MATHEMATICAL MODELING AND COMPUTING, Vol. 10, No. 2, pp. 557–565 (2023)

Toward new data for IT and IoT project management method prediction

Merzouk S.[1], Gandoul R.[2], Marzak A.[1], Sael N.[1]

[1] *Laboratory of Information Technology and Modeling, Faculty of Sciences Ben M'Sik, Hassan II University of Casablanca, Casablanca, Morocco*
[2] *Laboratory of Sciences and Technology of Information and Education, Faculty of Sciences Ben M'Sik, Hassan II University of Casablanca, Casablanca, Morocco*

(Received 13 February 2023; Accepted 21 April 2023)

Selecting the best project management method at the workplace helps to deliver a high-quality product to the customer. Hence, the need for good knowledge of management methods, their characteristics, advantages, and disadvantages, is necessary to be able to select the best for the specific project. However, until now, no large dataset for Machine Learning and decision-making model, model or system has been proposed to help project managers to the most efficient method adapted to the constraints of their projects. This work develops the construction of the dataset for agile and IoT project management method based on the real experiences. In this paper, our objective is to propose a criteria-based model that allows the choice of the best management method to adopt for such an IT or IoT project according to a set of criteria.

Keywords: *agile method; information technology (IT); internet of things (IoT); project management; dynamic system development method (DSDM).*

2010 MSC: 68P99 **DOI:** 10.23939/mmc2023.02.557

1. Introduction

The Internet of Things (IoT) industry is expanding daily as more devices connect to the Internet. Its typical applications are expanding in a variety of fields, including e-health, smart homes and buildings, smart meters, smart grid, among other areas [1,2]. The IoT project is the combination of software and hardware in IoT project unlike the IT project which is based on software and less complex.

Managing an IT or IoT project is a significant challenge for any organization. Because the management method enables teams to self-organize, operate effectively, and complete each phase of the project while adhering to the set criteria. Selecting a project management method is not an easy task. The method used has an impact on the ultimate output, either positively or badly, depending on the project and personnel criteria [3]. There is a diversity of agile and IoT methods like Scrum [4], Extreme Programming (XP) [5], SAFe [6], Ignite [7], IoT Methodology [8], and more. Each of them has its own particularities, even if they are based on the same values and principles as the agile methods.

According to our state-of-the-art studies [9, 10], each method has its own singularity in terms of project/team size, iteration time, change, and so on. The non-existence of a platform or a model that allows the selection of the right method for a specific IT or IoT project makes project managers search and compare methods in order to choose the right method or to adopt the method already used in their previous project even if it is not adapted to their current project. This can pose a time-wasting problem in the search or the risk of failure of the project. Table 1 shows some management methods (agile/IoT methods) for IT and IoT projects.

Table 1. Agile/IoT methods for IT/IoT projects.

Method	IT project	IoT project
Scrum	x	x
eXtreme Programming	x	
Ignite		x
SAFe	x	x
IoT Methodology		x
Kanban	x	x
FDD	x	
DSDM	x	

For this reason, we aim to develop a choice model that allows us to detect the most adapted management method for IT or IoT projects. To achieve this goal, we first realize state of art analysis and investigation to collect the research datasets offered for this object, however, we did not find any dataset available, and which offers a description of software projects in this regard. To face this challenge, we propose in this work paper the construction of a Project management dataset containing the necessary information for the methods used in real projects.

The remainder of this paper is organized as follows: Section 2 presents State-of-art of IT and IoT project management. Section 3 define data construction approach. Then the results and analysis of data construction is left for Section 4. Section 5 presents association rule for management method selection using the Rapidminer software tool. The discussion and conclusion will present in Section 6.

