006-2016, the program that use for the analysis is KNIME, the methodology shown in Fig. 2. The results of the questionnaire shown in Table 1 KNIME program KNIME (pronounced/nam/), the konstanz information miner is anopen source for data analysis, reporting and integration platform. Several mechanisms can be merging using KNIME indata mining and machine learningby using the concept of data pipelining modular. Agraphical user interfacepermits assembly of nodes for the preprocessing of data (ETL: Extraction, Transformation, Loading) for data analysis, modeling and visualization. Since, 2006, pharmaceutical research was the area of KNIME, however, it can use in different areas like CRM customer data analysis, financial data analysis and business intelligence (Tiwari and Sekhar, 2007). The measurement that used to show the effect on the quality and these measurement are giving by the respond as shown in Table 2. The effect was taken by dividing the additional period on the contract periods.

Factors effect on cost of soil investigation:
- F1: lack of topographic and geological maps of the region
- F2: personal relationship the site
- F3: type of Soil and rock
- F4: natural moisture content
- F5: machinery, equipment used, types and models thereof
- F6: soil layers and their different types

Table 1: The norms and standard

Scales	Numericals
Very high	5
High	4
Medium	3
Low	2
Very low	1

Table 2: The norms and standard for factors effect on cost of soil investigation

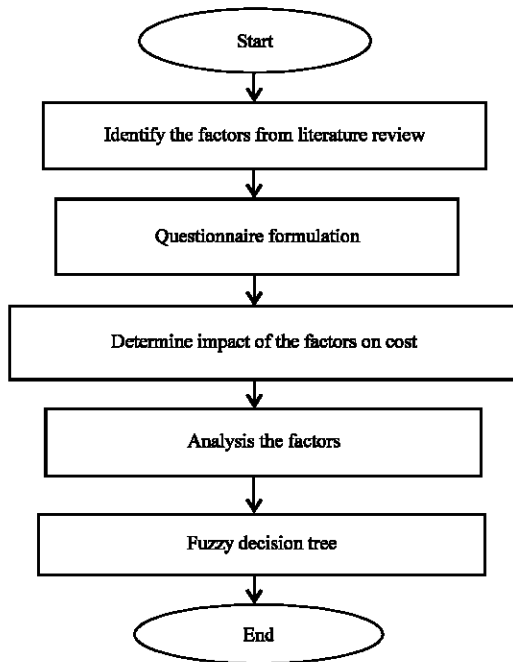| Cost	Effects on cost
More than 217500000	Very high
217500000-6000000	High
5000000-3250000	Medium
3000000-1000000	Low
>1000000	Very low

Fig. 2: The methodology of the research

- F7: the project site
- F8: project type
- F9: type of foundation used and quantity of load allowed
- F10: safety available at the project site
- F11: the absence of previous soil studies, agricultural soil reports and other important studies conducted on
- F12: type of the test

RESULTS AND DISCUSSION

The first step of factors effect on soil investigation analysis and identify them using the questionnaire, it's was distributed over 30 experts who previously work in the projects. The questionnaire was gathered and then formulation of the data, a 5-point scale as follows: very high, high, medium, low and very low. The impacts were calculated based on Eq. 1:

$$\text{Mean}\left(\overline{X}\right) = \sum_{i=1}^{h} x_i \times \frac{f_i}{n} \qquad (1)$$

Where:
$\overline{X}$ = Mean
x_i = Class center
f_i = The No. of iterations for each class
n = Total sample size or duplicates of the varieties
i = Sequence of class
h = Number of class

As a result of the questioner of the impact of the factors were determined. The results of the questioner are shown in the appendix. The second step involve entering the data to the program and start the analysis process for the analysis fuzzy decision tree (gradient boosted trees) techniques were used in KNIME program, this involve the following steps.

First (acess to the program): When you run the program, the program appears as shown in Fig. 3.

Second (the program selection list): The second step to start new workflow and select the node to insert the data from the IO node select the fuzzy decision tree in KNIME and start the process analysis to determine the impact on the cost of soil investigation (Fig. 4-6). The learner statistic used for the fuzzy as.

