



Comparative Comparison of Contractors' Evaluation Criteria, Sub-Criteria and Indicators in Water Industry Tender Using Questionnaire and Text Mining

M. Parchami Jalal*, H. Shakiba Zahed, M. A. Ghaaderi

Faculty of Architecture, University of Tehran, Tehran, Iran

ABSTRACT: This research investigates the collapse responses of a concrete moment frame considering modeling uncertainties. These modeling uncertainties are considered for evaluating a collapse response related to the modified Ibarra-Krawinkler moment-rotation parameters for beam and column elements of a given structure. To analyze these uncertainties, the correlations between the model parameters in one component and between two structural components were considered. Latin Hypercube Sampling (LHS) method was employed to produce independent random variables. Moreover, Cholesky decomposition was adopted to produce correlated random variables. Performing 281 simulations for the uncertainties involved considering their inter-correlations, incremental dynamic analysis (IDA) was done using 44 far-field accelerograms to determine structural collapse responses. Collapse responses of each simulation, including mean collapse capacity, mean collapse drift and mean annual frequency, were obtained. Then, the collapse responses were predicted using the response surface method and artificial neural network. The results show that the Correlation coefficients (R) between the target data resulted from incremental dynamic analysis (IDA), output data resulted from response surface method (RSM), and artificial neural network (ANN) were obtained for the collapse responses above 0.98. The maximum prediction errors for mean collapse capacity and mean collapse drift are less than 5% and for mean annual frequency less than 10% under the response surface method (RSM) and artificial neural network (ANN).

Review History:

Received: Nov. 15, 2019
Revised: Jan. 21, 2020
Accepted: Feb. 05, 2020
Available Online: Mar. 02, 2020

Keywords:

Construction Evaluation Criteria
Text Mining
Data Mining
Clustering
Tender

1- Introduction

Today, choosing the most suitable contractor is one of the most important decisions for employers. Contractor evaluation requires evaluation criteria and indicators [1]. In most of the researches used to identify contractors 'evaluation criteria, a questionnaire tool has been used to identify contractors' evaluation criteria [2] [3]. One of the disadvantages of identifying contractors' evaluation criteria using the questionnaire is that the way the questionnaire is designed and the factors that are mentioned may influence the audience's orientation and affect the way the questionnaire is filled. Compensating for this weakness led the researchers to identify contractors' evaluation criteria and indicators using two complementary tools, a questionnaire and text mining. The overall process of the research is shown in Fig. 1.

2- Methodology

This research consists of three parts. In the first part, contractors' evaluation criteria were identified using library studies and questionnaire tools and In the second part, water industry experts were interviewed and text mining operations were conducted on the Interview transcripts. Then they were compared, and finally, the findings of the research were

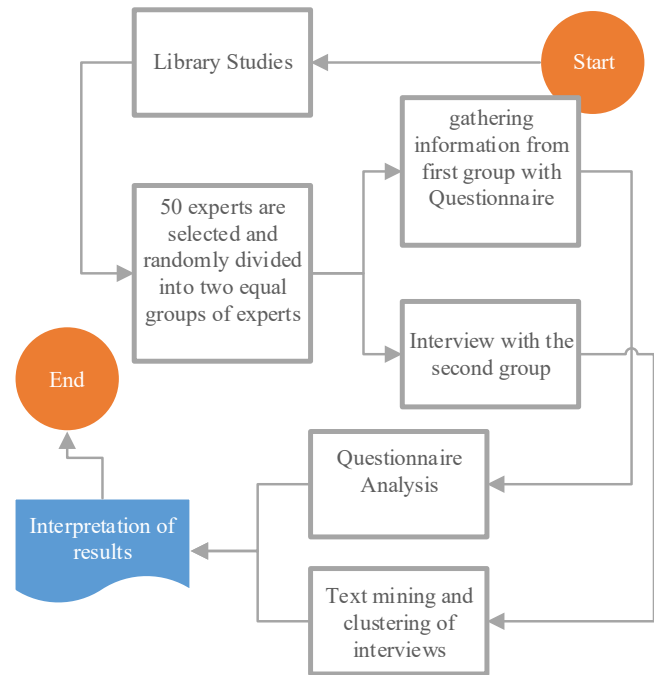


Fig. 1. Research process

*Corresponding author's email: parchamijalal@ut.ac.ir

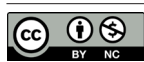


Table 1. Evaluation main-criteria and sub-criteria and their weight

Main-Criteria	Weight	Sub Criteria	Weight
Work Capacity	14	-	-
Technical skills and key staff skills	14	Technical Skills	36
		Key Staff	32
		Understanding and regulations	32
Performance in past projects and Organizational experience	14	Performance in past projects	50
		Organizational experience	50
Claim Management	13	-	-
Financial situation	12	-	-
Management skills	12	Project Management Skills	50
		Innovative technical solutions	50
Good reputation	11	-	-
HSE and Quality Control	10	HSE	50
		Quality Control	50

compared with the bidding law implementing regulations in Iran. In other words, contractors' evaluation criteria and indicators are identified using a questionnaire, statistical analysis, and text mining, and after comparing the results, the final criteria are presented. The statistical population consists of 78 people working in the marketing offices of a consulting firm. 64 industry experts from two different companies (to ensure the generalizability of the research results) were identified and invited, and 50 of them agreed to participate in the research. The experts were randomly divided into two equal groups. Opinions of the first group were collected through questionnaires and statistical analysis, and the second group's opinions were collected by interview and text mining tools.

3- Discussion and Results

In this study, the Delphi method was used to design the questionnaire. Pairwise comparisons were used to determine the significance ratio of each of the main criteria in Expert Choice software. To determine the significance of the sub-criteria, respondents were asked to assign a number from 1 to 7 to each sub-criteria. The number 1 represents the least important, and the number 7 represents the most important. Then the mean scores assigned to each sub-criteria were calculated and considered as the significance index. Seven main evaluation criteria and 14 sub-criteria and their weight were identified by questionnaire

In the data preprocessing step, stop words were removed. To perform text mining operations, the textual data must first be converted to numerical vectors to be processed by software [4]. This way, unstructured data is converted to structured data. With the "Generate TFIDF" operator, words are converted to numerical vectors, and the text of the interviews is converted to a numerical matrix. TF-IDF is a statistical proposition which means the weight of the words that represents the importance of the words in the text. After calculating the TF-IDF matrix, a clustering operation was performed on this matrix using the K-means algorithm. The value of K, which represents the number of clusters, was

determined from 2 to 10, respectively, and each clustering result was recorded, and by comparing them, it was concluded that $K = 6$ produced the most appropriate cluster composition [5]. After performing the clustering operation, the data were subdivided into 6 different clusters, and the keywords in each cluster were identified using two criteria, TF-IDF value and cluster center value for each word [6]. The topic of each cluster was considered as the main criteria and it was found that the topics of the 5 clusters are highly correlated with the 5 main criteria identified by the questionnaire, and the sixth cluster can be considered as new criteria. In the next step, by analyzing the keywords of each cluster, the evaluation indices for each of the criteria and sub-criteria are determined.

The high consistency and consistency of the results of the first tool (questionnaire and statistical analysis) and the second tool (interview and text mining) indicate that contractors' evaluation criteria, sub-criteria, and indices are extracted with acceptable accuracy and generalizability.

4- Conclusions

In this study, the main criteria, sub-criteria, evaluation Indicators, and their weights were identified using two text mining and interviewing tools (Table 1). The main criteria identified in this study are Work Capacity, Technical skills, Performance on previous projects, Claim Management, Financial situation, project management skills, reputation, and HSE management. The results of the study, in addition to covering all the criteria presented in the bidding law implementing regulations in Iran, have added the following items:

- Claim Management and HSE management as the two main criteria
- Three evaluation Indicators to assess claims management ability and eight evaluation Indicators to assess safety and health management ability
- Determine the weight of the main criteria, sub-criteria, and evaluation Indicators specifically for the water bidding industry.

References

- [1] M. Sönmez, J.B. Yang, G.D. Holt, Addressing the contractor selection problem using an evidential reasoning approach, *Eng. Constr. Archit. Manag.* 8 (2001) 198–210.
- [2] D.J. Watt, B. Kayis, K. Willey, Identifying key factors in the evaluation of tenders for projects and services, *Int. J. Proj. Manag.* 27 (2009) 250–260.
- [3] J. Yang, W. Wang, Contractor selection by the most advantageous tendering approach in Taiwan, *J. Chinese Inst. Eng.* 26 (2003) 381–387.
- [4] J.P. J. Han, M. Kamber, *Data Mining concepts and techniques*, Elsevier, 2012.
- [5] N.T. Ratrou, Subtractive clustering-based k-means technique for determining optimum time-of-day breakpoints, *J. Comput. Civ. Eng.* 25 (2010) 380–387.
- [6] M. Al Qady, A. Kandil, Automatic clustering of construction project documents based on textual similarity, *Autom. Constr.* 42 (2014) 36–49.

HOW TO CITE THIS ARTICLE

M. Parchami Jalal, H. Shakiba Zahed, M. A. Ghaaderi, *Comparative Comparison of Contractors' Evaluation Criteria, Sub-Criteria and Indicators in Water Industry Tender Using Questionnaire and Text Mining. Amirkabir J. Civil Eng., 53(6) (2021)521-524*

DOI: [10.22060/ceej.2020.17388.6547](https://doi.org/10.22060/ceej.2020.17388.6547)





مقایسه تطبیقی معیارها، زیرمعیارها و شاخص های ارزیابی پیمانکاران در مناقصات صنعت آب با استفاده از ابزار پرسشنامه و متن کاوی

مجید پرچی جلال^{*}، هادی شکبیا زاهد، محمد امین قادری

دانشکده مهندسی معماری، دانشگاه تهران، تهران، ایران

تاریخچه داوری:

دریافت: ۱۳۹۸/۰۸/۲۴

بازنگری: ۱۳۹۸/۱۱/۰۱

پذیرش: ۱۳۹۸/۱۱/۱۶

ارائه آنلاین: ۱۳۹۸/۱۲/۱۲

کلمات کلیدی:

معیارهای ارزیابی پیمانکاران

متن کاوی

داده کاوی

خوشه بندی

مناقصه

خلاصه: همواره کارآمدی روش واگذاری طرح های زیر بنایی به پیمانکاران یکی از دغدغه های اصلی کارفرمایان می باشد. هدف از این پژوهش، پیشنهاد روشی جدید برای اختصاص نمودن معیارهای ارزیابی، با توجه به نوع پروژه های تحت واگذاری و اولویت بندی و تعیین امتیاز پیشنهادی معیارهای ارزیابی پیمانکاران با استفاده از دو روش مختلف می باشد، روش اول شامل تحلیل های آماری بر روی پرسشنامه ها و روش دوم شامل متن کاوی مصاحبه ها است. برای این منظور، پس از مطالعات کتابخانه ای، میدانی و طراحی پرسشنامه، پرسشنامه ها در اختیار خبرگان این صنعت قرار گرفت و با تحلیل آماری آن ها، معیارها و زیرمعیارهای مهم تعیین شد و سپس با گروهی دیگر از خبرگان به صورت باز مصاحبه انجام شد. با متن کاوی متون مصاحبه ها و خوشه بندی آن ها، معیارها و شاخص های ارزیابی پیمانکاران شناسایی شده و نتایج این دو روش با آیین نامه اجرایی بند "ج" ماده ۱۲ قانون برگزاری مناقصات، مقایسه گردید. برای تحلیل پرسشنامه ها از نرم افزار SPSS و برای متن کاوی مصاحبه ها از الگوریتم K-means بهره گرفته شده است. یافته های پژوهش بیانگر همسویی نتایج متن کاوی و پرسشنامه می باشد. از همسویی نتایج متن کاوی و پرسشنامه می توان نتیجه گرفت که معیارهای شناسایی شده از دقت و قابلیت تعمیم قابل قبولی برخوردارند و سازمان ها می توانند از این معیارها و شاخص های ارزیابی برای انتخاب مناسب ترین پیمانکار در مناقصات استفاده کنند. معیارهای شناسایی شده توسط ابزار متن کاوی و پرسشنامه، علاوه بر پوشش تمامی معیارهای معرفی شده در آیین نامه اجرایی قانون برگزاری مناقصات، دو معیار جدید تحت عنوان مدیریت دعاوی و بهداشت، ایمنی و محیط زیست و کنترل کیفیت را نیز شامل می شوند.

