Iranian Journal of Economic Studies, 8(2) 2019, 375-396

Shiraz University

Iranian Journal of Economic Studies

Journal homepage: ijes.shirazu.ac.ir

Identification and Prioritization of Public-Private Partnership Indicators in Iran's Water and Wastewater Industry via Data Mining Algorithms

Malihe Eskandary*, Mohammad Taghi Taghavifard, Iman Raeesi Vanani, Soroush Ghazi Noori,

College of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran.

Article History

Received date: 27 November 2019
Revised date: 08 June 2020
Accepted date: 09 June 2020
Available online: 24 June 2020

JEL Classification:
L32
E22
R42

Keywords:
Public-Private Partnerships
Investment
Key Performance Indicator
Water and Wastewater Industry

Abstract

The restrictions of government resources and the recent alterations in the economy have prompted government agencies to employ the capacities of private sector in all infrastructures. In this regard, a variety of financing methods, including the participatory models, have been applied for many years in the water and wastewater industry of Iran. The aim of this study is to identify and prioritize the Public-Private Partnership (PPP) indicators in the water and wastewater industry of Iran via machine learning techniques. To this end, after collecting, preparing and preprocessing the data, weighted indexing techniques including information gain and Gini index were utilized to prioritize the PPP factors. The results indicated that 93% of the indicators were effective in predicting the success of the projects. To compare the two methods, the precision of Naïve Bayes and Random Forest classifiers were taken into account and the information gain method yielded more reasonable findings with one percent difference. The evaluation of indicators elucidated that "complaints about service quality," "contract type," and "Conventional tariffs" revealed a huge impact on the success of collaborative projects. Among the 15 indicators, eight were directly pertinent to the project financing which is the main concern in this industry.

1. Introduction

One of the pillars of each country's development is access to safe drinking water and sanitary sewer (Company, 2009). Given the environmental priorities and the inherent value of wastewater in the context of promoting the health of the community, it is essential to implement these projects at high priority and in the shortest possible time. According to World Bank reports, sewage treatment's coverage in developing countries is less than 40% (Dorian, 2006). The shortage of funds is one of the issues that undermine the completion of sewage projects. According to World Bank, World Health Organization and UNICEF documents,

* eskandary@nww.ir
DOI: 10.22099/ijes.2020.35590.1625

about $300 billion are required to provide sanitary sewer in the Asian countries (Hall, & Lobina, 2005) and due to the dependence of these projects on the governmental resources, this problem might not be solved easily. In fact, the high volume of financial resources required for achieving the scheduled development goals and the limitation of government funding over the past few years, make it essential to use modern methods for funding, and to promote partnership and investment of the nongovernmental sector for financing the water and wastewater projects, thereby enabling the public sector to benefit from the economic, managerial and executive ability of non-governmental sector (Company, 2009). Figure 1 shows the private-sector investment in the water and wastewater industry of developing economies. Figure1 shows the investment commitments in the infrastructure projects with private participation in Emerging and Developing Economies. Examples of emerging economies include many countries in Africa, most countries in Eastern Europe, some countries of Latin America, some countries in the Middle East, Russia and some countries in Southeast Asia.

According to Public- Private- Investment (PPI) database in World Bank, only about 10 percent of these projects are involved in the water and wastewater section but the trend of investment in this industry is the same as the other infrastructures. Higher investment levels are typically recorded in the latter half of each year and PPI investment seems to be showing signs of recovery since 2017.

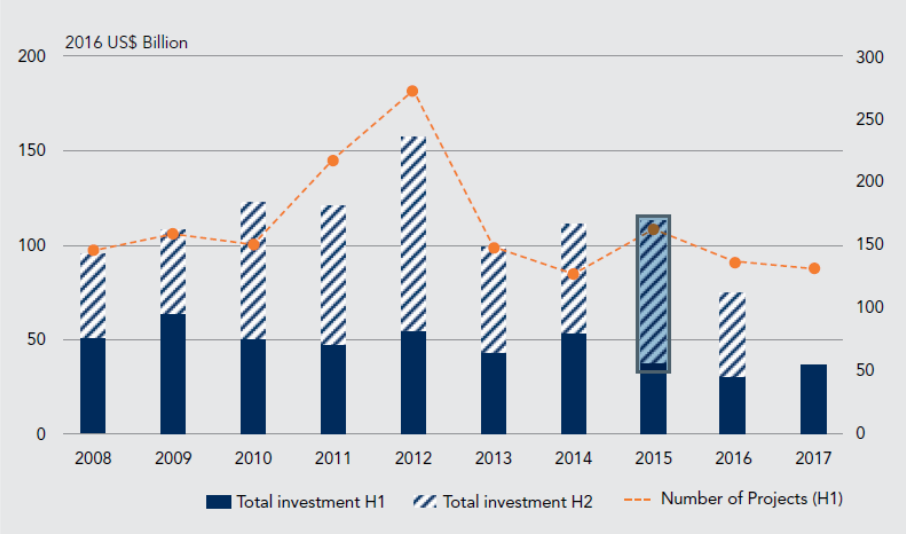

Figure 1. Investment commitments in infrastructure projects with private participation in EMDE[1]s, 2008 H1-2017 H1 (PPI database, World Bank report, 2017)

[1]- Emerging and Developing Economies

Iran is located in an arid area, and therefore, safe drinking and sanitary water are scarce. As a result, it is essential to pay more attention to the implementation of the water and wastewater projects due to the limitation of water resources, especially regarding the recent year's droughts which led to the decline of groundwater aquifers and the spread of deserts in Iran (NWWEC website, 2018). According to their main mandate, the water and wastewater companies in Iran are responsible for the supply of safe water and sanitary collection of the wastewater. However, they have always encountered problems for financing the projects. In fact, the high cost of the water and wastewater projects, as presented in Table 1,(Company, 2009), limited financial, executive and specialized resources in water and wastewater companies and the other constraints, have prolonged the time required for the implementation of such projects by the public sector, thereby imposing additional costs on the plans.

