https://doi.org/10.26682/csjuod.2023.26.2.20
Journal of University of Duhok., Vol. 26, No.2 (Pure and Engineering Sciences), Pp 201-215, 2023
4th International Conference on Recent Innovations in Engineering (ICRIE 2023) (Special issue)

A DATA-DRIVEN APPROACH TO SUPPLIER SELECTION IN INDUSTRIAL CONSTRUCTION PROJECTS

ALI AHMAD, FAYEZ JRAD, and SAMAH MAKIA
Dept. of Engineering and Construction Management, Faculty of Civil Engineering, Tishreen

(Accepted for Publication: November 27, 2023)

ABSTRACT

The process of selecting suppliers is a formidable challenge in managing industrial construction projects. Construction companies generate a huge amount of data that is spread across multiple databases, but this data is not used to support decisions for future projects. To address this issue, a data warehouse was developed specifically for construction companies which are specializing in industrial projects. The proposed data warehouse utilizes operational databases from past projects, including planning and execution details, to organize and analyze data for supplier selection decisions in ongoing and future projects. The warehouse's dimensional model was built to meet the requirements of construction enterprises and the available data. The methodology employs Online Analytical Processing (OLAP), which enables direct queries and generates relevant reports to support construction management decisions and evaluate suppliers for different types of work. By adopting this approach, construction companies can make more informed decisions about which suppliers to choose for their projects, ultimately improving their economic performance. By leveraging the data available to them, these companies can enhance the quality of their decisions and improve outcomes for their industrial construction projects.

KEYWORDS: Business Intelligence; Construction companies; Data warehouse; Industrial construction; Online Analytical Processing; Project management; Supplier selection

1. INTRODUCTION

The supplier selection process is classified as one of the most important decisions in industrial project management to achieve project success and completion within the specified deadline, as procurement is the largest and most important part of this type of projects. The process of supplier selection primarily involves evaluating all possible alternatives based on various criteria and factors, most notably quality, price, as well as manufacturing feasibility (Wang, et al., 2018), and then selecting the best alternative. This significantly reduces the final costs of the projects and improves the company's competitiveness in achieving projects with better performance. At the local level, the current stage, which involves the necessity of reconstruction and rehabilitation of infrastructure, is considered one of the most important priorities and challenges facing construction companies in the Syrian situation, which calls for changing the approach followed in managing construction projects, including the process of supplier selection in industrial projects, by investing in the historical data of companies and utilizing it to manage future projects in a modern scientific manner, relying on business intelligence techniques. The process of selecting suppliers plays a crucial and important role in increasing the competitiveness of construction companies