۱- مقدمه

با توجه به ترجیح کارفرمایان به واگذاری طرح از طریق مناقصه، انتخاب مناسب ترین پیمانکار، یکی از مهم ترین تصمیم های کارفرمایان است. امروزه به جز موارد معدودی، واگذاری طرح به پیمانکار بر اساس انتخاب حداقل قیمت پیشنهادی از بین پیمانکاران واجد شرایط انجام می شود و این روش واگذاری یکی از دلایل اصلی مشکلات تحویل پروژه در مدت و مبلغ مصوب است، زیرا پیمانکاران در هنگام مناقصه با پیشنهاد قیمت کمتر در مناقصه برنده شده و از طریق ادعا و در مواردی کاهش کیفیت اجرا، سعی می کنند درآمد بیشتری کسب کنند. از آنجایی که مسئله ارزیابی و انتخاب پیمانکار با ورودی های

نامطمئن، فشار زمان و کمبود اطلاعات همراه است [۱]، ماهیتا یک مسئله تصمیم گیری غیر قطعی است بنابراین ارزیابی پیمانکاران بیشتر یک هنر است تا علم [۲].

به طور خلاصه می توان گفت اهداف ارزیابی پیمانکاران عبارت است از:

- ۱- حذف پیمانکارانی که بر اساس ارزیابی سوابق کاری، توانایی اجرای پروژه را نداشته باشند و پاسخگو، مسئول و شایسته نباشند
- ۲- انتخاب پیمانکاران واجد شرایط با در نظر گرفتن معیارهای ارزیابی خاص نوع پروژه
- ۳- ایجاد رقابتی سالم که منجر به انجام سریع، اقتصادی و با کیفیت

مطلوب پروژه شود

^{*} نویسنده عهده دار مکاتبات: parchamijalal@ut.ac.ir



۴-انتخاب پیمانکار برتر، در نتیجه تأمین رضایت کارفرما از طریق حذف یا حداقل کردن ریسک شکست

مهم‌ترین معیارها برای انجام موفق پروژه‌های ساخت؛ رسیدن به هدف‌های کیفی پروژه و رعایت زمان و هزینه تعیین شده در قرارداد می‌باشد. بنابراین روش مقایسه پیمانکاران و معیارهایی که برای ارزیابی آن‌ها به کار گرفته می‌شود باید طوری طراحی شود که معیارهای اصلی موفقیت پروژه را با اطمینان پوشش دهد. عدم موفقیت پیمانکاران در دستیابی به اهداف اصلی پروژه در زمان مشخص، با هزینه معین و کیفیت مطلوب، کارفرمایان را با ریسک بیشتری در نیل به آرمان‌های پروژه مواجه می‌کند. بنابراین کارفرما باید معیارهای مهم در انتخاب پیمانکاران را شناسایی کرده و در پیشنهاد قیمت پیمانکاران تأثیر دهد تا از این طریق قیمت پیشنهادی نیز متناسب با شرایط و مشخصه‌های هر پروژه تراز شود. در میان تمام ریسک‌های مثبت (فرصت‌ها) و ریسک‌های منفی (تهدیدها)، ریسک تاخیر در پروژه‌های زیربنایی صنعت آب که نسبت سود به هزینه بالایی دارند، اهمیت زیادی دارد. شاید بتوان گفت مهم‌ترین رسالتی که بر عهده محققان است، مانع از به تاخیر افتادن این پروژه‌ها است زیرا تاخیر در اتمام پروژه توأم با هزینه است. هزینه‌هایی از قبیل:

- ۱- هزینه دیر رسیدن به بهره‌برداری یا سود از دست رفته به خصوص در پروژه‌های زیر بنایی؛
- ۲- هزینه ناشی از گران شدن منابع مصرفی؛
- ۳- افزایش هزینه منابع کاری؛
- ۴- هزینه تمدید مجوزها؛
- ۵- هزینه بهره سرمایه صرف شده؛
- ۶- ضرر از دست دادن بازار رقابتی؛
- ۷- ضرر ناشی از عدم اشتغال‌زایی در کشور؛
- ۸- ضرر ناشی از کاهش درآمد دولت و رفاه اجتماعی مردم و...

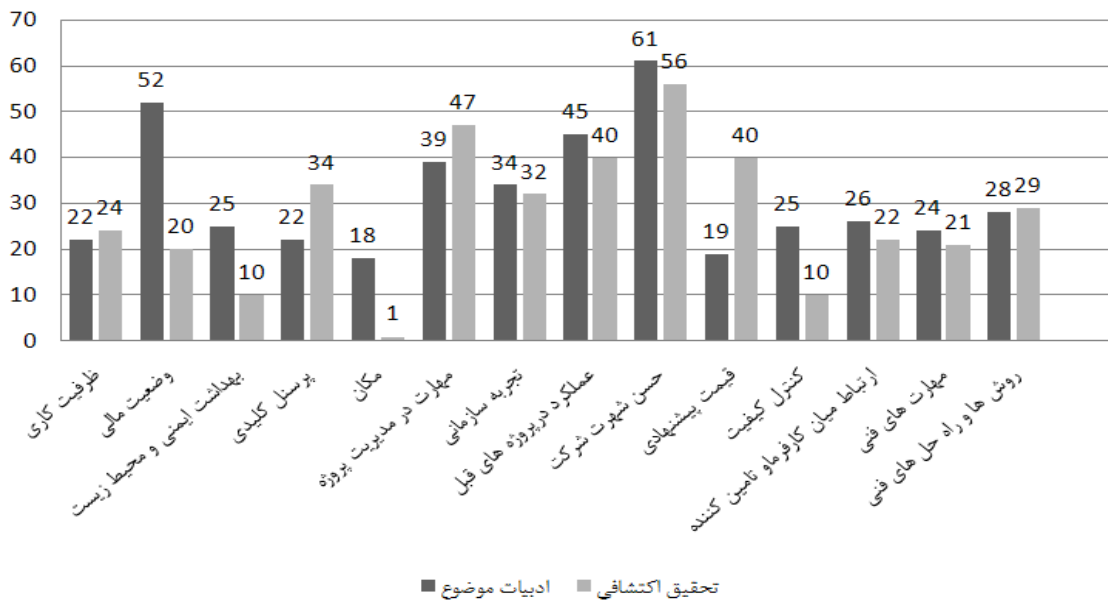
لذا در پژوهش حاضر با مطالعات کتابخانه‌ای، تحلیل آماری پرسشنامه‌ها و متن کاوی مصاحبه‌های باز، معیارها، زیرمعیارها و شاخصهای ارزیابی پیمانکاران به درستی شناسایی و اولویت‌بندی میشوند. برای دستیابی به اهداف پژوهش، تحقیق جامعی در حوزه معیارهای ارزیابی پیمانکاران، داده کاوی^۱ و متن کاوی انجام شده است و در دو بخش زیر خلاصه‌ای از آن آمده است.

1 Data Mining

۱-۱- مروری بر مطالعات گذشته در حوزه معیارهای ارزیابی پیمانکاران

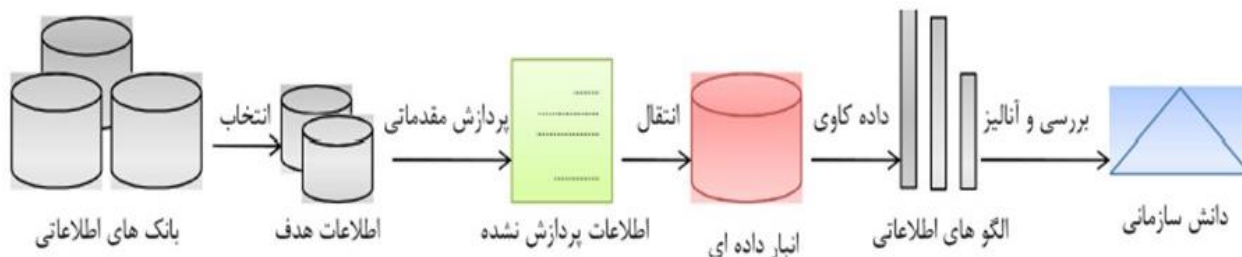
وات و همکاران [۳] برای جمع‌آوری معیارهای ارزیابی، تحقیق جامعی انجام داده‌اند؛ آن‌ها در مجموع ۹۰۱ معیار را گردآوری کردند (۴۷۲ معیار را با بررسی کلیه تحقیقات گذشته و ۴۲۹ معیار را با استفاده از مطالعات میدانی). سپس با انجام آزمون‌هایی همچون تست‌آستانه، اشتراک‌سنجی، تفاوت مطلق و... آن‌ها را غربال‌گری کرده و نهایتاً نتایج تحقیق خود را به صورت جدولی با ۱۶ دسته معیار منتشر نمودند، شکل ۱ تناوب تکرار معیارها در ادبیات موضوع و پیمایش در تحقیقات وات و همکاران را نشان می‌دهد. همچنین تحقیقی درباره‌ی تعداد معیار مناسب برای ارزیابی انجام دادند، نتیجه تحقیق آن‌ها نشان داد؛ ۷۶ درصد از پرسش‌شوندگان ۳ تا ۹ معیار را برای ارزیابی پیمانکاران با توجه به اهمیت پروژه مناسب دانستند. حالت [۴] در مقاله‌ای که درباره‌ی معیارهای انتخاب پیمانکاران نوشته است، تجربه پیمانکار در پروژه‌های مشابه، حجم پروژه‌های در حال اجرا، شایستگی‌های مدیریت و... را حاصل تحقیقات خود می‌داند. بررسی دیگری که توسط حالت صورت گرفته روشی برای اندازه‌گیری توانایی پیمانکار پیشنهاد می‌دهد که در آن پیمانکار از نظر ساختار سازمانی، توانایی‌های مدیریتی، توانایی‌های بانکی، توانایی مالی، شایستگی‌های مدیریتی و تجربه عملکرد در گذشته، ارزیابی می‌شود. وانگ [۵] با بررسی معیارهای متفاوت در انتخاب پیمانکار، فلسفه انتخاب پیمانکار بر مبنای پایین‌ترین قیمت را مورد بررسی قرار می‌دهد. بررسی‌های وی نشان می‌دهد که قیمت پایین‌تر لزوماً پارامتر متمایزکننده‌ی پیمانکار برنده از سایر پیمانکاران نیست و کارفرمایان تمایل دارند در انتخاب و ارزیابی پیمانکار، افق وسیع‌تری را مدنظر قرار دهند و به همین دلیل است که کارفرما مبنای ارزیابی خود را ارزش پیمانکار می‌داند و نه قیمت، و از تصمیم‌گیری چند شاخصه برای تعیین این ارزش استفاده می‌کند. هاتوش و اسکیتور [۶] در تحقیقات خود برای تعیین معیارهای انتخاب پیمانکار، توانایی‌های مدیریتی، توانایی مالی، نیروی انسانی مجرب، عملکرد گذشته، تجارب پیشین، ساختار مدیریت پروژه و توانایی بر عهده گرفتن شرح کار معین شده را به عنوان پارامترهای مهم معرفی می‌کنند. این دو محقق در مقاله دیگری، دیدگاه خود را در انتخاب پیمانکار، در قالب یک انتخاب چند معیاره بیان می‌کنند. این معیارها شامل توانایی‌های فنی،

