Int. J. Nonlinear Anal. Appl. 14 (2023) 8, 83–94
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2022.28571.3929

# Mining association rules for identifying critical factors affecting the implementation of business intelligence systems through WST-WFIM algorithm

Saeed Nalousi[a], Yousef Farhang[b,*], Amin Babazadeh Sangar[a], Kambiz MajidZadeh[a]

[a]Department of IT and Computer Engineering, Urmia Branch, Islamic Azad University, Urmia 57169-63896, Iran
[b]Department of IT and Computer Engineering, Khoy Branch, Islamic Azad University, Khoy 58135-175, Iran

*(Communicated by Majid Eshaghi Gordji)*

---

**Abstract**

In spite of the various advantages of Business Intelligence Systems (BIS), implementing them brings different challenges. Implementing BIS without considering the related challenges and determinants will increase the total cost and decrease added value for the organization. In this study, a questionnaire is developed to identify the critical factors affecting the implementation of BIS in automotive parts manufacturing companies and analyzed through a data mining technique, namely association rules, and the WST-WFIM algorithm on weighted data. The algorithm aims to extract sets of frequent rules and their weights (determined according to their importance) obtained from an expert panel. Taguchi method was adapted to design the algorithm's optimized parameters in order to obtain more effective rules. After applying the new algorithm to effective weighted factors, the relation and importance of each collection of effective factors are analyzed. The findings showed that, from the experts' viewpoint, the most important factors for successful implementation of BIS include (1) Removing potential negative resistances and barriers in spite of the various advantages of the Business Intelligence Systems (BIS), implementing them brings different challenges. Implementing BIS without considering the related challenges and determinants will increase the total cost and decrease added value for the organization. In this study, a questionnaire is developed to identify the critical factors affecting the implementation of BIS in automotive parts manufacturing companies and analyzed through a data mining technique, namely association rules, and the WST-WFIM algorithm on weighted data. The algorithm aims to extract sets of frequent rules and their weights (determined according to their importance) obtained from an expert panel. Taguchi method was adapted to design the algorithm's optimized parameters in order to obtain more effective rules. After applying the new algorithm to effective weighted factors, the relation and importance of each collection of effective factors are analyzed. The findings showed that, from the experts' viewpoint, the most important factors for successful implementation of BIS include (1) Removing potential negative resistances and barriers to implement BIS, (2) alignment between business strategy and BIS characteristics; and (3) system reliability, flexibility, and scalability. implement BIS, (2) alignment between business strategy and BIS characteristics; and (3) system reliability, flexibility, and scalability.

Keywords: business intelligence, data mining, weighted association rules, WST-WFIM algorithm, taguchi method
2020 MSC: 62R07

---

*Corresponding author
*Email addresses:* saeed.nalousi@yahoo.com (Saeed Nalousi), yfarhang@yahoo.com (Yousef Farhang), bsamin2@liveutm.onmicrosoft.com (Amin Babazadeh Sangar), k.majidzadeh@iaurmia.ac.ir (Kambiz MajidZadeh)

*Received:* November 2021 *Accepted:* February 2022

## 1 Introduction

Business intelligence (BI) is one of the key methods for understanding the business world. Neglecting the effective potential of BI in the organization makes discovering the business strengths and weaknesses in the past and making efficient decisions for the future impossible (or [at least] difficult) for managers and policy-makers. Although the role of management information systems (MIS), which were introduced to the business world in the early 1970s, is undeniable, they are unable to meet the needs of organizations, alone. Managers of organizations seek options for instant and accurate monitoring of the past and current events in the organization transparently in order to make constructive decisions for the future of the organization [19]. Therefore, business intelligence systems (BIS) are becoming more and more popular, and organizations allocate significant budgets towards implementing and maintaining them every year [22].

BIS will help managers to make informed business decisions by analyzing a vast amount of the organization's information. This analysis adopts intelligent data mining (DM) techniques for using and extracting/mining knowledge. These techniques can address two objectives: (1) describe the current situation, and (2) predict the potential situation in the future [8]. Extracting association rules (AR) through data is one of the key applications of DM that is useful for finding sets of items (itemsets) and seeks to identify the situations that explain information items co-occurring in large amounts of data [2, 9].

There exists a variety of algorithms to find and analyze RA. The APriori algorithm was proposed by Agrawal and Srikant in 1994 as a basic algorithm for association rule mining (ARM) [1]. This algorithm helps to find itemsets that repeat with a certain probability within the future. These results are obtained through repeated scanning of the data and produced temporary results at each stage. However, the challenges of the APriori algorithm, including the high complexity of the computations, high allocation of memory, and low speed of running, led Han, Pei, and Yin (2000) to propose a new algorithm, namely FP-Growth. FP-Growth algorithm uses a tree structure, called FP-Tree, in order to find AR and includes a fewer number of scans for each transaction than APriori.

