Research project leadership stipulation system

Mona Taghavi
Department of Electrical, Electronics and Systems Engineering, Universiti Kebangsaan Malaysia, UKM, Bangi, Selangor, Malaysia, and
Hamed Taghavi and Milad Taghavi
Department of Information Technology Engineering, Islamic Azad University, Parand Branch, Tehran, Iran

Abstract

Purpose – Due to the increased research funding which academic institutions receive to perform advanced R&D, there is an indispensable need to have a systematic approach for selecting competent academicians capable of leading such projects. This paper aims to propose an approach to develop a system for Research Project Leadership Stipulation in finding such academicians.

Design/methodology/approach – The system was developed using a decision tree model for leader selection and a neural network model for leader performance prediction, and validated through quantitative and empirical analysis by exercising it on a Research University's Human Resource dataset.

Findings – In contrast with common perception, the results showed that the level of an academic leadership expertise does not alone determine R&D project success. Managerial and intellectual competencies complemented by soft skills are more influencing factors on the success of a research project.

Originality/value – This paper provides a comparative analysis of selection criteria and influencing factors on research project leader's performance in terms of their hard and soft skills. The developed system by this research, selects and *intelligently* predicts the performance of an academician who possesses optimum ability to lead projects with a high confidence level of successful delivery.

Keywords R&D project, Academic research, Project management, Data mining, Performance evaluation, Leadership skills.

Paper type Research paper

1. Introduction

Presently, universities and academic institutes receive huge budgets from various government and industrial sponsored programs to perform advanced academic research on the premise of developing the skill sets or knowledge to transfer to industry. There is a worldwide interest in translating the technology product of university research project into economic gains through academic entrepreneurship and generating wealth. As such, conducting these research projects successfully is becoming more significant as countries perceived R&D as one of the key factors for sustainable and successful operations of organizations and institutes aiming to improve their national competitiveness (Kao and Pao, 2012). The interrelated dependencies between the academic projects, market needs, national competitiveness, and research potentials require a systematic approach for selecting the academicians who conduct and lead such R&D projects.

It is certified by many researchers that the role of a project leader and his leadership abilities is a critical success factor in projects via direct and indirect means (Geoghegan and Dulewicz, 2008; Muller and Turner, 2007; Sumner *et al.*, 2006). This relationship between project success and project leadership is correlated by other

Campus-Wide Information Systems
Vol. 30 No. 4, 2013
pp. 266-287

1065-0741
DOI 10.1108/CWIS-03-2013-0015

research showing that "projects often fall short of achieving their anticipated results, not due to a lack of project management, but rather from a lack of project leadership" (Lee-Kelley and Leong, 2003). Furthermore, the non-technical issues and neglecting the behavioral and social factors are counted as the main reasons of the project's failure (Sumner *et al.*, 2006; Thite, 1999/2002).

The obscurity and vagueness of the leadership concept has enticed researchers to interpret, capture, and analyze the essence of leadership in universities and academic institutes from different perspectives. These studies further identified leadership as a tangible and observable phenomenon, while there is no consensus on the exact characteristics of a successful leader in universities and academic institutes (Koen and Bitzer, 2010). All of the conducted researches in this area are about the top leadership position which a chancellor or the dean of a faculty holds rather than the research project leader (Deem, 2011; Fullan and Scott, 2009; McCaffery, 2010; Neary, 2012).

In summary, the problem to be solved is to select the qualified academicians who can lead R&D projects with high level of confidence in executing projects successfully. The common perception is that academicians with strong knowledge and academic expertise are capable of leading research projects productively and successfully. This paper is to prove or reject the hypothesis that the "level of academic leader expertise figures out the research project success." The research questions which should be answered are as follows:

(1) Is there a link between academic soft skills and the projects' success?

(2) What are the most important criteria in nominating an academician for a project leadership position?

(3) How can we select a qualified academician of maximum leadership performance for executing research projects successfully?

To answer these research questions, this paper proposes a new approach to develop a system for research project leadership stipulation (RPLS) in order to provide a systematic way of selecting research project leaders and to help decision makers in this process. Through this approach, hidden information relevant to academicians' skills and success of research projects can be extracted. The RPLS system defines the possibility of selection and prediction of the selected leaders' performances to provide the opportunity of choosing the best qualified candidate. As it has the ability of machine learning, it can improve its prediction output method over time.

2. State of the art – human resource selection techniques

The process of organizing, managing, and leading a project team is defined as human resource management (HRM) (Indelicato, 2009). Personnel selection plays a decisive role in HRM since it determines the input quality of personnel (Chien and Chen, 2008). It contributes to the success of the project and creates a competitive advantage for the organization (Huemann *et al.*, 2007). Performance improvement is a fundamental concern of management in HRM, and the final goal of personnel selection is to increase the performance by placing the most suitable person (Purcell and Wright, 2007). Various factors affect the performance of projects, of which leadership is considered as one of the most significant ones (Boyatzis, 2008; Geoghegan and Dulewicz, 2008).

Professional soft skills and abilities which are the basic requirements of human resource selection information system are qualitative and cognitive factors. These characteristics motivate researchers to use the fuzzy decision-making methods to cope with fuzzy sets (Chen and Cheng, 2005). As shown in Table I, quite a number of researchers have developed intelligent and fuzzy systems to contribute to decision makers to employ the most appropriate human resources. Though these researchers developed intelligent systems by taking qualitative criteria into account, they used collected data through the interview or assumed that the hiring organizations have the qualitative data about their soft skills. Besides, the competency criteria hierarchies that they studied are mostly for general employees, and academician leadership is not investigated in any of them. In these systems, past experience of performance of the employee is not considered which means that they do not consider the possibility of identifying the characteristics of the top individuals. Therefore, they cannot improve the selection based on their performance and learn the best criteria to judge.

According to the existing literature, there are two major problems in employee selection: the decision-making criteria hierarchy, and the methodology to be used. The most popular methods to address the methodology problem are fuzzy sets and numbers, expert systems, and Artificial Neural Networks (ANNs). Apparently, most of them use fuzzy numbers to assess the performance of alternatives in the particular human characteristics and skills criteria. Current HRMSs include the quantity values which decision makers should use to select the best candidates among the current academicians without any stored information about their characteristics.

Most of the above researches developed intelligent systems to select human resources for industry based on their qualitative criteria. Before an organization or company selects an employee, an interview is conducted with the candidate to source out his/her personal characteristics. The interviewer then can feed the system with the acquired qualitative input data about their personnel soft skills. These systems cannot deal with the existing quantity values stored in the HRMS to select the best among the current employees without human interference. They need input or instructions from experts to define fuzzy rules and parameters, besides, the process of tuning of the parameters (e.g. membership functions) of the fuzzy system requires a long time, especially if the number of fuzzy rules in the system is high. Interpretation of the fuzzy rules and combining the outputs of several fuzzy rules with defuzzification of the output is complex but can be conducted in various ways depending on the expert.