نمودار فراوانی معیارها در تحقیقات وات، کایس و وایلی



شکل ۱. تناوب تکرار معیارها در ادبیات موضوع و پیمایش در تحقیقات وات و همکاران

Fig. 1. Frequency of criterion repetition in subject literature



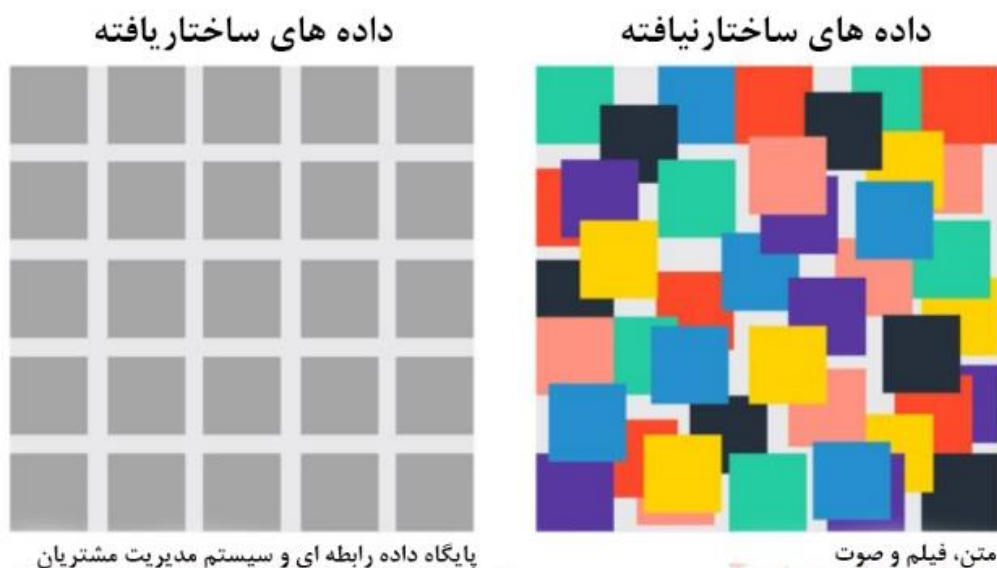
شکل ۲. روند داده کاوی

Fig. 2. Data Mining Process

مرحله است، ارزیابی مقدماتی پیمانکاران و انتخاب پیمانکار شایسته از بین پیمانکاران واجد شرایط. در این مدل معیارها به صورت سلسله مراتبی دسته بندی شده اند. با توجه به اهمیت زمان و کیفیت برای کارفرمایان، در این مدل از معیارهای زمان و کیفیت در گام اول و از معیارهای هزینه که مرتبط به قیمت پیشنهادی است در گام دوم استفاده می شود.

در اکثر تحقیقات ذکر شده از ابزار پرسشنامه برای شناسایی معیارهای ارزیابی پیمانکاران استفاده شده است. یکی از نقاط ضعف شناسایی معیارهای ارزیابی پیمانکاران با استفاده از پرسشنامه این

ایمنی و بهداشت، اعتبار و توانایی های مدیریتی و هزینه می شوند [۷]. آن جی و اسکیت مور، روشی برای تعیین رتبه پیمانکاران بر اساس آنالیز سود هزینه در مرحله تعیین صلاحیت ارائه دادند. پالانیسواران و کوماراسوامی [۸] مدلی جامع برای ارزیابی پیمانکاران بر اساس تحقیقات و مطالعات مستمر در هنگ کنگ طراحی نمودند، به طوری که در این مدل فرض بر این است که پیمانکاران پیشنهاد دهنده، اطلاعات تفصیلی خود را در ارتباط با معیارهای مورد نیاز جهت ارزیابی ارسال نمایند. مدل پیشنهادی برای انتخاب پیمانکاران ساخت و نصب در ترکیه که در مقاله تاپکو [۹] منتشر شده است، دارای دو



شکل ۳. تفاوت شماتیک داده های ساختاریافته و داده های ساختاریافته [۱۱]
 Fig. 3. Schematic difference between structured data and unstructured data

جدول ۱. اهداف روش متن کاوی

Table 1. Goals of Text Mining Method

اهداف	تعریف
استخراج ویژگی	منظور از استخراج ویژگی، مشخص کردن تمایز میان عبارت اسمی، شخص، مکان و یا سازمان از سایر اشیا است. عملیات مرتبط با این هدف، شامل استخراج کلمات و محاسبه تعداد دفعاتی است که هر کلمه در متن تحلیل شده است.
هدایت مبتنی بر آزمون	هدایت بر اساس آزمون، استخراج کننده‌ی متن را قادر می‌سازد تا کلمات مرتبط را در متن بیابد و روابط مهم بین آن‌ها را به هم ارتباط دهد.
جستجو و بازیابی	جستجو و بازیابی، کاربر را بر آن می‌دارد تا بر اساس معیارهای جستجوی از پیش تعیین شده به جستجو و بازیابی اطلاعات بپردازد.
خوشه‌بندی	خوشه‌بندی، عملیات گروه بندی کلمات کلیدی بر اساس میزان شباهت ها و یا عدم تشابه می‌باشد.
طبقه بندی	طبقه بندی، عملیات تعریف مجموعه‌ای از کلمات مخصوص دامنه و روابط بین آن‌ها از طریق الگوریتم‌های طبقه بندی است.
خلاصه سازی	خلاصه سازی عبارت است از عملیات کاهش مقدار داده‌های متنی با حفظ عناصر اصلی.
تجزیه و تحلیل روند	تجزیه و تحلیل روند فرآیندی است که طی آن برای کشف روند از داده‌های متنی وابسته به سازمان استفاده می‌شود.
قواعد انجمنی	مرتبط ساختن الگوی استخراج شده با الگوهایی است که از قبل کشف و استخراج شده‌اند.
مصورسازی	در مصورسازی از استخراج ویژگی و کلمات کلیدی فهرست شده برای ایجاد یک نمایش گرافیکی استفاده می‌شود که می‌تواند کاربر را برای شناسایی موضوعات اصلی با مفاهیم مهم‌شان یاری دهد.

بر آن داشت تا معیارها و شاخصهای ارزیابی پیمانکاران را با استفاده از دو ابزار مختلف و مکمل، یعنی پرسشنامه و متن کاوی مصاحبه با خبرگان مورد شناسایی قرار دهند.

است که شیوه طراحی پرسشنامه و عواملی که در آن ذکر شده‌اند، ممکن است باعث جهت گیری مخاطب شود و شیوه پر کردن پرسشنامه را تحت تاثیر قرار دهد. جبران این نقطه ضعف، محققان را

۲-۱- مروری بر مطالعات گذشته در حوزه داده کاوی، متن کاوی و خوشه بندی

داده کاوی عبارت است از فرایند استخراج اطلاعات معتبر، معانی از پیش ناشناخته، قابل فهم و قابل استفاده از پایگاه‌های داده و استفاده از آن در تصمیم‌گیری‌ها [۱۰]. روند داده کاوی در شکل ۲ نمایش داده شده است.

داده‌ها می‌توانند به صورت ساختاریافته^۱ یا غیرساختاریافته^۲ باشند. داده‌های ساختاریافته معمولاً به صورت داده‌های کمی و در قالب سطر و ستون هستند، درجه بالایی از سازمان دهی دارند، از داده‌های ساختاریافته متداول‌تر هستند و به راحتی با زبان ماشین قابل فهم هستند، مانند: داده‌های ذخیره شده در پایگاه‌های داده‌ی رابطه‌ی ای و صفحات گسترده مانند نرم افزار اکسل. داده‌های غیرساختاریافته معمولاً به صورت داده‌های کیفی هستند، تجزیه و تحلیل این داده‌ها دشوار است، زیرا هیچ مدل از پیش تعریف شده‌ای ندارند و نمیتوانند در پایگاه‌های داده رابطه‌ی ای مانند اکسس سازماندهی شوند، مانند: متون، تصاویر و فایل‌های صوتی. داده‌های ساختاریافته و غیرساختاریافته تفاوت‌های اساسی با هم دارند و معمولاً در مدل‌های یادگیری ماشین^۳ و داده کاوی سعی بر این است که داده‌های غیرساختاریافته به داده‌های ساختاریافته تبدیل شوند تا برای رایانه قابل فهم باشند. تفاوت این دو نوع داده در شکل ۳ نمایش داده شده است.

داده‌کاوی در دو نوع نظارتی^۴ و غیرنظارتی^۵ ظاهر می‌شود. داده‌کاوی نظارتی دارای هدفی خاص و از پیش تعیین شده است که به دنبال الگویی خاص می‌گردد، در حالی که هدف داده کاوی غیرنظارتی یافتن الگوها یا تشابهات بین گروه‌هایی از اطلاعات بدون داشتن هدفی خاص و یا مجموع‌های از دسته‌ها و الگوهای از پیش تعیین شده میباشد. در داده کاوی نظارتی همیشه یک متغیر هدف وجود دارد، چیزی که باید دسته بندی^۶، تخمین یا پیشبینی شود. در داده کاوی غیرنظارتی، متغیر هدف وجود ندارد و وظیفه‌ی داده کاوی این است که الگوهای کلیدی را بیابد که متعلق به یک متغیر خاص نیستند. مدل‌های داده کاوی شامل دسته‌بندی، تخمین،

پیشبینی (داده کاوی نظارتی)، خوشه‌بندی، قوانین انجمنی^۷ (داده کاوی غیرنظارتی) و توصیف مفهومی^۸ (نظارتی و غیرنظارتی) می‌باشد [۱۱].

متن کاوی را میتوان شاخه‌ای از داده کاوی دانست که عبارت است از کشف الگوهای ناشناخته از بین داده‌های متنی. این رویکرد به دنبال آن است که اطلاعات مفید را از داده‌های متنی غیرساختاریافته استخراج کند. متن کاوی شباهت زیادی با داده کاوی دارد با این تفاوت که ابزارهای داده کاوی برای مدیریت داده‌های ساختاریافته طراحی شده است اما متن کاوی می‌تواند بر مجموعه‌ای از داده‌های غیرساختاریافته نیز برای استخراج الگوهای ناشناخته‌ی جدید تمرکز کند [۱۲]. اهداف روش متن کاوی را می‌توان در ۹ گروه عملیاتی طبقه بندی کرد که شرح هر یک از این اهداف در جدول زیر آورده شده است [۱۳].