a-ahmad@tishreen.edu.sy
, fayezalijrad@gmail.com
samahsm2016@hotmail.com

in the future. Therefore, many studies have been conducted that have studied the issue of selecting suppliers in projects based on multiple criteria, the most important of which are cost, quality, time, geographical location, as well as technical capabilities, capacity, and other criteria (Kumar, 2017). The environmental criterion has also been added in later studies, which means achieving sustainability and considering environmental conditions within the mechanism of selecting suppliers (Konys, 2019). Over the past years, many researchers have developed various techniques and methods to solve the problem of selecting suppliers more efficiently. Different methodologies and methods have been used to solve the problem, where linear programming (LP) was used by (Kılıç, 2013) and (Ghodsypour and O'Brien, 1998), while the Analytic Hierarchy Process (AHP) was used by (Yahya and Kingsman, 1999) and (Deng, et al., 2014). Recently, several researchers have resorted to using Fuzzy Logic in solving this problem (Dalalah, et al., 2011), (Kannan, et al., 2014), (Kannan, et al., 2015), and (Gündüz Ç, 2019). While most of the previous studies focused on the limited scope of the problem and its restriction to specific conditions, and the inability to apply it to real projects, this study developed a suitable data warehouse for the reality of construction companies specializing in industrial projects. It was used to benefit from business intelligence techniques in supporting the decision of selecting suppliers for future projects by choosing the best suppliers in terms of commitment to delivering the required supplies and equipment on time, as specified in the project schedule. It is worth mentioning that this methodology was not used in previous studies to solve this problem, while this technique was used to solve many problems in the field of construction management. (Chau, 2002) developed a decision support system for construction management using a data repository with online analytical processing (OLAP) to obtain information and predict disputes for new construction projects. The researcher found that the method of collecting data in the data warehouse provides project managers with a powerful tool to make decisions and find appropriate solutions to construction problems. It converts operational data into summarized information for making strategic decisions. In another study, (Zhang et al., 2003) developed a data warehouse system according to company requirements and used Data Mining (DM) technology to obtain useful information and necessary knowledge from this repository. The repository included human resources, materials, equipment, schedule, quality, and cost. Special query tools were developed for direct query by the user using Online Analytical Processing (OLAP). A series of data mining models were built to support decision-making in the project planning phase, and the results proved its effectiveness in construction companies to support decision-making in this phase of the project life cycle. A decision support system was developed using data warehouse technology to assist contractors in selecting suitable construction sites for residential buildings. The data warehouse was built within an Access environment and was analyzed using the Analytic Hierarchy Process model (Rujirayanyong and Shi, 2006). Based on the concept that construction companies produce a large amount of operational data distributed across various databases, Rujirayanyong and Shi developed a directed data warehouse for construction projects at the company level, using a dimensional model. The basic information in each task included planning and execution information related to the plan, cost, resources, and change orders that occurred during implementation. The study recommended the importance of data warehousing in extracting useful information and improving the quality of decisions made. Another study, (Moon, 2007) presented a study aimed at improving the effective use of historical data for construction project costs through data warehousing and OLAP. A model was proposed to support OLAP processing functions, and a Cost Data Management System (CDMS) was developed to manage cost data and demonstrate the benefits of the OLAP environment in understanding the uncertainty in estimating construction costs. The experimental results showed that the proposed model provided a more reliable method for estimating construction costs. (Desai, 2010) proposed a methodology for applying the Decision Tree to analyze the productivity of construction workers. The methodology presented an improved version of the decision tree creation process that linked statistical methods and human expertise to create a more accurate decision tree in the forecasting process. The results of calculating labor needs compared to standard worker productivity showed that the proposed methodology was more realistic than traditional decision tree methods. The researcher concluded that this methodology could be applied to other areas of construction management for engineering projects according to the user's desire and the patterns to be predicted. (Makia et al., 2012) developed a data warehouse for construction equipment in engineering companies within the reality of the Syrian public sector for the Water Projects Company in the Lattakia branch. This warehouse allows for effective analysis and can search for information and perform online analytical processing (OLAP) for interactive analysis of construction equipment data and to achieve the required queries for obtaining information and knowledge quickly and directly. Data mining was also used to explore hidden knowledge in the data and to predict future faults, readiness, or productivity over the coming months or years. A new and highly efficient algorithm was presented for discovering clusters and outliers effectively within a data set. The recommendations of this study included the development of an integrated data warehouse for construction projects and researching to study the use of this warehouse in supporting the necessary decisions for managing various construction operations. (Zakaria, 2015) presented a study on selecting the best team (partners) for construction projects according to the Malaysian financial reality (project manager - surveyor - consultant - contractor, etc.). A relatively high percentage of construction projects in Malaysia experience delays, cost overruns, and poor execution due to poor choice of partners and neglect of reliability. Therefore, a new predictive model structure was built for selecting teams that give greater reliability using data mining technology. The data was collected from a public sector company, and the data mining model was developed and evaluated on real government projects. The model proved its ability to select the optimal team for the construction project with the lowest possible risk in terms of reliability at work. As a result, this model will effectively contribute to improving the government's savings system and will contribute to raising the level of construction projects for construction companies in the public sector. (Botha, 2018) evaluated the application of data mining in the construction sector to improve project management and applied it to a construction company in the public sector, the General Company for Transport and Public Works. The databases of these projects were collected, and the work was done to predict the planning stage on the number of job opportunities provided by the projects. This is related to the state's policy of combating unemployment, where the models were built in the Python environment. The recommendations of this study included studying how to apply data mining in each stage of the project in a way that helps to adopt it in improving construction project management. (Rong et al., 2020) used historical data, technical specifications, and financial statements for electricity network construction projects to determine the key factors affecting the cost to assist in making accurate decisions in electricity network construction projects and improving investment efficiency for electricity network companies by building a project cost analysis and data mining model to predict the planning stage through importing technical parameters such as the voltage level for the substation project and the type of distribution devices and the transmission capacity, so that the model can predict the financial investment value of the project. Previous studies have not addressed data mining techniques to support decision-making in the selection of suppliers for industrial projects. Most previous studies aimed to solve the problem of selecting suppliers for small projects that are unrealistic and cannot be applied to large projects with different conditions and constraints. Therefore, this research will use a multidimensional model to build a suitable data warehouse for construction companies specializing in industrial projects by relying on historical data. Basic information will be included in each task related to planning and implementation information regarding the plan and resources during the project implementation period to contribute to supporting the supplier selection decision, as well as predicting the expected delay in each system of the project according to the suppliers selected to provide materials and procurement for the project. This plays a role in improving the quality of ongoing projects and those planned for the future. At the end of the research, an evaluation of the proposed method will be presented, along with recommendations on the possibility of developing it in scheduling issues for construction projects. Supplies are an essential part of industrial projects, where the wrong choice of project suppliers and the lack of a clear method to rely on in the selection process can lead to a deviation from the project timeline as a result of the supplies not arriving on time, which can cause these projects to fail. Construction companies have thousands of paper and electronic files that contain data about previous projects, and therefore, this data must be harnessed and utilized to solve the problem of selecting suppliers for these projects. The current method relies on experience and individual decisions, which may cause significant losses that are invisible to companies that urgently need to adopt modern methods in all areas of project management. The study aims to build a suitable data warehouse for construction companies specializing in industrial projects that allows for effective analysis, the ability to search for information, and the exploration of hidden knowledge in the massive amount of data available to support decision-making in the supplier selection process for projects using online analytical processing tools. Additionally, the study aims to predict the expected delay rate in each section of the project according to the selected suppliers.

2. RESEARCH METHODOLOGY AND MATERIALS

The research methodology combines the use of online analytical processing (OLAP) and data mining techniques to analyze a dataset. OLAP focuses on analyzing multidimensional data typically stored in data warehouses, it is utilized for aggregating and summarizing data across different dimensions, while data mining involves employing neural networks to discover patterns, relationships, and trends within large datasets. The methodology includes data collection, OLAP analysis, data mining using neural networks, data analysis, and considering the acknowledgment of limitations. the study will rely on the descriptive method and content analysis to determine the theoretical framework for building data sources for the data warehouse to be constructed in order to be utilized to support decision-making in the supplier selection process for industrial projects through the application of OLAP on the designed data warehouse cubes. The study will also rely on the predictive analytical methodology by adopting one of the data mining techniques, which is Artificial Neural Network (ANN), valuable insights can be extracted from the dataset, contributing to supporting decisions about supplier selection problems in industrial projects and predicting the expected delay rate in each section of the project according to the selected suppliers. Fig (1) clarifies the flowchart of the working mechanism for the research.