2. IT and IoT project management methods state of art

Since the emergence of the agile approach known by the agile methods such as Scrum, XP, Kanban, etc. Information Technology (IT) companies adopt this approach to manage their projects because it solves the problems caused by the traditional approach as the length of the life cycle. Agile methods are iterative and incremental that allows changes during the project realization. They are based on four values and twelve principles agreed upon in the manifesto in 2001 [11]. The evolution of technology contributes to the emergence of the internet of things (IoT) which implies the emergence of a new type of project which is IoT projects that integrate hardware and software layers. The managers of IoT projects propose their own methods (Ignite and IoT Methodology) based on their experiences after experimenting with agile methods and scalable methods like SAFe. The proposed methods specific to IoT projects do not prevent managers from using agile methods such as Scrum or Dynamic System Development Method (DSDM [12]) or scalable methods such as SAFe. There are also other methods proposed in the IoT context. The variety of these methods for neither IT nor IoT projects makes the choice of the best technique adapted for a software project an important challenge in the field of software development.

Until now, variant initiatives are conducted to study and analyze the most adapted techniques to IT or IoT projects. Pouyandeh et al. [13] conducted a study that determined the multiple construction industry success criteria, and as a consequence, the latter discovered the appropriate agile method to use by merging the Delphi and AHP methodologies. This research investigates how using agile approaches influences how construction project success criteria are improved and compared. Then, in order to select the best agile method to adopt, prioritize agile methods and criteria.

Moreover, we found another initiative to build a dataset concerning agile methods. The dataset found on Kaggle [14] contains 30 lines and information about the methods used in the companies, the reasons for their choice, whether they have made changes to their chosen method, whether they want to migrate to agile methods in case of using traditional methods, and so on. However, this data does not allow good visualization of the criteria that allow the choice of the method according to the project and company specificities. As seen before, no system or model facilitates the choice of the most adapted method to adopt for an IT or IoT project. To overcome this challenge, we aim in this work to construct a dataset that contains numerous project descriptions and the method which has been chosen and which has proven its performance for its management.

We compare according to the state-of-the-art [9, 10] the most agile method depending on several criteria viz; project size, team size, iteration length, Roles and responsibilities, daily meetings, risk mitigation, virtual team support, documentation, information sharing via document, Meeting face-to-face, and model-based. As well as the most IoT management method viz; methodologist, organization, approach (iterative or no), team size, roles and responsibilities, artifacts, meeting, rhythm, change philosophy, top priority, hardware, and software. The agile/IoT methods selected through the study that are based on the state of the art are as follows: DSDM (Dynamic System Development Method), Kanban, FDD (Feature Driven Development), Scrum, Crystal Family, LSD (Lean Software Development), XP (Extreme Programming), ASD (Adaptive Software Development), Ignite, IoT Methodology and SAFe (Scaled Agile Framework).

Based on the analysis of the obtained results, we identified the most important criteria that meet the professional requirements and preferences allowing the selection of the appropriate method. Table 2 and 3 show the valid criteria such as Large/small project (LP/SP), Roles and responsibilities (R&R), Process/People-centric (Proc/PPL centric), High/Medium risk mitigation (High/Med RM), daily meetings (Dly meet), iterative, Small team, Large team, and Hardware.

Table 2. Management method selection for IT projects.

	LP	SP	R& R	Proc centric	PPL centric	High RM	Med RM	Dly meet
DSDM	X		X	X		X		
Kanban	X		X	X			X	X
FDD	X		X	X			X	
Scrum	X		X		X			
Crystal Family	X				X			
LSD	X			X				
XP		X	X		X		X	
Scrum		X	X		X	X		
DSDM		X	X	X		X		
Kanban		X	X	X			X	
LSD		X		X				
Crystal Family		X			X			X
ASD		X			X			X

As shown in Table 2, the main criteria for choosing the FDD method for instance are project size is large, definition of roles and responsibilities, process centric and based on Medium risk mitigation. There are methods that are valid for both large and small projects at the same time, using different criteria. In the case of the Scrum method, for example, in order to choose it for large projects, it is necessary to validate two other criteria: Roles and responsibilities, and people-centric. if you want to choose it for small projects, it is necessary to validate the same two criteria adding the Risk mitigation criterion which is high.

As shown in Table 3, Ignite is the only method that can choose it if we do not want an iterative process contrary to others. So to choose other method it is necessary to validate other criteria than the first. The main criteria for choosing the IoT methodology for instance it is necessary to validate three criteria: large team, Roles and responsibilities, and hardware.