خوشه‌بندی یک تابع کاوشی غیرنظارتی داده کاوی به منظور کشف گروه بندی طبیعی درون داده‌ها است. یک خوشه به مجموعه‌ای از داده‌ها اطلاق می‌شود که از جهاتی شبیه به همدیگر هستند. الگوریتم‌های خوشه‌بندی به طور خودکار ویژگی‌های متمایز کننده‌ی زیر گروه‌ها را تعریف میکنند و آن‌ها را سازماندهی می‌نمایند. به بیان دیگر می‌توان گفت که خوشه‌بندی، قراردادن داده‌ها در گروه‌هایی است که اعضای هر گروه از زاویه‌ی خاصی به یکدیگر شباهت دارند و با اعضای خوشه‌های دیگر هیچ شباهتی ندارند یا حداقل نسبت به اعضای خوشه‌ی خود از شباهت بسیار کمتری با اعضای دیگر خوشه‌ها برخوردارند. یک الگوریتم خوشه‌بندی میتواند دارای مشخصات مطلوب زیر باشد: [۱۰]

- قابلیت مقیاس پذیری - عدم حساسیت به ترتیب ورود
- توانایی مواجهه با انواع داده‌ها
- داده‌ها - عدم نیاز به پارامترهای ورودی
- استخراج خوشه‌هایی - پذیرش داده‌هایی با ابعاد بالا
- به هر شکل دلخواه - قابل فهم بودن نتایج الگوریتم
- توانایی مقابله با داده‌های نویز، نادرست و ناقص

- 1 Structured Data
- 2 Unstructured Data
- 3 Machine Learning
- 4 Supervised
- 5 Unsupervised
- 6 Classification

7 Association Rules
8 Conceptual Description

یادگیری غیرنظارتی و با استفاده از خوشه‌بندی برای شناسایی نواقص دستگاه‌های هواساز ارائه کردند. آن‌ها در این پژوهش از داده‌های قبلی تولید شده توسط دستگاه هواساز به عنوان ورودی مدل استفاده کردند [۱۶]. ناگانانان و همکاران با استفاده از یادگیری غیرنظارتی و خوشه‌بندی، چارچوبی برای بهینه‌سازی و خودکار کردن فرآیند شناسایی عوامل هدر رفتن انرژی در مدل‌سازی انرژی ساختمان^۱ توسعه دادند [۱۷]. کاشانی و گراتینگر با استفاده از خوشه‌بندی، مدلی برای شناسایی آسیب‌های باد به پوشش‌های بام ارائه کردند [۱۸] رودریگز و همکاران مدلی برای خوشه‌بندی خروجی‌های طراحی مولد^۲ به منظور کمک به تصمیم‌گیری در انتخاب بهترین طرح، ارائه کردند [۱۹]. سایتا و همکاران از الگوریتم K-Means برای محدود کردن تعداد گزینه‌های مناسب مدل‌های سازه‌ای به منظور انتخاب بهترین مدل سازه‌ای استفاده کردند [۲۰]. بریلاکیس و همکاران با استفاده از خوشه‌بندی، چارچوبی برای مدیریت تصاویر کارگاه‌های ساختمانی توسعه دادند. این چارچوب با شناسایی مواد و مصالح، عکس‌ها را بر اساس نوع مواد و مصالح در خوشه‌های مختلفی قرار می‌داد [۲۱]. یارمحمدی و همکاران چارچوبی برای ساخت مدل سه بعدی از طریق پردازش داده‌های متنی ساختاریافته توسعه دادند [۲۲]. دینگ و همکاران چارچوبی برای پردازش متون قوانین شهری و انطباق آن‌ها با مدل اطلاعاتی ساختمان ارائه دادند [۲۳].

۲- روش‌شناسی انجام پروژه

با توجه به هدف اصلی پژوهش که شناسایی معیارها و شاخصهای ارزیابی پیمانکاران در صنعت آب است، از نتایج تحقیقات گذشته، ابزار پرسشنامه، مصاحبه با خبرگان و متن کاوی برای رسیدن به هدف پژوهش بهره گرفته شده است. این پژوهش شامل سه بخش است، در بخش اول با استفاده از مطالعات کتابخانه‌ای و ابزار پرسشنامه، معیارهای ارزیابی پیمانکاران شناسایی شده، در بخش دوم با خبرگان صنعت آب مصاحبه شده و مصاحبه‌ها مورد متن کاوی قرار گرفتند، در بخش سوم نتایج بخش اول و دوم مقایسه شدند و در نهایت یافته‌های پژوهش با آیین نامه اجرایی بند "ج" ماده (۱۲) قانون برگزاری مناقصات، مقایسه شدند. به عبارت دیگر معیارها و شاخصهای ارزیابی پیمانکاران با استفاده از ابزار پرسشنامه، تحلیلهای آماری و متن کاوی

استراتژیهای خوشه‌بندی بسیار زیاد هستند و راهکارهای بسیار زیادی برای محاسبه‌ی شباهت میان نمونه داده‌ها در خوشه‌ها وجود دارد. برای روشن شدن موضوع فرض کنید شما با n نمونه از داده‌ها روبرو هستید و قصد دارید این نمونه‌ها را در تعداد c خوشه قرار دهید. تعداد حالت‌های ممکن برای خوشه‌بندی را می‌توانید از رابطه ۱ محاسبه کنید:

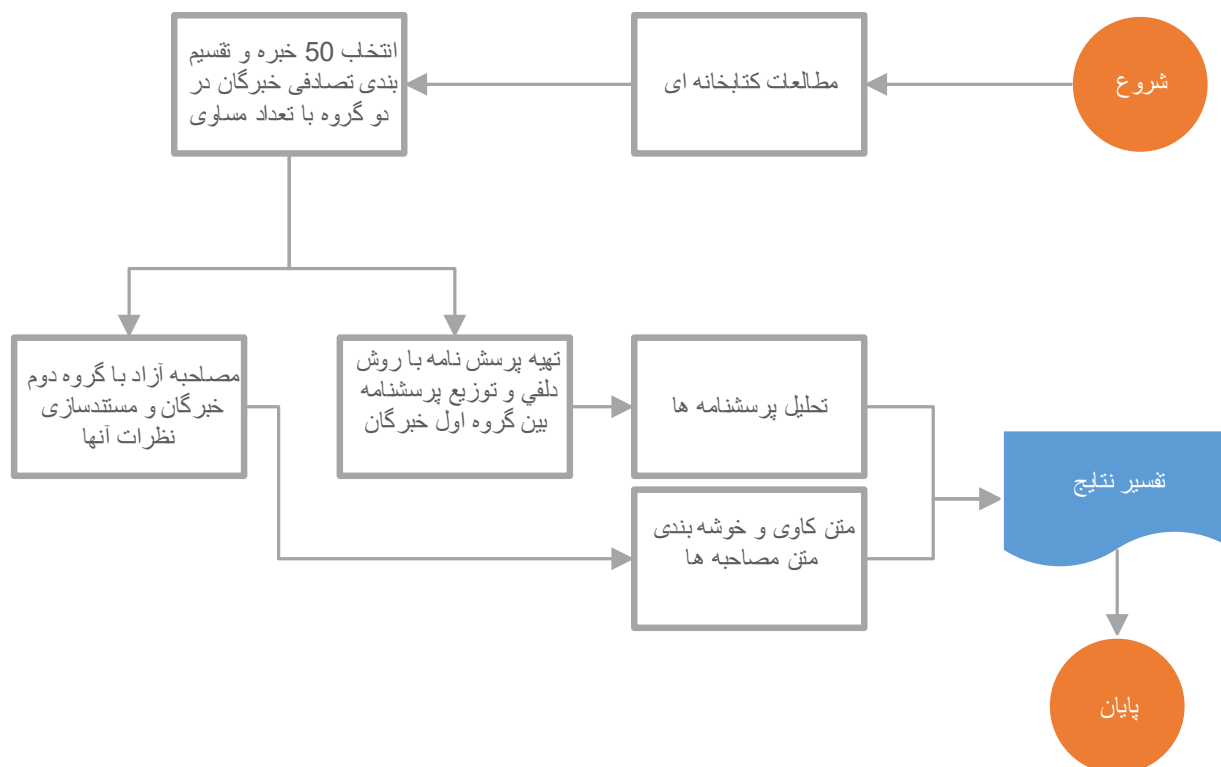
$$\frac{1}{c!} \sum_{i=1}^c (-1)^{c-i} \binom{c}{i} \times i^n \quad (1)$$

بر اساس فرمول ذکر شده، حتی با تعداد نمونه‌های کم، تنوع حالت‌های ممکن خوشه‌بندی یک عدد نجومی است. بنابراین نیاز به استفاده از الگوریتمهای خوشه‌بندی به خوبی احساس میشود چرا که ارزیابی حالت‌های خوشه‌بندی برای یک مسئله‌ی کوچک نه تنها زمان‌گیر بلکه غیرممکن است.

یکی از محبوب‌ترین الگوریتمهای داده کاوی غیرنظارتی که قادر به حل مسائل خوشه‌بندی می‌باشد، الگوریتم K-Means است [۱۴]. ورودی این الگوریتم n نمونه داده و مقدار K که تعداد خوشه‌های خروجی را مشخص می‌کند، می‌باشد. در ابتدا تعداد K نمونه به صورت اتفاقی از میان کل نمونه‌ها انتخاب می‌شوند. این نمونه‌ها به عنوان نماینده‌ی K خوشه شناخته خواهند شد. گاهی به آن‌ها مرکز ثقل یا مرکز خوشه نیز اطلاق می‌شود. هر یک از نمونه‌های باقی‌مانده عضوی از خوشه‌ای خواهند بود که یکی از نماینده‌ها متعلق به آن است. به عبارت دیگر با کمک معیارهایی چون فاصله‌ی اقلیدسی تشابه هر یک از نمونه‌های باقی‌مانده را با K نماینده محاسبه می‌کنیم و نمونه‌ی مورد نظر به هر یک نزدیک‌تر بود، به عضویت آن خوشه در می‌آید. پس از آن برای هر خوشه، با محاسبه‌ی میانگین میان اعضای خوشه، نماینده‌ی جدیدی انتخاب می‌گردد. این فرآیند تا پوشش معیاری جهت خاتمه‌ی کار تکرار می‌شود [۱۰]. القادی و کندیل مدلی برای خوشه‌بندی خودکار اسناد و مدارک پروژه‌های صنعت ساخت، بر پایه‌ی شباهت‌های متنی ارائه کردند. از نظر آن‌ها تکنیک‌های یادگیری نظارتی مانند دسته‌بندی به دلیل نیاز به داده‌های جامع برای یادگیری دارای محدودیت‌هایی می‌باشند و برای رفع این محدودیت، از خوشه‌بندی به عنوان یکی از تکنیک‌های یادگیری غیرنظارتی استفاده کردند [۱۵]. یان و همکاران مدلی بر پایه

1 Building energy modeling (BEM)

2 Generative design



شکل ۴. نمای کلی تحقیق

Fig. 4. Research process

شماره		۱	۲	۳	۴	۵	۶	۷
	معیار تصمیم گیری	حسن شهرت	عملکرد گذشته	ظرفیت کاری	وضعیت مالی	مدیریت ایمنی	مهارت مدیریت پروژه	مهارت های فنی
۱	حسن شهرت	۱						
۲	عملکرد گذشته		۱					
۳	ظرفیت کاری			۱				
۴	وضعیت مالی				۱			
۵	مدیریت ایمنی					۱		
۶	مهارت مدیریت پروژه						۱	
۷	مهارت های فنی							۱

شکل ۵. پرسشنامه ی مقایسه زوجی معیارهای اصلی

Fig. 5. Paired Comparison Questionnaire for main Criteria

تمامی کارشناسان پانل طراحی پرسشنامه قرار گرفت و برای نمونه ی آماری (گروه اول شامل ۲۵ نفر از خبرگان) فرستاده شد. ابتدا با روش مقایسه زوجی نسبت اهمیت هر از معیارهای اصلی تعیین گردید. مقایسه زوجی کمک می کند تا اهمیت تعدادی از گزینه های نسبت به یکدیگر سنجیده شود. نحوه ی اقدام به این ترتیب بود که پرسشنامه ی مقایسه زوجی برای اعضای پانل و پرسش شوندگان ارسال و نتیجه ی خرد جمعی وارد نرم افزار Expert Choice گردید. شکل ۵ نشان دهنده ی پرسشنامه ی مقایسه زوجی و شکل ۶ نشان دهنده ی خروجی نرم افزار و وزن هر یک از معیارهای اصلی است.