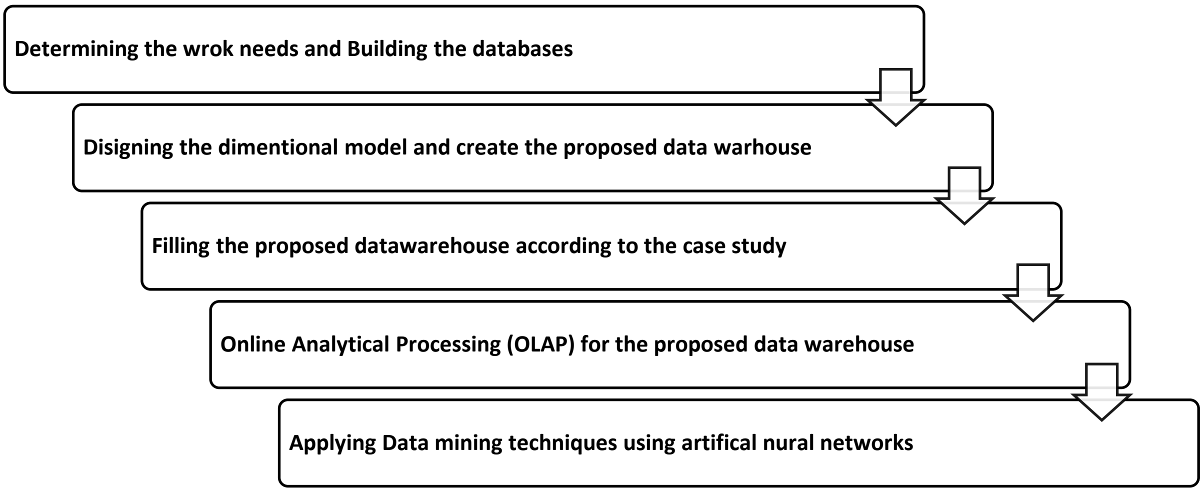

Fig.(1):- The flowchart of the working mechanism

3. ARTIFICIAL NEURAL NETWORK (ANN)

Neural networks, which emulate the functioning of the human brain, are computational models designed to process information by recognizing patterns and understanding the relationships between various system parameters (Kohzadi et al., 1995). These artificial neural networks draw inspiration from the structure and behavior of biological neural networks found in the human brain. The organization of artificial neural networks mirrors that of the human brain, with neurons distributed across different layers as shown in Fig. 2. including input, hidden, and output layers. In the artificial neural network, input values are transmitted through interconnected nodes known as neurons via weighted connections called synapses. Each input value is multiplied by its corresponding connection weight, representing the strength of that particular node. The weighted inputs are then propagated through the network, ultimately reaching the output layer. To make sense of these numerical values, an activation function is applied to each node, transforming the input signal into an output signal that serves as input for the subsequent layer. The training of neural networks involves an iterative process of adjusting the connection weights to minimize the difference between the actual output and the predicted output. This process relies on a training dataset, which is typically divided into three sets training, validation, and test sets, with the training set's error used to update weights. The validation set's error aids in determining when to halt training as the error begins to rise, while the test set's error facilitates model comparisons (Beale et al., 2010). Overall, artificial neural networks offer a powerful approach to processing information and recognizing patterns by simulating the behavior of biological neural networks. They have found applications in various fields, ranging from image and speech recognition to financial forecasting and medical diagnosis, contributing to advancements in machine learning and artificial intelligence research. Given the extensive amount of data collected and organized in our data warehouse, we have chosen to employ Artificial Neural Network (ANN) for our analysis. This powerful computational model offers an ideal solution for handling a large volume of information. By training the neural network, we aim to harness its predictive capabilities to anticipate delays in each system of the project. Our focus lies particularly on the suppliers responsible for providing the necessary materials and equipment for the studied system. Through the ANN's ability to learn from the data patterns and relationships, we anticipate gaining valuable insights that will aid in understanding the factors contributing to project delays. With this understanding, we can make well-informed decisions to address and mitigate potential delays effectively.

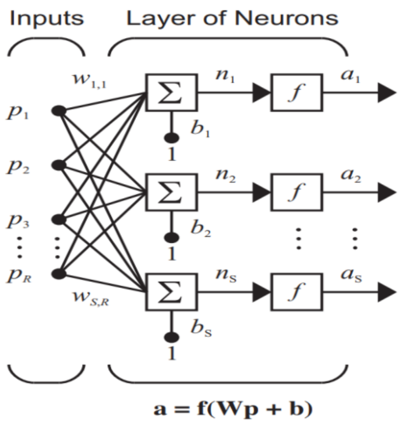

Fig.(2):- Schematic Diagram of the neural network architecture (Beale et al., 2010)

4. BUILDING DATABASES AND IDENTIFYING BUSINESS NEEDS

5.

With the maturation of relational databases in the 1980s, Online Transaction Processing (OLTP) systems were built using relational systems to automate business operations and manage daily processes. However, these systems were unable to meet decision-makers requirements in terms of speed, accuracy, and data volume that they could handle. Hence, the need arose for Decision Support Systems (DSS), a special type of computerized information system that support decision-making activities in organizations and companies. DSS differ from operational databases in their ability to analyze data and look beyond the data to obtain information and draw conclusions and identify new trends. Data Warehouses are the mainstay of DSS (Inmon, 2005). The necessary information and business topics to be focused on in the proposed data warehouse for tracking projects of specialized industrial companies have been identified. The general template for the necessary databases to build a suitable data warehouse for tracking projects of specialized industrial companies was designed and built accordingly. The previous projects' data of MAPNA GROUP of Iran were taken as a case study in this research. This company specializes in engineering and construction and developing thermal power plants, renewable energy power plants, and synchronized (shared) power generation facilities for electricity production. The company holds a 50% stake in expanding the capacity of electricity generation plants in Iran, in addition to effective participation in oil and gas projects, transportation by rail, health, and electricity. It has become one of the leading gas turbine owners worldwide, and Table (1) shows the company's project list that was relied upon in building the proposed data warehouse, while the number of the projects is fifteen projects divided between Iraq and Iran, and it should be noted that the data of "Kahnuj Combined Cycle Power Plant" project has been imported for only the first four months of the project because the original data of this project will be compared with the results of the proposed model during the testing stage. And the total activities of these projects are 32,515 activities distributed across 1,223 systems, and each system from these systems will be utilized later as an input for the proposed neural network. The data sources that were collected were in the form of Xlsx and Xer:

5.1 Xlsx tables

They include databases for the company's projects, project owners, secondary suppliers and contractors, evaluation, weather conditions, units, regions, and systems, in addition to monthly evaluation of secondary contractors and suppliers for each department of the company throughout the entire project life cycle.

5.2 XER tables

They include the company's project databases stored in the Primavera P6 environment, including databases of completed tasks in the company's projects, in addition to monthly tracking data for each of the previous projects. It is worth noting that the company mainly relies on the Primavera P6 program to plan and track projects.