Table 3. Management method selection for IoT projects.

	Iterative	Small team	Large team	R&R	Hardware
Ignite					
Scrum	X	X		X	
Kanban	X	X			
Kanban	X		X		
IoT Methodology	X		X	X	X
SAFe	X		X	X	

The results obtained offer an idea on selecting the appropriate method for a project that can be extracted from the choice rules. In the next section, we develop a process flow with the aim of collecting the companies' experience of the choice of the method to manage a project.

3. Dataset construction approach

To construct our dataset based on companies experiencing, we followed a set of steps as explained in Figure 1.

Table 4. Specification table of data.

Specification	Details
Subject area	Project management method
More specific Subject area	IT and IoT Project management method
Type of data	Table
How data was acquired	Field survey
Data format	Raw, filtered, and analyzed data

The data construction is based on a questionnaire specific to the software engineering and IoT domain. This questionnaire contains a set of questions containing criteria extracted from our state-of-the-art studies [9, 10]. Table 4 presents the specifications of the constructed data.

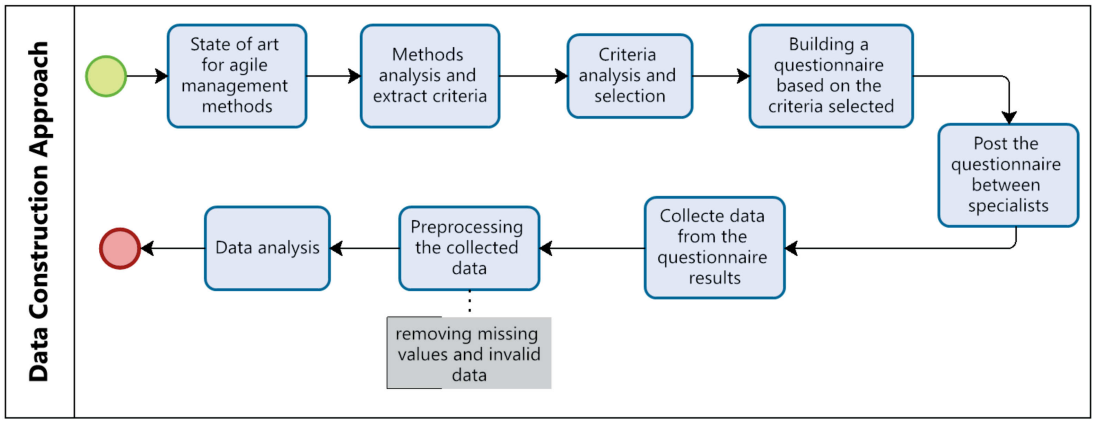

Fig. 1. The approach followed to construct dataset.

The people were asked about their experiences in project management, specifically, which method they used to manage their projects and what were the basic criteria for making this choice.

We first conducted state-of-the-art studies for agile management methods for IT and IoT projects. This work allowed us to understand the IT and IoT project management concepts and to know the criteria of each method. Then, we analyzed these criteria one by one to select the most relevant criteria for the choice of the method. Then, these criteria were used in the construction of a questionnaire intended for specialists in the field. The purpose was to see the companies' reality in the use of project management methods as well as the validity of these theoretically extracted criteria and to discover additional criteria if available. Finally, after gathering the data, we started by filtering the missing and invalid data. Then, we realize an exploratory data analysis to explore and examine this data. Next, we propose labeling scenario to prepare the dataset for the first prediction models.

Our current dataset contains the characteristics of more than 112 project members from different countries. These characteristics contain member identity (country, company name, company position), project data (project name, project description), and project criteria (project size, roles and responsibility, process-centric, people-centric, high-risk mitigation, medium risk mitigation, daily meetings).