نرخ ناسازگاری^۲ نشان می دهد تا چه اندازه می توان به داده های گردآوری شده از دیدگاه هر یک کارشناسان اعتماد کرد. اساس محاسبات فرآیند تحلیل سلسله مراتبی بر اساس قضاوت اولیه تصمیم گیرنده که در قالب ماتریس مقایسه زوجی ظاهر می شود، صورت می پذیرد. بنابراین هرگونه خطا و ناسازگاری در مقایسه عناصر، نتیجه نهایی به دست آمده از محاسبات را تحت تاثیر قرار می دهد. مطالعات نشان داده است که اگر نرخ ناسازگاری کمتر از ۰/۱ باشد سازگاری مقایسه ها قابل قبول بوده و در غیر این صورت مقایسه ها باید تجدید نظر شود. در این پژوهش، نرخ ناسازگاری مقایسه زوجی معیارها ۰/۰۷ می باشد که نشان می دهد مقایسه زوجی سازگار می باشد.

برای تحلیل پرسشنامه از نرم افزار SPSS^۳ استفاده شد و معیارها زیر معیارها اولویت بندی شده و برای آن ها امتیاز پیشنهادی از ۱۰۰ تعیین گردید. بر اساس نتایج این پژوهش برای پروژه های زیر بنایی، ارزیابی پیمانکاران در ۷ معیار اصلی مناسب ترین حالت است. نتیجه این بررسی در شکل ۶ آورده شده است. با محاسبه ضریب آلفای کرونباخ توسط نرم افزار، پایایی داده ها در حد عالی (۰/۹۱) به دست آمد.

برای تبدیل شاخص اهمیت معیارهای اصلی به درصد، شاخص اهمیت هر یک از معیارهای اصلی به مجموع شاخص اهمیت همه معیارهای اصلی تقسیم شد.

$$(۲) \quad \text{وزن معیار اصلی} = \frac{\text{شاخص اهمیت معیار}}{\text{مجموع شاخص اهمیت همه معیارها}}$$

برای تعیین درصد اهمیت زیر معیارها، از پرسش شوندگان خواسته

شناسایی شده و پس از مقایسه نتایج، معیارهای نهایی ارائه می شوند. جامعه آماری شامل ۷۸ نفر از افراد شاغل در دفاتر بازاریابی یک شرکت مشاور و یک شرکت پیمانکار دارای رتبه ۱ در گروه آب می باشد. بر اساس فرمول کوکران، حداقل حجم نمونه برابر با ۶۴/۹۸ است. از روش نمونه گیری طبقه ای برای نمونه گیری استفاده شده است، زیرا این روش امکان بررسی گروه هایی از جامعه را به وجود می آورد که ممکن است در نمونه گیری تصادفی ساده در نظر گرفته نشوند. ۶۴ نفر از خبرگان این صنعت از دو شرکت متفاوت (برای اطمینان از تعمیم پذیری نتایج تحقیق) شناسایی شده و برای آن ها دعوت نامه ارسال شد و ۵۰ نفر از آن ها موافقت خود را برای مشارکت در پژوهش اعلام کردند. خبرگان به صورت تصادفی به دو گروه با تعداد مساوی تقسیم شدند. نظرات گروه اول با ابزار پرسشنامه و تحلیل های آماری و نظرات گروه دوم با ابزار مصاحبه و متن کاوی جمع آوری شد. شمای کلی تحقیق در شکل ۴ نمایش داده شده است.

۳- بحث

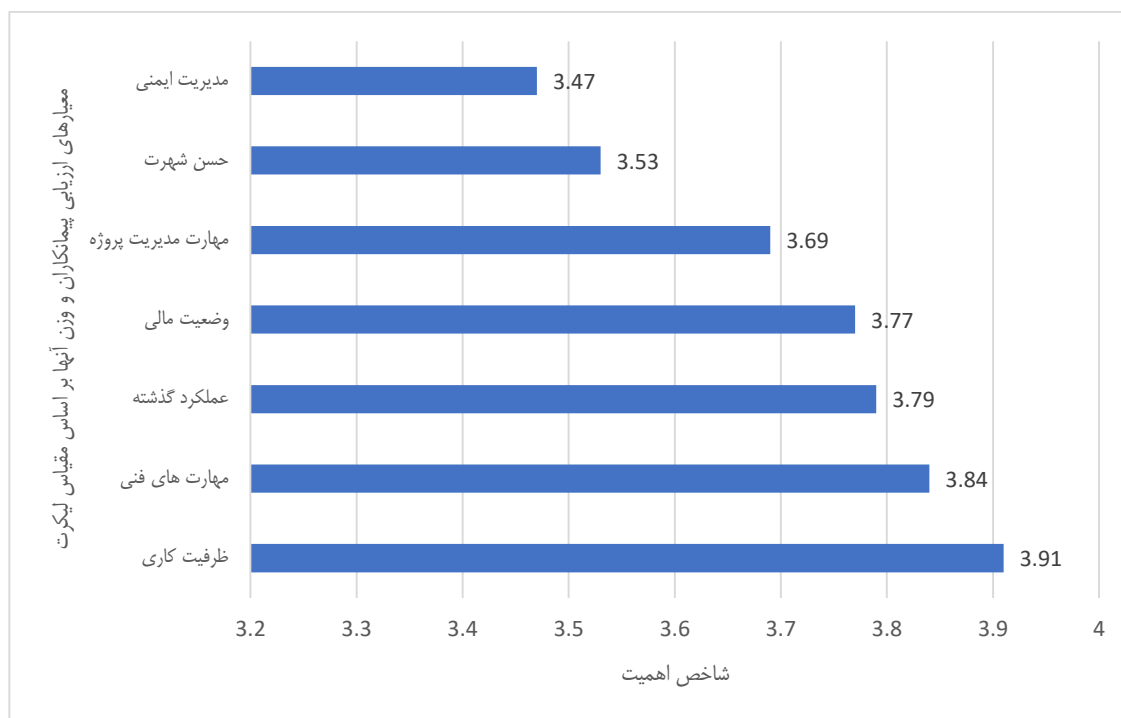
۳-۱- بخش اول (شناسایی معیارهای ارزیابی پیمانکاران با ابزار پرسشنامه و تحلیل های آماری)

در بخش اول با استفاده از مطالعات کتابخانه های، معیارهای اولیه ارزیابی با بررسی مطالعات گذشته در زمینه ارزیابی پیمانکاران به خصوص تحقیقات جامع وات، کاپس و وایلی استخراج شده و پیش پرسشنامه آماده گردید. سپس با بهره گیری از روش دلفی تیمی از کارشناسان خبره صنعت تحت عنوان پانل طراحی پرسشنامه، تشکیل و معیارهای ارزیابی با توجه به نوع و نیاز پروژه های زیر بنایی، بومی سازی و نهایتاً پرسشنامه نهایی مورد تایید قرار گرفت. در طراحی پرسشنامه از روش دلفی استفاده گردید و گزینه ها بر مبنای مقیاس لیکرت^۱ تدوین گردیدند و پس از تعیین تعداد جامعه آماری، برای افزایش روایی و پایایی داده ها در نتایج پرسشنامه ها درباره ی افرادی که پرسشنامه برای آن ها فرستاده می شود با اعضای پانل بحث و تبادل نظر شد تا افرادی مورد پرسش قرار گیرند که کارشناس خبره باشند. لذا با توجه به تمرکز روی صنعت آب کشور سعی شد افراد خبره این صنعت انتخاب شوند و پرسشنامه به صورت خطاب به شخص پرسش شونده فرستاده شود. در انتها پرسشنامه ی نهایی مورد تایید

2 Inconsistency

3 Spss Statistics

1 Likert scale



شکل ۶. معیارهای اصلی ارزیابی پیمانکاران، مستخرج از تحلیل پرسشنامه‌ها

Fig. 6. The main criteria for contractor evaluation derived from questionnaire analysis

جدول ۲. معیارها و زیر معیارها و امتیازات آنها

Table 2. Criteria, sub-criteria and their weight

وزن زیر معیار (درصد)	زیرمعیار ارزیابی به ترتیب اهمیت و اولویت	وزن (درصد)	معیار اصلی
۳۳	۱- ظرفیت آزاد برای در دست گرفتن پروژه‌های جدید	۱۵	۱- ظرفیت کاری
۳۳	۲- ظرفیت نیروی انسانی قابل تخصیص به پروژه		
۳۳	۳- ظرفیت ماشین آلات قابل تخصیص به پروژه		
۳۶	۱-۲- مهارت های فنی	۱۵	۲- مهارت های فنی، داشتن کارکنان کلیدی و آشنایی با قوانین
۳۲	۲-۲- کارکنان کلیدی		
۳۲	۳-۲- آشنایی با قوانین		
۵۰	۱-۳- عملکرد در پروژه‌های گذشته	۱۵	۳- عملکرد در پروژه‌های گذشته و تجربه سازمانی
۵۰	۲-۳- تجربه سازمانی و گواهی نامه‌ها		
-	دارایی‌های جاری، میزان اعتبارات و گردش مالی	۱۵	۴- وضعیت مالی
۵۰	۱-۵- مهارت در مدیریت پروژه و داشتن کارشناسان فنی مرتبط	۱۴	۵- مهارت در مدیریت پروژه و روش‌ها و راه حل‌های فنی و ابتکاری
۵۰	۲-۵- روش‌ها و راه حل‌های فنی و ابتکاری		
-	گواهینامه‌های حسن انجام کار	۱۴	۶- حسن شهرت شرکت و ارتباطات
۵۰	۱-۷- بهداشت، ایمنی و محیط زیست	۱۲	۷- بهداشت، ایمنی و محیط زیست و کنترل کیفیت
۵۰	۲-۷- کنترل کیفیت		

کلیدی است، روش TF-IDF صرفاً میزان تکرار یک کلمه کلیدی را نشان نمی‌دهد، بلکه هدف آن نشان دادن اهمیت کلمه کلیدی مورد نظر از طریق مقایسه تعداد تکرار کلمه در متن با تکرار آن کلمه در مجموعه‌ای بزرگ‌تر از مستندات می‌باشد. به طور مثال کلمه ای که توسط ۵ نفر از خبرگان بیان شده باشد، نسبت به کلمه ای که ۵ بار توسط یک خبره تکرار شده، اهمیت بیشتری دارد. TF-IDF شامل دو بخش است، بخش اول (TF^5) نمایانگر مقدار نرمال شده تکرار کلمات در متن و بخش دوم (IDF^6) نمایانگر اهمیت یک کلمه نسبت به سایر کلمات در متن می‌باشد. با ضرب مقادیر TF در IDF برای هر کلمه مقدار TF-IDF محاسبه شده و این مقادیر در کنار هم ماتریس کل متن را شکل می‌دهند. بخشی از متن تعدادی از مصاحبه‌ها به همراه تشریح روش TF-IDF در پیوست ۱ آورده شده است.