Table(1):-List of the projects in the case study

Project. ID	Project name	Project short name	Country	City
1	Al Sadr Simple Cycle Power Plant	SSCPP	Iraq	Baghdad
2	Chabahar Gas Power Plant	CGPP	Iran	Baluchistan
3	Heidarieh (Najaf) Power Plant	HPP	Iraq	Najaf
4	Kahnuj Combined Cycle Power Plant	KCCPP	Iran	Kahnuj
5	Kashan Gas Power Plant	KGPP	Iran	Kashan
6	Kermanshah Zagros Power Plant	KPP	Iran	Kermanshah
7	Semnan Gas Power Plant	SGPP	Iran	Semnan
8	Yazd Combined Cycle Power Plant	YCCPP	Iran	Yazd
9	Soltaniyeh Gas Power Plant	SGPP	Iran	Zanjan
10	Khorramshahr Gas Power Plant	KGPP	Iran	Khorramshahr
11	Sanandaj Gas Power Plant	SGPP	Iran	Sanandaj
12	Isfahan II Combined Cycle Power Plant	ICCPP	Iran	Zavareh
13	Urmia Gas Power Plant	UGPP	Iran	Urmia
14	Shirvan Combined Cycle Power Plant	SCCPP	Iran	Shirvan
15	Shahrud Thermal Power Plant	SPP	Iran	Semnan
Number of Activities			32,515	

6. DESIGNING THE DIMENSIONAL MODEL AND BUILDING THE PROPOSED DATA WAREHOUSE

It is a specific model that collects data in a summarized format and consists of fact and dimension tables. All detailed data that needs to be retained is stored in the data warehouse in fact tables. Tracking tables, the performance of secondary contractors and suppliers, schedules in primavera format, progress tables, project performance tables, and system tables were adopted as fact tables for the data warehouse to be built. These tables are linked to many smaller dimension tables that contain descriptive characteristics of the collected data. The fact and dimension tables in the data warehouse were organized using star schemas, while the proposed “Bus Architecture Matrix” by (Kimball, 2002) was used to gather the company's operational processes and obtain the main system that describes the data warehouse from the inside. Fig. 3. shows the dimensional model that was built. The Data Source was built by exporting the collected databases to Microsoft SQL Server, as it provides optimal analysis speed from multiple analyses to data mining. The built data warehouse was imported into Microsoft Power BI, which provides a suitable environment for Business Intelligence. Thr

ough the dimensional structure of the data, it allows for data review and analysis in a flexible manner.

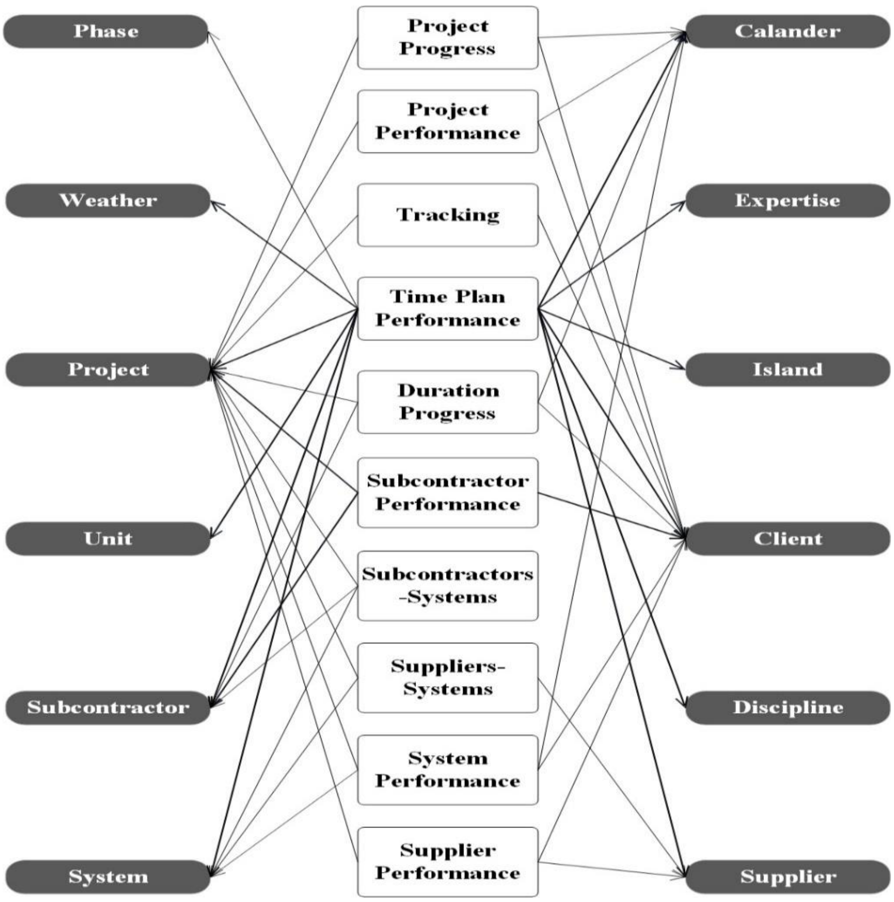

Fig.(3):-The dimensional model for the proposed data warehouse

7. RESULTS AND DISCUSSION

7.1 Online analytical processing (OLAP) on the

a-ahmad@tishreen.edu.sy
, fayezalijrad@gmail.com
samahsm2016@hotmail.com

designed data warehouse

The application of Online Analytical Processing (OLAP) in the proposed data warehouse aims to provide a new and effective tool that allows for flexible and easy visual analysis of data. This will enable decision-makers to identify problem areas and reduce errors in decision-making at the management level when selecting suppliers for each department of ongoing projects. Online Analytical Processing (OLAP) will be applied in this research paper to the suppliers' performance cube, where the data warehouse was exported to Microsoft Power BI software to display numerical values in visual charts in a Business Intelligence environment, while Microsoft Power BI is a business intelligence software that empowers organizations to visualize and analyze their data effectively. It provides a comprehensive suite of tools and features for data modeling, data preparation, interactive data visualization, and collaborative sharing of insights. it enables businesses to unlock the full potential of their data, make data-driven decisions, and communicate insights effectively across the organization. This will help project management at all levels to easily evaluate previous projects and make optimal decisions. Fig. 4. shows a comparison of the average delays for the electrical procurement suppliers in the turbine and generator island, while Fig. 5. shows a comparison of the average delays for the instrument and control procurement suppliers in the turbine and generator island in terms of delivery delays. It can be noted that “B. Iran Transfo” is the most committed supplier in delivering electrical equipment for the turbine and generator island with the lowest average delays in delivery on the site, meanwhile, “Elektro Kavir” was the most delayed supplier as shown in Fig. 4. On the other hand, Fig. 5. Shows that Tose'eh Nirou (T.N.C) was the most delayed supplier for the instrument and control equipment in turbine and generator island, while “Payand Paya Parseh” was the best one.