The above mentioned deficiencies and limitations motivated this research to use the data mining approach to overcome these problems, and to develop a model which will retain academics' history for project leadership positions and selection criteria through the existing data in the universities' HRMS. Some researchers claimed that the HRM activities are very complicated and only a few quantitative approaches have been employed in practice (Chien and Chen, 2008). Measuring the soft skills through the available data in HMRS which are in quantity values, is a very difficult job.

3. Data mining in HRM

Data mining techniques have been developed to explore and analyze large quantities of data in order to discover meaningful patterns and rules. Data mining involves various techniques including Genetic Algorithm, Decision Tree (DT), ANNs, Statistics, and Visualization techniques. During the last decade, a lot of effort has been put forth to apply a wide variety of data mining techniques to different types of data sets in broad applications (Triantaphyllou, 2009).

Table I. Applied intelligent techniques in human resource selection from published literature

Techniques	Field of study	Demonstration	Year	Remarks (pros and cons)
Ordered weighted averaging (OWA) operators (Carlsson *et al.*, 1997)	Doctoral student selection	No	1997	Define the importance of criteria (weight) Handle linguistic values No defined selection criteria
Fuzzy technique for order preference by similarity to ideal situation (TOPSIS) (Chen, 2000)	General	System analysis engineer selection in a company	2000	Handle linguistic values Apply very limited criteria
Fuzzy numbers (Butkiewicz, 2002)	General	Staff selection in a tourism agency	2002	Handle linguistic values Apply limited criteria Considering performance in the selection process
Discriminate analysis, Decision Trees, Artificial Neural Networks (ANN) (Cho and Ngai, 2003)	Insurance sales agents selection	No	2003	Flexible to cope with quantity and quality values No defined selection criteria
Fuzzy neural networks, fuzzy analytic hierarchy process (AHP), simple additive (Huang *et al.*, 2004)	Middle manager selection	No	2004	Support group decision making Define the importance of criteria Handle linguistic values No defined selection criteria
Fuzzy numbers (Chen and Cheng, 2005)	IS personnel hiring	IS project manager recruitment	2005	Support group decision making No defined selection criteria Handle linguistic values
Fuzzy TOPSIS (Saghafian and Hejazi, 2005)	General	University professor hiring	2005	Support group decision making Handle linguistic values
AHP, TOPSIS, Borda function (Shih *et al.*, 2005)	General	Online manager recruitment of a chemical company in southern Taiwan	2005	Support group decision making Handle linguistic values No defined selection criteria Define the importance of criteria
Fuzzy multiple objective mathematical programming (Baykasoglu *et al.*, 2007)	General	Project team members selection	2007	Handle linguistic values Consider project environment factors (time and budget)

(*continued*)

Table I.

Techniques	Field of study	Demonstration	Year	Remarks (pros and cons)
Fuzzy logic and decision-making trial and evaluation laboratory (Wu and Lee, 2007)	Competency development of global managers	Manager competency in a high-tech company in Taiwan	2007	Segment required competencies Handle the linguistic values Apply limited criteria
Expert systems (Saidi Mehrabad and Fathian Brojeny, 2007)	Intelligent selection in an R&D organization	No	2007	No defined selection criteria Generating rules by human experts
Fuzzy set, OWA (Canós and Liern, 2008)	General	No	2008	No defined selection criteria Define the importance of criteria Handle linguistic values
Decision Trees (Chien and Chen, 2008)	General	Engineers and managers selection in a semiconductor company	2008	Flexible to cope with quantity and quality values Predict performance in selection processes Apply general and limited criteria
Fuzzy TOPSIS (Mahdavi *et al.*, 2008)	General	System analyst selection in a software company	2008	Support group decision making Handle linguistic values No defined selection criteria Define the importance of criteria
Fuzzy set (Shipley and Johnson, 2009)	Selection of employees for a project	No	2009	Handle linguistic values No defined selection criteria
Fuzzy TOPSIS (Saremi *et al.*, 2009)	TQM consultant selection	No	2009	Flexible to cope with quantity and quality values No defined selection criteria Define the importance of criteria
Fuzzy TOPSIS, fuzzy AHP (Celik *et al.*, 2009)	Academic personnel recruitment in Maritime Education and Training (MET) institutions	No	2009	Various detailed criteria Define the importance of criteria Handle linguistic values

(*continued*)

Techniques	Field of study	Demonstration	Year	Remarks (pros and cons)
Fuzzy TOPSIS (2-tuple linguistic variables), OWA (Dursun and Karsak, 2010)	General	No	2010	Define the importance of criteria Flexible to cope with quantity and quality values Apply general and limited criteria
Analytic network process (ANP) and TOPSIS (Dağdeviren and Yüksel, 2010)	Personnel selection in manufacturing systems	Personnel selection for a factory operating in Turkey	2010	Define the importance of criteria Flexible to cope with quantity and quality values Apply limited criteria
ANN, fuzzy set (Goonawardene *et al.*, 2010)	General	Software development project team member selection and performance prediction in Sri Lanka	2010	Deal with imprecise data, difficulties in retrieving information and expressing an explicit opinion
Multi-criteria decision making (MCDM), fuzzy TOPSIS with veto threshold (Kelemenis *et al.*, 2011)	Manager selection	Selection of middle level manager in a large IT Greek firm	2011	Support group decision making Define the importance of criteria and threshold of decision making Apply limited and general criteria Handle linguistic values
Fuzzy AHP, adaptive network-based fuzzy inference system (ANFIS), ANN (Shahhosseini and Sebt, 2011)	Selection and assignment of human resources to construction projects	No	2011	Evaluate various and detailed competency criteria Able to cope with high volume of data Handle the linguistic values Define the importance of criteria
Fuzzy TOPSIS (Fathi *et al.*, 2011)	General	No	2011	Handle the linguistic values Define the importance of criteria Apply limited and general criteria
Fuzzy ANP, fuzzy TOPSIS, fuzzy ELECTRE (Kabak *et al.*, 2012)	Sniper selection	No	2012	Apply limited criteria Define the importance of criteria Flexible to cope with quantity and quality values Support group decision making

Table I.

In HRM activities, data mining techniques can help to understand the characteristics of the top performing individuals. Through the understanding of their characteristics such as education, skills, personality traits, and years of experience, a profile for hiring can be established in order to help the recruitment of the individuals who possess similar characteristics as their best performing individuals. The one who employs this technique should realize that it is based upon historical data, and may not remain in the same performance and capability through changes in economic, social, and environmental conditions. Any one of these changes can result in inefficiency of data mining techniques to indicate the future top performing individuals.