پس از محاسبه ماتریس TF-IDF، عملیات خوشه‌بندی با الگوریتم K-means (واقع در گروه Segmentation) بر روی این ماتریس انجام شد. مقدار K که بیانگر تعداد خوشه‌ها می‌باشد به ترتیب از ۲ تا ۱۰ تعیین شد و هر بار نتیجه خوشه‌بندی ثبت شد و در نهایت با مقایسه آن‌ها این نتیجه کسب شد که $K=6$ مناسب‌ترین ترکیب خوشه‌ها را ایجاد می‌کند. پس از اجرای عملیات خوشه‌بندی، داده‌ها در ۶ خوشه مختلف قرار گرفته و با استفاده از دو معیار، مقدار TF-IDF و مقدار مرکز خوشه برای هر کلمه، کلمات کلیدی در هر خوشه مشخصه شده و در جدول ۳ نمایش داده شده‌اند.

۳-۳-۳- بخش سوم (مقایسه نتایج ابزار پرسشنامه، ابزار متن کاوی و ماده ۱۲ قانون برگزاری مناقصات)

در این پژوهش، معیارهای ارزیابی پیمانکاران در مناقصات صنعت آب توسط دو ابزار مختلف شناسایی شدند، ابزار پرسشنامه و تحلیل‌های آماری و ابزار مصاحبه و متن کاوی. با استفاده از ابزار پرسشنامه و تحلیل‌های آماری، ۷ معیار اصلی و مهم‌ترین زیرمعیارهای ارزیابی پیمانکاران در مناقصات صنعت آب تعیین شد و با انجام متن کاوی مصاحبه‌ها، کلمات کلیدی متن مصاحبه‌ها با توجه به دو معیار، مقدار TF-IDF و مقدار مرکز خوشه در ۶ خوشه مختلف قرار گرفتند که کلمات کلیدی قرار گرفته در یک خوشه، بیشترین شباهت را با

شد که به هر یک از زیر معیارها عددی بین ۱ تا ۷ اختصاص دهند. عدد ۱ نشان دهنده کمترین اهمیت و عدد ۷ نشان دهنده بیشترین اهمیت می‌باشد. سپس میانگین امتیازات اختصاص داده شده به هر یک از زیر معیارها محاسبه شده و به عنوان شاخص اهمیت در نظر گرفته شد. برای تبدیل شاخص اهمیت زیرمعیارهای هر معیار اصلی به درصد، شاخص اهمیت هر یک از زیرمعیارها به مجموع شاخص اهمیت زیر معیارهای همان معیار اصلی، تقسیم شد. خلاصه نتایج به دست آمده از تحلیل پرسشنامه‌ها در جدول ۲ آورده شده است.

۳-۲- بخش دوم (شناسایی معیارها و شاخصهای ارزیابی پیمانکاران با مصاحبه و متن کاوی)

همانطور که گفته شد، یکی از نقاط ضعف استفاده از پرسشنامه برای شناسایی مهم‌ترین معیارهای ارزیابی این است که شیوه طراحی پرسشنامه و عواملی که در آن ذکر شده‌اند، ممکن است باعث جهت‌گیری مخاطب شود و شیوه پرسشنامه را تحت تاثیر قرار دهد. برای پوشش این نقطه ضعف تصمیم گرفته شد که با گروه دوم شامل ۲۵ نفر از خبرگان به صورت باز مصاحبه شده و متن مصاحبه‌ها مورد متن کاوی قرار گیرند، بدین منظور از نرم افزار Rapidminer استفاده شده است. با توجه به مشکلات موجود برای متن کاوی متون فارسی، ابتدا مصاحبه‌ها به زبان انگلیسی ترجمه شده و سپس در یک فایل متنی ساختار نیافته^۱ جمع آوری شده و عملیات متن کاوی بر روی این فایل انجام شده است. در مرحله پیش پردازش^۲ داده‌ها، کلمات بی اهمیت مانند حروف ربط، حروف اضافه و... (از، و، و...) حذف شدند. برای انجام عملیات متن کاوی باید ابتدا داده‌های متنی به بردارهای عددی تبدیل شده تا توسط نرم افزار قابل پردازش باشد. بدین ترتیب داده‌های ساختار نیافته به داده‌های ساختار یافته^۳ تبدیل میشوند. با استفاده از عملگر Generate TFIDF، کلمات به بردارهای عددی تبدیل میشوند و متن مصاحبه‌ها به یک ماتریس عددی تبدیل میشود. TF-IDF یک گزاره آماری است که به معنای فراوانی وزنی کلمات کلیدی است که نمایانگر میزان اهمیت کلمات در متن می‌باشد. برخلاف روشهای دیگر مانند یواست^۴ که تمرکز اصلی آن‌ها صرفاً روی تعداد دفعات تکرار کلمه

- 1 Unstructured
- 2 Pre-processing
- 3 Structured
- 4 Yoast

5 Term Frequencies
6 Inverse Document Frequencies

مشاهده شد که موضوعات ۵ خوشه قرابت و همبستگی زیادی با ۵ معیار اصلی ارزیابی پیمانکاران (ظرفیت کاری، وضعیت مالی، مهارت در مدیریت پروژه، حسن شهرت و بهداشت و ایمنی) دارند و خوشه شماره ۶ (سابقه دعاوی) را میتوان به عنوان یک معیار جدید در نظر گرفت. در گام بعدی با تحلیل کلمات کلیدی هر یک از خوشه ها، شاخصهای ارزیابی برای هر یک از معیارها و زیر معیارها تعیین شده

یکدیگر و کمترین شباهت را با کلمات سایر خوشهها دارند. سه نفر از هر گروه خبرگان به صورت تصادفی انتخاب شده و نتایج به دست آمده از متن کاوی برای آنها تشریح شد و از آنها خواسته شد که با توجه به کلمات کلیدی هر خوشه، موضوع خوشه ها را تعیین کنند که نتایج آن در جدول ۴ نمایش داده شده است. موضوع هر خوشه به عنوان معیار اصلی در نظر گرفته شد و

جدول ۳. خوشه ها و کلمات کلیدی آنها
Table 3. Clusters and their keywords

کلمات کلیدی	شماره خوشه
Status-Financial-Stability-Profit-Estimate-Payment-Income	۱
Manpower-Resource-Machinery-Equipment-Material-System-Technical	۲
Safety-Plan-Demolition-HSE-Accident-Certificate-Insurance-Rate	۳
Satisfaction-Relation-Reputation-Past-Client-Longtime-communication	۴
PMO-Maturity-Methodology-Knowledge-Scheduling-Risk	۵
Claim-Method-Negotiation-Compensation-Disputes-Arbitration-Number- Delay	۶

جدول ۴. خوشه ها، کلمات کلیدی و موضوع هر یک از آنها
Table 4. Clusters, Keywords and Their Subjects

موضوع	کلمات کلیدی	شماره خوشه
وضعیت مالی	Status-Financial-Stability-Profit-Estimate-Payment-Income	۱
ظرفیت کاری	Manpower-Resource-Machinery-Equipment-Material-System-Technical	۲
بهداشت و ایمنی	Safety-Plan-Demolition-HSE-Accident-Certificate-Insurance-Rate	۳
حسن شهرت	Satisfaction-Relation-Reputation-Past-Client-Longtime-communication	۴
مدیریت پروژه	PMO-Maturity-Methodology-Knowledge-Scheduling-Risk	۵
سابقه دعاوی	Claim-Method-Negotiation-Compensation-Disputes-Arbitration-Number- Delay	۶

جدول ۵. معیارها، زیرمعیارها، شاخص های ارزیابی پیمانکاران و وزن آن‌ها

Table 5. Criteria, sub-criteria, contractors evaluation indicators and their weight

وزن(درصد)	شاخص های زیرمعیار ارزیابی به ترتیب اهمیت و اولویت	وزن(درصد)	زیرمعیار ارزیابی	وزن(درصد)	معیار اصلی
۳۵	۱-۱-ظرفیت ماشین آلات قابل تخصیص به پروژه				
۳۴	۲-۱-ظرفیت نیروی انسانی قابل تخصیص به پروژه	-	-	۱۴	۱-ظرفیت کاری
۳۱	۳-۱-ظرفیت آزاد برای در دست گرفتن پروژه‌های جدید				
۲۷	۱-۱-۲-داشتن مدیران مجرب با بیش از ۱۰ سال سابقه کار مرتبط (ارائه گواهی بیمه و دارایی)				
۲۶	۲-۱-۲-داشتن کارشناسان فنی با بیش از ۵ سال سابقه کار مرتبط (ارائه گواهی بیمه و دارایی)	۳۶	۱-۲-مهارت‌های فنی		
۲۵	۳-۱-۲-متوسط سابقه کار پرسنل دفتر فنی پیمانکار			۱۴	۲-مهارت های فنی، داشتن کارکنان کلیدی و آشنایی با قوانین
۲۲	۴-۱-۲-داشتن تکنسین فنی با بیش از ۵ سال سابقه کار مرتبط				
۳۵	۱-۲-۲-سابقه کار کارکنان کلیدی	۳۲	۲-۲-کارکنان کلیدی		
۳۳	۲-۲-۲-صلاحیت و تجربه کارکنان کلیدی				
۳۲	۳-۲-۲-ارتباط تخصص کارکنان با نوع پروژه (گواهی تحصیلی کارکنان)				
۴۵	۱-۳-۲-آشنایی با قوانین و بخشنامه‌ها				
۵۵	۲-۳-۲-آشنایی با قوانین داخلی و خارجی مالی	۳۲	۳-۲-آشنایی با قوانین		
۲۶	۱-۱-۳-تعداد و مبلغ کل پروژه‌های انجام شده در ۵ سال گذشته				
۲۵	۲-۱-۳-متوسط درصد تاخیر کل(میزان تاخیر به زمان پروژه) در پروژه‌های ۵ سال گذشته	۵۰	۱-۳-عملکرد در پروژه‌های گذشته		
۲۵	۳-۱-۳-ارائه گواهی حسن انجام کار و تشویق				
۲۴	۴-۱-۳-متوسط افزایش مبلغ (نسبت کل مبالغ افزایش هزینه به کل مبالغ پروژه) در پروژه‌های ۵ سال گذشته			۱۴	۳-عملکرد در پروژه های گذشته و تجربه سازمانی
۲۵	۱-۲-۳-عدم سوء پیشینه				
۲۵	۲-۲-۳-ارائه گواهی تجربه کاری مرتبط				
۲۵	۳-۲-۳-ارائه گواهی تعداد پروژه‌های انجام شده مرتبط	۵۰	۲-۳-تجربه سازمانی		
۲۵	۴-۲-۳-مبلغ کل پروژه‌های مشابه				