Table 1 Projections of rural and urban requirements by 2025 (Company, 2009)

Type of cost	Urban sector cost (billion dollar)	Rural sector cost (billion dollar)
Operation Cost	40688	3430
Reinvestment Cost	42861	3793
Service Cost	122947	18171
Total	206495	25394

Most countries have fully accepted the principle that the public sector should refrain from entering into executive activities and focus more on governance and protecting the rights of the people (Ibid).

In recent years, the water and wastewater industry of Iran has been successful in the field of private sector partnership and the use of modern financial resources. One of the issues in this industry is to identify and prioritize the PPP indicators so they can improve the most effective indicators for achieving the better results. In this research, after extracting all PPP indicators from the literature review and having the comments of experts from all areas related to this field, 33 indicators were extracted. In the data collection phase, a total of 176 public-private partnership contracts of the water and wastewater industry of Iran (between the years 2009 to 2019) were utilized. Next, in order to prioritize the indices, the weight indexation methods of data mining were applied to investigate the influential factors. The CRISP method was also used for data mining process.

2. Literature Review

Considering the main purpose of the research which is identifying and prioritizing Public-Private Partnership indicators in the water and wastewater industry of Iran via data mining algorithms, this study reviewed prior investigations in two areas.

2.1 Public-Private Partnership and Its Different Modes

There are different definitions for PPPs across the world. International Monetary Fund defined this item as the arrangements, where the private sector is responsible for supplying infrastructure assets and services traditionally provided by the governmental sector (Brune, et al. 2004). According to Asian Development Bank and the World Bank, PPPs refer to the arrangements between government and private sector in the field of infrastructures and the other services (Commission, 2003). The Organization for Economic Cooperation and Development defines PPPs as a long-term contractual agreement between the government and a private partner, where the private sector delivers and funds public services through a capital asset, sharing the associated risks (Golabchi & Nourzaye, 2015). According to the European Union, PPPs are the contracts concluded between two or more entities allowing collaboration to achieve collective and compatible goals. In these contracts, there are always parts of common authority and responsibility, joint resource investment, joint risk-taking and mutual benefits (Commission, 2003).

According to various definitions, different countries based on their experiences and backgrounds of using the private sector in the infrastructure projects, have utilized different names for the PPP spectrum.

In the view of studies, public-private partnerships projects can be classified into three generations (Tang et al., 2010)

- The first Generation of Partnership: The projects carried out in this period often encounter many problems due to the low experience of both sectors and consultants.
- The second Generation of Partnership: In developed countries, companies often define specialized projects in the urban area. To implement the projects, they employ specialists in the field of planning to manage the public-private interactions.
- The third Generation of Partnership: Given the social capacities created by the implementation of the second generation of projects, in the third, developers propose the project and in order to implement it, they try to find the private sector to benefit from its capacities. Accordingly, there are various public-private projects that are implemented in different fields.

In the early 1990s, the developing countries applied the public-private partnerships increasingly, especially for the infrastructure projects, while the number of such contracts fell sharply after the economic crisis in East Asia (Zhang, 2005). Some years later, the number of contracts increased with greater trust in the private sector. This approach was more welcome by the energy sector, from 1990 to 2011. This sector accounted for 41.3% of the total projects (Hodge, et al. 2010).

According to surveys conducted on the water and wastewater industry of Iran, the categorization of the financing modes depends on the source and method of supply. Based on the source of supply, the financing modes are divided into two types of domestic (i.e. the origin of money is from the country) and international (i.e. the source of money is from abroad) (Savas, 2000). In the water and wastewater industry of Iran, the financing modes can be divided into two sections according to the method of supply:

1. Mobilizing financial resources through loan procedures, including:
   - Using the facilities of development banks (World Bank, Islamic Development Bank and ECO Trade and Development Bank).
   - Types of financing procedures (international finance, domestic finance and current finance)
   - Issue and sale of corporate bond

2. Development of private sector partnerships, including:
   - Contract arrangements in the form of Build- Operate- Own (BOO) and Build- Operate- Transfer (BOT)
   - Buyback[1]

The main element to separate the two methods is contingent upon the acceptance of various investment risks such as technical, production, marketing, exploitation, contingencies, management, technical knowledge risks, and the risk of equity and income. In the borrowing method, the foreign country will supply the capital without accepting any risk and the government guarantees everything. All risks are borne by the country, which receives the capital. In the methods of partnership development, the risks are largely divided and the government's guarantees are provided to ensure a secure environment for the foreign capital investment (Wei, 2011 & Miller, 2000).

2.2 Key Performance Indicators of Public-Private Partnerships

Developed and developing countries have used various methods of public-private partnerships in different projects. In some cases, these methods have been successful and in the others, the projects have failed. One such unsuccessful project was the Malaysian sewerage project in 2001 (Abdul-Aziz, 2001). These projects are often confronted with many problems due to the long duration of the contracts, the low experience in this area and the extremely high risk factors (Zhang, 2005). Given the high willingness of the international community to implement such projects through public-private partnerships, the identification of the key performance indicators is a crucial issue for the researchers since the measurement and management of these indicators can be the best way to assess the success rate of the projects, and hence to select the appropriate approach to future projects (Yu, 2007).