The automobile industry is one of the strategic industries in Iran and plays an important role in the economic development of the country [7, 16]. Considering the significant investments in this industry and its importance to the country, it seems using BI tools is required for analyzing information and making critical decisions in active companies in the Iranian automobile industry and automotive parts manufacturing domain. Nevertheless, identifying and analyzing the success factors affecting BI development is required for implementing BI in an automotive parts manufacturing company.

According to the above explanations, it is clear that most major decisions in automotive parts manufacturing companies are made based on analyzing various information sources and considering all areas that can be affected by these decisions. Therefore, it is vital to adopt BIS for making managerial decisions and to analyze the success factors affecting the implementation of BIS from managers' and information technology (IT) experts' viewpoints. Following this approach requires analyzing collected data through scientific procedures and up-to-date and efficient algorithms and methods.

To achieve this objective, the current literature on AR algorithms is reviewed in the next section; then, in section three, the research problem is stated. Section four includes an improved algorithm and section five explains the results of applying this algorithm to analyze the collected data. Finally, section six covers the conclusions and recommendations for future studies.

## 2 Background

Taking informed, fast, and consistent decisions based on real data is a key for success in the current information age. Inappropriate business decisions based on minimum or incomplete information can cause many damages to a company and its market share, and even predict the company's bankruptcy [24]. On the other hand, it's impossible to process big data (including a large amount of transactional data) through traditional processing software and tools. BI tools and techniques are capable of the in-depth process of large amounts of data produced by the business. These techniques include strategies and technologies that are used by different companies to analyze business data. BI technologies can provide historical, current, and predictive views of business activities and operations [4].

### 2.1 Success factors of BI

The complex business operations, due to the increasing volume of information as a major competitive advantage in today's business world, create a large amount of data in organizations and impose new challenges for managers and managerial decision-making. Therefore, meeting the needs of the organization for making informed decisions requires modern tools and techniques [14, 18].

BI, as a competitive advantage, is required for strategic decision-making in organizations and analysis of business data. BI is defined in several ways; some of these definitions have focused on the goals of BI and the others consider its goals, structure, and processes [6]. Since business decisions are related to the economy as well as the processing and using knowledge, BI is critical for representing information assets in organizations that are required for making informed decisions. In the current information age, the quality and readiness of BI in organizations can predict profitability or loss, as well as viability or failure. BI tools and techniques not only enable organizations to save, modify, model, and analyze a vast amount of information about their business activities and operations but also help them to improve their performance through collecting and examining costumer' feedback. Accordingly, it can be argued that BI includes a wide range of technologies and applications for collecting, retrieving, and analyzing data in order to help make decisions by managers and decision-makers [17, 5, 25]. BI is divided into three high-level processes: (1) information collection, (2) information processing, and (3) information distribution. In addition, each process includes separate sub-processes [21].

### 2.2 Association rules

In 2004, Antonie and Zaiane proposed a new algorithm in order to generate both positive and negative AR, simultaneously. The process of ARM includes two phases: (1) mining for frequent itemsets; and (2) generating strong AR from the discovered frequent itemsets [3]. The negative association rule (NAR) from frequent itemsets [11] is one of the first algorithms for mining NAR. This algorithm is based on APriori, which discovers both positive and negative AR and enhances APriori by considering the absence of itemsets. Those patterns in that their minimum support is greater than a threshold are labeled as extraction rules. Therefore, this algorithm is able to extract the majority of existing rules among the data. Discovering NAR was implemented for the first time in the tourism industry in 2012 by Rong, Vu [23] in order to draw required information. This algorithm follows several steps; the first step is to generate candidate patterns. The next step is to calculate their support and lever values; if a certain pattern satisfies both values, it is added to the interesting candidate pattern itemsets. However, if the pattern satisfies the lever value only, it is added to the interesting negative patterns itemsets [26].

### 2.3 Questionnaire and data mining

Since the traditional methods for analyzing questionnaires include various limitations in terms of distribution, data collection, and statistical analysis, they have limited generalizability. Therefore, Lai, Zhang [12] proposed adopting DM and text mining (TM) methods for analyzing a questionnaire to study the use of new energy vehicles. Park, Lee [20] conducted a survey in order to derive safety implications that are useful for developing policies to enhance taxi safety based on analyzing intrinsic characteristics underlying the roots and causes of traffic accidents. This study was conducted an in-depth questionnaire survey to gather useful data representing the intrinsic characteristics. In their survey, 781 corporate taxi drivers participated in Korea. The proposed method for analyzing the collected data consists of two-stage DM techniques, including a random forest method (RFM), with data that explains the working condition and welfare environment of participants. The first stage includes deriving the participants' intrinsic characteristics in order to classify four types of taxi drivers: (1) unspecified normal, (2) work-life balanced, (3) overstressed, and (4) work-oriented. The second stage includes determining the priority for classifying high-risk taxi drivers based on various factors derived from the first stage. In addition, Li, Fei [13] examined burnout of construction project managers (CPMs) to measure their chronic stress and aimed to explore the key effective factors on burnout from individual, job-related, organizational, and social aspects, systematically. Considering the domestic and classic measurement scales, the study was conducted a cross-sectional survey on 634 Chinese CPMs. Their study ignores the traditional research paradigm (i.e. "hypothesis-test") and adopted a DM method and AR analysis in order to identify the key factors and mechanisms of Chinese CPM burnout. In addition to the traditional mechanisms, complex interactive networks among the key factors, job burnout, and the reversed effects of burnout were acquired.