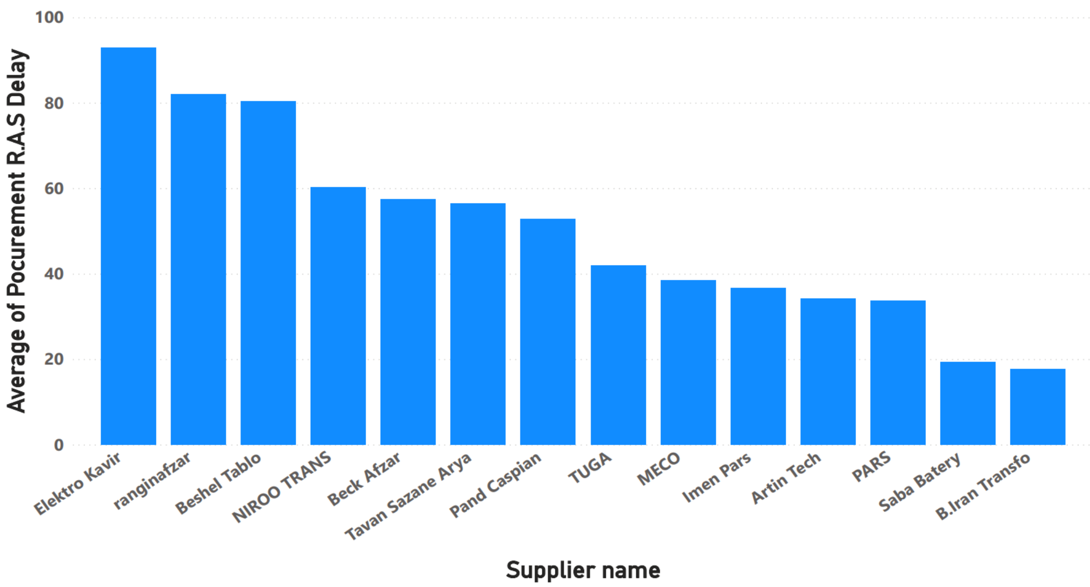

Fig.(4):- Average of procurement receive at site delay for electrical suppliers in turbine and generator island

Journal of University of Duhok., Vol. 26, No.2 (Pure and Engineering Sciences), Pp 201-215, 2023
4th International Conference on Recent Innovations in Engineering (ICRIE 2023) (Special issue)

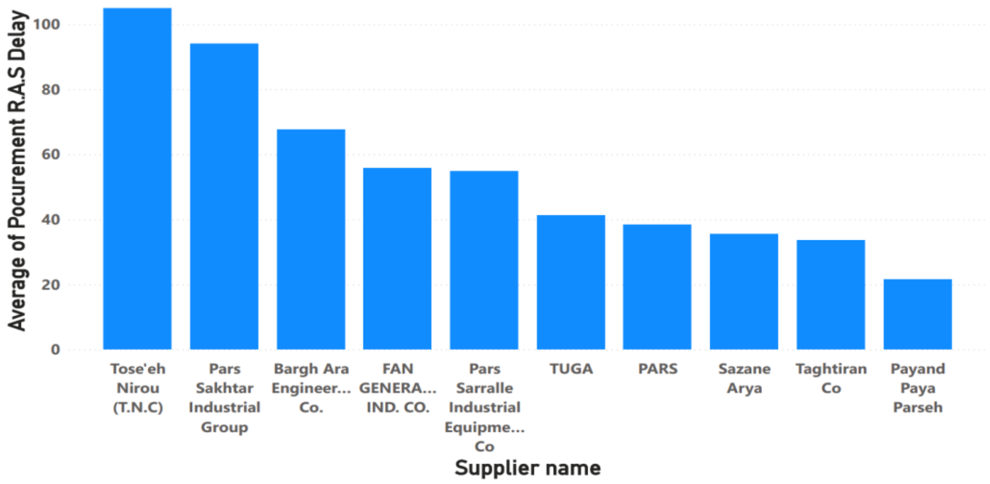

Fig.(5):- Average of procurement receive at site delay for instrument & control suppliers in turbine and generator island

Fig. 6. shows a comparison of the average delays for the mechanical procurement suppliers in the control island, and shows that "Pars Sarralle Industrial Equipment Co" is the most committed supplier in delivering mechanical equipment for the control island with the lowest average delays in delivery on the site, while "FAN GENERATOR IND. CO" was the most delayed supplier in this regard.

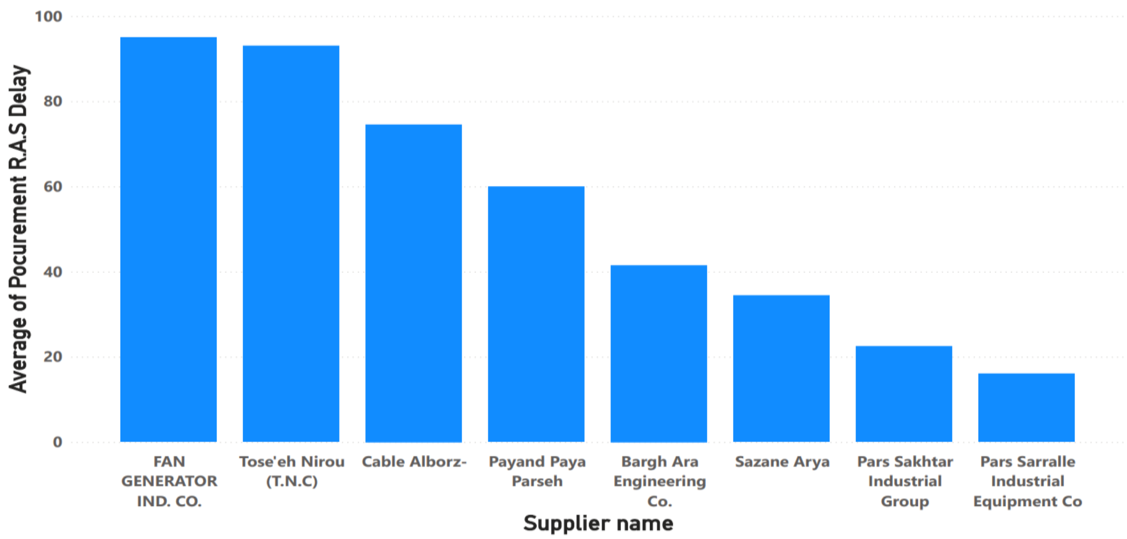

Fig.(6):- Average of procurement receive at site delay for mechanical suppliers in control island

Fig. 7. shows a comparison between the quality evaluation for mechanical and electrical suppliers in the main and auxiliary cooling island, and it can be noted that "B. Iran Transfo" had the best quality evaluation for the electrical procurement, while "Artin Tech" and "Taghtiran Co" had the worst evaluation for the electrical and mechanical respectively.