[1]. Buyback projects in the wastewater sector, established in Iran for the first time in the world

In the early years of public-private partnerships, the most important indices of the success for such projects were the completion of the project within a specified time and the accomplishment of maximum results (Ibid). However, given the subsequent experiences, especially in the field of urban infrastructure development, researchers shifted their attention to the choice of the optimal method for the successful implementation of the projects. Several indicators were analyzed to evaluate the success rate of project implementation in different ways. Given the increasing trend of public-private partnerships since 2000, it was taken as the base for the studies which tried to identify the effective key indicators in the public-private partnership. All the studies conducted on this subject were considered in this research and the relevant indicators were extracted.

Among the studies undertaken to identify the partnership indices, Tiong et al. (2001) studied the Critical Success Factors (CSFs) of Build-Operate-Transfer (BOT) projects and found 6 key performance indicators of innovation and entrepreneurship, selecting the right project, forming a strong combination of stakeholders, appropriate technical solutions, strong financial proposition, and specific bidding conditions as critical success factors (Qiao, 2001).

In the same manner, Akintoye (2003) identified the most significant indicators of the optimal PPP in project implementation as improving quality, increasing efficiency and effectiveness, optimized using of financial resources, and benefiting from better performance standards (Akintoye, 2003). In another study by Zhang (2005), after reviewing and aggregating all the studies undertaken, he identified the main CSFs in the public-private partnerships in the infrastructure projects as a conducive environment for investment, viability, risk sharing and the existence of reliable contracts, the involvement of influential and high-tech investors, and the provision of appropriate financial packages. Along the same line, by reviewing the partnership projects conducted in 2005, Li et al. identified important CSFs in implementing the public-private partnership projects as a strong and robust mix in the private sector, realistic cost-benefit assessment, availability of financial markets, sharing appropriate risk, technical feasibility of the project, political support, competitive procurement process, transparency in the procurement process, multidimensional goals and benefits, public and private sector commitment and responsibility, good governance, government participation in providing guarantees, macroeconomics, the existence of an appropriate legal framework, political and economic condition, the existence of appropriate private skills, governance split between the public and private sectors, social support, and technology transfer (Li, 2005). Likewise, Böhl (2007) conducted a research on the public-private partnerships in the water and wastewater industry to identify these indicators in Germany. Identifying the goals of the public-private partnerships in the industry, the group identified 36 indicators as the key performance indicators of the PPP projects' performance and success and weighed the indicators through reviewing projects in some countries.

In 2007, Sash et al. explored the opportunities and impacts of political risks on the public-private partnership projects in the field of infrastructure in China and 14 other Asian countries. Many of these projects were in the field of water and wastewater, particularly, the construction of public-private water and wastewater treatment plants during which the impact of six factors and the political risk of exchange non-ease and restrictions, expropriation and government monopolies, lack of commitments, threats and political problems, legal and bureaucratic risks and risks unaffected by the government were scrutinized. This study then examined the status of each country in the six areas mentioned and proposed appropriate solutions in each country (Sachs et al., 2007). In the same year, Salman and his colleagues conducted surveys on the public-private partnership projects, and in particular BOT projects to complete the key infrastructure; they extracted and categorized the most significant indicators of the effective decision-making and success in different areas. The group put forward 21 indicators in three categories: environmental and legal factors, technical factors, and financial and business factors (Salman et al., 2007). In 2009, Ian and colleagues also considered the selection of a public-private partnership model from the perspective of all project stakeholders. The stakeholders' survey included the main sectors of the society, the private sector, the public and governmental sector, and the academic community. The research identified 15 indicators as the main performance indicators in these projects and identified the importance of these indicators in each of the stakeholder groups. He then uses the fuzzy methods to weigh the mentioned indices and determine their significance in each of the stakeholder groups (Yuan et al., 2010). In 2012, Yuan et al. (2012), by reviewing the public-private partnership projects, divided the key performance indicators into five main areas, using a questionnaire, submitting it to stakeholders involved in the public-private partnerships, and examining thereof; they identified a total of 48 indicators as the key performance indicators in the collaborative projects; after identifying them, the Likert-scale method was used to determine the importance of each indicator.

In 2015, Golabchi and Noorzaei conducted a research on selecting the appropriate method of public-private partnership to finance Iranian freeway projects and after presenting two questionnaires to the experts and analyzing the results using the t-test method, they presented 19 important indicators. Similarly, in 2015, the World Bank stated in its report on creating a framework for public-private partnerships, the main constraints on the development of partnership projects to include the statutory restrictions, and the key factors in the form of 15 items (Delmon, 2015). Moreover, Marzouk and Fayez (2018) identified 46 indicators in 6 areas through a public-private partnership questionnaire.

Although much research has been done on the PPP contracts in the recent years, studies on this domain have shown that identifying and evaluating the effective criteria for selecting the type of contract used to finance the water and wastewater projects in the world, in particular in Iran, have received less attention.

2.3 Data Mining Algorithms

Data mining combines the techniques like machine learning, pattern recognition, statistics and databases for summarizing and linking the concepts in order to extract the interesting patterns automatically from the corporate databases. In 1996, Osama-Fayad identified the data mining as a concept of the discovery of knowledge from the database. To extract knowledge from the data, data mining consists of 5 steps according to Figure 2:

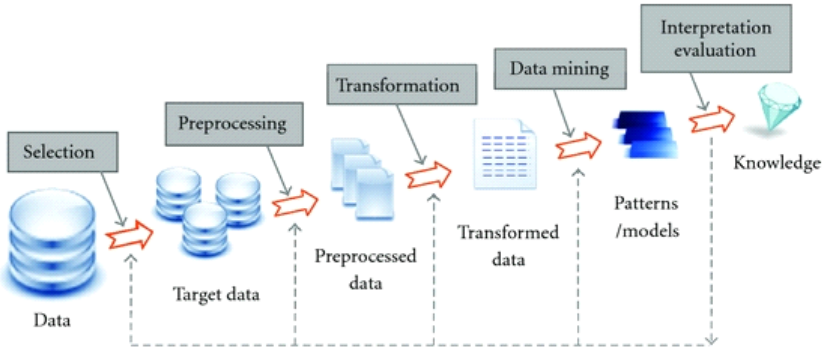

Figure 2. The data mining or knowledge discovery process (Gullo, 2015)