According to the above explanations, the role of BI analysis in different industries is undeniable and the traditional methods for analyzing questionnaires are not effective in the current businesses. Therefore, the current study uses BI analysis in the automobile industry, and instead of traditional statistical analysis, adopts AR and intelligent algorithms for analyzing the collected data.

### 2.4 Problem statement

The structure and characteristics of a certain industry and company can impact the success/failure and acceptance/rejection of the implementation of BI in that industry and company. Therefore, the next section explains the characteristics of automotive parts manufacturing companies and the importance of implementing BI in these companies, as well as the proposed algorithm.

### 2.5 Case study: automotive parts manufacturing holding

The case of the current study is an automotive parts manufacturing holding in Iran that includes 15 companies and begins its business in the early 1990s. The companies are located in different regions around the country with thousands of employees. This holding has earned a variety of key global certifications and supports the Iranian automobile industry by manufacturing important and excellent automotive parts. The holding's main strategy is an active presence in the market through establishing various sites to produce automotive parts, manufacturing excellent goods by earning key global certifications and technical cooperation with other companies, taking advantage of the latest technologies. This holding aims to adopt entrepreneurship and flexible and agile structures.

During the past years, the holding has had its own enterprise resource planning (ERP) systems and in recent years, it has established BIS and made a team for maintaining these systems. One of the basic needs of this holding is to identify success factors of BIS and this research aims to address this need.

### 2.6 The characteristics of the problem

According to [27], the critical success factors (CSFs) of BIS include three categories:

- Organizational factors (e.g. 'committed management support and sponsorship' and 'clear vision and well-established business case');
- Process-related factors (e.g. 'business-centric championship and balanced team composition', 'business-driven and iterative development approach', and ''user-oriented change management');
- Technical factors (e.g. 'business-driven, scalable and flexible technical framework' and 'sustainable data quality and integrity').

In addition, environmental factors have an important role in implementing BIS. Therefore, through desk research (or library research), 145 CSFs for implementing BIS were identified. Then, 10 experts reviewed these CSFs in order to exclude, merge, and verify them. Eventually, 40 CSFs were verified by the expert panel, which were categorized into seven dimensions:

- Impact of 'human resources' factor on system productivity;
- Impact of 'human resources' factor on senior managers' decision making;
- Impact of 'information technology' factor on system productivity;
- Impact of 'information technology' factor on senior managers' decision making;
- Impact of 'organizational factor' on system productivity;
- Impact of 'environmental factor' on system productivity;
- Impact of 'environmental factor' on senior managers' decision-making.

All the above factors were used to design a questionnaire for surveying different groups. These groups and the number of items developed for each group are shown in Table 1.

After designing the questionnaire, participants were selected from BIS experts and stakeholders, including managers at the top and middle levels of management in different companies of the mentioned holding, information technology (IT) managers and experts, and experienced people and influencers in the organizational decision-making process. CSFs in the questionnaire were scaled by 5-point Likert scales (from 1 [=Not important at all] to 5 [=very important]). The questionnaire was distributed online among 250 samples and 193 completed questionnaires were collected.

Table 1: Surveyed groups in the study and number of items developed for each group

| Groups | reliminary information | General | Organization | Management | Project Team | Technology | Vision | Users | Data |
|---|---|---|---|---|---|---|---|---|---|
| Number of Items | 2 | 1 | 9 | 5 | 11 | 8 | 2 | 3 | 2 |

## 3 Solving method

This section describes the ARM algorithm and the proposed algorithm for solving the stated problem.

### 3.1 Association rules

In a large amount of transactional data ($T$), there are some hidden relations among transactions' data that discovering them can be useful for predicting systems' behavior. Given $X$ and $Y$ are two individualsamples in items from transaction $t$ in $T$. The law $X \to Y$ is considered as a rule that satisfies the following conditions:

$$\mathrm{Sup}\,(X \to Y) \geq \mathrm{MinSup} \;\text{ that }\; \mathrm{Sup}\,(X) = |\{t \in T;\; (X \cup Y) \in t\}|\,/\,|T| \tag{3.1}$$

$$\mathrm{Conf}\,(X \to Y) \geq \mathrm{MinConf} \;\text{ that }\; \mathrm{Conf}\,(X \to Y)\,\mathrm{Sup}\,(X \cup Y)\,/\,\mathrm{Sup}(X) \tag{3.2}$$

These two conditions determine the interestingness of the discovered law for users. Minimum support (MinSup) and minimum confidence (MinConf) are realistic criteria in order to define the interestingness. These two criteria in the AR algorithms are determined by the user and according to the characteristics of the problem. Changes to these variables allow users to make different assessments of the status of their data.