Learner statistics:
- Number of epochs: 3
- Number of classes: 2
- Number of rules learned per class (in total 4)
- Hight: 1
- Medium: 3
- Number of training instances per class (in total 45)
- High: 26
- Medium: 19

The accuracy of the technique as follow Fig. 6 depending on the above and the ROC curve its obviously that the medium class effect on the accuracy, however, 94 consider a very good one but to be more specific a researcher flow is made to show the variable importance. The workflow above show the effect of each variable on the cost on several of iteration in Fig. 7, Table 4 and 5. From the above its clearly show that the personal relationship has direct impact on the cost which case reducing the accuracy.
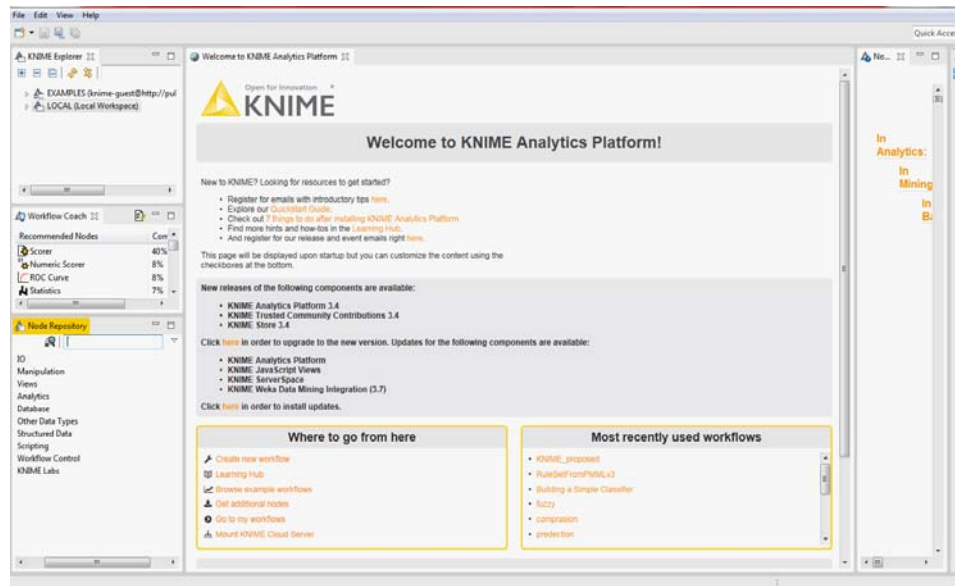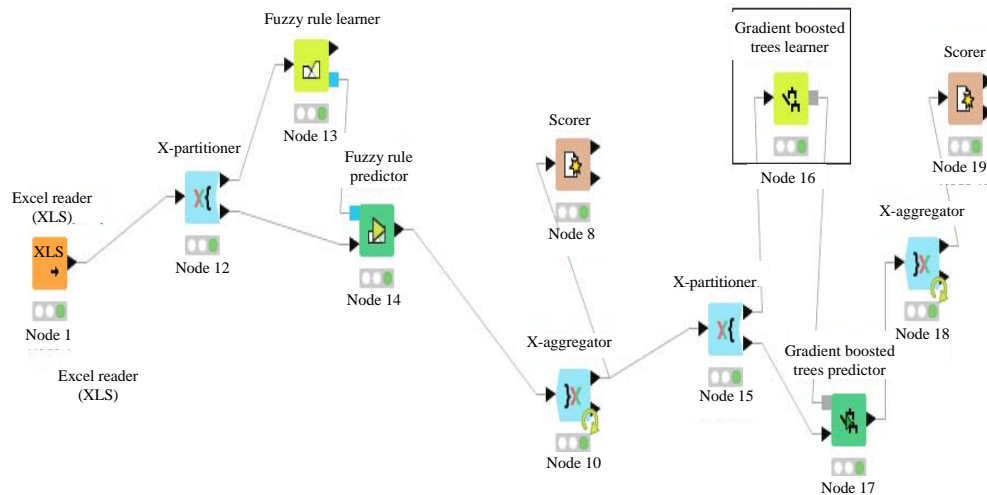
Fig. 3: The main interface of the program