Although the data including personnel data can provide a rich resource for knowledge discovery and decision support, the data mining application in HRM was found rarely in the existing literature (Chien and Chen, 2008). In particular, Cho and Ngai (2003) used data mining to develop a decision support system to predict the length of service, sales premiums, and persistence indices of insurance agents. Another study by Chien *et al.* (2005), employed data mining to analyze disoperation behaviors of operators. In 2008, Chien and Chen (2008) further developed a data mining framework based on DT and association rules in order to generate useful rules for personnel selection. The results from their research provided decision rules relating to personnel information with work performance and retention with an empirical study in a semiconductor company.

In the human resource selection area, two data mining techniques of DT and ANN were used by previous studies. Therefore, both techniques are adopted in our proposed approach and tested and systematically evaluated in term of their efficiency in the implemented system.

4. The proposed approach to develop a RPLS system

The aims of the human resource selection system are to improve the personnel performance and increase the probability of success through sound stipulation criteria and factors. We propose a system composes of two models for selection and prediction of academic project leadership performance in research projects to have a comparison between the performance and selection criteria, and to improve the selection process with performance outcomes over time.

At the highest overview level, the stages of the data mining processes are:

(1) exploring the problem scope and defining expected outcomes;

(2) data collection and preparation;

(3) mining the data and constructing the model; and

(4) interpreting and using the discovered knowledge.

Detailed processes of developing the RPLS system based on data mining technique applied in this research are shown in Figure 1.

There are huge amounts of rigorous research on technical aspects of data mining that made it mature fast (Pechenizkiy *et al.*, 2008). Many organizations implemented or are planning to implement data mining projects. But the collection of data mining algorithms is not enough to implement those projects successfully. The considered problems and the proposed solutions should be selected carefully in a way that supports other human and organizational aspects. Therefore, they should be employed practically to solve the real-world problems and benefit from the discovered knowledge to achieve the full competitive advantages. The maturation of data mining supporting processes that deal with organizational and human aspects is still living on its childhood (Pechenizkiy *et al.*, 2008). To validate the proposed system and to overcome the above mentioned issue, this research implemented the processes depicted in Figure 1 by an empirical application.

5. The RPLS system development processes through a practical case

To investigate the proposed approach in a practical setting, the RPLS system development processes were applied through a research university's human resource data set. This research university is located in a non-English-speaking country, even though English is used as a medium of instruction and communication. The data were anonymized to ensure privacy of individuals under existing legislations. The system development processes are explained as follows.

5.1 Exploring the problem space and defining the objectives

The first step in data mining is to identify the exact problem and specify the objectives which should address the problem. In this research study, the problem was defined after reviewing relevant literature in academic societies and discussing it with academics in managerial and non-managerial positions. One of the draw back issues of executing research projects in the studied university in this paper, is that students do not deliver their contributions according to their deal with their academic project leaders. Consequently, the leaders fail to provide project deliverables on time; even in some cases they exceed four or five years. Some of the leaders do work in a group with students consistently and run their projects successfully. The data mining through the project records and academics' characteristics can help to discover the relation between the academic leaders' skills and the research projects' success or failure.

There were many constraints to fulfill the objectives desirably. The targeted data set was limited to one university and some of the patterns which were discovered through the mining were not general enough to be applicable to other academic institutes. Despite this drawback, all aspects where considered to construct a general model, but each institute has its own standards and criteria. The data set were collected over 15 years from the existing HRMS, thus lacking important attributes, such as academic training courses, key performance indicators, personality evaluation remarks, and social and scientific committees' membership information.

5.2 Data collection and preparation

Human resource data are usually stored in separate databases for privacy reasons. The related data had to be combined and prepared before further analysis. Since the purpose of RPLS system is to select project leaders with maximum performance, two different data sets for two different models were organized: one for selecting the leader and the other for evaluating his/her performance by comparing with the previous leaders skills. The first data set with 3,331 instances consisted of the entire university academicians to investigate who can be a leader. The second data set with 1,129 instances composed of only project leaders to evaluate their performance and skills and which had more influence on the projects' success.

The evaluation criteria for the academic leaders are organized based on their leadership and expertise competencies as shown in Table II in terms of their