- Data selection: In this step, based on the type of goals, data are selected or created from different sources.
- Data preprocess: At this point, invalid data are separated from the educational set. Noise and incomplete data are examples of the data that need to be cleared. This step can take up to 60% of the time of the data mining.
- Data transformation: This phase is the process of transforming and consolidating the data into different forms that are suitable for mining. Data transformation normally involves normalization, aggregation, generalization, etc.
- Data mining: Data mining is the core process where several complex and intelligent methods are applied to extract some patterns from the data. Data mining process includes several tasks such as association, classification, prediction, clustering, and time series analysis.
- Pattern evaluation: This step identifies the truly interesting patterns representing the knowledge based on different types of interestingness measures. A pattern is considered to be interesting if it is potentially useful, easily understandable by humans, validates some hypotheses

that someone wants to confirm or validate on the new data with some degree of certainty.

In the water and wastewater industry, the data mining techniques have been used for prediction, including the research by Van et al (2013), which predicted the future consumption of the water via clustering methods (Wen et al., 2013). Sun et al. (2014) also analyzed the fractures and events occurring in the water networks through spatial clustering, and then, by density-based clustering, they identified the areas of the water supply network that are exposed to the events. Yan et al. (2018) designed and implemented a data mining system that optimizes the design and construction of the sewage treatment plants in China. Moreover, the water use pattern in all areas in 2016 was identified with the help of the data mining and K-Medoids algorithms (Dehghani et al., 2016).

According to prior research, studies more likely predicted consumption, and there was no evidence of using data mining to prioritize the public-private partnership indicators in the water and wastewater industry.

According to foreign and Iranian studies, the public-private partnership indicators in the water and wastewater industry (except for Bohl's study in 2007), have not been examined. Due to the fundamental differences between the structure of the public-private partnership in different industries and the differences between these indicators in Iran and the other countries, in this research, it was sought to extract and prioritize the public-private partnership indicators in Iran's water and wastewater sector so as to increase the probability of success in participatory projects.

3. Methodology

To identify the key performance indicators in the field of public-private partnership, it is necessary to investigate all aspects of the partnership from all stakeholders' viewpoints. Accordingly, the goals of public-private partnership were first identified and then the available indicators to realize each objective were studied and specified. This approach allowed all indicators to be fully considered from the stakeholders' viewpoints. According to World Bank studies (Penelope & Cowen, 1997), the governments seeking private-public partnership in the water and wastewater projects are pursuing one or more of the following objectives:

- Introducing knowledge, technical and managerial experiences or new technology into the water and wastewater sector
- Increasing economic efficiency in the industry in both areas of operational efficiency and using investment development
- Injecting funds and capital into the water and wastewater sector or accessing capital markets
- Reducing the available subsidies in the industry or directing them from the current groups to target ones

- Keeping safe industry from short-term political interventions and influences of high-profile people
- Increasing industry responsiveness to the needs and demands of subscribers

Based on the previous studies conducted on the water and wastewater industry and the approvals of the industry's experts (Tang et al., 2010), seven main objectives were presented for measuring the performance and efficiency of the public-private partnerships in the water and wastewater sector:

Objective 1: Technical experience and knowledge
Objective 2: Operating effectiveness
Objective 3: Customer Responsiveness
Objective 4: Economic Efficiency
Objective 5: Financing and attracting funds from the private sector
Objective 6: Management of knowledge and experience
Objective 7: Safeguard against political interference

The results of the survey on the indicators (described in section 2-2), along with their corresponding targets are presented in Table 2. In the next step, when the indicators were extracted, the experts were interviewed again to determine whether all aspects and indicators were fully considered, and to avoid double counting of indicators.

Experts have been selected so that they can examine all areas related to public-private partnerships. In this regard, the experience and opinions of 8 experts were taken into account with the following combination:

- 2 representatives from the investors of the water and wastewater projects
- 2 representatives among the water and wastewater companies as the investee's
- 2 representatives from the national water and wastewater engineering company
- 1 representative from the consultants
- 1 representative from the management and planning organization

After this step and according to the information available in the water and wastewater industry, the indices that were not relevant to this industry were eliminated and based on the nature of the projects, some indicators were added. The indicators were examined by experts in the water and wastewater industry and, based on the seven objectives, the public-private partnership indicators of this industry were extracted as Table 2.

Table 2. Key performance indicators of public-private partnerships and the measured targets

No.	Indicator	Definition of indicator	measured targets
1	Wastewater treatment	Volume of wastewater treated by modern methods.	Technical experience and knowledge
2	Non-Revenue Water	The difference between the amount of water entering the system and the amount of water that customers receive and pay the bill for (with two aspects of physical loss and government loss).	Technical experience and knowledge
3	The ratio of network accidents to the network length	It describes the extent of the accidents that occurred in the water and wastewater networks due to various causes to create the problems.	Operating effectiveness
4	The ratio of connections accidents to the total connections	It describes the extent of the accidents that occurred in the water and wastewater connections due to various causes to create the problems.	Operating effectiveness
5	Complaints about service quality	The total number of subscribers' complaints about service quality	Operating effectiveness
6	Water quality	To study the quality of supplied water, its turbidity, microbiological and the physical and chemical tests.	Operating effectiveness
7	Optimizing construction time	It investigates how private sector partnership matches the construction time with the scheduled time.	Operating effectiveness
8	Population covered by water services	The percentage of the population covered by the water sector services	Customer Responsiveness
9	Population covered by sewage treatment service	The percentage of the population covered by the sewage treatment services	Customer Responsiveness
10	The ratio of water cut-offs to the number of connections	It describes the water cut-offs and the extent of non-availability of the water services.	Customer Responsiveness
11	Low water pressure	It investigates the time span and the number of customers who face low water pressure and disruptions of services.	Customer Responsiveness
12	Conventional tariff	It investigates the suitability of water tariff.	Economic Efficiency
13	Operating revenue (loss)	It investigates the organization's profitability and its ability to control the operating costs.	Economic Efficiency