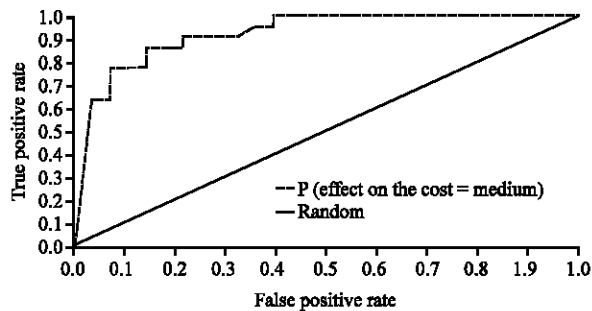Fig. 4: The work flow of fuzzy decision tree

Fig. 5: The ROC curve of fuzzy decision tree for medium class of time; P (effect on the cost = medium) (0.923)
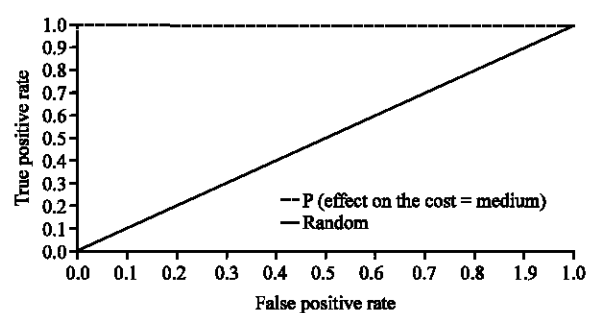
Fig. 6: The ROC curve of fuzzy decision tree for high class of time; P (effect on the cost = high) (1)
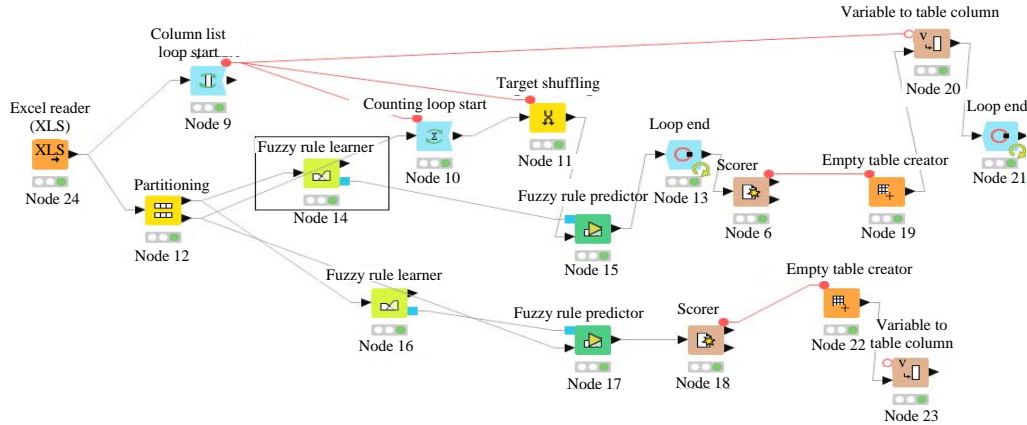
Fig. 7: The variables importance

Table 4: The accuracy if the fuzzy decision tree

Row ID	True (Po)	False (Po)	True (Ne)	False (N)	Recall	Precision	Sensitivity	Specifty	F-means	Accuracy	Cohen
High	27	1	20	2	0.931	0.964	0.931	0.952	0.947	?	?
Medium	20	2	27	1	0.952	0.909	0.952	0.931	0.930	?	?
Overall	?	?	?	?	?	?	?	?	?	0.94	0.878

