Journal of Optimization in Industrial Engineering
Vol.14, Issue 2, Summer & Autumn 2021, 115-125
DOI: 10.22094/JOIE.2020.575685.1589

Research Article

Identifying and Evaluating Effective Factors in Green Supplier Selection using Association Rules Analysis

Nima Esfandiari [a], Mohammad Amin Adibi [*,a]
Faculty of Industrial and Mechanical Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran
Received 02 October 2018; Revised 10 September 2020; Accepted 20 September 2020

Abstract

Nowadays companies measure suppliers on the basis of a variety of factors and criteria that affect the supplier's selection issue. This paper intended to identify the key effective criteria for selection of green suppliers through an efficient algorithm called iterative process mining or i-PM. Green data were collected first by reviewing the previous studies to identify various environmental criteria. Then, the suppliers were evaluated and ranked on the basis of those criteria. The score table derived for the green criteria was one of the inputs to the algorithm. Moreover, membership functions and minimum support values were specified for each criterion as another input to the algorithm. The supplier ranking index was also obtained based on the score assigned to supplier's performance. Then, the hidden relationships between data were discovered and association rules were achieved and analyzed to identify the most important green criterion for selecting green suppliers.

Keywords: Supplier Selection; Association Rules Analysis; Iterative Process Mining Algorithm; Fuzzy Logic.

1. Introduction

Traditionally, suppliers in supply chain management (SCM) are selected based on their capability to respond to qualitative requirements, delivery schedule, prices, and services. In modern chain management, however, there are many other factors involved in development of long-term relationships with the suppliers. Disregarding the capability of suppliers to comply with environmental regulations, a company may be more risk-taking or disrupting the supply chain (Hsu and Hu, 2009, Sahebjamnia et al., 2020). In the past few years, green movements have compelled numerous institutions, governments and companies to improve their environmental performance. Furthermore, many companies have found growing enthusiasm about the natural environment, building integrated relationships with suppliers to design new green products (Noci, 1997). Just as global awareness is raised about environmental protection, green production has become an important trend, guaranteeing the manufacturer’s sustainability over the long-term. Therefore, the performance evaluation system for green suppliers is crucial in determining the desirable suppliers for cooperation with the company (Lee et al., 2009). Selection of an appropriate supplier can be a key strategic road toward elimination of environmental impacts in SCM for manufacturing companies (Tseng and Chiu, 2013). A supplier should be selected based on two essential parameters; First, the criteria for selection of suppliers, and, second, the methods for comparison of suppliers. The former underpins the subject of this research explained in methods for comparison and selection of suppliers, a few of which are mentioned below. Chen and Wang (2009) adopted a fuzzy VIKOR method to select suppliers. In fact, the fuzzy VIKOR serves to provide a more effective delivery method for evaluation and analysis of suppliers. Using a fuzzy VIKOR, they proposed a logical, systematic process to develop the best fitting options and solutions under any selection criteria. The findings suggested an important reference to solve fuzzy multi-criteria decision-making problems. Lin et al. (2009) used Association Rules for selection of suppliers. The new technique identified the key suppliers by combining the generalizable association rules algorithm for data mining with the set theory. Gou et al. (2009) used the decision tree for selection of suppliers. They introduced a new Support Vector Machine (Adibi and Shahrabi, 2015) technology known as a Potential Support Vector Machine (P-SVM). Then, they combined the new technique with the decision tree to solve the problem of supplier selection. This method has a better overall performance and less computation than a standard SVM. Chang and Hung (2010) adopted the Rough Set Theory to solve the supplier selection problem. In fact, they provided a supplier selection model to better encourage organizational capability and competitiveness, while solving practical problems. The new model could also be customized to meet real needs and changes in a competitive environment. Lin (2012) used the Analytic Network Process (ANP) to select suppliers. He proposed a Fuzzy Analytic Network Process (FANP) technique, where top suppliers are identified by taking into account the effects of interdependence between the selection criteria and adoption of inconsistent, unknown

*Corresponding author Email address: maadibi@aut.ac.ir

judgments. As noted earlier, there are two important parameters to consider in selection of suppliers, one of which is the criteria used in supplier selection. The present study mainly intended to identify and evaluate the most effective criteria for green supplier selection. For this purpose, we first reviewed the previous studies on green supplier selection criteria. Several papers have discussed the green criteria, a few of which are explored in the following. Tuzkaya et al. (2009) proposed green process management, environmental costs, pollution control, main purpose, green image in society, legislative and environmental management, and green (biocompatible) product for supplier selection. Lee et al. (2009) proposed quality, technological capabilities, overall cost of product life cycle, green image pollution control, environmental management, green product, and green competencies for the environmental evaluation of suppliers. Yeh and Chuang (2011) proposed green image, product recycling, green design, green supply chain management, pollution purification cost, and environmental performance assessment for evaluation of suppliers. Prosman and Sacchi (2018) proposed usability, transport, affected handling activity as environmental criteria for supplier selection in circular supply chains. Having reviewed the criteria in various papers, we proposed a new data mining algorithm dubbed iterative Process Mining (i-PM) to identify the most effective criteria for supplier evaluation leading to an efficient selection. In practice, the i-PM combines the association rules in data mining with fuzzy sets to extract association rules from the data. The i-PM basically extracts association rules by combining fuzzy rules with association rules. This was first experimented by Hong et al. (2003). According to Lau et al. (2009) which proposed the algorithm, i-PM serves to identify the improvement measures for optimization of supply chain network under fuzzy rules. As mentioned earlier, this algorithm was first proposed by Leo et al. as an adaption of fuzzy generalized search algorithm presented by Hong et al. (2003). It actually combines the data mining technologies with fuzzy sets to identify valuable association rules in quantitative information. The i-PM is less time-consuming than other fuzzy methods, because it finds the most important fuzzy item for each parameter, thus mitigating time complexity (Hong et al. 2003). In this algorithm, a lower bound support threshold is set as a key parameter directly determining the quality of association rules. Relying on fuzzy concepts, the i-PM employs verbal phrases in natural and colloquial language, allowing experts to conduct more relevant, accurate analyses on the subject of research.