Table 2(Continued). Key performance indicators of public-private partnerships and the measured targets

No.	Indicator	Definition of indicator	measured targets
14	Collection period	It illustrates the ability of organization to collect the debt.	Economic Efficiency
15	Working ratio (cost coverage ratio)	It illustrates the organization cost coverage ratio.	Economic Efficiency
16	Construction cost optimization	It investigates how private sector partnership matches the construction costs with the assessed costs.	Economic Efficiency
17	Operating cost optimization	It illustrates how private sector partnership matches the operating costs with the assessed costs.	Economic Efficiency
18	Annual investment to sale ratio	The ratio shows how investments made during the year, constitute a percentage of the adjusted sales (revenues of the organization).	Financing and attracting funds from the private sector
19	Fixed assets turnover ratio	The ratio shows the extent to which the company has been successful in using its fixed capacity.	Financing and attracting funds from the private sector
20	Indebtedness level	It illustrates the debt coverage ratio in the organization.	Financing and attracting funds from the private sector
21	Measurement level	The water level is a tool to measure the volume of the water released from a specific outlet.	Management knowledge and experience
22	Cost of water at break-even point	It shows the final cost of the water at the delivery point.	Management knowledge and experience
23	Cost of disposed wastewater at break-even point	It includes the final cost for disposing and treatment of each unit of sewage.	Management knowledge and experience
24	Workforce productivity (gross)	This indicator shows the workforce productivity.	Management knowledge and experience
25	Workforce effectiveness (gross)	It shows how much wealth is produced by workforce	Management knowledge and experience
26	Per capita workforce cost	It shows that how much the staff costs per personnel	Management knowledge and experience
27	Level of professionalism in the company	To measure the human resource's abilities in an organization in the water and wastewater industry, the levels of staff education and skills are considered jointly and an indicator is presented as the degree of professionalism of the company.	Management knowledge and experience

Table 2(Continued). Key performance indicators of public-private partnerships and the measured targets

No.	Indicator	Definition of indicator	measured targets
28	Absence	It shows the absence of personnel during their business hours.	Management knowledge and experience
29	Education	The training of personnel at different stages of work is assessed by this index.	Management knowledge and experience
30	Political appointment	It shows the impact of political issues on industry's decisions.	To keep safe from political interference
31	Bureaucracy	The level of bureaucracy in the industry is shown by this indicator.	To keep safe from political interference
32	Obtaining a license	It shows the speed and ease of obtaining the necessary license in the industry.	To keep safe from political interference
33	Corruption	The existence and impact of various types of corruption in the industry is shown by this indicator.	To keep safe from political interference

Next, in order to prioritize the indices, weight indexation methods of data mining were used to investigate the influential factors. The CRISP method was also applied for the data mining process.

In the data collection phase, the total of 176 public-private partnership contracts (these contracts were provided by Bureaus for Procurement of Funds & Development of Non- Governmental Participation of the water and wastewater companies all over the country) of the water and wastewater industry in all provinces (between the years 2009 to 2019) was selected. The indicator's data were gathered based on the annual reports of the water and wastewater companies. The completed information for each of the projects of all provinces was sent to the National Water and Wastewater Engineering Company and then double-checked and approved by the experts of the company as the head of the associations of all companies.

The distribution of the projects in the water and wastewater sector based on the contractual templates is shown in Table 3.

Table 3. Distribution of projects in the water and wastewater sector

Contract Field		Contract type	Number of Contracts
Wastewater	Treatment Plant & Pumping Station	BOT	18
		Buyback	31
	Network	Current Finance	73
Water	Treatment Plant & Transmission Line	BOT	6
	Desalination Projects	BOO	48

Figure 3 shows the frequency of each of the contract types of PPP projects in the water and wastewater industry (BOO, BOT, Buyback, Current finance) for different groups of successful projects (over 75% success), semi-successful projects (50-75% success), and unsuccessful projects (success below 50%) (In all the shapes, the blue column is for the successful projects, the red for the failures and the green column for the semi-successful ones). As it is shown in this Figure 3, in the forms of BOT and Buy back contracts, the projects are successful or semi successful and no failure is observed in the projects. Failed projects are related to the Current finance and BOO contracts.