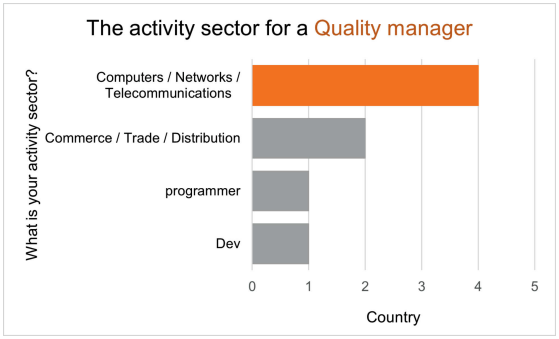

Fig. 2. Quality manager's activity sector.

The interviewed experts belonged to different activity sectors (i.e. Commerce/Trade/Distribution, Building materials, Computers/Networks/Telecommunications, etc.) from different countries. The majority of the respondents were from the Computers/Networks/Telecommunications sector (68%) followed by Commerce/Trade/Distribution (5%). Furthermore, in the case of the Computers/Networks/Telecommunications activity sector, the IT department position that most often appear from the respondents are Service Manager (13%) and Company manager (12%). Figure 2 illustrates the activity sectors of the responding managers.

The most activity sector responding belongs to the Quality manager in this questionnaire is Computers/Networks/Telecommunications by 4 countries out of 8 (50%).

4. Dataset construction results

The data based on companies experiencing remains under construction. Until now, the data contains in total of 101, 92 IT Projects and 9 IoT Projects. Table 5 shows the number of respondents for each country.

The country with the highest response rate to the survey to determine the criteria used for the decision on the appropriate method to manage a project is Morocco (59) followed by India (22). Figure 3 illustrates the percentage of response by country.

Table 5. Number of respondents per country.

Country	number of respondents
Morocco	59
India	22
France	4
US	3
China	3
Burkina Faso	3
France	2
Iran	2
Indonesia	2
Philippines	1
Uzbekistan	1
South Korea	1
Algeria	1
Canada	1
Chile	1
Saudi Arabia	1
IRAQ	1
Congo	1
Japan	1
India	1
Mauritanie	1
Grand Total	112

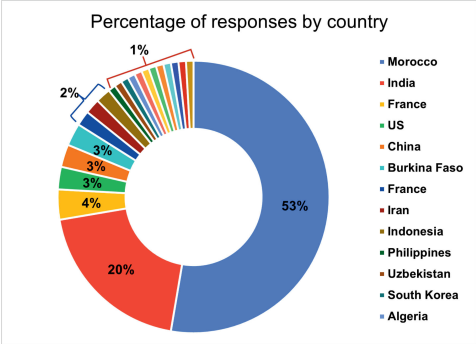

Fig. 3. Percentage of responses to the questionnaire by country.

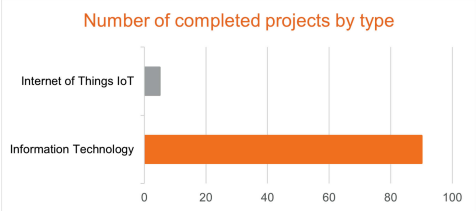

Fig. 4. The number of projects completed by type and size of project.

After filtering data obtained, we find 95 projects 90 for IT projects and 5 for IoT projects. The most of the projects achieved are of Information Technology type (as shows in Figure 4).

Table 6 presents the number of IoT and IT projects according to the project size.

Table 7 shows the use of centric criterion by managers according to project size. The most used is process-centric.

Table 6. Number of IT or IoT projects according to project size.

	Project size		Grand Total
	Large	Small	
Information Technology	63	27	90
Internet of Things IoT	1	4	5
Grand Total	64	31	95

Table 7. Number of IT or IoT completed projects by 'Project size' and 'Centric'.

	Centric		Grand Total
	People	Process	
Large	30	34	64
Small	15	16	31
Grand Total	45	50	95

Table 8 provides an example of the most appropriate method to apply to a project, based on two criteria: project size and process-centric. In this case, the most used methods are Scrum and SAFe.

Table 8. Number of completed projects with Large size and process-centric.