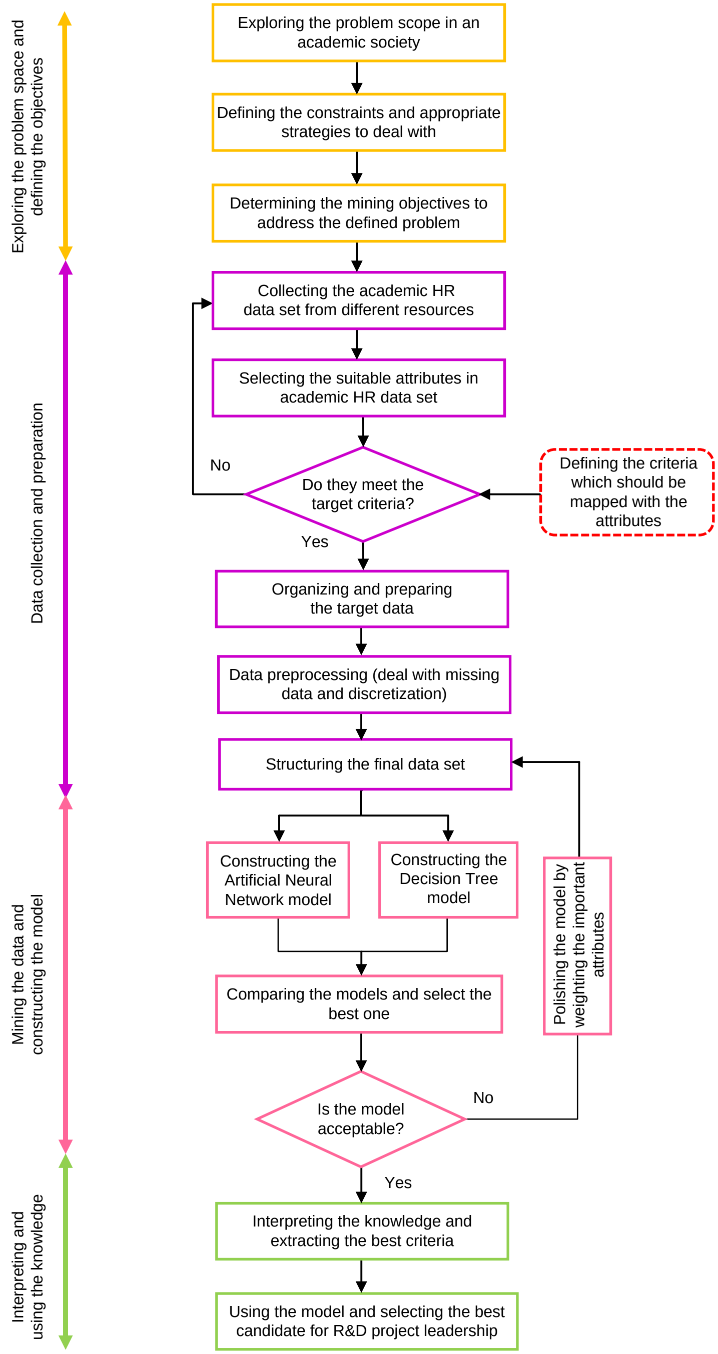

Figure 1.
The RPLS system development processes

Importance levels	Fundamental	Essential	Distinguishing features of an extra ordinary person
Competency clusters			
Emotional competencies	Self-awareness and self-confidence Integrity	Motivational Conscientiousness	Influence Intuitiveness
Managerial competencies	Engaging communication Managing resources Achieving scientific and technical (S&T) results Encouraging	Developing R&D workforce Empowering Mentoring/coaching skills Inspiring	Excellent personal relations Foresight and far-sight
Intellectual competencies	Safe tactical decision making with self-belief Sound knowledge with critical analytical skills	Strategic decision making perspective with perceptive intellectual abilities Critical creative analysis and analytical skills and sound judgment	Vision and imagination, self-realization and fulfilment Creative thinking with exceptional judgment taking abilities
Expertise competencies			
Knowledge perspective	Education (same area as R&D field/subject)	Teaching and supervision (same area as R&D field/subject)	Evaluation and grading Training courses Having practical experience in the same R&D field/subject area
Research perspective	Publications (in recognized indexed proceedings and secondary ISI/SCOPUS indexed journal) Presentations at workshop and conference Develop products from the achieved R&D results	Publications (primary (high indexed) ISI/SCOPUS listed journal, books, news clipping, encyclopaedia) Serving on at least one ISI/SCOPUS indexed journal editorial boards Keynote talks on R&D at workshops and conferences, international R&D gatherings	National and international awards and appreciation Serving on ISI/SCOPUS indexed journal editorial boards Serving as editor-in-chief of at least one ISI/SCOPUS indexed journal Talks and input at R&D programs and initiative meetings and international fora/workshops
Administration perspective	Membership position – deputy or subordinate Membership of professional societies (IEEE, ACM, etc. Membership of R&D committee	Managerial position – head or divisional research head	Leadership position – chief or executive role Vice/deputy chancellor R&D directive

Table II.
Academic leadership competencies and evaluation criteria

importance level. Three clusters of managerial (MQ), emotional (EQ), and intellectual (IQ) competencies which are leadership competencies proposed by Geoghegan and Dulewicz (2008), were combined with expertise competency as a composite academic competency scheme. Each competency has its own evaluation criteria and measurement metrics. The university human resource data set did not include detail information, especially on cognitive and soft skills of academicians. The majority of target attributes were selected based on the professional identity of academicians and we attempted to evaluate their personal characteristics and leadership abilities through their academician profile. For example, we defined the metric of "the number of keynotes and speeches in scientific societies, conferences and workshops" to measure the level of the academics' self-confidence and motivation. This shows their ability of communicating and sharing their ideas and knowledge with the other scientists. The selected attributes and the purpose of their selection are provided in Table III. The last two attributes (19-20) were available only for project leaders' data set.

The final collected data often included noisy, missing, and inconsistent data. Therefore, the data preparation processes consisted of checking the data distribution, dealing with empty or missing values, and transforming data into analyzable formats which were utilized to improve the data quality capable of conducting effective data mining processes. Finally, the data were discretized to improve the quality and increase the accuracy of the proposed model.

5.3 Mining the HRMS data and constructing the models

Considering all the techniques which were used in developing the human resource selection system by the researchers, two techniques of ANN and DT were employed to mine the data of both data sets by EasyNN-plus (software, 2011) and WEKA machine learning toolkit, respectively (Hall *et al.*, 2009).

The learning algorithm of EasyNN-plus is back-propagation with feed-forward network architecture. The popular and accurate algorithm of C4.5 was used as a DT method in WEKA. The obtained models were compared in order to select the best model with highest accuracy. The constructed models were evaluated by conducting several tests and splitting the data to train and test the data to validate the model.

After selecting the best model, it was evaluated against its accuracy, meaning, and ease of use by decision makers. For further improvement, the feature weighting method was employed for classification tasks by extracting relevant information from a trained ANN. This method weights an attribute based on strengths (weights) of related links in the ANN, in which an important feature is typically connected to strong links and has more impact on the outputs. This method showed significant reduction in the size of constructed DTs and slight improvement in accuracy in a study by Zeng and Martinez (2004).

5.4 Interpreting and using the discovered knowledge

Data mining results should be analyzed, interpreted, and assessed according to the experience and body of knowledge of domain experts in order to justify the purpose and meaning of extracted knowledge. Interpretation of the model was done in two basis of project leader selection and leader performance prediction which are explained in the next section. The performance prediction was utilized to obtain better criteria in hiring the academic leaders. This useful information or patterns were extracted and

Table III. Utilized features in model construction based on HRMS

Attribute number	Attribute name	Meaning	Representing of (relevant evaluation criteria)	Competency cluster
1.	Awards	Number of the awards	Self-awareness and self-confidence	Emotional
2.	Journal publications	Number of published papers in journal	Inspiring, achieving scientific and technical results, publication	Managerial, intellectual, expertise (research)
3.	Publications	Number of publications in form of conference proceedings, journal papers, book reviews, etc.	Motivational, achieving scientific and technical results, publication	Emotional, intellectual, expertise (research)
4.	Leadership position	Years of experience in leadership position	Strategic decision-making perspective with perceptive intellectual abilities	Intellectual
5.	Project research co-operation	Number of conducted research projects in co-operation with other academics	Self-realization and fulfillment, sound knowledge	Intellectual
6.	Students	Number of students under supervision	Conscientiousness, engaging communication, inspiring	Emotional
7.	Publications in English	Number of publications in English language	Fluency in English to be able to communicate, scientific publication in international level	Managerial, expertise (research)
8.	Degree	The latest academic degree	Education	Expertise (knowledge)
9.	Publications in high-indexed journals	Number of publications in high indexed journals	Publications	Expertise (research)
10.	Keynote talks	Number of keynote speeches	Motivational, influence, self-confidence, keynote talks on R&D programmes	Emotional, managerial, expertise (research)
11.	Presents	Number of talks and presentations in conferences and workshops	Engaging communication, presentations at workshop and conference	Managerial, expertise (research)