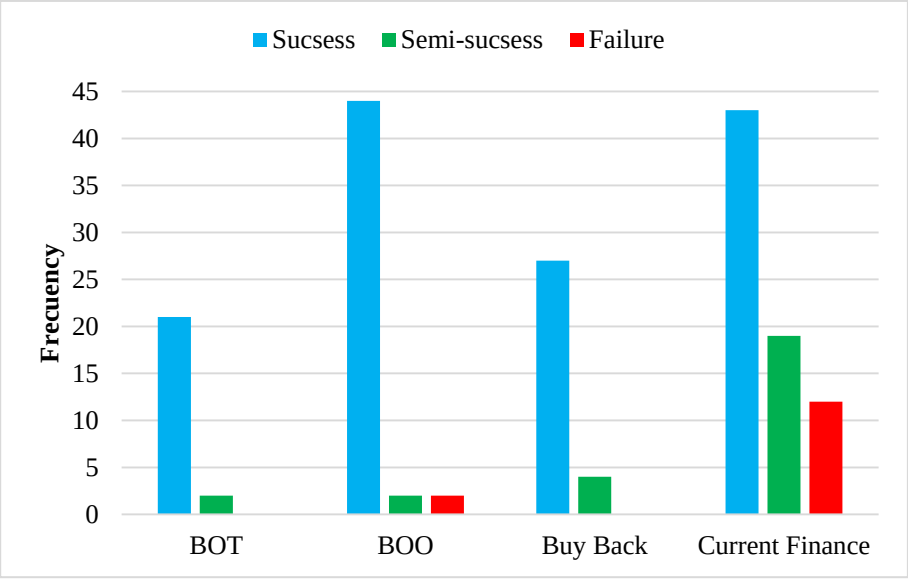

Figure 3. Frequency diagram of contract forms in different groups of successful, semi-successful and Failed Projects (Blue column is for successful, Red for failures and the Green column for semi-successful projects)

According to Table 2, each of the projects includes a data indicator as well as the success rate of the project. As it is not possible to collect data for the indicators related to the goal of "Safeguard against political interference", a total of 29 indicators were identified as public-private participation indices in the water and wastewater projects and their data were collected in order to prioritize the indicators in the next steps.

After collecting the data from the whole country, the data format, the number of variables, the type of variables (continuous and discrete) and the concept of variables were examined and the main data specifications were extracted using the statistical methods. According to the information obtained, if there are lost data, noise and discarded data, the duplicate lines should be deleted and the cleared data should be prepared for modeling.

In the modeling stage, the two techniques of information gain index (Entropy) and Gini index were applied to extract the pattern (influential factors on the data).

3.1 Information Gain Index (Entropy)

The amount of information in a given set of data can be defined as (1 – entropy). If any observation about the given data is made, new information can then be recomputed. The difference between the two information values is the "information gain". (Han, et al. 2011).

The equation for calculating the information gain on the basis of entropy is defined as Equation (1).

$$informationGain(A) = Entropy(D) - Entropy_A^D \quad (1)$$

In this formula "A" is the feature and "D" denotes the set of training data. The entropy formula is defined as Equation (2).

$$Entopy(D) = -\sum_{i=1}^{c} P_i \times lo\,g(pi)$$
$$Entopy_A(D) = \sum_{j=1}^{v} \frac{|Dj|}{|D|} \times Entopy(D_j) \quad (2)$$

In Equation (2), "c" is the number of tags for the classes in the training data; "P_i" is the probability that an instance of the data belongs to the i^{th} class; "v" is the number of the domain members of the attribute A and "D_j" is the part of initial data whose attribute value is v_j. Also "D_j" is the size of D.

3.2 Gini Index

Gini index is a measure of how often a randomly-chosen element from a set of elements would be incorrectly labeled if it was randomly labeled according to the distribution of the labels in the subset. The probability of the correct labeling can be computed by summing the probability of choosing each item multiplied by the probability of correctly labeling it. In this setting, the probability of correctly labeling an item is equal to the probability of choosing that item. Therefore, Gini Index can be computed as in Equation (3) (Han, et al. 2011) (Alizadehsani et al., 2013).

$$Gini(D) = 1 - \sum_{i=1}^{c} p_i^2 \quad (3)$$

In this equation, "c" is the number of classes in the data set and P_i is equal to the probability that a sample of data belongs to the class "i".

4. Results and Discussion

According to the mentioned methods and using the Rapid Miner software version 9, the weights of indices were calculated whose results are shown in Tables 5 and 6. Weights take numbers between 0 and 1. The closer to one, the more significant impact on the success or failure of the PPP project.

Table 4 shows the weighing of indices (variables) for the implementation of the Information Gain index method.

Table 4. weights of indices in method of information gain index

Indicator	Weight	Indicator	Weight
Complaints about service quality	1	Absence	0.584
Contract type	0.834	Operating revenue (loss)	0.554
Conventional tariff	0.763	Low water pressure	0.452
Working ratio (cost coverage ratio)	0.737	Annual investment to sale ratio	0.409
Cost of disposed wastewater at break-even point	0.731	Project's year	0.355
Collection period	0.631	Water quality	0.328
Fixed assets turnover ratio	0.626	Education	0.316
Population covered by water services	0.608	Non-Revenue Water	0.314
Cost of water at break-even point	0.608	Workforce productivity (gross)	0.256
Measurement level	0.251	The ratio of water cut-offs to the number of connections	0.201
Workforce effectiveness (gross)	0.249	Per capita workforce cost	0.176
The ratio of network accidents to the network length	0.158	Ability to sell the product in the province	0.104
Population covered by sewage treatment service	0.140	The ratio of connections accidents to the total connections	0.080
Indebtedness level	0.080	Wastewater treatment	0
Level of professionalism in the company	0.043		

- With regard to the weight of the features in the information gain method, only did the indices of “Wastewater treatment" and " Level of professionalism in the company" have no influence on the success of PPP projects and consequently, more than 93% of the indicators had great influence on the classification and prediction of the failure or success of the projects.
- "Complaints about service quality", directly, with the highest weight 1, affected the prediction of the projects' success.
- The "Contract type", "Conventional tariff" and "Working ratio (cost coverage ratio)" were the most influential variables.
- "Wastewater treatment", "Level of professionalism in the company," and "Indebtedness level" had extremely little effect on predicting the failure or success of the projects.

Table 5 shows the weights of indices (variables) for the implementation of Gini index method.