The current paper employed i-PM to detect the hidden relationships between green criteria data and supplier ranking index. It identifies the key criteria improving the index. The evaluation and scoring procedures in this research involved Iran Khodro's Supplying Auto Parts Co. (SAPCO). Due to various factors and criteria, the supplier selection leads to excessive computational and time complexity. So, the goal is to find a mechanism to reduce time and computational complexity in the supplier selection process and one simple strategy is to pre-qualify the most important criteria effective in selection of an efficient supplier. Hence, this paper revolved around identification of the most important criteria. The results were obtained by extraction of rules that can be analyzed to identify the most important criteria. at the end, the effective and important criteria to select green suppliers were identified by implementing the algorithm to discovery latent relationships. The criteria identified in this study are among important, useful and effective criteria. In the conclusion section, more details are given about the three criteria identified as the most important criteria. Another advantage of this research is that, in other methods, which provide the most important environmental criteria for the supplier selection, membership functions were found using the opinion of experts and elites which may not always be available, While, in this research membership functions were obtained using an effective method which its mechanism is such that it can uses data to find membership functions. In the discussion section, the advantages of the algorithm are given in more detail. The rest of the paper is organized as follows: In Section 2 the relevant literature is reviewed and the exploration of the criteria, the collected information and also obtaining the solution and rules are explained in Section 3. Discussion is presented in Section 4 and the Conclusions and suggestions are given in Section 5.

2. Literature Review

Liu et al. (2009) presented a new technique to solve the problem of extracting association rules. This technique allows the user to specify the value of multiple supports, thereby to demonstrate the nature of items and their various iterations in the database during extraction of rules. Delgado et al. (2001) presented a new method to extract association rules from the quantitative values in dependent databases. They improved the meaning of such rules by introducing inaccurate conditions before and after. They experimented the new method on accurate data. So, the fuzzy rules extracted by them. were more meaningful than rules with accurate values. Hong et al. (2003) developed a technique for extraction of association rules based on a fuzzy set for quantitative values. The new algorithm could focus on the most important linguistic terms, while reducing the time complexity. The new algorithm combined the fuzzy transaction data mining algorithm with the generalized association rule extraction algorithm. Tsay and Chiang (2005) extracted association rules by proposing an effective method known as Cluster-Based Association Rule (CBAR). The CBAR could find cluster tables by reviewing the database once, and then cluster the transaction data into k number of cluster tables. Aouissi et al. (2007) adopted the genetic algorithm to extract quantitative association rules. They provided a system called QUANTMINER for

extraction of association rules. The QUANTMINER was developed based on a genetic algorithm dynamically detecting the good intervals in association rules by improving the support value and the confidence value. In 2009, Lau et al. proposed a process extraction system to support knowledge detection in the supply chain network. The extraction system provided by Lau et al. is able to detect a set of fuzzy association rules based on operational data of daily logistics obtained within the network. Zawaidah et al. (2011) presented an improved algorithm to extract association rules from large-scale databases. The new method could effectively identify association rules from large-scale databases. Foguem and Mauget (2013) extracted association rules to improve the quality of manufacturing process. In the new method, they extended the application of association rule extraction procedures to extract knowledge from operations and information management. The significance of that research lies in improvement of operational processes. Jain et al. (2014) extracted the association rules for evaluation of criteria during identification of supplier capabilities and the selection times. They adopted an iterative process mining algorithm to discover hidden relationships between the data of supplier capability identification and supplier ranking index. Chen et al. (2016) obtained the fuzzy temporal association rules through item lifespan. In their research, Chen et al. proposed a data mining algorithm for extraction of fuzzy temporal association rules. Yuan (2017) proposed an improved Apriori algorithm for extraction of association rules. After a series of studies, Yuan realized that traditional Apriori algorithms have two major bottlenecks: first, frequent browsing of database and second, generation of numerous candidate sets. Based on the inherent defects of the Apriori algorithm, Yuan made certain improvements. Firstly, he used a new database roadmap to avoid frequent browsing of database. Secondly, he pruned the iterative item set and candidate item set to improve the efficiency of connection. Thirdly, he used an overlapping strategy to calculate the support value to achieve high efficiency. Under the same conditions, the results suggested that the new algorithm improved by Yuan outperforms other improved algorithms in terms of operational efficiency.
Wang et al. (2018) Presented a new framework for mining temporal association rules by discovering item sets with frequent item sets tree. The experimental results of their research showed that the algorithm proposed by them can provide better efficiency and interpretability in mining temporal association rules in comparison with other algorithms. Haeri and Rezaei (2019) proposed a grey-based green supplier selection model for uncertain environments. For criteria selection, they only used cases which were listed in the literature. As another example of uncertain environments, well defined by Alinezhad et al. (2019), Gupta et al. (2019) used environment management system, pollution control, quality, and green image to evaluate suppliers in a fuzzy multi-criteria decision making method. Kumar and Mishra (2020) used multi-criteria COPRAS (complex proportional assessment) method based on parametric measures for intuitionistic fuzzy sets in green supplier selection problem incorporating criteria evaluation. Artificial Neural Network (ANN) was proposed by Manohar and Kumar (2020) to important criteria in green supplier selection process. Then, they used the selected criteria to build a model applicable in green supplier selection.

3. Material and Methods

As discussed in the Introduction, the criteria proposed in the literature for environmental evaluation of suppliers were reviewed and the most common and significant of them were selected for this study.