Method used	Count of Projects
Scrum	13
Scaled Agile Framework (SAFe)	12
Dynamic Systems Development Method	3
Lean software development	2
Extreme Programming	2
Kanban	1
Kind of mixture management tools/approaches between Feature Driven Development, Agile, and some others related to security.	1
Grand Total	34

Figure 5 clearly shows the result of Table 8.

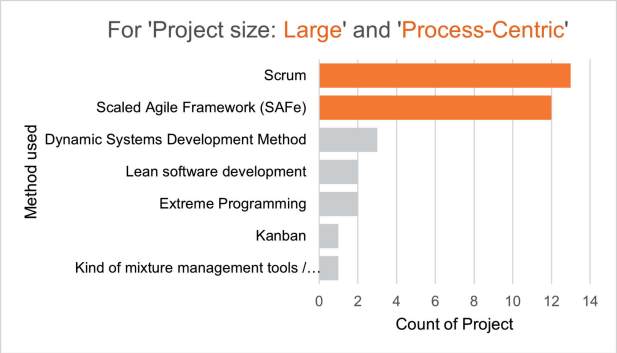

Fig. 5. The appropriate method when the project size is large and process-centric.

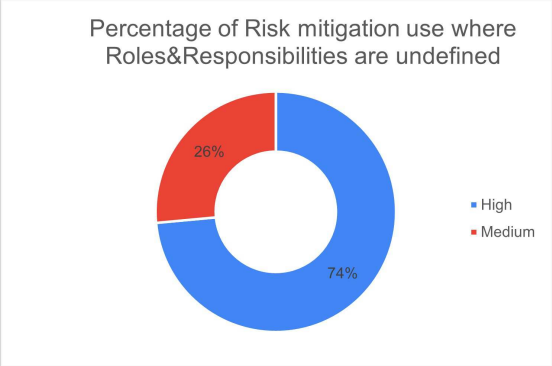

Fig. 6. The appropriate method when the project size is large and process-centric.

Table 9. Number of completed projects with Undifined Roles & Responsibilitie and Medium Risk Mitigation.

Method used	Count of Projects
Extreme Programming	4
Scrum	3
Dynamic Systems Development Method	1
Lean software development	1
Grand Total	9

Table 9 provides another example of the most appropriate method to apply to a project, based on two criteria: Roles & Responsabilities and Risk Mitigation. In this case, the most used methods where Roles & Responsibilities is undefined with Medium Risk Mitigation are Extreme Programming and Scrum.

When the roles and responsibilities criterion are undefined, the most projects are based on high risk mitigation as shown in Figure 6.

In the case of IoT projects the most used method with People-centric criterion is IoT Methodology as shown in Figure 7.

Figure 8 presents the used criteria to choose a management methodology when the risk mitigation is medium. There are two most used cases. In the first case only the definition of roles & responsibilities is used, while in the second case there is a definition of roles & responsibilities with daily meetings as well as the project's size.

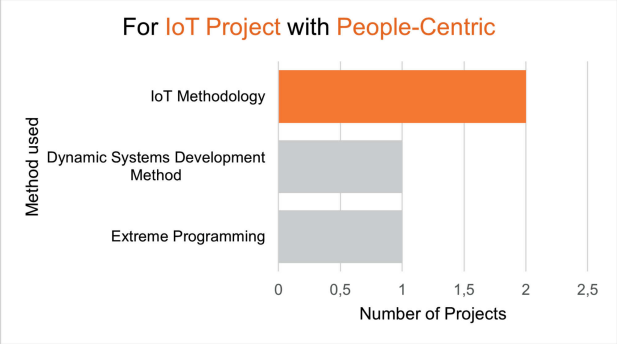

Fig. 7. The appropriate method for IoT Project focusing on people-centric.

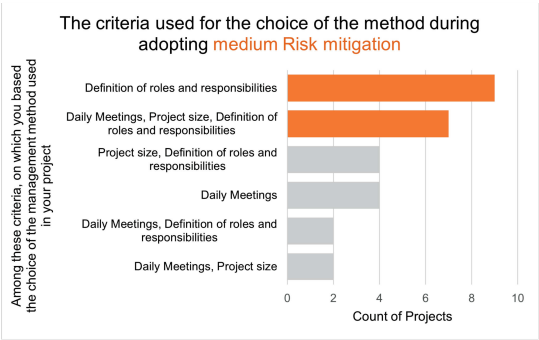

Fig. 8. Criteria used to choose a management method when adopting medium-risk mitigation.

