Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 35 (2014) 1586 – 1595
Procedia Computer Science

18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems - KES2014

The use of intelligent systems for planning and scheduling of product development projects

Marcin Relich[a,*], Wojciech Muszyński[b]

[a]University of Zielona Gora, Licealna 9, Zielona Gora 65-216, Poland
[b]Wrocław University of Technology, Institute of Computer Engineering, Control and Robotics, Wybrzeże Wyspiańskiego 27, Wrocław 50-370, Poland

Abstract

The paper investigates the use of intelligent systems to identify the factors that significantly influence the duration of new product development. These factors are identified on the basis of an internal database of a production enterprise and further used to estimate the duration of phases in product development projects. In the paper, some models and methodologies of the knowledge discovery process are compared and a method of knowledge acquisition from an internal database is proposed. The presented approach is dedicated to industrial enterprises that develop modifications of previous products and are interested in obtaining more precise estimates for project planning and scheduling. The example contains four stages of the knowledge discovery process including data selection, data transformation, data mining, and interpretation of patterns. The example also presents a performance comparison of intelligent systems in the context of variable reduction and preprocessing. Among data mining techniques, artificial neural networks and the fuzzy neural system are chosen to seek relationships between the duration of project phase and other data stored in the information system of an enterprise.

Peer-review under responsibility of KES International.

Keywords: data mining; knowledge acquisition; pattern recognition; fuzzy neural system; artificial neural networks; business intelligence systems

* Corresponding author. Tel.: +48-68-328-2426; fax: +48-68-328-2554.
E-mail address: m.relich@wez.uz.zgora.pl.

Peer-review under responsibility of KES International.
doi:10.1016/j.procs.2014.08.242

1. Introduction

Present information and communication technologies have become one of the most important factors, conditions and opportunities of the company development. These technologies enable the collection, presentation, transfer, access and use of an enormous amount of data. The data is a potential source of information that, in connection with management skills and experience, may support development of a new product. The advancement of information technology in business management processes has increased interest in the use of integrated information systems, such as an Enterprise Resource Planning (ERP) system. ERP systems operate, collect and store data connected with the daily activities of enterprises (e.g. client orders) as well as information concerning previous product development projects[1]. However, ERP software does not include data mining techniques that can discover the complex and potentially useful relationships contained within an ERP database, and support planning and scheduling of product development projects.

New product development (NPD) is a crucial process in maintaining a company's competitive position[2]. Because of its inherent features, NPD is a relatively risky activity[3], as market competition and product technology advancement are often intense[4]. It is reported that more than half of the 20% of successful cases fail to return investment costs and become profitable, with more cost and time than expected having been consumed to achieve the project goals[5,6]. The main reasons why most companies have failed in the development of new products concerns an increase of time and costs in the phases of NPD due to their sequential processes and difficulties in constructing reasonable development schedules and resource distribution plans[6]. These reasons indicate a need for the development of an intelligent system that uses a database of past product development projects to estimate the duration of project phase and improve the quality of planning and scheduling.

The traditional approaches to project scheduling are the well-known CPM (Critical Path Method) and PERT (Program Evaluation and Review Technique). The hypothesis made in CPM is that activity durations are deterministic and known is rarely satisfied in real life where tasks are often uncertain and variable[7]. The inherent uncertainty and imprecision in project scheduling has motivated the proposal of several fuzzy set theory based extensions of activity network scheduling techniques[9-11]. Considerable research effort has also been recently focused on the application of constraint programming frameworks in the context of project scheduling[12-14]. The above-mentioned approaches are based on the durations of project activities, neglecting their estimation from the source data. This provides the motivation to develop an intelligent system that is able to evaluate the duration of project phase, and as a result, support planning and scheduling of new product development.

As real world data is inherently nonlinear, traditional linear tools may suffer from significant biases in data mining[15]. Therefore, this paper considers intelligent systems (neural networks and hybrid fuzzy neural system) for modeling complex data mining problems. These systems are able to solve problems that have imprecise patterns or data containing incomplete and noisy information with a large number of variables. The goal of this paper is to present the use of intelligent systems to identify the relationships in an ERP database between the ERP attributes (e.g. duration of delivery, number of subcontractors or team members) and the duration of a project phase. The sought relationships can lead to more relevant estimates and, as a consequence, improve project planning and scheduling. It is unrealistic to expect very accurate estimations because of the inherent uncertainty in product development projects and the complex and dynamic interaction of factors that influence their development. However, even a small improvement in the estimation quality can have a positive influence on planning project schedule, cost, and resource allocation. The novelty of this research can be considered in the context of identifying the impact of variable reduction and preprocessing on estimation quality in the context of different data mining techniques (artificial neural networks, fuzzy neural system, linear regression). Moreover, the proposed approach uses the firm's experiences concerning past projects which are stored in the ERP database.

The remaining sections of this paper are organised as follows: Section 2 presents a model of knowledge discovery in databases. A method using intelligent systems for data mining in the context of an ERP database is shown in Section 3. An example of the proposed approach, which includes a comparison between intelligent systems and traditional linear tools for estimating the duration of project phase, is illustrated in Section 4. Finally, some concluding remarks are contained in Section 5.