Table 1
Criteria and sub-criteria for identifying the capabilities of green suppliers with related symbols and ranges

row	Criteria	sub-criteria	range	symbol
1	pollution control	hazardous air pollution , water and wastewater , solid waste energy consumption , waste	60-0	A
2	green image	social responsibility , employee training program for green awareness	25-0	B
3	environmental management and legislation costs	Environment-related certificates, Green process management	15-0	C
4	environmental costs	The cost of green products and other Environmental costs	5-0	D
5	green products	Materials used in the supplied components for green production, green packaging, waste management	35-0	E

Table 1 shows criteria and sub-criteria for identifying the capabilities of green suppliers along with related symbols and ranges. Next, suppliers are rated based on the criteria. For this purpose, questions will be designed as measures for each criterion and sub-criterion and the suppliers are rated using these questions. The supplier ranking index is calculated by rating the performance of the suppliers. Rating is carried out using performance measurement criteria. Here, quality, cost, delivery time, and service level are used. The final ranking index for suppliers is obtained from the sum of the scores calculated by these criteria.

$$Q_R = \sum_{i=1}^{4} w_i \, . \, Q_i \tag{1}$$

$$SRQ = \frac{Q_R}{Q} .100 \quad (2)$$

$$SRD = \frac{Q_1}{Q} \times \frac{T}{T \times P + T_1 \times q} \times 100 \quad (3)$$

$$SRP = \frac{P_U - P}{P_U - P_L} \times 100 \quad (4)$$

$$SRI = Wa.SRQ + Wb.SRD + Wc.SRP + Wd.SRS \quad (5)$$

Quality scores can be calculated using Eq. 1, where Q_1 represents the accepted value according to the specifications, Q_2 shows the accepted value with deviation, Q_3 shows the accepted value after correction, Q_4 shows the rejected value, W_i shows the Q_i weights, and Q is the total value of inspection (total Q_i). The final quality score is calculated by Eq. 2.

Equation 3 is used to calculate the delivery time score where Q represents the order value, Q_1 shows the order delivered on time, T represents the agreed delivery time, T_1 is the real time spent on preparation, P represents Q_1/Q, and q is $1-p$. Moreover, Eq. 4 is used for calculating the price score. Here, P_L is the lowest price given for a specific product, P_U is the highest price, and P is the price quoted by the supplier. Furthermore, the service level was rated using the IS 12040:2001 standard. Each supplier's service is rated using parameters such as willingness to accept returns for replacement, the speed of response, and so on. The SAPCO Sales Department and some other departments were asked to rate suppliers according to these parameters. Finally, the suppliers were ranked using the five criteria.

Table 2 shows the scores of suppliers based on environmental criteria and supplier ranking index. The minimum support value is then determined for each criterion.

Table 2
Scores of suppliers based on environmental criteria and supplier ranking index

supplier number	A	B	C	D	E	F	G
1	20	7.5	10	0	20	0	84.13
2	15	0	0	0	10	5	85.52
3	17.5	12.5	10	0	25	10	130.79
4	32.5	12.5	15	2.5	25	20	121.78
5	30	10	10	0	25	20	117.08
6	30	22.5	10	0	25	5	129.73
7	32.5	0	5	0	20	10	113.33
8	27.5	7.5	10	2.5	35	15	117.03
9	20	7.5	10	2.5	20	25	114.16
10	35	17.5	15	0	20	15	134.29
11	37.5	17.5	5	2.5	20	5	137.73
12	22.5	0	0	0	15	0	114.35
13	20	5	10	2.5	30	5	107.25
14	40	12.5	10	2.5	25	10	161.85
15	20	5	5	0	20	15	109.81
16	32.5	22.5	15	2.5	35	15	126.68
17	25	12.5	15	0	15	5	117.3
18	37.5	10	10	0	30	10	111.51
19	50	17.5	15	2.5	25	15	110.53
20	25	7.5	10	0	20	10	115.04
21	25	5	15	0	5	10	117.02
22	15	12.5	15	2.5	10	5	124.53
23	42.5	12.5	5	0	30	15	122.33
24	12.5	7.5	0	0	10	5	119.32
25	15	17.5	10	0	30	15	121.33
26	55	22.5	15	2.5	35	20	72.57
27	25	10	15	0	30	15	128.19
28	22.5	17.5	15	2.5	20	15	114.1
29	10	17.5	0	2.5	5	5	72.5
30	12.5	17.5	15	0	25	10	62.46

Table 3 shows the minimum support value for each criterion. Membership functions were calculated for each criterion using the method proposed by Hong and Chen, (1999). Then, the membership functions were plotted using MATLAB. The membership functions for the criteria A to G are shown in Figure 1.