Figure 9 shows the management methods used in IT and IoT projects. The most used method is Scrum.

We share the most used methods as shown in Figure 9 (Scrum 55.79%, SAFe 14.74%, DSDM 9.47%, etc.), the data contains variant methods applied in IT or IoT projects according to a set of criteria. The main objective of the data construction is to develop a model based on business reality. This model will help to choose an appropriate method to adopt for a project. The current data is unbalanced as seen in Figure 9, from this perspective, in our first experiment, we used the statistics of the obtained

results to reduce the number of classes to 4 classes. Figure 10 shows the methods used to manage the project after reducing the number of classes to four.

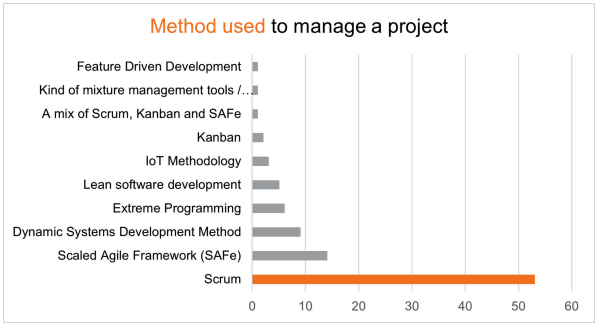

Fig. 9. Project management method used on real projects.

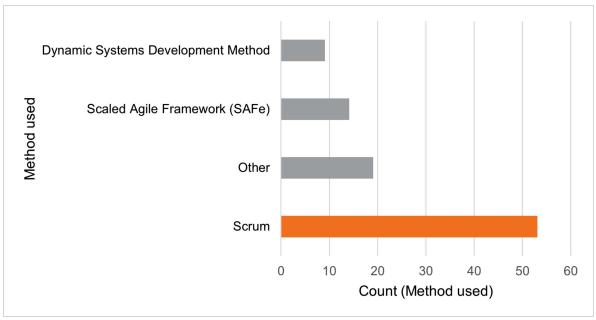

Fig. 10. The four classes of Project management methods used.

This dataset allows us to discover the most management methods used in reality for each project in companies. The result is compared to the theoretical studies, which shows that the most used methods are those of Figure 10, up from this conclusion, we decide in the next section to develop a predictive model of the best management method according to a set of predefined criteria based on a Machine Learning (ML) approach.

5. Association rules mining for project management selection

5.1. Association rules mining background

Machine learning (ML) is a branch of artificial intelligence that focuses on creating algorithms that depend on a large collection of data in order to be effective. It is used to educate machines how to better process data and used by many sectors to retrieve pertinent data [15–17]. Learning can be supervised, semi-supervised, unsupervised, or by reinforcement.

In this section, we aim to apply one of the known unsupervised machine learning techniques which is the association rules with the purpose of exploring the potential of the dataset created in the generation of association rules. This later is typically thought of as a two-step process that begins with finding all of the frequent item sets and ends with generating strong association rules from the frequent item sets. These rules must, by definition, fulfill a minimal support and a minimal confidence. This later is based on application of FP-Growth algorithm. In a set of element $E = \{e_1, e_2, \ldots, e_n\}$ as well as transactions $T = \{t_1, t_2, \ldots, t_n\}$. Every transaction T includes a subset of the elements in E. The association rule is the implication $A \implies B$, where $A \subseteq E$, $B \subseteq E$ and $A \cap B = \varnothing$ [18, 19].

Jiawei Han, Jian Pei, and Yiwen Yin propose FP-growth. Meanwhile, they are the distinguishing traits of Frequent Pattern Growth Algorithm known by FP-growth. First, an extended prefix tree structure known as a frequent pattern tree, or FP-tree for short, is utilized to compress the necessary information in the database. Second, FP-growth, an FP-tree-based pattern fragmentation growth mining approach, is created. Starting with a common length-1 pattern, FP-growth analyzes just its conditional pattern base, builds its conditional FP-tree, and mines recursively on it [19].