(*continued*)

Attribute number	Attribute name	Meaning	Representing of (relevant evaluation criteria)	Competency cluster
12.	Editor of a quality journal	Serving on editorial board of ISI or SCOPUS indexed journal	Serving on at least one ISI/ SCOPUS indexed journal editorial boards	Expertise
13.	Employment date	Start date of working or employment	Teaching and supervision	Expertise (knowledge)
14.	Promotions	Number of promotions	Conscientiousness, achieving scientific and technical results	Emotional, managerial
15.	Degree date	The date of graduating from the latest degree	Self-awareness, motivational	Emotional
16.	Joint research with other organizations	Number of conducted research projects for industry or organizations outside of the university	Engage communication	Managerial
17.	Position	Academic position	Evaluation and grading	Expertise
18.	Research area	Number of research areas	Creative thinking	Intellectual
19.	Promotion date	Date of first promotion	Motivational	Emotional
20.	Lead research projects	Number of granted research projects' position as leader	Managerial or leadership position	Expertise (administration)

Table III.

summarized into decision support rules, and discovered knowledge formed the basis for decision support to generate academic leader hiring strategies which could also be used to improve related management activities.

6. Experimental results

This section provides the results of explained processes of developing the RPLS system. The results are explained based on the two models of research project leader selection and research project leader performance prediction which were derived from the two data sets of academicians and project leaders, respectively.

6.1 Research project leader selection model

The attributes number (1-18) provided in Table III, with the supervised class of "Project leadership" were used to construct the selection model and to analyze the specific features of project leaders who were selected by the exercised university. The supervised class values consisted of "yes" for those who posed as project leaders, and "no" for those who were not granted this position.

The results of the data mining are compared in Table IV. The DT model showed higher performance in terms of accuracy, and was easier to understand and use. The model was written in decision rules to be used in development of the final system. However, such a large tree was likely to be over fit. WEKA has a number of tunable parameters which allows for pruning of the tree to reduce over fitting, but still the model was complex and large. Also, if the significance level is set too small then the accuracy may decrease markedly. So, ideally pruning the tree should be to the point just before accuracy starts to deteriorate. Consequently, through the weighting and choosing the most important attributes, the final model of this research was restructured and polished to have more stable and shorter rules.

For further investigation on selected criteria and attributed weights, two factors of importance and sensitivity were assessed using the ANN technique. The importance is sum of the absolute weights of the connections from the input neurons to all the neurons in the first hidden layer. It can be justified that since they have more influence on classification decisions, the important features should have strong links along the nodes correlated to these features. The inputs are shown in Figure 2 in descending order of importance from the most important input to the least important input.

This provides ANNs with the potential to better capture the relevance of features related to classification. Therefore, these attributes were selected to use in building the final data mining model. The second run of modeling was conducted with the same accuracy holding less attributes and lower size and complexity.

Sensitivity analysis is an analysis of how values of the outcomes and probabilities can affect the final decision. Sensitivity analysis methods estimate which input parameter is more important or sensible to achieve accurate output values (Saltelli *et al.*, 2008). It is also used to understand the system behavior which is being modeled,

Table IV. Comparative results of applied classification techniques in research project leader selection model

Classification technique	Accuracy (%)	Size
DT	88.8622	Number of leaves: 72 Size of the tree: 109
ANN	85.4698	One hidden layer with 24 nodes

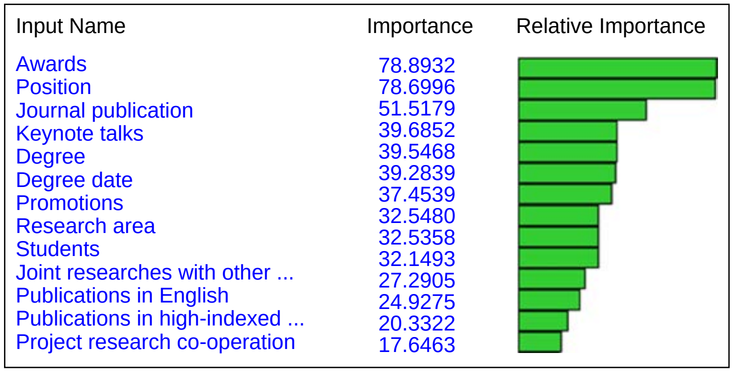

Figure 2.
Importance of attributes in research project leader selection model

and to verify if the model is applicable and is doing what it is supposed to do. Figure 3 shows the only sensitive attribute in the research project leader selection model which is "position."

6.2 Research project leader performance prediction model

To construct this model, all 20 attributes (refer to Table III) of project leader data set were utilized. The supervised class of "Successfulness" was set up which represented the success rate of the projects under their leadership. The success rate of the project leaders was classified into three categories:

- successful: the leaders who conducted their projects successfully and delivered the project deliverables;
- midrate: the leaders who experienced both successful and unsuccessful projects; and
- unsuccessful: this was of concern to project leaders who could not close their projects successfully.

According to Table V, the accuracy of the ANN model was better than the DT model. In reference to this, the NN model was selected to develop the final RPLS system. Table VI provides an insight specification of the NN model.

Figure 4 estimates the relative importance of each feature in the research project leader performance prediction model in descending order from the most important to

Figure 3.
Sensitivity of attribute in research project leader selection model

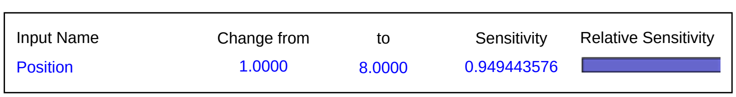

Table V.
Comparative results of applied classification techniques in research project leader performance prediction model

Algorithm	Accuracy (%)	Size
DT	76.3508	Number of leaves: 63 Size of the tree: 101
ANN	86.3060	One hidden layer with 11 nodes

the least important. Analysis of these important attributes provides an in-depth investigation on their performance criteria.

As previously explained, the term of sensitivity shows the intensity of output changes with changing the inputs. The inputs are all set to the median values, and then each, in turn, is increased from the lowest value to the highest value. The change in the output is measured as each input is increased from the lowest to the highest to establish the sensitivity to change. The inputs are shown in Figure 5 in descending order of sensitivity from the most sensitive input to the least sensitive input.

Figure 6 shows a snapshot of the final developed RPLS system using two selected models of DT and ANN. This system stipulates the academicians for leadership position and predicts their performance, and using its self-learning ability improves its selection based on the performance output or feedback over time.