### 3.2 Basic algorithm for association rules mining

There are various algorithms for ARM and APriori is one of the commons algorithms in this domain. APriori is a method that is applied to records in a dataset, especially transactional datasets or records containing a certain number of fields or items. APriori uses a bottom-up approach that (gradually) makes comparisons between complex records. This algorithm is an effective method for solving current complicated problems in DM and machine learning (ML). This algorithm adopts a pruning tree mechanism for comparing huge amounts of transactions in order to make problem smaller. If an itemset that its support measure is greater than MinSup will be considered as a frequent itemset, pruning algorithm will be performed through two key principles [1]:

- All subsets of a frequent itemset must also be frequent.
- All supersets of an infrequent itemset are infrequent.

APriori is efficient for a dataset containing a small number of transactions and items, but it's very expensive in terms of computation. Even if APriori decreases the number of candidate items for investigation, when the number of transactions is large or the value of MinSup is low, the volume of computations will still be large. An alternative solution to this problem is to reduce the number of comparisons through advanced data structures (for example, hash tables) for sorting candidate items in a more effective way.

Various challenges of APriori algorithm, including high complexity of the computations, high allocation of memory, and low speed of running, Han, Pei [10] to propose a new algorithm, namely FP-Growth. FP-Growth algorithm uses a tree structure, called FP-Tree, in order to mine AR and includes a fewer number of scans for each transaction than APriori. However, FP-Growth requires more memory allocation. APriori and FP-Growth do not consider the items' weights in the transactions for mining AR, and this issue is not promising in the practice. For example, to analyze the profit of products in a store, the profit of each product in the invoice and the frequency of products together will be checked. Then, the profit can be considered as the weight and importance of a certain product in the invoice. It can be argued that those algorithms that seek such patterns assign a certain weight to each item in each transaction and aim to mine weighted AR (WAR).

### 3.3 Key algorithms for weighted association rules mining

One of the key algorithms for WAR mining is WST-WFIM. This algorithm uses the weight of each transaction's item FP-Tree - a tree structure in the FP-Growth algorithm to calculate the average weight of discovered rules. By keeping information about each group's weight in FP-Tree, a new tree (i.e. Partial Weighted FP-Tree [PWFP-Tree]) is created and the required calculations will be performed based on PWFP-Tree. If a set of 'n' items of participated items in transactions be $I = \{i_1, i_2, i_3, i_4, \ldots, i_n\}$, a transactional weighted dataset is a set of transactions, which is defined as $T = \{t_1, t_2, t_3, t_4, \ldots, t_m\}$. Then, each transaction is a set of binary attributes as follows:

$$T_k = \{(i_1, w_1), (i_2, w_2), (i_3, w_3), \ldots, (i_n, w_n) | \forall x \neq y \ = \gg i_x \neq i_y\} \tag{3.3}$$

This set of binary attributes includes the weight (importance) of each item. Table 2 illustrates an example of a weighted dataset. This table includes the responses of 193 participants (TID) to 40 items (ItemID) that are scaled by 5-point Likert scales (from 1 [=Not important at all] to 5 [=very important]).

Table 2: Weighted transactional dataset

| TID | $I = \{1, 2, 3, \ldots, 40\}$ |
|---|---|
| | Transaction (ItemID, Weight) |
| 1 | $\{(1,2),(2,3),(3,4),(5,2),\ldots,(38,3)\}$ |
| 2 | $\{(2,5),(4,1),(5,4),(7,2),\ldots,(40,1)\}$ |
| 3 | $\{(1,1),(2,4),(6,3),(8,5),\ldots,(39,5)\}$ |
| ... | $\{\ldots\}$ |
| 193 | $\{(1,5),(2,5),(4,4),(5,4),\ldots,(40,5)\}$ |

By describing the item weight per transaction, two definitions can be provided:

**Definition 3.1.** The weight of each transaction ($T_k$) is equal to the sum of the weights of all participating items in the transaction divided by the number of participating items.

$$W(T_k) = \frac{\sum_{y=1}^{|T_k|} (w_y)}{|T_k|} \tag{3.4}$$

**Definition 3.2.** The weight of each item set ($\mathrm{IS}_i$) in each transaction ($T_k$) is equal to the sum of the weights of the relevant items in the transaction divided by the number of participating items in the item set.

$$\begin{aligned} &W(\mathrm{IS}_i\ , T_k) = \frac{\sum_{y=1}^{|\mathrm{IS}_i|} (w_y)}{|\mathrm{IS}_i|}\ , \text{Where } \mathrm{IS}_i \in T_k \\ &W(T_2) = \frac{5+1+4+2+\cdots+1}{31} = 3.66 \\ &W(\{2,4,5\}, T_2) = \frac{5+1+4}{3} = 3.33 \end{aligned} \tag{3.5}$$

The tree structure in PWFP-Tree is created through developing the tree structure in FP-Tree. Therefore, PWFP-Tree has not only the same basic structure but also the following components:

- Like in the basic tree, PWFP-Tree includes an item number (ItemID) and an item frequency (support) for each node;
- An ID, namely NodeUniqueID, is assigned to each node in the tree;
- A certain array structure, namely TranListArr, is added in order to keepinformation about transactions and weights in each node, including the transaction ID (TID) and the weight of the item in that transaction.