Table 3
Minimum support value for each criterion

parameter	A	B	C	D	E	F	G
Minimum support value	0.8	1	1.2	1.3	1	1.1	0.6

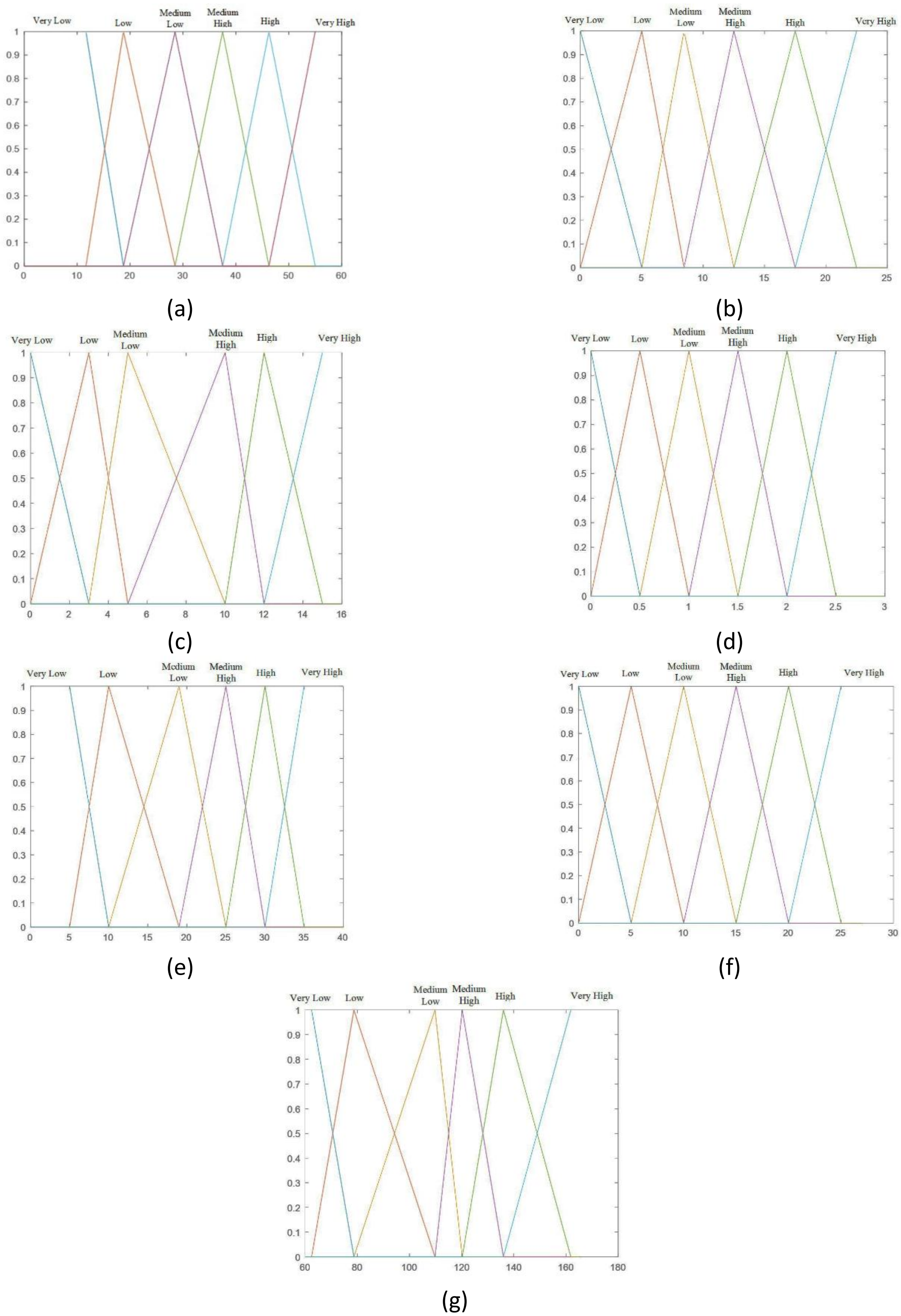

Fig. 1. (a) membership functions for the criteria A; (b) membership functions for the criteria B; (c) membership functions for the criteria C; (d) membership functions for the criteria D; (e) membership functions for the criteria E; (f) membership functions for the criteria F; (g) membership functions for the criteria G

The i-PM algorithm or iterative data mining algorithm discussed in the Introduction is used to solve the problem. The symbols employed in this algorithm are introduced in Table 4 and the algorithm steps are as follows3

Table 4
Symbols used in the i-PM algorithm

Q_{ij}	Data item i for supplier j $i \in (1, n)$ $j \in (1,30)$
F_{ij}^k	Fuzzy value for Q_{ij} with linguistic value k $k \in (1,6)$
R_{ij}^k	Linguistic notation of Q_{ij}
R_i^k	Linguistic notation after combining all data
$Count_i^k$	Fuzzy count for R_i^k
$\propto_i$	Minimum support value for data item i
$Max - Count_i^k$	Maximum count for data item i with the fuzzy linguistic value $\acute{k}$
CV_p	Confidence value of rule p

Step One**:** The quantitative values of criteria and supplier ranking index are converted to a fuzzy set using the membership functions calculated in the previous section. For example, Table 5 shows how to convert values to a fuzzy set for suppliers 1 through 10. In fact, Q_{ij} is converted to F_{ij} in this step where i in F_{ij}^k is the counter for the criteria ranging from 1 to 7 and j represents supplier counter ranging from 1 to 30 and k is the numeration of the descriptive values of Very Low to Very High ranging from 1 to 6.

$$Count_i^k = \sum_{j=1}^{30} F_{ij}^k \ \forall i \in (1, n), \forall k \in (1, m) \tag{6}$$

Step 2: Fuzzy count is calculated for the fuzzified values and the results are placed in the List C_1 shown Table 6. The fuzzy count is calculated using Eq. 6 where n represents the number of criteria and m is the number of the descriptive values of Very Low to Very High.

Step 3: The maximum fuzzy count for each criterion is calculated from the Eq. 7.

$$Max - Count_i^{\acute{k}} = \max(Count_i^1, \dots, Count_i^m) \tag{7}$$

For example, the values 4.625, 7.987, 9.0982, 5.5317, 1.4285 and 1.4286 are available for the criterion A in the List C1. Max-Count^k_i for the criterion A is equal to the maximum value of 9.0982. In the fuzzy count maximum formula, $\acute{k}$ is the numeration of the descriptive value of k for the criterion selected as the maximum fuzzy count. For the criterion A, $\acute{k}$ equals 3. For other criteria, the maximum fuzzy count is calculated in the same way and the results are presented in Table 7.

Table 5
Value converted to the fuzzy set (the criterion A for suppliers 1 through 10)

Criteria	Linguistic notation	1	2	3	4	5	6	7	8	9	10
A	Very Low	0	0.5289	0.1763	0	0	0	0	0	0	0
A	Low	0.1282	0.4711	0.8237	0	0	0	0	0.1026	0.8718	0
A	Medium Low	0.8718	0	0	0.5556	0.8333	0.8333	0.5556	0.8974	0.1282	0.2778
A	Medium High	0	0	0	0.4444	0.1667	0.1667	0.4444	0	0	0.7222
A	High	0	0	0	0	0	0	0	0	0	0

Table 6
The List C_1 for fuzzy counts

Linguistic notation	Count	Linguistic notation	Count	Linguistic notation	Count
A.Very Low	4.526	D.Very Low	18	G.Very High	1.7577
A.Low	7.987	D.Low	0	G.Low	2.9301
A.Medium Low	9.0982	D.Medium Low	0	G.Medium Low	8.2952
A.Medium High	5.5317	D.Medium High	0	G.Medium High	11.3956
A.High	1.4285	D.High	0	G.High	4.5564
A.Very High	1.4286	D.Very High	12	G.Very High	1.0646
B.Very Low	3	E.Very Low	2		
B.Low	4.3665	E.Low	3.8888		
B.Medium Low	5.4809	E.Medium Low	7.7776		
B.Medium High	7.1526	E.Medium High	8.3336		
B.High	7	E.High	4		
B.Very High	3	E.Very High	4		
C.Very Low	4	F.Very Low	2		
C.Low	0	F.Low	8		
C.Medium Low	4	F.Medium Low	7		
C.Medium High	12	F.Medium High	9		
C.High	0	F.High	3		
C.Very High	10	F.Very High	1		

Step 4: Max-Count^k_i obtained in Step 3 is compared with the minimum support value α_i for the data item i , If Max-Count^k_i is greater than α_i, it is placed in the List L_1 shown in Table 8. For example, the count value for the criterion B in Table 7 is 7.1526. The minimum support value equals 1 according to Table 3. Since the count value is greater than the minimum support value, B. Medium High with a count value of 7.1526 is placed in the List L_1. All items are greater than their minimum support values, thus they are placed in the List L1.