Table 5. weighing of indices in method of Gini index

Indicator	Weight	Indicator	Weight
Complaints about service quality	0.660	Absence	1
Contract type	0.682	Operating revenue (loss)	0.672
Working ratio (cost coverage ratio)	0.920	Annual investment to sale ratio	0.575
Conventional tariff	0.672	Low water pressure	0.485
Fixed assets turnover ratio	0.960	Education	0.389
Population covered by water services	0.393	Non-Revenue Water	0.307
Cost of disposed wastewater at break-even point	0.455	Project's year	0.195
Collection period	0.402	Water quality	0.179
Population covered by sewage treatment service	0.045	The ratio of connections accidents to the total connections	0.105
Measurement level	0.130	The ratio of water cut-offs to the number of connections	0.100
Workforce effectiveness (gross)	0.426	Per capita workforce cost	0.085
Indebtedness level	0.028	Wastewater treatment	0.061
The ratio of network accidents to the network length	0.234	Ability to sell the product in the province	0.042
Cost of disposed water at break-even point	0.367	Workforce productivity (gross)	0
Level of professionalism in the company	0.234		

Based on the results, it was determined that:

- "Absence" with the weight 1 was one of the most influential factors that directly affected the projects' success.
- "Fixed assets turnover ratio" and "Working ratio (cost coverage ratio)" with an average weight of above 0.91 were the most influential factors in the weighting indices with the Gini index method.
- "Workforce productivity (gross)" and "Ability to sell the product in the province" were evaluated by calculating the Gini coefficient of the least influential factors.

In the following section, the data mining techniques were applied to compare these two methods.

4.1 Comparison of Two Methods

According to the weighing results and the differences in the prioritization in the two methods of the weighing index, classification algorithms were utilized to compare the methods. After applying different classifiers and considering the resulted accuracy, Random forest and Naïve Bayes classifiers were selected to compare the feature selection methods. The comparison procedure was that each feature selection method chose high-weight indexes applied to each classifier.

The precision of the classifier was the criterion for assessing and comparing the method to examine the factors that affect.

The accuracy of the classifier was calculated through the classification matrix (Tina et al., 2013). The benchmark for this study was the accuracy or classifying rate. This accuracy was calculated based on the concepts of the matrix according to Equation (4) (where TP=true positive, TN= true negative, FP= false positive, and FN= false negative).

$$Accuracy = \frac{Tp+TN}{Tp+TN+Fp+FN} \quad (4)$$

The two classifiers of Random forest and Naïve Bayes were utilized. Given the relatively high weight of the variables, after the simulation, the highest accuracy was reported with the 27 selectable variables.

In the Information gain method, the two "Wastewater treatment" and "Level of professionalism in the company" indicators were removed in the Gini index due to the extremely low influence. The results of the total accuracy for each classifying method are described in Table 6.

Table 6. Results of total accuracy for each classifying method:

	Information gain	Gini index
Random Forest	85.68%	84.85%
Naïve Bayes	73.86%	73.86%

The results of Table 6 reveal that the weighing of indices with the information gain method was more accurate. That is, the selected variables of this method and the effect of each parameter had less error in the prediction model of the failure or success of the projects, and therefore, the ranking of the indicators in terms of the impact on the success of public-private partnership projects was conducted based on the method of information gain. The first 15 indicators that were effective in the public-private partnership in the priority order are presented in Table 7.

Table 7. Ranking of public-private partnership indicators in terms of impact on the success of PPP projects in the water and wastewater industry

Indicator	Priority	Indicator	Priority
Complaints about service quality	1	Population covered by water services	9
Contract type	2	Absence	10
Conventional tariff	3	Operating revenue (loss)	11
Working ratio (cost coverage ratio)	4	Low water pressure	12
Cost of disposed wastewater at break-even point	5	Annual investment to sale ratio	13
Collection period	6	Project's year	14
Fixed assets turnover ratio	7	Water quality	15
Cost of disposed water at break-even point	8		

As Table 7 indicates, in the water and wastewater industry of Iran, the success of public-private partnership projects was most affected by the "Complaints about service quality" indicator. Given that more than 40 percent of the existing contracts were of "Current finance" type and executed with the participation of the people in financing the plan, this issue suggested the effect of the satisfaction of people concerning the quality of the service on the success and continuity of the project.

Based on the data from Table 7, among the 15 key indicators of the public-private partnership projects in the water and wastewater industry, eight of them were directly related to the financing of the projects. This suggested that, unlike other developed countries, which are moving towards PPP projects for increasing the productivity, the qualitative design, and the use of knowledge and experience of the private sector, in Iran the project financing is still the main issue.

According to the results, the "contract type" was considered as the second most important factor in the success of the collaborative project, and as a result, the water and wastewater companies can easily evaluate the result of the implementation of the project in each of the contract types and choose the most appropriate type for the project implementation.

"Conventional tariff" is the third most crucial factor in the success of a collaborative project. This implied that in the event of the elimination of tariff duties and by creating a competitive environment based on the actual costs of the water and wastewater production, many of the projects can be completed on a self-contained basis and with the participation of the private sector; consequently, the industry would be largely away from dependence on the governmental resources.

5. Conclusion and Recommendations

Identifying public-private partnership indicators and prioritizing them in all sectors, including the water and wastewater industry is very important and it can greatly improve the success of public-private partnership projects. In this research, it was sought to identify and prioritize such indicators. Identification of public-private partnership's contract types and its indicators in the industry were performed with a full review of the industry and projects carried out in it, with the participation of all PPP projects' stakeholders.

In the next step, by reviewing the literature in this field and receiving the views of the interested experts in these projects, 33 indicators were identified as the indicators of public-private partnership, and the data of the indicators for all cooperative projects of this industry were collected throughout the country.