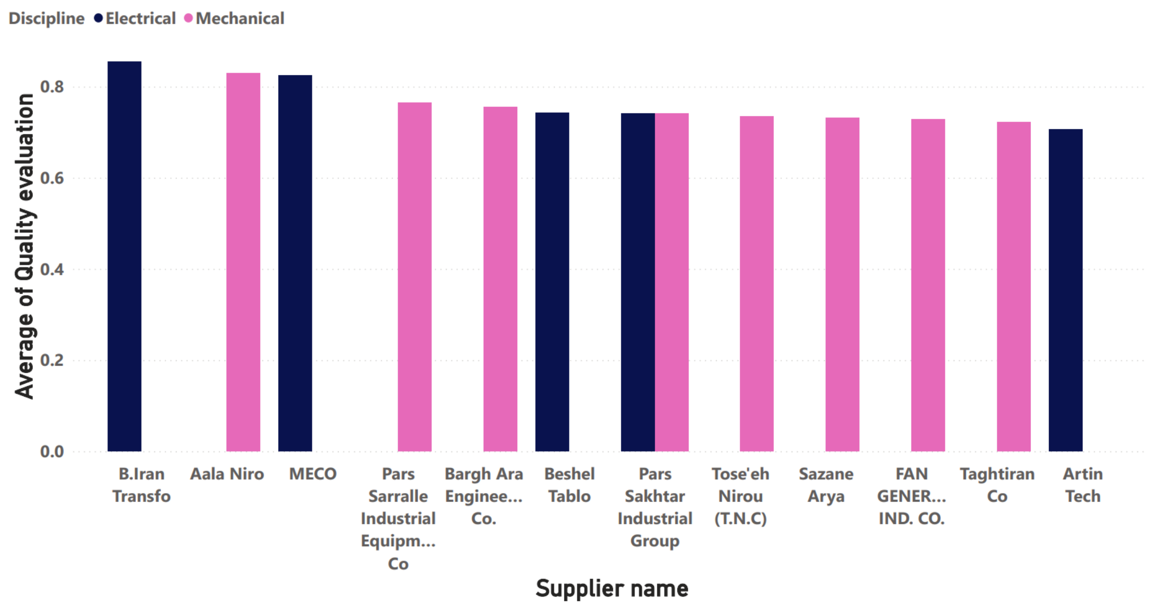

Fig.(7):- Average of quality evaluation for mechanical and electrical suppliers in the main and auxiliary cooling island

Fig. 8. shows a comparison between the islands regarding the average of the delays of the procurement receive at the site for mechanical and electrical works in previous projects. The chart shows that the water and steam cycle island, which is one of the most important areas in power generation stations, was the most delayed island in delivery on the site for the electrical and mechanical procurement in the previous projects. High efficiency is needed to carry out these works in future projects, which supports the decision to focus on contracting with reliable suppliers and contractors who are highly efficient in implementing these works.

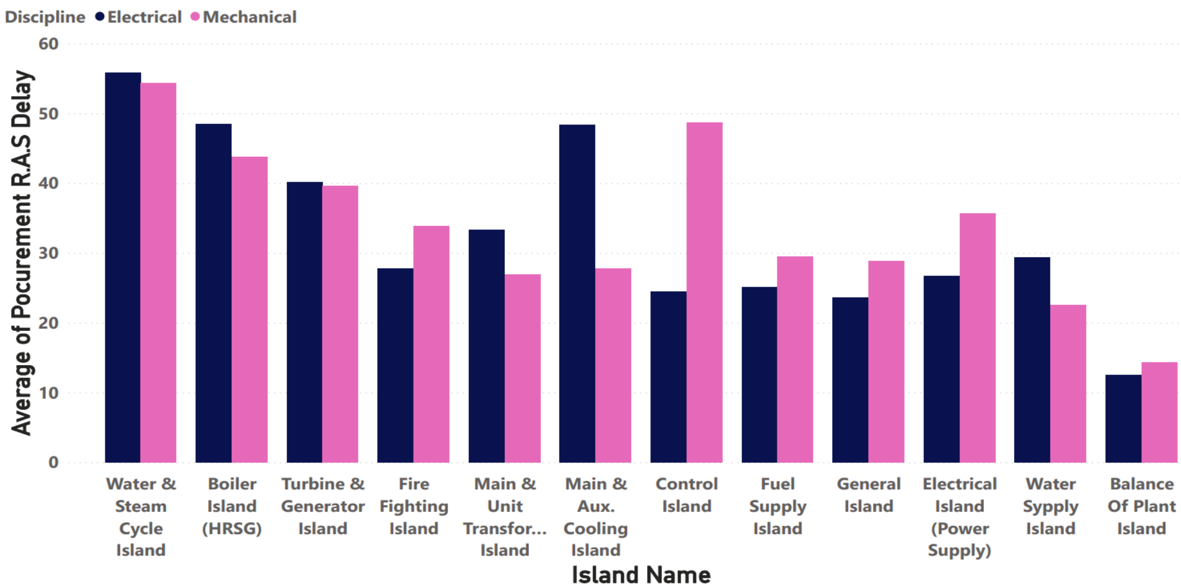

Fig.(8):- Average of procurement receive at site delay for mechanical and electrical works in each Island

7.2 Data mining of the proposed data warehouse

Data mining techniques aim to explore useful information within largely automated data sources by finding hidden patterns that may remain unknown without them. In addition to their ability to predict the future, the value of data mining lies in understanding the reasons for past events and the ability to predict what will happen in the future. The data mining approach relies on many machine learning algorithms for cognitive discovery, and one of the algorithms used in the data mining process is Artificial Neural Networks (ANNs) for its role in predicting the future based on stored massive data sources. The data mining model will be applied to the performance cube of resources for each project system to predict the expected delay in each project system, which contributes to improving the process of selecting suppliers for each project system.