Step 5: The data items in the List L_1 are combined to obtain the binary item dataset. For example, B. Medium High and A. Medium Low can be combined to form the {A. Medium Low, B. Medium High} set. Then, the minimum combined support value or the maximum value for α_1 and α_2 is calculated. The binary count for each binary set is found using Eq. 8. Binary counts and binary support values are shown in the List C_2 in Table 9.

$$Combined\_Count_{i1,i2}^{k1,k2} = \sum_{j=1}^{30} \min(F_{i1j}^{k1}, F_{i2j}^{k2}) \quad (8)$$

Table 7
Maximum fuzzy count for each criterion

Linguistic notation	Count
A.Medium Low	9.0982
B.Medium High	7.1526
C.Medium High	12
D.Very Low	18
E.Medium High	8.3336
F.Medium High	9
G.Medium High	11.3956

Table 8
L_1 List

Linguistic notation	Count
A.Medium Low	9.0982
B.Medium High	7.1526
C.Medium High	12
D.Very Low	18
E.Medium High	8.3336
F.Medium High	9
G.Medium High	11.3956

Step 6: The binary counts that pass the binary support values are placed in the List C_2 in Table 10.

Table 9
List C_2 of binary item dataset with fuzzy counts and binary support values

binary item dataset	Count	Sup
{ A.Medium Low, B.Medium High }	1.965	1
{ A.Medium Low, C.Medium High}	4.333	1.2
{ A.Medium Low, D.Very Low }	6.448	1.3
{ A.Medium Low, E.Medium High }	3.312	1
{ A.Medium Low, F.Medium High }	2.885	1.1
{ A.Medium Low, G.Medium Low }	6.524	0.8
{ B.Medium High, C.Medium High }	2.768	1.2
{ B.Medium High, D.Very Low }	4.153	1.3
{ B.Medium High, E.Medium High }	3.384	1
{ B.Medium High, F.Medium High }	1.384	1.1
{ B.Medium High, G.Medium Low }	4.483	1
{ C.Medium High, D.Very Low }	8	1.3
{ C.Medium High, E.Medium High }	5.5	1.2
{ C.Medium High, F.Medium High }	2	1.2
{ C.Medium High, G.Medium Low }	4.13	1.2
{ D.Very Low, E.Medium High }	4.834	1.3
{ D.Very Low, F.Medium High }	5	1.3
{ D.Very Low, G.Medium Low }	7.584	1.3
{ E.Medium High, F.Medium High }	1.5	1.1
{ E.Medium High, G.Medium Low }	3.18	1
{ F.Medium High, G.Medium Low }	4.178	1.1

Table 10
The List L_2

binary item dataset	Count
{ A.Medium Low, B.Medium High }	1.965
{ A.Medium Low, C.Medium High}	4.333
{ A.Medium Low, D.Very Low }	6.448
{ A.Medium Low, E.Medium High }	3.312
{ A.Medium Low, F.Medium High }	2.885
{ A.Medium Low, G.Medium Low }	6.524
{ B.Medium High, C.Medium High }	2.768
{ B.Medium High, D.Very Low }	4.153
{ B.Medium High, E.Medium High }	3.384
{ B.Medium High, F.Medium High }	1.384
{ B.Medium High, , G.Medium Low }	4.483
{ C.Medium High, D.Very Low }	8
{ C.Medium High, E.Medium High }	5.5
{ C.Medium High, F.Medium High }	2
{ C.Medium High, G.Medium Low }	4.13
{ D.Very Low, E.Medium High }	4.834
{ D.Very Low, F.Medium High }	5
{ D.Very Low, G.Medium Low }	7.584
{ E.Medium High, F.Medium High }	1.5
{ E.Medium High, G.Medium Low }	3.18
{ F.Medium High, G.Medium Low }	4.178

Step 7: The data items in the List L_2 are categorized into groups of three and the support values are found for each group. The groups are put in the List C_3. The support values for ternary groups are also calculated as for binary groups. Ternary counts and ternary support values are shown in the List C_3 in Table 11.

Step 8: The process is iterated in Steps 5 and 6 until the list is empty and other data items cannot be categorized into larger groups

Step 9: The fuzzy count values of ternary groups are compared with the ternary minimum support values and items larger than the minimum support values are placed in the List L_3 in Table 12. Now the List C_3 is checked. Since quaternary groups cannot be made from this list, the algorithm stops. So the List L_3 is the final list and association rules are extracted using this list. The formula presented by Leo et al. (2009) is used to find the confidence value (CV) of each associate rule. For example, Eq. 9 is