As a benchmark for comparing the methods of assessing the impact of PPP indicators, the Random forest classifier was applied, providing the high accuracy of 85.6% when using the information gain method (because the accuracy of the classification methods was to compare the methods of choosing the feature); consequently, the information gain index (entropy) reported the weight of

effective indicators with a fairly higher accuracy than the Gini index. Both methods also proved that out of a total of 29 indicators, 98 percent of the indicators had a high weight on the prediction of the failure and success of the projects. Finally, by weighing the indicators, 15 indicators were reported as priorities. The "Complaints about service quality" index was the most effective indicator in promoting collaborative projects, indicating the effect of the people's satisfaction of the quality of services on the success and continuity of projects.

The public-private partnership in the water and wastewater industry is an option to be taken into consideration in developing countries. In the public-private partnership, the realization of the interests of both public and private parties is very important in the final outcome of the projects and should be considered in the design of the contracts. However, increasing efficiency and improving service quality are important reasons for the cooperation with the private sector, but improving financial flows and increasing investment are the main reasons for such types of cooperation contracts in Iran. The public-private partnership in the water and wastewater industry is being designed and implemented in different parts of the world, and various results have been achieved. Examining the implementation models of this kind of partnership in different countries and analyzing their strengths and weaknesses can be important in designing a suitable model for Iran. As public-private partnerships are being pursued very seriously in the water and wastewater industry, it is suggested that, in the future research, a public-private partnership model be developed in this industry in order to extract the appropriate model for participation in each region.

References

Abdul-Aziz, A. R., (2001). Unraveling of BOT scheme: Malaysia's Indah water konsortium. *Journal of construction engineering and management*, 127(6), 457-460.

Akintoye, A., (2003). Achieving best value in private finance initiative project procurement. *Construction Management and Economics*, 21(5), p. 461-470.

Böhl, C. G. P. (2007). Development of a knowledge based decision support system for private sector participation in water and sanitation utilities. Oldenbourg Industrieverlag, 189.

Brune, N.G., Garrett, & Kogut, B. (2004). The International Monetary Fund and the global spread of privatization. *IMF Staff Papers*, 195-219.

Commission, E. (2003). *Guidelines for successful public-private partnerships*. DG Regional Policy.

Company, i. (2009). *National Wastewater Finanving and investment strategy Report*.

Delmon, J. (2015). *Creating a Framework for Public-Private Partnership Programs*.

Dorian, J. P. (2006). Central Asia: A major emerging energy player in the 21st century. *Energy Policy*, 34(5), 544-555

Golabchi, M., Nourzaee. (2015). Selecting the best PPP method in rail projects by using AHP methods. *Journal of Transportation Engineering*, No.6

Hall, D., Lobina. E. (2005). *The relative efficiency of public and private sector water*. PSIRU Reports.

Han, J., Jian, P., & Kamber, M. (2011). *Data mining: concepts and techniques*, Elsevier.

Hodge, G. A., Greve, C., & Boardman, A.E. (2010). *International Handbook on Public-Private Partnership*. Edward Elgar Publishing

Li, B. (2005). Critical success factors for PPP/PFI projects in the UK construction industry. *Construction management and economics*, 23(5), 459-471.

Liu, Haoyue, MengChu, Zhou, Xiaoyu, Sean Lu. & Cynthia Yao. (2018). Weighted Gini index feature selection method for imbalanced data. In 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC). 1-6. IEEE.

Marzouk, M., Fayez, E. (2018). Public Private Partnership Projects: Concessionaire Performance Measurement. *Journal of Al Azhar University Engineering Sector*.

Penelope, B.J., Cowen, B. (1997). The private sector in water and sanitation–how to get started. *Washington: Bank*, p. 89-92.

Qiao, L. (2001). Framework for critical success factors of BOT projects in China. *The Journal of Structured Finance*, 7(1). p. 53-61.

Sachs, T., Tiong, R.. & Qing, S. (2007). Analysis of political risks and opportunities in public private partnerships (PPP) in China and selected Asian countries: Survey results. *Chinese Management Studies*, 1(2). p. 126-148.

Salman, A. F., Skibniewski. M. J., & Basha, I. (2007). BOT viability model for large-scale infrastructure projects. *Journal of Construction Engineering and Management*, 133(1), 50-63.

Savas, E. S. (2000). *Privatization and public-private partnerships*. Chatham House New York.

Sun, J., Wanga, R., Wanga, X., & Pinga. J. (2014). Spatial cluster analysis of bursting pipes in water supply networks. *Procedia Engineering* , 70,1610 – 1618

Tang, L., Shen, Q., & Cheng, E.W. (2010). A review of studies on public–private partnership projects in the construction industry. *International Journal of Project Management*, 28(7), 683-694.

Tina, R., Patil, S., & Sherekar, S. (2013). Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification

Wei, Y. (2011). Regulating Municipal Water Supply Concessions. p. 1-0.

Wen,Y., Huang, W., & Wu, J. (2013). Water Consumption analysis System Based On data Mining, Mechanics and Materials, 241-244, 1093-1097

Yong Qiu., Ji. Li., Xia, Huang., & Hanchang, Shi. (2018). A Feasible Data-Driven Mining System to Optimize Wastewater Treatment Process Design and Operation.

Yu, I. (2007). Comparable performance measurement system for construction companies. Journal of Management in Engineering, 23(3), 131-139.

Yuan, J., Skibniewski, M. J., Li, Q. and Zheng, L. (2010). Performance objectives selection model in public-private partnership projects based on the perspective of stakeholders. *Journal of Management in Engineering*, 26(2).

Yuan, J., Wang, C., Skibniewski, M. J., and Li, Q. (2012). Developing key performance indicators for public-private partnership projects: questionnaire survey and analysis. *Journal of Management in Engineering*, 28(3),. 252-264.

Zhang, X. (2005). Critical success factors for public–private partnerships in infrastructure development. *Journal of construction engineering and managemen*, 131:1(3),