5.2. Association rules mining process

Once the dataset has been created, the task of exploration begins. Using the Rapidminer software tool, to create and execute the process of association rule creation for management method selection.

Figure 11 shows the process followed for the creation of association rules. Starting by applying a transformation of the information of the dataset to the binominal type so as to apply the FP-Growth algorithm which is used to represent the dataset in the form of an FP-tree. Closing with the creation of the association rules.

Figure 12 shows the process followed for the creation of association rules in Rapidminer software. The Nominal to Binominal operator is used for the transformation of dataset data in the first step

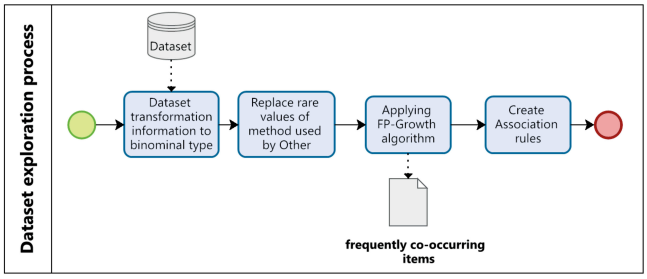

Fig. 11. The proccess followed to explore the created dataset.

using all in the attribute filter type and the binomial transform parameter. Next by using the previous step results in the application of the FP-Growth algorithm with support in min requirement which is equal to 0.95. Finally, the frequency calculation result of the algorithm will be the input of the association rules creation operator which generates these rules according to the confidence criterion with min confidence is 0.8.

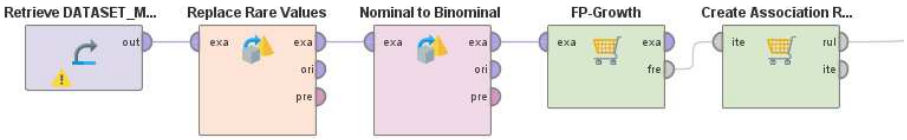

Fig. 12. The Process of generating association rules using the FP-Growth algorithm.

5.3. Association rules mining results

This sub-section will present the results of applying the process of creating association rules for the selection of the appropriate method to manage a project. Nominal to Binominal Model covering 7 attributes: Project Type, Project size, Roles & Responsibilities, Centric, Risk Mitigation, Daily Meetings, and Method used.

[Project size = Large, Centric = People] –> [Method used = Scrum] (confidence: 0.800, support 0.340)

[Project size = Large, Centric = People] –> [Project Type = Information Technology, Method used = Scrum] (confidence: 0.800, support 0.255)

[Project Type = Information Technology, Method used = Scaled Agile Framework (SAFe)] –> [Project size = Large, Risk Mitigation = High] (confidence: 0.923, support 0.211)

[Daily Meetings = D, Method used = Scrum] –> [Project Type = Information Technology] (confidence: 1.000, support 0.415).

6. Conclusion

Choosing a good method to manage a project is a major challenge because it affects the outcome of the project. For this reason, we decided to propose a predictive model that allows this choice by following a set of criteria specific to the project and the company based on the reality of the company. For this purpose, we built our own dataset because of the lack of datasets concerning this problem. The data is still being improved. To sum up, this paper presents the steps of construction of the dataset and the analysis of the results in order to use it in the construction of a predictive model.

[1] Patil M., Suresh M. Modelling the enablers of workforce agility in IoT projects: a TISM approach. Global Journal of Flexible Systems Management. **20** (2), 157–175 (2019).

[2] Hurtoi V., Avadanei D. IoT Project Management. Informatica Economică. **24** (3), 75–80 (2020).

[3] Budiman R. Utilizing Skype for Providing Learning Support for Indonesian Distance Learning Students: A Lesson Learnt. Procedia – Social and Behavioral Sciences. **83**, 5–10 (2013).