7.2.1 Determining the inputs and outputs of the network

A part of the data warehouse was used by calling a query and using the query results as data for the proposed network. The project type, system type, and suppliers who carried out the supply operations were used as inputs for the network, and they were entered as a binary matrix {0-1} to remove the importance of order for the network. A dataset comprising 1,223 cases was utilized to train and test the network, aligning with the count of systems distributed among the fifteen studied projects after cleaning the data and removing outliers and this is a sufficient number to obtain highly reliable results. The data mining model was built in the SQL Server environment using Microsoft SQL Server software, and several neural networks with different numbers of hidden layer neurons were designed and compared with each other in order to select the optimal network that gives the best performance in predicting the expected delay in each system.

7.2.2 Assessing the accuracy and reliability of the proposed model

A test ratio of 30% of the total samples was adopted to verify the accuracy of the built model. To simulate the performance of the model and verify it, statistical indicators were used to test the suitability of the network for solving the problem and to determine the best and most accurate models. The Root Mean Square Error (RMSE) was used, which is given by equation (1).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{i=N}(Oi-Pi)^2}{N}} \qquad (1)$$

Where:

RMSE: Root Mean Square Error

Oi: Measured value

Pi: Predicted value

N: Number of samples

In addition, the Mean Absolute Error (MAE) was used as a standard for comparison between models and is given by equation (2):

$$MAE = \frac{\sum_{1}^{N}|Oi-Pi|}{N} \qquad (2)$$

The researcher has developed a suitable Data Warehouse for the reality of construction companies specialized in industrial projects to be a central data source and a knowledge base for decision-makers in the company. In addition to the analytical additions which have been provided at the company level, such as ranking, organizing, and sorting, it also enables effective analysis and direct information search to support decision-making in the process of selecting suppliers for industrial projects using Online Analytical Processing (OLAP) tools. These tools allowed for flexible and dynamic interactive analysis at different levels of management, which made it possible to perform the required queries to obtain the necessary information and knowledge to support decision-making in a fast and direct way, while providing results according to the decision-makers desired schemes, giving a clear and easy-to-understand picture in a short time, while obtaining such

a-ahmad@tishreen.edu.sy
, fayezalijrad@gmail.com
samahsm2016@hotmail.com

Journal of University of Duhok., Vol. 26, No.2 (Pure and Engineering Sciences), Pp 201-215, 2023
4th International Conference on Recent Innovations in Engineering (ICRIE 2023) (Special issue)

results may take several days relying on relational databases and referring to various sources. Data mining techniques were also used to predict the future delay rate in each project system according to the selected suppliers, which gives an idea of the expected delay rate resulting from selecting a group of suppliers for the project. This process is a tool that can help decision-makers improve the process of selecting suppliers for future projects for each project system, compatible with the project's baseline plan and does not exceed the planning period for project activities. Table (2) shows the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values for the Artificial Neural Networks models that were built for the performance cube of resources based on the number of hidden layer neurons, and the network with a hidden layer of 10 neurons gave the best performance in predicting the expected delay rate in each project system with the lowest Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values. And as a result, Fig. 9. Shows a data mining scatter plot for the delay network of the case study, which in turn reflects the accuracy of the predicted values of the delay in each system of the project based on the project type, system type, and suppliers who carried out the supply operations in each system, compared with the real delays in these systems.

Table (2):- The performance evaluation criteria for the constructed network models

Network No.	Number of neurons in the hidden layer	RMSE	MAE
1	4	0.2665	0.3479
2	6	0.2473	0.3154
3	8	0.231	0.3048
4	10	0.2243	0.2884
5	11	0.2303	0.2995
6	12	0.2279	0.2922
7	15	0.2259	0.2961
8	20	0.232	0.2989