Number of input neurons	20
Number of output neuron	1
Number of neurons in hidden layer	11
Learning rate	0.6
Momentum	0.8
Average training error	0.0002
Average validation error recorded	0.1362
Number of cycles	112,833
Accuracy (%)	86

Table VI. Specification of ANN research project performance prediction model

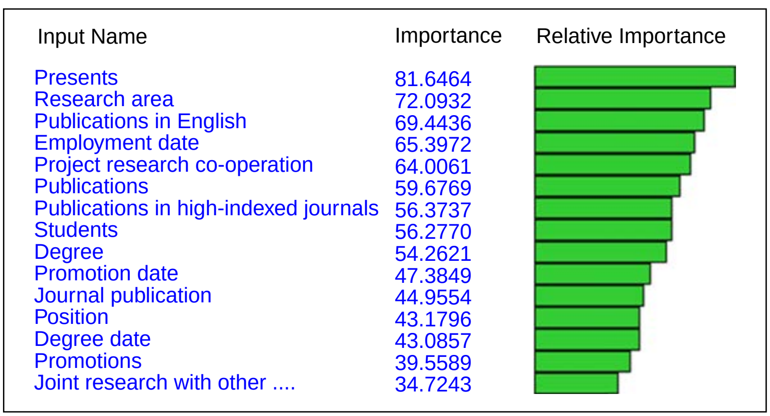

Figure 4. Importance of attributes in research project leader performance prediction model

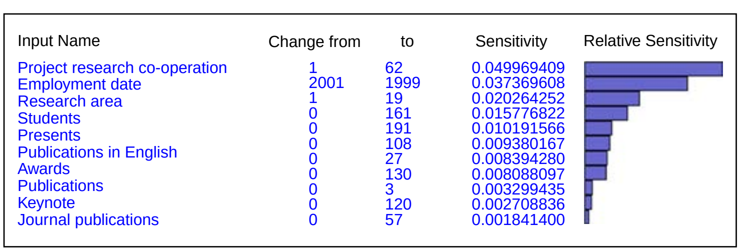

Figure 5. Sensitivity of attributes in research project leader performance prediction model

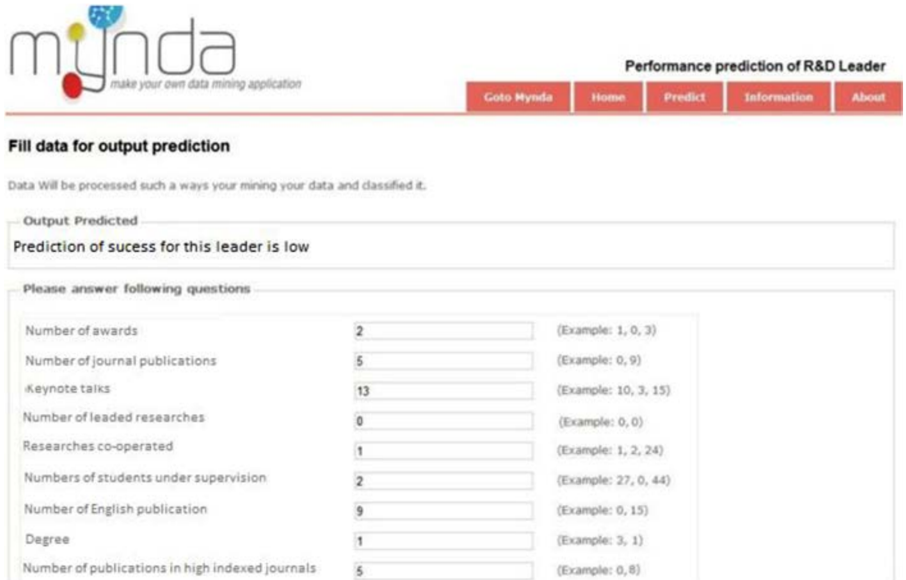

Figure 6.
A snapshot of RPLS system

6.3 Discussion and analysis of the results

To interpret the knowledge extracted from this experiment, the results are discussed based on the two constructed models.

Research project leader selection model. The constructed model for selecting the project leaders shows the project leaders' characteristics. Based on the analyses of the decision rules from the DT model and sensitivity diagram, position was the most deterministic criteria in selecting these project leaders for leadership positions (refer to Figures 2 and 3). The top five deterministic criteria are as follows: position at number one, followed by journal publications, keynote talks, and degrees acquired. These five criteria allow decision makers of the university to stipulate which academicians are suitable for leading research projects. Referring to our defined competencies in Table III, these criteria mainly represent the expertise competencies, which mean that decision makers give priority to academic skills to appoint a project leader.

Research project leader performance prediction model. Number of research presentations in conferences or workshops – shown in Figure 4 as "presents" – received the top most important level among the set of attributes in the project leader performance. It can be interpreted as the most common performance of presenting the project results which determined the criteria of success in research projects ranging from the input named attributes as given in Figure 4. In addition, according to the defined evaluation criteria having a larger number of presentations and talks prove higher communication ability and stronger self-confidence in the leader (referring to managerial and emotional competencies). All of these criteria are good reasons of making the presentations as the most important attribute deserving the top most (first) rank position.

The variety of research areas in which the project leader has worked is the second important attribute in the leadership performance ranking hierarchy. Although it has a negative effect, leaders with less number of research areas act more successfully than the leaders with more domains as their previous research outputs can contribute to their new researches. The academics who work on limited research areas may have better concentration, can harmonize their knowledge and past experience or research outputs by handling and sharing their new projects/research works more efficiently among their students and colleagues in the same field. The number of published publications, especially in the English language, shows a strong positive effect on the performance of project leaders even though it is an adopted success measure for R&D projects by universities. As co-operation and publication with international researchers need communication through the same medium, leaders fluent in English can share their ideas and communicate proficiently with each other.

The next important criteria are academic experiences and the number of research projects in which the project leader is involved in the studied university. According to Figure 5, these attributes are the most sensitive attributes in the leader performance model. It can be concluded that experience has a strong influence in the successful execution of a research project by the leader. Thus, an experienced academic has the potential to handle research project challenges and achieve its final project objectives in a consistent, logical, and professional manner.

Comparing the models and answering the research questions. In contrast to the research project leader selection model, performance prediction model showed position alone is not an influencing factor as an academic leader performance measure. Although most of the research project leaders have higher third-level education degrees, those who hold lower positions and fewer degrees tend to deliver more effectively as shown in the DT model of performance prediction. It is apparent that academics in lower positions are more motivated and determined to work consistently harder to climb up the achievement ladder rapidly compared to their seniors. In the performance model, managerial and intellectual competencies gained the most priority, whereas in the selection model, expertise competencies received more attention from the decision makers in the university. Given these facts, we should reject the commonly established research hypothesis which states that the "level of academic leader expertise figures out the research project success," since the academic expertise individually does not determine the success or failure of a research project. In turn, his/her abilities in engaging communication, fluency in English as a communication enabler, self-realization and fulfillment, and having sound knowledge with analytical skills are most important factors for a project leader in delivering research project objectives successfully. The influence of these factors which represent the academic soft skills on project success also answer the first research question: "is there a link between academic soft skills and the projects' success?"