Table 11
The List C_3 for ternary data items and fuzzy counts and ternary support values

ternary data items	Count	Sup
{A.Medium Low,B.Medium High,C.Medium High}	0.384	1.2
{A.Medium Low,B.Medium High,D.Very Low}	1.409	1.3
{A.Medium Low,B.Medium High,E.Medium High}	0.94	1
{A.Medium Low,B.Medium High,F.Medium High}	0.384	1.1
{A.Medium Low,B.Medium High,G.Medium High}	1.965	1
{A.Medium Low,C.Medium High,D.Very Low}	3.179	1.3
{A.Medium Low,C.Medium High,E.Medium High}	2.128	1.2
{A.Medium Low,C.Medium High,F.Medium High}	0.897	1.2
{A.Medium Low,C.Medium High,G.Medium High}	2.415	1.2
{A.Medium Low,D.Very Low,E.Medium High}	2.462	1.3
{A.Medium Low,D.Very Low,F.Medium High}	1.047	1.3
{A.Medium Low,D.Very Low,G.Medium High}	4.207	1.3
{A.Medium Low,E.Medium High,F.Medium High}	0.462	1.1
{A.Medium Low,E.Medium High,G.Medium High}	2.392	1
{A.Medium Low,F.Medium High,G.Medium High}	2.245	1.1
{B.Medium High,C.Medium High,D.Very Low}	1.768	1.3
{B.Medium High,C.Medium High,E.Medium High}	2.384	1.2
{B.Medium High,C.Medium High,F.Medium High}	0	1.2
{B.Medium High,C.Medium High,G.Medium High}	0.879	1.2
{B.Medium High,D.Very Low,E.Medium High}	1.384	1.3
{B.Medium High,D.Very Low,F.Medium High}	1.384	1.3
{B.Medium High,D.Very Low,G.Medium High}	2.848	1.3
{B.Medium High,E.Medium High,F.Medium High}	0	1.1
{B.Medium High,E.Medium High,G.Medium High}	1.623	1
{B.Medium High,F.Medium High,G.Medium High}	1.254	1.1
{C.Medium High,D.Very Low,E.Medium High}	4.333	1.3
{C.Medium High,D.Very Low,F.Medium High}	1	1.3
{C.Medium High,D.Very Low,G.Medium High}	3.023	1.3
{C.Medium High,E.Medium High,F.Medium High}	0	1.2
{C.Medium High,E.Medium High,G.Medium High}	1.762	1.2
{C.Medium High,F.Medium High,G.Medium High}	1.626	1.2
{D.Very Low,E.Medium High,F.Medium High}	0.333	1.3
{D.Very Low,E.Medium High,G.Medium High}	1.875	1.3
{D.Very Low,F.Medium High,G.Medium High}	2.415	1.3
{E.Medium High,F.Medium High,G.Medium High}	0.346	1.1

used to find the CV of the rule If {A. Medium Low, C. Medium High} then {D. Very Low}.

$$CV_P = \frac{Count(\{\text{A. Medium Low, C. Medium High, D. Very Low}\})}{Count(\{\text{A. Medium Low, C. Medium High}\})} \quad (9)$$

The count value for all items of the rule is shown in the numerator. In this case, the count value is tripled, because of two items in the If section and an item in the Then section. The sum will be a ternary group. In this example, {A. Medium Low, C. Medium High, D. Very Low} equals 3.179. The count value for *Then* section is placed in the denominator of this formula. Here, it is a binary count and

.

{A. Medium Low, C. Medium High} equals 4.333. So the CV for the rule is 73%. Other rules may be extracted and calculated with their corresponding confidence values.

Step 10: A limitation is defined for CV to optimize the rules. Here, the rules with a CV greater than 60% are accepted and placed in Table 13.

Table 12
The List L_3

ternary data items	Count
{A.Medium Low,B.Medium High,D.Very Low}	1.409
{A.Medium Low,B.Medium High,G.Medium High}	1.965
{A.Medium Low,C.Medium High,D.Very Low}	3.179
{A.Medium Low,C.Medium High,E.Medium High}	2.128
{A.Medium Low,C.Medium High,G.Medium High}	2.415
{A.Medium Low,D.Very Low,E.Medium High}	2.462
{A.Medium Low,D.Very Low,G.Medium High}	4.207
{A.Medium Low,E.Medium High,G.Medium High}	2.392
{A.Medium Low,F.Medium High,G.Medium High}	2.245
{B.Medium High,C.Medium High,D.Very Low}	1.768
{B.Medium High,C.Medium High,E.Medium High}	2.384
{B.Medium High,D.Very Low,E.Medium High}	1.384
{B.Medium High,D.Very Low,F.Medium High}	1.384
{B.Medium High,D.Very Low,G.Medium High}	2.848
{B.Medium High,E.Medium High,G.Medium High}	1.623
{B.Medium High,F.Medium High,G.Medium High}	1.254
{C.Medium High,D.Very Low,E.Medium High}	4.333
{C.Medium High,D.Very Low,G.Medium High}	3.023
{C.Medium High,E.Medium High,G.Medium High}	1.762
{C.Medium High,F.Medium High,G.Medium High}	1.626
{D.Very Low,E.Medium High,G.Medium High}	1.875
{D.Very Low,F.Medium High,G.Medium High}	2.415

Step 11: The rules showing the relationship between the criteria and the supplier ranking index are listed in Table 14. As is clear from the rules in Table 12, the criteria A, B and F respectively show pollution control, green image and green process management criteria. These criteria show a more significant relationship with the supplier ranking index or G. Given the higher CV of these three criteria with G, the quality of the association rule is higher. The criteria C and E respectively show environmental management and legislative costs, and green product. These criteria show a weak correlation with G than the previous three criteria. The criterion D or environmental costs shows the weakest correlation with G than other criteria. Among binary rules, the criteria A and B show the most significant correlation with the index G.

Table 13
Association rules with a CV greater than 60%

Association rules	Conf
{A.Medium Low,B.Medium High} then {G.Medium High}	1.00
{B.Medium High,F.Medium High} then {D.Very Low}	1.00
{B.Medium High,F.Medium High} then {G.Medium High}	0.91
{D.Very Low,E.Medium High} then {C.Medium High}	0.90
{B.Medium High,C.Medium High} then {E.Medium High}	0.86
{C.Medium High,F.Medium High} then {G.Medium High}	0.81
{C.Medium High,E.Medium High} then {D.Very Low}	0.79
{A.Medium Low,F.Medium High} then {G.Medium High}	0.78
{E.Medium High,G.Medium High} then {A.Medium Low}	0.75
{A.Medium Low,E.Medium High} then {D.Very Low}	0.74
{A.Medium Low,C.Medium High} then {D.Very Low}	0.73
{C.Medium High,G.Medium High} then {D.Very Low}	0.73
{A.Medium Low,E.Medium High} then {G.Medium High}	0.72
{A.Medium Low} then {G.Medium High}	0.72
{A.Medium Low,B.Medium High} then {D.Very Low}	0.72
{A.Medium Low} then {D.Very Low}	0.71
{B.Medium High,E.Medium High} then {C.Medium High}	0.70
{B.Medium High,D.Very Low} then {G.Medium High}	0.69
{C.Medium High} then {D.Very Low}	0.67
{G.Medium High} then {D.Very Low}	0.67
{E.Medium High} then {C.Medium High}	0.66
{A.Medium Low,D.Very Low} then {G.Medium High}	0.65
{A.Medium Low,G.Medium High} then {D.Very Low}	0.64
{A.Medium Low,E.Medium High} then {C.Medium High}	0.64
{B.Medium High,C.Medium High} then {D.Very Low}	0.64
{B.Medium High,G.Medium High} then {D.Very Low}	0.64
{B.Medium High} then {G.Medium High}	0.63