In the WST-WFIM algorithm the weight of items in each transaction can be defined as an independent numerical variable (less than or greater than zero).The key features of the proposed algorithm are:

- In each transaction, in addition to item ID, there is the item weight;

- The FP-Growth is considered as the basic algorithm;
- All concepts and operations for mining frequent items are the same as in the basic algorithm;
- The proposed algorithm calculates the weight of frequent item sets that are discovered by FP-Growth.

To obtain the algorithm's purpose, an array structure is added to the FP-Tree. The array structure keeps the TID and nodes' weights in the original tree and sub-trees of conditional FP-Tree, in addition to the main information (ItemID and support). After the creation of the original weighted tree, the operation of creating conditional FP-Tree and processing them are conducted in order to discover the frequent item sets. During the mentioned operation, the following steps are performed in order to apply proper weights to items.

- NodeUniqueID is transferred from the original tree;
- If the generated conditional FP-Tree for the node '$i$' be Tr $i$, during the generation of Tr $i$, the average weight of participating items will be calculated as conditional FP-tree weight ($\mathrm{CFPT}_W(\mathrm{Tr}_i, j)$). The obtained results are stored in the $\mathrm{FP}_{\mathrm{Header}}$ associated with that tree in order to calculate the final weights of the discovered frequent itemsets;
- $\mathrm{CFPT}_W(\mathrm{Tr}_i, j)$ is calculated based on the weight of the item '$j$' in joint transactions with the item '$i$' (Definitions 3.3 and 3.4).

**Definition 3.3.** Joint transactions in a tree, which include the nodes '$i'$ and '$j$', in a tree are those that create the path between the two nodes in the original PWFP-Tree ($\cap\mathrm{tr}(i, j)$).

**Definition 3.4.** The average weight of each item '$j$' in $\mathrm{Tr}_i$ is calculated as follows:

$$\mathrm{CFPT}\_W\left(\mathrm{Tr}_i, j\right) = \frac{\sum_{k\in\cap\mathrm{tr}(i,j)} w_k}{\mathrm{Sup}\left(\mathrm{Tr}_i, \mathrm{j}\right)}\ ,\text{where } \mathrm{Sup}\left(\mathrm{Tr}_i, j\right) \text{ is Support } (j)\ in\ \mathrm{Tr}_i \tag{3.6}$$

## 4 Experimental results

This section demonstrates the results of executing the proposed algorithm. Key parameters of the algorithm will be set according to Taguchi method.

### 4.1 Algorithm execution

In order to identify the CSFs for implementing BIS through the results of online survey, in the WST-WFIM algorithm, MinSup parameter was set to 95%, 90%, 85%, and 80%. Since the lower the selected value for MinSup, the more rules is extracted, selecting the proper value for the key parameters of algorithm is critical. Although it is possible to select different values for MinSup, if these values are less than the mentioned values, the number of datasets (factors group) will be more than expected results and useless for the evaluation. Table 3 shows the extracted rules based on different MinSup values.

According to Table 3, the number and weights of extracted rules are dependent on the key parameters of the algorithm. Therefore, selecting the proper value for the key parameters of the algorithm, according to the problem, is critical.

### 4.2 Setting parameters based on taguchi method

The performance of an algorithm is dependent on its key parameters. This section explains the results of the investigation of the proper value for key parameters of the WST-WFIM algorithm. Taguchi experimental design method was adopted to conduct this phase. Key parameters in the proposed algorithm are:

- MinSup: is the minimum support threshold; the transactions' values greater than MinSup are important and smaller values will be pruned from a certain tree;
- MinConf: is the minimum confidence threshold; the values greater than MinConf are useful for extracting rules and smaller values are not important;

Table 3: Extracted rules from the results of the online survey based on different MinSup values