[4] Schwaber K., Sutherland J. The Scrum Guide The Definitive Guide to Scrum: The Rules of the Game. `https://scrumguides.org/scrum-guide.html`.

[5] Dudziak T. Extreme programming an overview. Methoden Werkzeuge Softwareproduktion WS. 1–28 (1999/2000).

[6] What is SAFe®? Scaled Agile', SAFe® Enterprise Solutions. `https://www.scaledagile.com/enterprise-solutions/what-is-safe/`.

[7] Slama D., Puhlmann F., Morrish J., Bhatnagar R. M. Eds., Enterprise IoT: strategies and best practices for connected products and services. Beijing Boston Farnham Sebastopol Tokyo: O'Reilly (2016).

[8] IoT Methodology – The Internet of Things project lifecycle guide for creative, technical and business people. http://www.iotmethodology.com/.

[9] Merzouk S., Cherkaouii A., Marzak A., Nawal S. IoT methodologies: comparative study. Procedia Computer Science. **175**, 585–590 (2020).

[10] Merzouk S., Elhadi S., Ennaji H., Marzak A., Nawal S. A Comparative Study of Agile Methods: Towards a New Model-based Method. International Journal of Web Applications. **9** (4), 121–128 (2017).

[11] Manifesto for Agile Software Development. https://agilemanifesto.org/. Manifesto for Agile Software Development (2001).

[12] Dynamic Systems Development Method (1998). https://docplayer.fr/5488317-Dynamic-systems-development-method.html.

[13] Pouyandeh F., Golabchi M., Taghizadeh K. Providing a model to choose the most appropriate agile method in construction projects. Proceedings of the Institution of Civil Engineers – Management, Procurement and Law. **176** (1), 14–27 (2022).

[14] Software Development Methods. https://www.kaggle.com/datasets/mostafizmim/software-development-methods.

[15] Burkov A. The Hundred-Page Machine Learning Book. Polen (2019).

[16] Edgar T. W., Manz D. O. Chapter 6 – Machine Learning. Research Methods for Cyber Security. 153–173 (2017).

[17] Mahesh B. Machine learning algorithms – a review. International Journal of Science and Research. **9** (1), 381–386 (2020).

[18] Sulova S. Association rule mining for improvement of IT project management. TEM Journal. **7** (4), 717–722 (2018).

[19] Barik S., Mishra D., Mishra S., Satapathy S. K., Rath A. K., Acharya M. Pattern discovery using fuzzy FP-growth algorithm from gene expression data. International Journal of Advanced Computer Science and Applications. **1** (5), (2010).

До нових даних для прогнозування методів управління проектами IT та Інтернету речей

Мерзук С.[1], Гандул Р.[2], Марзак А.[1], Саель Н.[1]

[1] *Лабораторія інформаційних технологій та моделювання, Факультет наук Бен М'Сік, Університет Хасана II Касабланки, Касабланка, Марокко*

[2] *Лабораторія наук і технологій інформації та освіти, Факультет наук Бен М'Сік, Університет Хасана II Касабланки, Касабланка, Марокко*

Вибір найкращого методу керування проєктами на робочому місці допомагає надати замовнику високоякісний продукт. Звідси виникає необхідність добре знати методи керування, їх характеристики, переваги і недоліки, щоб мати можливість вибрати найкраще для конкретного проекту. Однак досі не було запропоновано жодного великого набору даних для машинного навчання та моделі прийняття рішень, моделі чи системи, яка б допомогла керівникам проєктів знайти найбільш ефективний метод, адаптований до обмежень їхніх проєктів. Ця стаття розробляє побудову набору даних для гнучкого методу керування проєктами та IoT на основі реального досвіду. Мета цієї статті полягає в тому, щоб запропонувати критеріальну модель, яка дозволяє вибрати найкращий метод керування для такого IT-проєкту або IoT-проекту відповідно до набору критеріїв.

Ключові слова: *гнучкий метод; інформаційні технології (IT); інтернет речей (IoT); керування проєктами; метод динамічного розвитку системи (DSDM).*