۳۴	۱-۴- تعداد کل دعاوی در ۵ سال گذشته				
۳۳	۲-۴- تعداد دعاوی ارجاع شده به داوری در ۵ سال گذشته	-	-	۱۳	۴- مدیریت دعاوی
۳۳	۳-۴- تعداد دعاوی ارجاع شده به دادگاه در ۵ سال گذشته				
۳۴	۱-۵- گردش مالی				
۳۳	۲-۵- دارایی‌های جاری	-	-	۱۲	۵- وضعیت مالی
۳۳	۳-۵- میزان اعتبارات				
۳۵	۱-۱-۶- ساختار مدیریت پروژه، مهارت و شایستگی‌های مدیریت پروژه				
۳۳	۲-۱-۶- به کارگیری تحلیل و مدیریت ریسک در سازمان و روش اجرا	۵۰	۱-۵- مهارت در مدیریت پروژه		
۳۲	۳-۱-۶- تعداد مدیران حرفه‌ای PMP				
۲۷	۱-۲-۶- استفاده از مدیریت ارزش و یا سایر تکنیک‌ها در اجرا، برای کاهش زمان و هزینه‌ها			۱۲	۶- مهارت در مدیریت پروژه و روش‌ها و راه حل‌های فنی و ابتکاری
۲۶	۲-۲-۶- مبلغ صرفه جویی شده از پیشنهادهای مصوب				
۲۵	۳-۲-۶- طرح‌ها و راه‌حل‌های پیشنهادی برای کاهش زمان و هزینه و بهبود کیفیت	۵۰	۲-۵- روش‌ها و راه‌حل‌های فنی و ابتکاری		
۲۲	۴-۲-۶- متوسط کارگاه‌های آموزشی برگزار شده برای توجیه کارکنان پروژه‌های ۵ سال گذشته				
۱۴	۱-۷- گواهی رضایت مشاور از انجام پروژه‌های ۵ سال گذشته				
۱۴	۲-۷- گواهی رضایت کارفرما از انجام پروژه‌های ۵ سال گذشته	-	-	۱۱	۷- حسن شهرت شرکت و ارتباطات

بازرگانی قرار بگیرد. نتایج پژوهش به صورت خلاصه موارد زیر را به آیین نامه اجرایی بند "ج" ماده (۱۲) قانون برگزاری مناقصات، اضافه نموده است:

- دو معیار اصلی "مدیریت دعاوی" و "مدیریت ایمنی و بهداشت"
- و دو زیرمعیار "بهداشت، ایمنی و محیط زیست" و "کنترل کیفیت"
- سه شاخص ارزیابی برای ارزیابی توانایی مدیریت دعاوی و هشت شاخص ارزیابی برای ارزیابی توانایی مدیریت ایمنی و بهداشت.
- تعیین وزن معیارهای اصلی، زیر معیارها و شاخصهای ارزیابی به صورت تخصصی برای مناقصات صنعت آب.

۴- نتیجه گیری

پیچیدگی و تنوع پروژه‌های صنعت ساخت باعث شده است که تاکنون روش‌های متفاوتی برای ارزیابی مقدماتی پیمانکار مطرح شود. مسئله انتخاب پیمانکار یک مسئله تصمیم‌گیری چند معیاره است و به دلیل ویژگی‌های خاص این نوع تصمیم‌گیری، یعنی اطلاعات نامطمئن و عدم قطعیت‌ها، فشار زمان، کمبود اطلاعات و... باید نگاهی دقیق‌تر به این موضوع داشت. همان گونه که در پژوهش حاضر تشریح شد، وابستگی روش‌های تصمیم‌گیری به قضاوت فردی همواره یک معضل محسوب می‌شود به همین دلیل باید به روش‌هایی روی آورد که این وابستگی را هدفمند کند و با تشکیل تیمی از کارشناسان برای تصمیم‌گیری، وابستگی به قضاوت فردی را کم‌رنگ و بر خرد جمعی و مستند سازی تکیه نماید.

بر اساس مطالعات انجام شده در ادبیات موضوع، استفاده از پرسشنامه و تحلیل‌های آماری بیشترین کاربرد را در شناسایی معیارهای ارزیابی پیمانکاران دارد اما یکی از نقاط ضعف استفاده از پرسشنامه، جهت‌گیری^۱ مخاطب تحت تاثیر شیوه طراحی پرسشنامه می‌باشد. محققان با درک این مشکل، از ابزار متن کاوی نیز در کنار تحلیل آماری پرسشنامه‌ها استفاده کردند. در این پژوهش از دو ابزار پرسشنامه و متن کاوی مصاحبه با خبرگان برای شناسایی معیارهای اصلی، زیرمعیارها و شاخصهای ارزیابی پیمانکاران در مناقصات صنعت آب بهره گرفته شده و نتایج پژوهش با فصل سوم آیین نامه اجرایی بند "ج" ماده (۱۲) قانون برگزاری مناقصات نیز مقایسه شده است. برای اطمینان از تعمیم پذیری نتایج تحقیق، ۵۰ نفر از

است. قرابت و انطباق زیاد نتایج حاصل از ابزار اول (پرسشنامه و تحلیل‌های آماری) و ابزار دوم (مصاحبه و متن کاوی) بیانگر این است که معیارها، زیر معیارها و شاخصهای ارزیابی پیمانکاران در مناقصات، با دقت و قابلیت تعمیم قابل قبولی استخراج شده است.

معیارها، زیر معیارها، شاخصهای ارزیابی و امتیازهای آنها که توسط دو ابزار متفاوت (پرسشنامه و متن کاوی) تعیین شده‌اند در جدول ۵ نمایش داده شده است.

در فصل سوم آیین نامه اجرایی بند "ج" ماده (۱۲) قانون برگزاری مناقصات که توسط سازمان مدیریت و برنامه‌ریزی کشور منتشر شده است، معیارهای ارزیابی پیمانکاران معرفی شده است. نتایج پژوهش انجام شده دو معیار اصلی "مدیریت دعاوی" و "مدیریت ایمنی و بهداشت" و دو زیر معیار "بهداشت، ایمنی و محیط زیست" و "کنترل کیفیت" را به آیین نامه مذکور، اضافه کرده است. این آیین نامه مصوبه هیئت وزیران در سال ۱۳۸۳ می‌باشد. حدوداً ۱۵ سال قبل و هنگام تصویب آیین نامه، مدیریت دعاوی و مدیریت ایمنی و بهداشت جز موضوعات مهم در ارزیابی پیمانکاران نبوده است اما امروزه بسیاری از پروژه‌ها به واسطه ضعف در مدیریت دعاوی و مدیریت ایمنی، زمان و هزینه زیادی را از دست می‌دهند و مطابق بررسیهای انجام شده در ادبیات تحقیق، اکثر پژوهشهایی که پس از سال ۲۰۱۰ در حوزه ارزیابی پیمانکاران در صنعت آب انجام شده‌اند، مدیریت دعاوی و مدیریت ایمنی را به عنوان معیارهای ارزیابی ذکر کرده‌اند. بنابراین مدیریت دعاوی و مدیریت ایمنی جز معیارهای مهم ارزیابی پیمانکاران در صنعت آب می‌باشند. برای ارزیابی دو معیار «مدیریت دعاوی» و «مدیریت ایمنی و بهداشت»، نیاز به شاخصهای ارزیابی می‌باشد. در این پژوهش ۳ شاخص برای ارزیابی توانایی مدیریت دعاوی پیمانکاران و ۸ شاخص برای ارزیابی توانایی مدیریت ایمنی و بهداشت پیشنهاد شده است. معیارهای معرفی شده در این آیین نامه به صورت عمومی برای مناقصات در تمامی حوزه‌ها می‌باشد، بنابراین در این آیین نامه وزن معیارها تعیین نشده است و تعیین وزن معیارها با توجه به ماهیت صنعتی که مناقصه در آن برگزار می‌شود، بر عهده کمیته فنی بازرگانی دستگاه مرکزی گذاشته شده است. از آنجا که تمرکز این پژوهش بر روی مناقصات صنعت آب بوده، وزن معیارها، وزن زیر معیارها و وزن شاخصهای ارزیابی به صورت اختصاصی برای مناقصات صنعت آب، محاسبه و ارائه شده است تا مورد استفاده کمیته فنی

تشکر و قدردانی

از همکاری پرسنل محترم شرکت مهندسی مشاور مهتاب قدس در انجام این تحقیق کمال تشکر و قدردانی را داریم.

پیوست

بخش‌هایی از متن مصاحبه‌ها

پرسش: با توجه به اهمیت ارزیابی پیمانکاران در مناقصات صنعت آب، به نظر شما چه عواملی در ارزیابی پیمانکاران باید مدنظر قرار بگیرد؟

جناب آقای سید علی اعتصامی

تحصیلات: کارشناسی

سابقه کار: بیش از ۳۰ سال

امروزه علاوه بر قیمت پیشنهاد شده از سوی پیمانکاران و پیشنهادات فنی، باید به سابقه دعاوی پیمانکاران نیز توجه شود، زیرا تعیین تاخیرهای مجاز و غیرمجاز، وقت و منابع زیادی طلب میکند. در بررسی سوابق دعاوی باید شیوهی مورد علاقهی پیمانکاران در حل دعاوی نیز توجه شود. بدیهی است که پیمانکارانی که برای حل دعاوی از روشهای غیر دادگاهی استفاده نموده‌اند، نسبت به سایر پیمانکاران ارجحیت دارند.

ترجمه‌ی استفاده شده در متن کاوی:

Today, in addition to the price offered by contractors and technical offers, one must also pay attention to contractors' claims history because It requires a lot of time and resources to determine the amount of unauthorised and unauthorised delays. The contractors' preferred approach to litigation should also be considered when reviewing litigation. Obviously, contractors who have used non-litigation methods have a higher preference for other contractors

خبرگان صنعت آب از دو شرکت متفاوت انتخاب شده و به صورت تصادفی به دو گروه با تعداد مساوی تقسیم شدند، ابتدا با استفاده از ابزار پرسشنامه و تحلیل‌های آماری، معیارها و زیر معیارهای ارزیابی پیمانکاران همراه با امتیازات آن‌ها شناسایی شده و سپس با گروه دوم خبرگان مصاحبه‌ی باز انجام شد و متن مصاحبه‌ها مورد متن کاوی و خوشه‌بندی قرار گرفت و نتایج دو روش مقایسه شد. موضوعات پنج عدد از خوشه‌ها انطباق زیادی با پنج معیار اصلی و زیر معیارهای آن‌ها داشتند و خوشه شماره شش نیز معیار مدیریت دعاوی را به معیارهای اصلی قبلی اضافه کرد. همچنین با تحلیل کلمات کلیدی هر یک از خوشه‌ها، شاخصهای ارزیابی برای هر یک از معیارها و زیرمعیارها معرفی شدند و در نهایت معیارها، زیرمعیارها و شاخصهای ارزیابی پیمانکاران همراه با امتیازات آن‌ها در قالب یک جدول نمایش داده شدند. یافته‌های این پژوهش نشان می‌دهد معیارهای؛ ظرفیت کاری، مهارت‌های فنی و داشتن کارکنان کلیدی خبره، عملکرد پیمانکار در پروژه‌های گذشته، دعاوی مطرح شده توسط پیمانکار در پروژه‌های قبلی، توان مالی، مهارت‌های مدیریت پروژه و سوابق حل خلاقانه مسائل و مهندسی ارزش، حسن شهرت و HSE مهم‌ترین معیارهایی است که باید در ارزیابی پیمانکاران مورد نظر قرار گیرد. نتایج پژوهش علاوه بر پوشش تمامی معیارهای معرفی شده در آیین‌نامه اجرایی بند "ج" ماده (۱۲) قانون برگزاری مناقصات، موارد زیر را به این آیین‌نامه، اضافه نموده است:

- دو معیار اصلی "مدیریت دعاوی" و "مدیریت ایمنی و بهداشت" به همراه دو زیرمعیار "بهداشت، ایمنی و محیط زیست" و "کنترل کیفیت"
- سه شاخص ارزیابی برای ارزیابی توانایی مدیریت دعاوی و هشت شاخص ارزیابی برای ارزیابی توانایی مدیریت ایمنی و بهداشت.
- تعیین وزن معیارهای اصلی، زیر معیارها و شاخصهای ارزیابی به صورت تخصصی برای مناقصات صنعت آب.
- مقایسه نتایج به دست آمده از دو ابزار مذکور نشان دهنده دقت و قابلیت تعمیم مناسب نتایج پژوهش می‌باشد. در نهایت پیشنهاد می‌شود که در پژوهش‌های آتی با استفاده معیارها، زیرمعیارها و شاخص‌های ارزیابی پیمانکاران که در این پژوهش معرفی شدند، به موضوع سیستم‌های پشتیبانی تصمیم‌گیری (DSS) پرداخته شود.