| MinSup | Rules Number | $K$-Itemset | Weighted Frequent Rules (Frequency , Weight) | |
|---|---|---|---|---|
| | | | Sort By Frequency | Sort By Weight |
| 95% | 3 | 1 | 29(184 , 3.76) | 19(183 , 3.80) |
| | | | 19(183 , 3.80) | 29(184 , 3.76) |
| | | | 25(183 , 3.63) | 25(183 , 3.63) |
| 90% | 35 | 2 | 1929(176 , 3.78) | 22 29(174 , 3.91) |
| | | | 1925(175 , 3.73) | 7 19(173 , 3.89) |
| | | | 2529(174 , 3.69) | 19 38(173 , 3.80) |
| | | | 2229(174 , 3.91) | 19 29(176 , 3.78) |
| | | | 1938(173 , 3.80) | 1925(175 , 3.73) |
| 85% | 84 | 5 | 19 29 30 34 38(154 , 3.79) | 19 29 30 34 38(154 , 3.79) |
| | | 4 | 4 19 25 29(159 , 3.68) | 7 22 29 34(154 , 3.93) |
| | | | 19 25 29 31(159 3.70) | 7 19 22 38(154 , 3.92) |
| | | | 19 22 25 29(159 , 3.82) | 7 22 29 38(154 , 3.91) |
| | | | 7 19 29 31(159 , 3.78) | 7 21 22 29(154 , 3.91) |
| | | | 19 21 29 37(159 , 3.81) | 7 19 22 29(158 , 3.90) |
| | | 3 | 19 25 29(168 , 3.74) | 7 15 22(154 , 4.00) |
| | | | 19 29 38(167 , 3.79) | 7 22 24(156 , 3.99) |
| | | | 19 29 31(167 , 3.72) | 7 20 22(156 , 3.98) |
| | | | 19 29 30(167 , 3.76) | 22 23 29(155 , 3.98) |
| 80% | 15755 | 8 | 4 19 22 25 29 30 34 38(135 , 3.77) | 7 19 22 29 30 34 37 38(135 , 3.81) |
| | | | 19 22 25 29 30 34 37 38(135 , 3.77) | 4 19 22 25 29 30 34 38(135 , 3.80) |
| | | | 7 19 22 29 30 34 37 38(135 , 3.81) | 19 22 25 29 30 34 37 38(135 , 3.79) |
| | | 7 | 19 22 25 29 30 34 38(141 , 3.81) | 7 19 22 24 29 30 34(136 , 3.87) |
| | | | 19 22 29 30 34 37 38(141 , 3.79) | 7 19 22 29 30 34 38(140 , 3.86) |
| | | | 19 22 25 29 30 37 38(140 , 3.76) | 7 19 21 22 29 30 34(136 , 3.86) |
| | | | 7 19 29 30 34 37 38(140 , 3.77) | 7 19 22 24 29 30 38(136 , 3.85) |
| | | 6 | 19 25 29 30 34 38(147 , 3.76) | 7 19 22 23 29 34(135 , 3.95) |
| | | | 19 22 29 30 34 38(147 , 3.84) | 19 22 23 24 29 34(135 , 3.94) |
| | | | 19 29 30 34 37 38(147 , 3.74) | 19 22 23 29 34 38(137 , 3.94) |
| | | | 7 19 22 29 30 34(146 , 3.87) | 7 22 23 29 30 34(136 , 3.93) |

- MinWeight: is the minimum weight(the importance of an item) threshold; the values greater than MinWeight are important and smaller values will be pruned from the FP-Tree;
- MinItemsCount: is the minimum number of items threshold; if the number of transactions is less than MinItemsCount, that transaction will be insignificant.

Experimental design is a scientific method proposed by Taguchi, which involves using orthogonal arrays to organize and manage the parameters of the algorithm affecting the process and the levels at which they should be varied to obtain desirable results. The experimental design aims to find a suitable combination of parameters' values in a reasonable time. Considering the number of selected parameters, Taguchi method uses the orthogonal array as a matrix experiment in order to check and test the experiment levels.

According to Taguchi method, the range of changes in defined function relative to the parameters' different values is illustrated by the signal-to-noise ratio (SNR). Then, those factors with the highest SNR will be selected as an optimal setting.

In the current study, Taguchi method is adopted in order to set the key parameters of the WST-WFIM algorithm, which can impact the outputs. Accordingly, three levels are defined for each parameter (Table 4).

Table 4: The key parameters of WST-WFIM algorithm and the defined levels

| Parameters | Level #1 | Level #2 | Level #3 |
|---|---|---|---|
| MinSup | 0.95 | 0.90 | 0.85 |
| MinConf | 0.6 | 0.7 | 0.8 |
| MinWeight | 3 | 3.5 | 4 |
| MinItemsCount | 20 | 30 | 35 |

Therefore, based on Table 4, the orthogonal table was constructed according to the number of parameters and the levels of the experiment. Since finding more and most frequent rules is helpful for discovering right and accurate rules, the number of extracted rules and the probability of occurrence of these rules are two important issues. Therefore, in the current study, the average weight of a rule and the probability of its occurrence are considered as criteria for selecting optimal parameters of the algorithm. Table 5 shows the impact of control factors on SNR figure. In addition, Figure 1 illustrates the results of adopting Taguchi method in Minitab, in which the smaller value is considered as the optimal value.

Table 5: The orthogonal table for four parameters and three levels

| Experiment number | MinSup | MinConf | MinWeight | MinItemsCount | The value of defined function |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 0.749 |
| 2 | 1 | 2 | 2 | 2 | 0.812 |
| 3 | 1 | 3 | 3 | 3 | 0.839 |
| 4 | 2 | 1 | 2 | 3 | 0.823 |
| 5 | 2 | 2 | 3 | 1 | 0.865 |
| 6 | 2 | 3 | 1 | 2 | 0.841 |
| 7 | 3 | 1 | 3 | 2 | 0.874 |
| 8 | 3 | 2 | 1 | 3 | 0.886 |
| 9 | 3 | 3 | 2 | 1 | 0.893 |