Table 14
Association rules representing the relationship between the criteria and the supplier ranking index

Association rules	Conf
{A.Medium Low,B.Medium High} then {G.Medium High}	1.00
{B.Medium High,F.Medium High} then {G.Medium High}	0.91
{C.Medium High,F.Medium High} then {G.Medium High}	0.81
{A.Medium Low,F.Medium High} then {G.Medium High}	0.78
{E.Medium High,G.Medium High} then {A.Medium Low}	0.75
{C.Medium High,G.Medium High} then {D.Very Low}	0.73
{A.Medium Low,E.Medium High} then {G.Medium High}	0.72
{A.Medium Low} then {G.Medium High}	0.72
{B.Medium High,D.Very Low} then {G.Medium High}	0.69
{G.Medium High} then {D.Very Low}	0.67
{A.Medium Low,D.Very Low} then {G.Medium High}	0.65
{A.Medium Low,G.Medium High} then {D.Very Low}	0.64
{B.Medium High,G.Medium High} then {D.Very Low}	0.64
{B.Medium High} then {G.Medium High}	0.63

4. Discussion

The iterative process mining algorithm is more efficient than the previous algorithms used for extraction of association rules, because it combines association rules with fuzzy sets. Consequently, the results are closer to human judgments than are often ambiguous. As another advantage of this algorithm compared to previous algorithms on extraction of fuzzy association rules, the proposed algorithm finds the maximum fuzzy count for each criterion. This reduces the runtime and volume of calculations as compared to previous algorithms. The iterative process mining algorithm was used to identify the most important criteria for selecting green suppliers. Clearly, this algorithm reaches the solution in lesser steps and finds the most important criteria for selecting green suppliers. Another advantage of this algorithm is that it uses the minimum value of multiple supports, because different parameters in the real world have different minimum support values. Membership functions were obtained using an effective method which uses data to find membership functions.

However, membership functions were found using the opinion of experts and elites in other methods which may not always be available. The effective and important criteria to select green suppliers were identified by implementing the algorithm to discovery latent relationships. The criteria identified in this study are among important, useful and effective criteria.

5. Conclusions and Suggestions

In this research, a new data mining algorithm was used to identify the most important and effective criteria for evaluating green suppliers. The algorithm used in this study is intended for the extraction of association rules to detect latent relationships between data. This algorithm was used to discover the latent relationships between the environmental criteria of the supplier selection and supplier ranking index. A number of criteria were chosen from the literature for evaluating green suppliers. Criteria include pollution control, green image, environmental management and legislation costs, environmental costs, green products and green process management. Using these criteria, 30 suppliers from SAPCO Company were surveyed and rated. Each criterion consists of several sub-criteria. Questions were designed for each sub-criterion to rate suppliers according to these criteria. Based on these questions, the suppliers were rated and evaluated and the scores were shown in a table. For each supplier, supplier ranking index which is an indicator for evaluating the supplier's performance was obtained by four criteria including quality, delivery, price, and service level. Membership functions were calculated for each criterion. Using these membership functions, the supplier scores were converted to fuzzy values to run the algorithm. Using the information obtained, the iterative data mining algorithm was implemented to discover latent relationships between data. With the help of the association rules obtained by this algorithm, it can be determined which of the environmental criteria affects the supplier ranking index. As mentioned, three criteria of pollution control, green image and green process management showed the most significant correlation with the supplier ranking index. So these three criteria are the most important criteria for choosing green suppliers. These criteria can be effective in choosing the efficient supplier. The criterion pollution control has five sub-criteria including air pollution, water and wastewater, solid waste, energy consumption and hazardous waste. Pollution control is an important criterion so that controlling its sub-criteria in a manufacturing company will help environmental protection. The criterion green image has two sub-criteria of social responsibility and employee training program for green awareness. The sub-criterion social responsibility refers activities that any company does to motivate customers to consume green products or support environmental festivals and conferences. The sub-criteria training program for green awareness concerns holding educational seminars regarding the importance and observance of environmental issues for employees or promoting the culture of paper consumption reduction and the use of office automation system among employees. The criterion green process management has five sub-criteria. The sub-criterion R & D projects for green production concerns scientific and technological measures to achieve new methods of production or structure of green products. The sub-criterion "design for the environment" means designing products and processes with the aim of reducing raw materials and energy consumption or reusing, rebuilding, and recovery of products. The sub-criterion "green production technology" means designing and producing products by wind energy, especially solar energy for power generation. In fact, it refers technologies which are less harmful to the environment than traditional ones. The sub-criterion "green distribution" concerns the effective use of vehicles and the selection of distribution and transportation networks with an emphasis on environmental criteria. So these three criteria are the most important and most effective criteria affecting the selection of green suppliers. These criteria were identified by the iterative data mining algorithm through extracting association rules. Triangular membership functions were used to convert the values into fuzzy numbers. It is therefore suggested to convert numbers to fuzzy ones using other membership functions such as trapezoidal, hyperbolic, and so on. Given the fact that this algorithm can show the relationship between the data and also show which criteria affect a particular criterion, it can be used in other areas to find latent relationships between data and criteria. It is also possible to use meta-heuristic algorithms for fuzzification.

References

Adibi, M. A., & Shahrabi, J. (2015). Online anomaly detection based on support vector clustering. International Journal of Computational Intelligence Systems, 8(4), 735-746.