The achieved evaluation criteria through our models were different in terms of their priority for academic hard and soft skills. To answer second research question "what are the most important criteria in nominating an academician for project leadership position?" we should consider both of the soft and hard skills. The academic soft skills mentioned in the answer of the first question belong to the managerial and emotional competencies, gained more weight for performance prediction model; meanwhile, we should also consider the university culture and the selection criteria that gained more priority by the university decision makers. As academic experience is a remarkable variable in the top five important criteria affecting a leaders' performance, we can state that involving university culture in executing R&D projects successfully is very important. Also the leaders with strong academic skills can leverage their knowledge to achieve research projects objectives more easily.

To answer the third research question "how can we select a qualified academician of maximum leadership performance for executing research projects?" we developed the RPLS system using data mining techniques consisting of two models which select the academic project leaders and predicts the performance of selected leaders for conducting future researches. This system improves the selection rules by receiving feedback from the performance predictions of the selected leader over a period of time. It is worth noting that although data mining provides the opportunity to analyze large quantities of the human resource data while extracting useful knowledge, it has its own drawbacks as any other techniques and technologies. Data mining products can be powerful decision-making tools but they are not fully fledged applications to determine absolute outcomes since it requires skilled analytical and technical experts who can analyze and interpret the outcomes. Despite identifying patterns and relations, data mining cannot value them or their significance because their validation depends on how accurately they reflect real-world scenarios and circumstances.

7. Conclusion

This paper developed a Research Project Leader Stipulation system through the data mining techniques to retain academics for future research projects' leadership. Two techniques of ANN and DT were chosen and compared. The DT model with a higher accuracy and more understandable result was chosen for the research project leader selection while ANN performed better in performance prediction. From the research findings, the established hypothesis (level of academic leader expertise figures out the research project success) was rejected since this research proved academic expertise individually does not determine the success or failure of a research project. The most deterministic factors on project research success are managerial and intellectual skills of academic project leaders, such as engaging communication, fluency in English as a communication enabler, self-realization and fulfillment.

Further research should be performed in the future to refine the model through more varied data sets and attributes that can be collected through more advanced human resource systems. The data which were mined in this research was limited to one university only, was not inclusive of all the attributes and many of the soft skills in leadership competencies. However, it should be taken into account that "the quality of a decision must be judged based on the information available when the decision is made, not on outcomes realized after the decision is made."

References

Baykasoglu, A., Dereli, T. and Das, S. (2007), "Project team selection using fuzzy optimization approach", *Cybernetics and Systems: An International Journal*, Vol. 38 No. 2, pp. 155-185.

Boyatzis, R.E. (2008), "Competencies in the 21st century", *Journal of Management Development*, Vol. 27 No. 1, pp. 5-12.

Butkiewicz, B.S. (2002), "Selection of staff for enterprise using fuzzy logic", paper presented at the International Conference on Systems, Man and Cybernetics, Vol. 4, Hammamet, October 6-9.

Canós, L. and Liern, V. (2008), "Soft computing-based aggregation methods for human resource management", *European Journal of Operational Research*, Vol. 189 No. 3, pp. 669-681.

Carlsson, C., Fullér, R. and Fullér, S. (1997), "OWA operators for doctoral student selection problem", *The Ordered Weighted Averaging Operators: Theory, Methodology, and Applications*, Kluwer Academic Publishers, Boston, MA, pp. 167-178.

Celik, M., Cebi, S., Kahraman, C. and Er, I.D. (2009), "Application of axiomatic design and TOPSIS methodologies under fuzzy environment for proposing competitive strategies on Turkish container ports in maritime transportation network", *Expert Systems with Applications*, Vol. 36 No. 3, pp. 4541-4557.

Chen, C.T. (2000), "Extensions of the TOPSIS for group decision-making under fuzzy environment", *Fuzzy Sets and Systems*, Vol. 114 No. 1, pp. 1-9.

Chen, L.S. and Cheng, C.H. (2005), "Selecting IS personnel use fuzzy GDSS based on metric distance method", *European Journal of Operational Research*, Vol. 160 No. 3, pp. 803-820.

Chien, C.F. and Chen, L.F. (2008), "Data mining to improve personnel selection and enhance human capital: a case study in high-technology industry", *Expert Systems with Applications*, Vol. 34 No. 1, pp. 280-290.

Chien, C.F., Wang, I. and Chen, L.F. (2005), "Using data mining to improve the quality of human resource management of operators in semiconductor manufactures", *Journal of Quality*, Vol. 12 No. 1, pp. 9-28.

Cho, V. and Ngai, E.W.T. (2003), "Data mining for selection of insurance sales agents", *Expert Systems*, Vol. 20 No. 3, pp. 123-132.

Dağdeviren, M. and Yüksel, İ. (2010), "A fuzzy analytic network process (ANP) model for measurement of the sectoral competititon level (SCL)", *Expert Systems with Applications*, Vol. 37 No. 2, pp. 1005-1014.

Deem, R. (2011), *Issues in Higher Education, Vol. 2*, Sense publisher, Rotterdam, pp. 159-176.

Dursun, M. and Karsak, E.E. (2010), "A fuzzy MCDM approach for personnel selection", *Expert Systems with Applications*, Vol. 37 No. 6, pp. 4324-4330.

Fathi, M.R., Zarei Matin, H., Karimi Zarchi, M. and Azizollahi, S. (2011), "The application of fuzzy TOPSIS approach to personnel selection for Padir Company, Iran", *Journal of Management Research*, Vol. 3 No. 2, pp. 1-14.

Fullan, M. and Scott, G. (2009), *Turnaround Leadership for Higher Education*, Jossey-Bass Inc, San Francisco, CA.

Geoghegan, L. and Dulewicz, V. (2008), "Do project managers' leadership competencies contribute to project success?", *Project Management Journal*, Vol. 39 No. 4, pp. 58-67.

Goonawardene, N., Subashini, S., Boralessa, N. and Premaratne, L. (2010), "A neural network based model for project risk and talent management", *Advances in Neural Networks-ISNN 2010*, Vol. 6064 No. 2010, pp. 532-539.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I.H. (2009), "The WEKA data mining software: an update", *ACM SIGKDD Explorations Newsletter*, Vol. 11 No. 1, pp. 10-18.