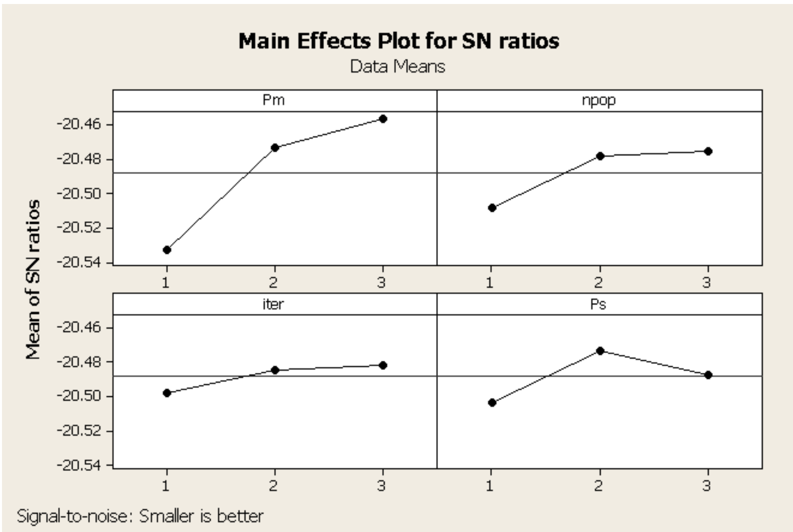


Figure 1: Comparing levels through Taguchi method

In Figure 1, the vertical axis represents SNR for parameters and the horizontal axis represents the levels. In general, the higher the SNR of the parameter, the closer the problem is to the defined objective. According to Figure 1 (first left raw, level #3), 0.85 is selected for MinSup. In addition, according to Figure 1 (second right raw, level #3), 0.8 is selected for MinConf. Table 6 shows the best values (among obtained results) for the key parameters.

Table 6: The orthogonal table for four parameters and three levels

| Parameters | Values |
|---|---|
| MinSup | 0.85 |
| MinConf | 0.8 |
| MinWeight | 4 |
| MinItemsCount | 30 |
| Defined function | 0.897 |

According to Table 6, extracted rules through parameters values are effective for identifying CSFs for implementing BIS in automotive parts manufacturing companies. The next section explains the key extracted rules.

## 5 Discussion and conclusion

The current study demonstrated the importance of implementing BIS in automotive parts manufacturing companies, and instead of traditional statistical analysis (such as defining hypothesis), adopts intelligent algorithms (based on DM techniques) for discovering the hidden rules in viewpoints of automobile industry experts. The basic AR algorithms are described in this paper. In addition, since the experts had different viewpoints on the weight and importance of CSFs for implementing BIS, WST-WFIM algorithm was adopted for identifying effective factors. Finally, in order to optimal discovery of rules, Taguchi method was used for setting key parameters of the WST-WFIM algorithm.

According to [28], the critical success factors (CSFs) of BIS include three categories: (1) Organizational factors (e.g. 'committed management support and sponsorship' and 'clear vision and well-established business case'); (2) Process-related factors (e.g. 'business-centric championship and balanced team composition', 'business-driven and iterative development approach', and 'user-oriented change management'); and (3) Technical factors (e.g. 'business-driven, scalable and flexible technical framework' and 'sustainable data quality and integrity'). In addition, environmental factors have an important role in implementing BIS. According to extracted rules, which were presented in the previous section, with the determined parameters by Taguchi method, the CSFs of implementing BIS in automotive parts manufacturing companies are identified. More than 85% of participants selected the following factors as the high priority CSFs for implementing BIS:

- Removing potential negative resistances and barriers to implement BIS;
- Alignment between business strategy and BIS characteristics;
- System reliability, flexibility, and scalability.

In addition, the following factors were identified as the second priority CSFs for implementing BIS in automotive parts manufacturing companies:

- System reliability, flexibility, and scalability;
- Transformational leadership in intelligent implementation of the system in the company;
- Strategic and long-term perspective for managing the company based on BIS.

In future studies, different weights and importance of CSFs can be investigated through selecting another expert panel with different expertise, experiences, organizational roles in the company, and patterns of interaction with the system. Worded differently, while top-level managers consider the financial dimension and performance of the company, IT experts are more experienced in implementing BIS. In addition, future studies can investigate the CSFs of implementing BIS in other companies and industries, including the pharmaceutical industry, banking industry, food industry, petrochemical industry, etc., and compare their findings to those of the current research.

## Acknowledgements

The authors would like to thank the reviewers for their constructive and valuable comments to improve the paper.

## References

[1] R. Agrawal and R. Srikant, *Fast algorithms for mining association rules*, Proc. 20th Int. Conf. Very Large Data Bases, VLDB, Citeseer, 1994.

[2] F.A. Amani and A.M. Fadlalla, *Data mining applications in accounting: A review of the literature and organizing framework*, Int. J. Account. Inf. Syst. **24** (2017), 32–58.

[3] M. L. Antonie and O.R. Zaïane, *Mining positive and negative association rules: An approach for confined rules*, Eur. Conf. Principles of Data Mining and Knowledge Discovery, 2004.