Keywords: technical, offers, technical, delays, litigation, methods

معمولا در پروژههای عمرانی از معیارهای معرفی شده در قانون برگزاری مناقصات برای ارزیابی پیمانکاران استفاده می شود. یکی از معیارهای معرفی شده توسط این قانون، حسن سابقه در کارهای قبلی میباشد. در اکثر موارد کارفرما به گواهی حسن انجام کاری که پیمانکار ارائه کرده بسنده میکند اما باید به این موضوع توجه کرد که اکثر کارفرماها این گواهی را پس از اتمام کار به پیمانکار ارائه میکنند. پس پیشنهاد میشود که برای بررسی حسن سابقه پیمانکار در کارهای قبلی - که از نظر اینجانب مهم ترین معیار کیفی ارزیابی پیمانکاران میباشد - حتما با کارفرماهای قبلی تماس گرفته شود و همچنین به روابط بلندمدت پیمانکار با کارفرماها توجه شود.

ترجمه‌ی استفاده شده در متن کاوی:

Construction projects use the criteria set out in the Tender Law to evaluate contractors. One of the criteria introduced by this law is good track record in previous work. In most cases the employer is satisfied with the certificate of good performance of the contractor, but it should be noted that most employers provide this certificate to the contractor after the job is completed. It is suggested that to check the contractor «s record in previous works, it must be contacted by previous contractors and will also address the long - term contractor» s long - term relationships with employers

Keywords: record, previous, satisfied, certificate, long-term, relationships

تشریح روش TF-IDF

TF-IDF از دو قسمت TF و IDF تشکیل شده است که هرکدام را با ذکر یک مثال محاسبه میکنیم. کلمه‌ی RISK، ۳ بار در خلال متن اول تکرار شده است. این متن دارای ۱۷ کلمه کلیدی است. اگر بخواهیم TF را برای این کلمه در این متن محاسبه کنیم، باید تعداد تکرار این کلمه در این متن را محاسبه کنیم و آن را بر

جناب آقای مهندس علی اصغر جلال زاده

تحصیلات: کارشناسی ارشد

سابقه کار: بیش از ۳۰ سال

مدیریت ایمنی و مباحث مرتبط با HSE نقش بسیار مهمی در موفقیت پروژهها دارد. در صورت بروز حوادث، هزینه‌های مستقیم و غیرمستقیم حوادث به پروژه تحمیل میشود که ارتباط بین آنها مانند کوه یخ شناور در آب میباشد! امروزه در ارزیابی کیفی پیمانکاران، مدیریت ایمنی یکی از نادیده گرفته شده ترین مسائل می باشد، پس بهتر است که علاوه بر معیارهای متداول مانند قیمت پیشنهادی، پیشنهاد فنی، وضعیت مالی و ... به سوابق ایمنی پیمانکاران نیز توجه شود. پیشنهاد میشود که از شرکتهای بیمه نیز استعلام گرفته شود.

ترجمه‌ی استفاده شده در متن کاوی:

Safety management and HSE related issues play a vital role in the success of projects. In the event of accidents, direct and indirect costs of the project are imposed on the project, which are interconnected, such as icebergs! Today in contractors 'quality assessment, safety management is one of the most overlooked issues, so it is best to take into account contractors' safety records in addition to common criteria such as bid price, technical proposal, financial status and ... It is also suggested that insurance companies be queried

Keywords: Safety, management, HSE, accidents, costs, direct, indirect, insurance, records, records, technical

جناب آقای داوود رحمانیان

تحصیلات: دکتری

سابقه کار: بیش از ۳ سال

- [5] J.B. Yang, W.C. Wang, Contractor selection by the most advantageous tendering approach in Taiwan, *Journal of the Chinese Institute of Engineers*, 26 (2003) 381-387.
- [6] Z. Hatush, M. Skitmore, Assessment and evaluation of contractor data against client goals using PERT approach, *Construction Management & Economics*, 15 (1997) 327-340.
- [7] S.T. Ng, R.M. Skitmore, Client and consultant perspectives of prequalification criteria, *Building and environment*, 34 (1999) 607-621.
- [8] E. Palaneeswaran, M. Kumaraswamy, Recent advances and proposed improvements in contractor prequalification methodologies, *Building and Environment*, 36 (2001) 73-87.
- [9] Y.I. Topcu, A decision model proposal for construction contractor selection in Turkey, *Building and environment*, 39 (2004) 469-481.
- [10] M.K. J. Han, J. Pei, Data Mining concepts and techniques, (2012) 1-2.
- [11] D. Pickell, Structured vs Unstructured Data – What's the Difference?, (2018).
- [12] M. de Miranda Santo, G.M. Coelho, D.M. dos Santos, L. Fellows Filho, Text mining as a valuable tool in foresight exercises: A study on nanotechnology, *Technological Forecasting and Social Change*, 73 (2006) 1013-1027.
- [13] N. Singh, C. Hu, W.S. Roehl, Text mining a decade of progress in hospitality human resource management research: Identifying emerging thematic development, *International Journal of Hospitality Management*, 26 (2007) 131-147.
- [14] N.T. Ratrou, Subtractive clustering-based k-means technique for determining optimum time-of-day breakpoints, *Journal of Computing in Civil Engineering*, 25 (2010) 380-387.
- [15] M. Al Qady, A. Kandil, Automatic clustering of construction project documents based on textual similarity, *Automation in construction*, 42 (2014) 36-49.
- [16] R. Yan, Z. Ma, G. Kokogiannakis, Y. Zhao, A sensor fault detection strategy for air handling units using cluster analysis, *Automation in Construction*, 70 (2016) 77-88.
- تعداد کل کلمات در همین متن تقسیم کنیم. تعداد تکرار کلمه‌ی ریسک در متن اول ۳ است، و تعداد کل کلمات در این متن هم ۱۷ است، پس با تقسیم عدد ۳ بر ۱۷ عدد ۰/۰۴ به دست می‌آید که مقدار TF برای کلمه‌ی ریسک در متن اول را نشان می‌دهد. به صورت شهودی هر چقدر تعداد تکرار یک کلمه در یک متن نسبت به تعداد کلمات آن متن بیشتر باشد، مقدار TF بیشتر می‌شود. در واقع مقدار TF نشان دهنده‌ی غلظت یک کلمه در متن است. برای محاسبه‌ی IDF باید تعداد کل متون محاسبه شود که در اینجا برابر با ۳ می‌باشد، این مقدار را باید بر تعداد متونی که کلمه‌ی ریسک در آن‌ها موجود است، تقسیم کنیم. همان‌طور که می‌بینیم کلمه‌ی ریسک در ۲ متن تکرار شده است. پس باید عدد ۳ (تعداد کل متون) را بر عدد ۲ (تعداد متونی که کلمه‌ی ریسک در آن‌ها آمده است) تقسیم کنیم. سپس یک لگاریتم بر روی حاصل این تقسیم اعمال می‌شود (لگاریتم موجب نرمال شدن نتیجه می‌شود). در نهایت مقدار TF را در IDF ضرب می‌کنیم تا مقدار TF-IDF برای کلمه‌ی ریسک در متن اول محاسبه شود. در زیر خلاصه‌ی روش مذکور برای کلمه‌ی ریسک در متن ۱ نشان داده شده است:
- TF-IDF (تعداد تکرار کلمه در متن / تعداد کلمات متن) * LOG (تعداد کل متون / تعداد متونی با کلمه مورد نظر)
- $$(0.007) = TF-IDF = (3/17) * LOG (3/2)$$

مراجع

- [1] M. Sönmez, J.B. Yang, G.D. Holt, Addressing the contractor selection problem using an evidential reasoning approach, *Engineering, Construction and Architectural Management*, 8 (2001) 198-210.
- [2] K.C. Lam, S. Thomas Ng, T. Hu, M. Skitmore, S.O. Cheung, Decision support system for contractor pre-qualification—artificial neural network model, *Engineering, Construction and Architectural Management*, 7 (2000) 251-266.
- [3] D.J. Watt, B. Kayis, K. Willey, Identifying key factors in the evaluation of tenders for projects and services, *International Journal of Project Management*, 27 (2009) 250-260.
- [4] G.D. Holt, Which contractor selection methodology?, *International Journal of project management*, 16 (1998) 153-164.

- [20] S. Saitta, P. Kripakaran, B. Raphael, I.F. Smith, Improving system identification using clustering, *Journal of Computing in Civil Engineering*, 22 (2008) 292-302.
- [21] I. Brilakis, L. Soibelman, Y. Shinagawa, Material-based construction site image retrieval, *Journal of computing in civil engineering*, 19 (2005) 341-355.
- [22] S. Yarmohammadi, R. Pourabolghasem, D. Castro-Lacouture, Mining implicit 3D modeling patterns from unstructured temporal BIM log text data, *Automation in Construction*, 81 (2017) 17-24.
- [23] Z. Ding, Z. Li, C. Fan, Building energy savings: Analysis of research trends based on text mining, *Automation in Construction*, 96 (2018) 398-410.
- [17] H. Naganathan, W.O. Chong, X. Chen, Building energy modeling (BEM) using clustering algorithms and semi-supervised machine learning approaches, *Automation in Construction*, 72 (2016) 187-194.
- [18] A.G. Kashani, A.J. Graettinger, Cluster-based roof covering damage detection in ground-based lidar data, *Automation in Construction*, 58 (2015) 19-27.
- [19] E. Rodrigues, D. Sousa-Rodrigues, M.T. de Sampayo, A.R. Gaspar, Á. Gomes, C.H. Antunes, Clustering of architectural floor plans: A comparison of shape representations, *Automation in Construction*, 80 (2017) 48-65.

چگونه به این مقاله ارجاع دهیم

M. Parchami Jalal, H. Shakiba Zahed, M. A. Ghaaderi, *Comparative Comparison of Contractors' Evaluation Criteria, Sub-Criteria and Indicators in Water Industry Tender Using Questionnaire and Text Mining. Amirkabir J. Civil Eng.*, 53(6) (2021) 2335-2352

DOI: [10.22060/ceej.2020.17388.6547](https://doi.org/10.22060/ceej.2020.17388.6547)