Huang, L.C., Huang, K.S., Huang, H.P. and Jaw, B.S. (2004), "Applying fuzzy neural network in human resource selection system", paper presented at the International Conference of the North American Fuzzy Information Processing, Banff, June 27-30, pp. 169-174.

Huemann, M., Keegan, A. and Turner, J.R. (2007), "Human resource management in the project-oriented company: a review", *International Journal of Project Management*, Vol. 25 No. 3, pp. 315-323.

Indelicato, G. (2009), "A guide to the project management body of knowledge (PMBOK® guide)", *Project Management Journal*, Vol. 40 No. 2, p. 104.

Kabak, M., Burmaoğlu, S. and Kazançoğlu, Y. (2012), "A fuzzy hybrid MCDM approach for professional selection", *Expert Systems with Applications*, Vol. 39 No. 3, pp. 3516-3525.

Kao, C. and Pao, H.L. (2012), "Predicting project approvals: a case of grants from the National Science Council of Taiwan", *Omega*, Vol. 40 No. 1, pp. 89-95.

Kelemenis, A., Ergazakis, K. and Askounis, D. (2011), "Support managers' selection using an extension of fuzzy TOPSIS", *Expert Systems with Applications*, Vol. 38 No. 3, pp. 2774-2782.

Koen, M. and Bitzer, E. (2010), "Academic leadership in higher education: a 'participative' perspective from one institution", *Academic Leadership Journal*, Vol. 8 No. 1, pp. 131-142.

Lee-Kelley, L. and Leong, K. (2003), "Turner's five-functions of project-based management and situational leadership in IT services projects", *International Journal of Project Management*, Vol. 21 No. 8, pp. 583-591.

McCaffery, P. (2010), *The Higher Education Manager's Handbook: Effective Leadership and Management in Universities and Colleges*, RoutledgeFalmer, London.

Mahdavi, I., Mahdavi-Amiri, N., Heidarzade, A. and Nourifar, R. (2008), "Designing a model of fuzzy TOPSIS in multiple criteria decision making", *Applied Mathematics and Computation*, Vol. 206 No. 2, pp. 607-617.

Muller, R. and Turner, J.R. (2007), "Matching the project manager's leadership style to project type", *International Journal of Project Management*, Vol. 25 No. 1, pp. 21-32.

Neary, M. (2012), "Leadership and learning landscapes in higher education: the struggle for the idea of the university", *Higher Education Quarterly*, Vol. 64 No. 4, pp. 333-365.

Pechenizkiy, M., Puuronen, S. and Tsymbal, A. (2008), "Towards more relevance-oriented data mining research", *Intelligent Data Analysis*, Vol. 12 No. 2, pp. 237-249.

Purcell, J. and Wright, P.M. (2007), *The Oxford Handbook Of Human Resource Management*, Oxford University Press, New York, NY.

Saghafian, S. and Hejazi, S.R. (2005), "Multi-criteria group decision making using a modified fuzzy TOPSIS procedure", *International Conference on Computational Intelligence for Modelling, Control and Automation, Vol. 2, Vienna, 28-30 November*, pp. 215-221.

Saidi Mehrabad, M. and Fathian Brojeny, M. (2007), "The development of an expert system for effective selection and appointment of the jobs applicants in human resource management", *Computers & Industrial Engineering*, Vol. 53 No. 2, pp. 306-312.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. (2008), *Global Sensitivity Analysis: The Primer*, Wiley Online Library, Chichester.

Saremi, M., Mousavi, S.F. and Sanayei, A. (2009), "TQM consultant selection in SMEs with TOPSIS under fuzzy environment", *Expert Systems with Applications*, Vol. 36 No. 2, pp. 2742-2749.

Shahhosseini, V. and Sebt, M. (2011), "Competency-based selection and assignment of human resources to construction projects", *Scientia Iranica*, Vol. 18 No. 2, pp. 163-180.

Shih, H.S., Huang, L.C. and Shyur, H.J. (2005), "Recruitment and selection processes through an effective GDSS*", *Computers & Mathematics with Applications*, Vol. 50 Nos 10-12, pp. 1543-1558.

Shipley, M.F. and Johnson, M. (2009), "A fuzzy approach for selecting project membership to achieve cognitive style goals", *European Journal of Operational Research*, Vol. 192 No. 3, pp. 918-928.

Sumner, M., Bock, D. and Giamartino, G. (2006), "Exploring the linkage between the characteristics of IT project leaders and project success", *Information Systems Management*, Vol. 23 No. 4, pp. 43-49.

Thite, M. (1999/2002), "Leadership: a critical success factor in IT project management", paper presented at the International Conference on Management of Engineering and Technology, 1999, Technology and Innovation Management, PICMET '99, Portland, OR, August.

Triantaphyllou, E. (2009), *Data Mining and Knowledge Discovery Approaches Based on Rule Induction Techniques*, 1st ed., Vol. 6, Springer-Verlag, Berlin and Heidelberg, 796pp.

Wu, W.W. and Lee, Y.T. (2007), "Developing global managers' competencies using the fuzzy DEMATEL method", *Expert Systems with Applications*, Vol. 32 No. 2, pp. 499-507.

Zeng, X. and Martinez, T.R. (2004), "Feature weighting using neural networks", paper presented at the Proceeding of 2004 IEEE International Joint Conference on Neural Networks, Budapest, 25-29 July, pp. 1327-1330.

About the authors

Mona Taghavi, received her MSc degree in Information System Management from UKM in 2012. Besides her involvement in several Iranian national ICT research projects, she had worked for an IT consulting and project managing company which was responsible for overseeing and preparing some of the technical reports for the Supreme Council of Information and Communication Technology (SCICT) of Iran. Currently, she is pursuing her PhD at UKM and undertaking research in cooperation with researchers from UKM (Malaysia), Vilnius University (Lithuania), Parand Azad Islamic University (Iran) in the area of expert systems, artificial intelligence, data mining, recommender system, IT outsourcing management, security and so on. She has published 13 papers and is a reviewer of papers for Information Science and Computer Standards & Interface Journals.

Hamed Taghavi received his BSc degree in Computer Engineering and his MSc in Information Technology Management (Advanced Information Systems) from Science and Research Branch, Islamic Azad University of Tehran in 2010. His research interests topics are in the area of Strategic management, risk assessment of IS/IT, Knowledge management and Business Intelligence. He worked as a Knowledge Management System developer in an IT consulting company and he is currently marketing director of MediTechSys Company. Besides, he is senior lecturer of Islamic Azad University and has published several papers.

Milad Taghavi is pursuing his BSc in Information Technology Engineering at Parand Islamic Azad University of Iran. His research interests are E-commerce, Artificial Intelligence and business applications of Information Technology. He is a member of "young research club" and has research collaboration with Universiti Kebangsaan Malaysia.