[4] E. Ayoubi and S. Aljawarneh, *Challenges and opportunities of adopting business intelligence in SMEs: collaborative model*, Proc. First Int. Conf. Data Sci. E-learn. Inf. Syst., 2018.

[5] M.K. Chen and S.C. Wang, *The use of a hybrid fuzzy-Delphi-AHP approach to develop global business intelligence for information service firms*, Expert Syst. Appl. **37** (2010), no. 11, 7394–7407.

[6] N. Dedić and C. Stanier, *Measuring the success of changes to existing business intelligence solutions to improve business intelligence reporting*, Int. Conf. Res. Practical Issues of Enterprise Inf. Syst., 2016.

[7] A. Fathali, *Examining the impact of competitive strategies on corporate innovation: An empirical study in automobile industry*, Int. J. Asian Soc. Sci.**6** (2016), no. 2, 135–145.

[8] A. Ghorbani and S. Farzai, *Fraud detection in automobile insurance using a data mining based approach*, Int. J. Mechatron. Elekt. Comput. Technol. **8** (2018), no. 27, 3764–3771.

[9] J. Han, M. Kamber and D. Mining, *Concepts and techniques*, Morgan Kaufmann **340** (2006), 94104–3205.

[10] J. Han, J. Pei and Y. Yin, *Mining frequent patterns without candidate generation*, ACM Sigmod Record **29** (2000), no. 2, 1–12.

[11] A.S.A. Kadir, A.A. Bakar and A.R. Hamdan, *Frequent absence and presence itemset for negative association rule mining*, 11th Int. Conf. Intel. Syst. Design Appl., 2011.

[12] X. Lai, S. Zhang, N. Mao, J. Liu and Q. Chen, *Kansei engineering for new energy vehicle exterior design: An internet big data mining approach*, Comput. Indust. Engin. **165** (2022), 107913.

[13] X. Li, Y. Fei, T.E. Rizzuto and F. Yang, *What are the occupational hazards of construction project managers: A data mining analysis in China*, Safety Sci. **134** (2021), 105088.

[14] Y.H. Lin, K.M. Tsai, W.J. Shiang, T.C. Kuo and C.H. Tsai, *Research on using ANP to establish a performance assessment model for business intelligence systems*, Expert Syst. Appl. **36** (2009), no. 2, 4135–4146.

[15] I. Martin-Rubio, J. Fombellida and D. Andina, *The evolution of business intelligence with neuroinformatics*, Int. Conf. Indust. Engin. Oper. Manag., 2018.

[16] A. Memarnezhad, S. Hosseini and S. Khatabi, *Evaluation of Structure and Performance of Automobile Industry in Iran*, Quart. J. Econ. Model. **3** (2010), no. 4, 103–120.

[17] A. Mikroyannidis and B. Theodoulidis, *Ontology management and evolution for business intelligence*, Int. J. Inf. Manag. **30** (2010), no. 6, 559–566.

[18] M. Nakayama, Ö. Isik, N. Sutcliffe and S. Olbrich, *Grassroots Business Intelligence as an Enabler of Change Management: A Case Study at a Large Global Manufacturing Firm*, Complex Syst. Inf. Model. Quart. **23** (2020), 1–11.

[19] S. Nanda, B.P. Bhol and S. Misra, *Business intelligence and decision making influence Bancassurance system*, Odisha J. Soc. Sci. **7** (2020), no. 1.

[20] J. Park, S. Lee, C. Oh and B. Choe, *A data mining approach to deriving safety policy implications for taxi drivers*, J. Safety Res. **76** (2021), 238–247.

[21] J. Reinschmidt and A. Francoise, *Business intelligence certification guide*, IBM International Technical Support Organisation, 2000.

[22] P. Rikhardsson and O. Yigitbasioglu, *Business intelligence & analytics in management accounting research: Status and future focus*, Int. J. Account. Inf. Syst. **29** (2018), 37–58.

[23] J. Rong, H.Q. Vu, R. Law and G. Li , *A behavioral analysis of web sharers and browsers in Hong Kong using targeted association rule mining*, Tourism Manag. **33** (2012), no. 4, 731–740.

[24] B. Sahay and J. Ranjan, *Real time business intelligence in supply chain analytics*, Inf. Manag. Comput. Secur. **16** (2008), no. 1, 28–48.

[25] M. Seah, M.H. Hsieh and P.D. Weng, *A case analysis of Savecom: The role of indigenous leadership in implementing a business intelligence system*, Int. J. Inf. Manag. **30** (2010), no. 4, 368–373.

[26] R. Sethi and B. Shekar, *Subjective interestingness in association rule mining: A theoretical analysis*, Digital Business, Springer, 2019.

[27] W. Yeoh and A. Koronios, *Critical success factors for business intelligence systems*, J. Comput. Inf. Syst. **50**

(2010), no. 3, 23-32.

[28] W. Yeoh and A. Popovič, *Extending the understanding of critical success factors for implementing business intelligence systems*, J. Assoc. Inf. Sci. Technol. **67** (2016), no. 1, 134–147.