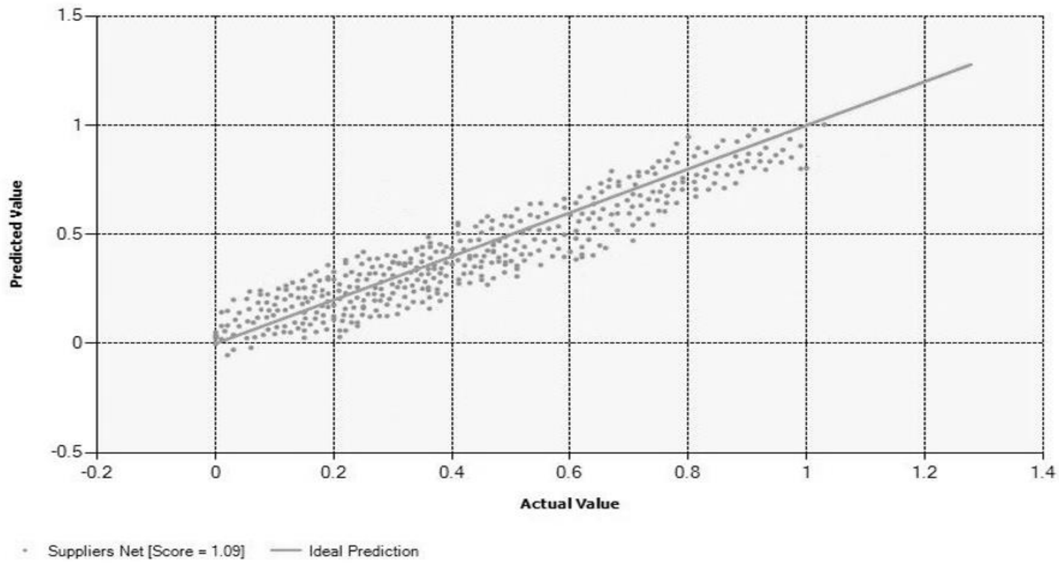

Fig.(9):- Data mining scatter plot for mining structure - Delay network

8. CONCLUSIONS AND RECOMMENDATIONS

The developed Data Warehouse has proven to be highly efficient and plays an important role in the decision-making process for selecting suppliers in industrial projects by predicting delay values in each project system according to the suppliers for each system with a good evaluation. This indicates the necessity of creating a project data warehouse in construction companies by identifying the set of work systems from which data will be extracted and then determining the dimensions of the model, the contents of the fact tables, and the linking foundations, in order to support the decisions related to supplier selection problem. The proposed methodology can be used and relied upon in subsequent specialized studies in the field of building data warehouses for construction companies, with a recommendation to provide additional capabilities to achieve a more comprehensive and detailed identification of all factors affecting the construction process and include them in the data warehouse. This allows for decision support related to various planning and tracking processes for construction projects. It is essential to adopt the proposed methodology and to apply it at the local level for construction companies due to its importance and role in supporting decisions at all administrative levels of the company regarding ongoing projects and future projects. This can be achieved by creating a data warehouse for previous projects of the companies and using Business Intelligence tools and online analytical processing, as well as adopting data mining techniques for future prediction and identifying hidden patterns to support various strategic decisions in the company. This leads to successful project management in the future and reduces the failure rates of industrial projects.

Acknowledgments

The authors would like to acknowledge the support provided by MAPNA GROUP company in facilitating access to the resources necessary for this research.

REFERENCES

Beale, M. H., Hagan, M. T., & Demuth, H. B. (2010). Neural network toolbox 7 user's guide (Version 7). MathWorks Inc.

Botha, L. J. (2018). Data mining construction project information to aid project management (Master's thesis). Stellenbosch University. https://scholar.sun.ac.za/handle/10019.1/103922

Chau, K. C. Y., Anson, M., & Zhang, J. (2003). Application of data warehouse and decision support system in construction management. Automation in Construction, 12(2), 213-224. https://doi.org/10.1016/S0926-5805(02)00057-9

Dalalah, D., Hayajneh, M., & Batieha, F. (2011). A fuzzy multi-criteria decision making model for supplier selection. Expert Systems with Applications, 38(7), 8384-8391. https://doi.org/10.1016/j.eswa.2010.12.102

Deng, X., Hu, Y., Deng, Y., & Mahadevan, S. (2014). Supplier selection using AHP methodology extended by D numbers. Expert Systems with Applications, 41(1), 156-167. https://doi.org/10.1016/j.eswa.2013.07.018

Desai, V. S. (2010). Application of decision tree technique to analyze construction project data. In A. Abraham, A.-E. Hassanien, & V. Snášel (Eds.), Computational social network analysis: Trends, tools and research advances (pp. 304-313). Springer. https://doi.org/10.1007/978-1-84882-229-0_14

Ghodsypour, S. H., & O'Brien, C. (1998). A decision system for supplier selection using an integrated analytic hierarchy process and linear programming. International Journal of

Production Economics, 56-57(1), 199-212. https://doi.org/10.1016/S0925-5273(97)00009-9

Güdüz, Ç., & Güdüz, G. Ş. (2019). Supplier selection under fuzzy environment. Tekstil ve Konfeksiyon, 29(4), 344-352.

Inmon, W. (1996). Building the data warehouse (2nd ed.). Wiley.

Hassan, B., Naja, H., Balouch, A., & Makiah, S. (2012). A decision support system for construction equipment management based on data warehousing technique. Tishreen University Journal-Engineering Sciences Series, 34(2), 203-223.

Kannan, D., Jabbour, A. B. L. D. S., & Jabbour, C. J. C. (2014). Selecting green suppliers based on GSCM practices: Using fuzzy TOPSIS applied to a Brazilian electronics company. European Journal of Operational Research, 233(2), 432-447. https://doi.org/10.1016/j.ejor.2013.07.023

Kannan, D., Govindan, K., & Rajendran, S. (2015). Fuzzy Axiomatic Design approach based green supplier selection: a case study from Singapore. Journal of Cleaner Production, 96, 194-208. https://doi.org/10.1016/j.jclepro.2014.01.070

Kohzadi, N., Boyd, M. S., Kaastra, I., Kermanshahi, B., & Scuse, D. (1995). Neural Networks for Forecasting: An Introduction. Canadian Journal of Agricultural Economics/Revue Canadienne D'Agronomie, 43(3), 463-474.

Konys, A. (2019). Green Supplier Selection Criteria: From a Literature Review to a Comprehensive Knowledge Base. Sustainability, 11(15), 4121. https://doi.org/10.3390/su11154121

Kumar, D. (2017). An analytical model for supplier selection. International Journal of Indian Culture and Business Management, 14(2), 131-140.

Kılıç, H. S. (2013). An integrated approach for supplier selection in multi-item/multi-supplier environment. Applied Mathematical Modelling, 37, (14-15), 7752-7763. https://doi.org/10.1016/j.apm.2013.02.032

Moon, S. K. J., & Kwon, K. (2007). Effectiveness of OLAP-based cost management in construction cost estimate. Automation in Construction, 16, 336-344. https://doi.org/10.1016/j.autcon.2006.06.005

Rong, P., Liu, G., An, P., Qi, X., Wang, Q., Yue, C., Zhang, W., Zhang, H., & Dai, Y. (2020). Construction and application of data mining model of unit scale cost of electric grid project. IOP Conference Series: Earth and Environmental Science, 446, 012021. https://doi.org/10.1088/1755-1315/446/1/012021

Rujiayanyong, T., & Shi, J. J. (2006). A project-oriented data warehouse for construction. Automation in Construction, 15, 800-807. https://doi.org/10.1016/j.autcon.2005.11.002

Wang, C. N., Dang, D. C., Vu, Q. Q., & Zeng, Y. X. (2018). Supplier selection for manufacturing industries. In International Conference on Advanced Manufacturing (ICAM) (pp. 1-6). IEEE.

Yahya, S., & Kingsman, B. (1999). Vendor rating for an entrepreneur development programme: a case study using the analytic hierarchy process method. Journal of the Operational Research Society, 50(9), 916-930.

Zakaria, N., Shaharnee, I. N. M., Jamil, J. M., & Nawi, M. N. M. (2015). Assessing stakeholder's credit risk using data mining in construction project. Advances in Environmental Biology, 9(5), 67-70.

Zhang, S., Zhang, C., & Yang, Q. (2003). Data preparation for data mining. Applied Artificial Intelligence, 17(5-6), 375-381.