Table 5: The variable importance of personal relationship

Row ID	Current	Cohen	Error	Accuracy	Iteration
9#3	f2qn	-0.043	0.302	0.698	3
15#3	f2qn	-0.043	0.302	0.698	3
20#3	f2qn	-0.043	0.302	0.698	3
23#3	f2qn	-0.043	0.302	0.698	3
33#3	f2qn	-0.043	0.302	0.698	3
37#3	f2qn	-0.043	0.302	0.698	3
42#3	f2qn	-0.043	0.302	0.698	3
51#3	f2qn	-0.043	0.302	0.698	3
57#3	f2qn	-0.043	0.302	0.698	3
63#3	f2qn	-0.043	0.302	0.698	3
70#3	f2qn	-0.043	0.302	0.698	3
74#3	f2qn	-0.043	0.302	0.698	3
75#3	f2qn	-0.043	0.302	0.698	3
76#3	f2qn	-0.043	0.302	0.698	3
84#3	f2qn	-0.043	0.302	0.698	3
92#3	f2qn	-0.043	0.302	0.698	3
101#3	f2qn	-0.043	0.302	0.698	3
102#3	f2qn	-0.043	0.302	0.698	3
104#3	f2qn	-0.043	0.302	0.698	3
111#3	f2qn	-0.043	0.302	0.698	3
112#3	f2qn	-0.043	0.302	0.698	3
119#3	f2qn	-0.043	0.302	0.698	3
127#3	f2qn	-0.043	0.302	0.698	3
137#3	f2qn	-0.043	0.302	0.698	3
138#3	f2qn	-0.043	0.302	0.698	3
141#3	f2qn	-0.043	0.302	0.698	3
146#3	f2qn	-0.043	0.302	0.698	3
148#3	f2qn	-0.043	0.302	0.698	3
152#3	f2qn	-0.043	0.302	0.698	3
154#3	f2qn	-0.043	0.302	0.698	3
158#3	f2qn	-0.043	0.302	0.698	3
168#3	f2qn	-0.043	0.302	0.698	3
177#3	f2qn	-0.043	0.302	0.698	3

CONCLUSION

The soil investigation consider very important in the construction projects have direct impact on the cost of the project as lack of topographic and geological maps of the region lead to misleading of what type of soil we dealing with hence thus will lead to inappropriate selection of the toll and equipment.

Type of soil and rock have direct impact on the investigation cost's if the soil contain rock require special attention and it was sandy require different equipment and depth on the other hand the project site play a vital role as if the location is far require extra cost for traveling the equipment and labor. Natural moisture content effect on the depth of borehole as if the moisture content is high it will require deeper depth. The fuzzy decision tree show the higher accuracy and the variable importance flow show that the personal relationship has direct impact on cost.

REFERENCES

Al-Hashemi, H.M.B., 2016. Literature review: Site investigation and boring layout. Master Thesis, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.

Al-Zubaidi, E.A., H.I. Naji and R.H. Ali, 2006. Descriptive classification of cost risks in construction projects. ARPN. J. Eng. Appl. Sci., 12: 4383-4389.

Anonymous, 1998. Major roads projects clock up £516M overspend. Tarmac International Inc., Lee's Summit, Missouri. https://www.newcivilengineer.com/major-roads-projects-clock-up-516m-overspend/842575.article.

Anonymous, 2010. Geotechnical site investigation guidelines for building foundations in permafrost. I. Holubec Consulting Inc., Oakville, Ontario. https://www.inf.gov.nt.ca/sites/inf/files/geotechnical_site_investigation_guidelines_for_building_foundations_in_permafrost.pdf.

Clayton, C.R.I., M.C. Matthews and N.E. Simons, 2000. Site Investigation. 2nd Edn., University of Granada Press, Granada, Spain.

Jones, M., 1998. Difficult ground: The biggest excuse in the book large cost overruns due to insufficient siteinvestigation still dog the construction industry. Tarmac International Inc., Lee's Summit, Missouri. https://www.newcivilengineer.com/difficult-ground-the-biggest-excuse-in-the-book-large-cost-overruns-due-to-insufficient-site-investigation-still-dog-the-construction-industry-matthew-jones-asks-why/84268.

Littlejohn, S., 1994. Ground: Reducing the risk briefing. Proc. Inst. Civ. Eng., 102: 3-4.

Naji, H.I. and R.H. Ali, 2017. Fuzzy decision tree of risks assessment generated from risk response. Intl. J. Appl. Eng. Res., 12: 10225-10232.

Ogallo, A.F., 2015. Geotechnical engineering investigation report for proposed office complex and slope stability assessment at Kakamega township for one acre fund: A case study. Ph.D Thesis, University of Nairobi, Nairobi, Kenya.

Simons, N.E., B.K. Menzies and M.C. Matthews, 2002. A Short Course in Geotechnical Site Investigation. Vol. 5, Thomas Telford Ltd., Westminster, England, UK., Pages: 357.

Tiwari, A. and A.K. Sekhar, 2007. Workflow based framework for life science informatics. Comput. Biol. Chem., 31: 305-319.