Alinezhad, A., Adibi, M. A., & Tohidi, A. (2019). Classification of Streaming Fuzzy DEA Using Self-Organizing Map. Journal of Optimization in Industrial Engineering, 12(1), 53-61.

Chang, B., & Hung, H. F. (2010). A study of using RST to create the supplier selection model and decision-making rules. *Expert Systems with Applications*, *37*(12), 8284-8295.

Chen, C. H., Lan, G. C., Hong, T. P., & Lin, S. B. (2016). Mining fuzzy temporal association rules by item lifespans. *Applied Soft Computing*, *41*, 265-274.

Chen, L. Y., & Wang, T. C. (2009). Optimizing partners' choice in IS/IT outsourcing projects: The strategic decision of fuzzy VIKOR. *International Journal of Production Economics*, *120*(1), 233-242.

Delgado, M., SáNchez, D., MartıN-Bautista, M. J., & Vila, M. A. (2001). Mining association rules with improved semantics in medical databases. *Artificial Intelligence in Medicine*, *21*(1-3), 241-245.

Guo, X., Yuan, Z., & Tian, B. (2009). Supplier selection based on hierarchical potential support vector machine. *Expert Systems with Applications*, *36*(3), 6978-6985.

Gupta, S., Soni, U., & Kumar, G. (2019). Green supplier selection using multi-criterion decision making under fuzzy environment: A case study in automotive industry. Computers & Industrial Engineering, 136, 663-680.

Haeri, S. A. S., & Rezaei, J. (2019). A grey-based green supplier selection model for uncertain environments. Journal of cleaner production, 221, 768-784.

Hong, T. P., & Chen, J. B. (1999). Finding relevant attributes and membership functions. *Fuzzy Sets and systems*, *103*(3), 389-404.

Hong, T. P., Lin, K. Y., & Wang, S. L. (2003). Fuzzy data mining for interesting generalized association rules. *Fuzzy sets and systems*, *138*(2), 255-269.

Jain, R., Singh, A. R., Yadav, H. C., & Mishra, P. K. (2014). Using data mining synergies for evaluating criteria at pre-qualification stage of supplier selection. *Journal of Intelligent Manufacturing*, *25*(1), 165-175.

Kamsu-Foguem, B., Rigal, F., & Mauget, F. (2013). Mining association rules for the quality improvement of the production process. *Expert systems with applications*, *40*(4), 1034-1045.

Kumari, R., & Mishra, A. R. (2020). Multi-criteria COPRAS Method Based on Parametric Measures for Intuitionistic Fuzzy Sets: Application of Green Supplier Selection. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, 1-18.

Lau, H. C., Ho, G. T., Chu, K. F., Ho, W., & Lee, C. K. (2009). Development of an intelligent quality management system using fuzzy association rules. *Expert Systems with Applications*, *36*(2), 1801-1815.

Lau, H. C., Ho, G. T., Zhao, Y., & Chung, N. S. H. (2009). Development of a process mining system for supporting knowledge discovery in a supply chain network. *International Journal of Production Economics*, *122*(1), 176-187.

Lee, A. H., Kang, H. Y., Hsu, C. F., & Hung, H. C. (2009). A green supplier selection model for high-tech industry. *Expert systems with applications*, *36*(4), 7917-7927.

Lin, R. H., Chuang, C. L., Liou, J. J., & Wu, G. D. (2009). An integrated method for finding key suppliers in SCM. *Expert Systems with Applications*, *36*(3), 6461-6465.

Liu, B., Hsu, W., & Ma, Y. (1999, August). Mining association rules with multiple minimum supports. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 337-341). ACM.

Manohar, H. L., & Kumar, R. G. (2020). A neural networks model for green supplier selection. International Journal of Services and Operations Management, 35(1), 1-11.

Noci, G. (1997). Designing green vendor rating systems for the assessment of a supplier's environmental performance. *European Journal of Purchasing & Supply Management*, *3*(2), 103-114.

Prosman, E. J., & Sacchi, R. (2018). New environmental supplier selection criteria for circular supply chains: Lessons from a consequential LCA study on waste recovery. *Journal of Cleaner Production*, *172*, 2782-2792.

Sahebjamnia, N., Goodarzian, F., & Hajiaghaei-Keshteli, M. (2020). Optimization of Multi-period Three-echelon Citrus Supply Chain Problem. Journal of Optimization in Industrial Engineering, 13(1), 39-53.

Salleb-Aouissi, A., Vrain, C., & Nortet, C. (2007, January). QuantMiner: A Genetic Algorithm for Mining Quantitative Association Rules. In *IJCAI*, 7, 1035-1040).

Tsay, Y. J., & Chiang, J. Y. (2005). CBAR: an efficient method for mining association rules. *Knowledge-Based Systems*, *18*(2-3), 99-105.

Tseng, M. L., & Chiu, A. S. (2013). Evaluating firm's green supply chain management in linguistic preferences. *Journal of cleaner production*, *40*, 22-31.

Tuzkaya, G., Ozgen, A., Ozgen, D., & Tuzkaya, U. R. (2009). Environmental performance evaluation of suppliers: A hybrid fuzzy multi-criteria decision approach. *International Journal of Environmental Science & Technology*, *6*(3), 477-490.

Wang, L., Meng, J., Xu, P., & Peng, K. (2018). Mining temporal association rules with frequent item sets tree. *Applied Soft Computing*, *62*, 817-829.

Yeh, W. C., & Chuang, M. C. (2011). Using multi-objective genetic algorithm for partner selection in green supply chain problems. *Expert Systems with applications*, *38*(4), 4244-4253.

Yuan, X. (2017, March). An improved Apriori algorithm for mining association rules. In AIP conference proceedings, 1820(1), 080005. AIP Publishing LLC.

Adibi, M., Esfandyari, N. (2021). Identifying and Evaluating Effective Factors in Green Supplier Selection using Association Rules Analysis. *Journal of Optimization in Industrial Engineering*, 14(2), 115-125.

http://www.qjie.ir/article_676303.html
DOI: 10.22094/JOIE.2020.575685.1589