2. Model of knowledge discovery in databases

Since relatively recently, an enormous amount of data has been routinely generated in various sectors of the economy, including business processes in companies. These data sets are not only huge but also often complex and unstructured. Analysis of this data and acquisition of knowledge with the use of manual methods is slow, expensive, subjective, and prone to errors. Hence, there is a need to automate the process through using data mining techniques. Knowledge discovery in databases (KDD) has evolved into a research direction in fields such as databases, machine learning, pattern recognition, statistics, artificial intelligence, reasoning with uncertainty, expert systems, signal processing, and information retrieval[16,17].

KDD is the process of discovering previously unknown and potentially interesting patterns in large databases. As the amount of available data in companies becomes greater and greater, companies have become aware of an opportunity to derive valuable information from their databases, which can then be used to improve their business[18]. KDD is concerned with the development of methods and techniques for making sense of data. KDD is the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data. The basic problem addressed by the KDD process is one of mapping low-level data (which is typically too voluminous to understand and digest easily) into other forms that might be more compact (e.g. a short report), more abstract (e.g. a descriptive approximation or model of the process that generated the data), or more useful (e.g. a predictive model for estimating the value of future cases). At the core of the process is the application of specific data-mining methods for pattern discovery and extraction[19].

Data mining tasks can be classified as descriptive and predictive[16]. While the descriptive techniques provide a summary of the data, the predictive techniques learn from the current data in order to make predictions about the behaviour of new data sets. The most commonly used tasks in data mining include classification, clustering, associations, visualization, summarization, deviation detection, link analysis, and estimation that is further considered.

Table 1 compares the steps according to the most used methodologies for developing data mining and knowledge discovery projects. CRISP-DM (Cross-Industry Standard Process for Data Mining) states which tasks have to be carried out to complete a data mining project[20]. Cios et al adapted the CRISP-DM model to the needs of the academic research community, providing a more general, research-oriented description of the steps model[21]. In turn, the main contribution to the KDD roadmap is the resourcing task.

Table 1. Comparison of DM & KD process models and methodologies[21]

<table>
<tr><th>Model</th><th>Fayyad et al.</th><th>Cabena et al.</th><th>Cios et al.</th><th>CRISP-DM</th><th>KDD Roadmap</th></tr>
<tr><td rowspan="11">Steps</td><td rowspan="2">Developing and Understanding of the Application Domain</td><td rowspan="2">Business Objectives Determination</td><td rowspan="2">Understanding the Data</td><td rowspan="2">Business Understanding</td><td>Problem specification</td></tr>
<tr><td>Resourcing</td></tr>
<tr><td>Creating a Target Data Set</td><td rowspan="5">Data Preparation</td><td rowspan="2">Understanding the Data</td><td rowspan="2">Data Understanding</td><td>Problem specification</td></tr>
<tr><td>Data Cleaning and Pre-processing</td><td>Data cleaning</td></tr>
<tr><td>Data Reduction and Projection</td><td rowspan="3">Preparation of the data</td><td rowspan="3">Data Preparation</td><td>Data cleaning</td></tr>
<tr><td>Choosing the DM Task</td><td rowspan="2">Pre-processing</td></tr>
<tr><td>Choosing the DM Algorithm</td></tr>
<tr><td>DM</td><td>DM</td><td>DM</td><td>Modeling</td><td>DM</td></tr>
<tr><td rowspan="2">Interpreting Mined Patterns</td><td rowspan="2">Domain Knowledge Elicitation</td><td rowspan="2">Evaluation of the Discovered Knowledge</td><td rowspan="2">Evaluation</td><td>Evaluation</td></tr>
<tr><td>Interpretation</td></tr>
<tr><td>Consolidating Discovered Knowledge</td><td>Assimilation of Knowledge</td><td>Using the Discovered Knowledge</td><td>Deployment</td><td>Exploitation</td></tr>
</table>

The steps of the KDD process in the above-presented models can be generalized into four further considered subprocesses: data selection, data transformation / preprocessing, data mining, and interpretation / evaluation of patterns. An ERP database contains hundreds of attributes and an enormous amount of data that can be irrelevant to the mining task or redundant. Fundamental pursuits of data analysis are data and dimensionality reduction[22]. Variable selection is intended to select the optimal subset of predictors, which allows improvements to be made to interpretability and model predictive ability, neglects insignificant effects thereby reducing noise, and speeds up modeling time. Variable selection aims at constructing the simplest possible model that predicts well and/or explains the relationships in the data. It is noteworthy that a variable selection method should take into account the nature of the problem because automatic variable selection is not guaranteed to be consistent with the assumed goals. The variable selection problem is most familiar in the linear regression context, where attention is restricted to normal linear models[23]. Among the best known dimension reduction methods are Mallows Cp, Akaike (AIC) and Bayesian Information Criterion (BIC), Principal Component Analysis, Factor Analysis and Projection Pursuit[22-25].

Knowledge acquisition requires some data mining techniques that can cope with the description of relationships among data and that can solve a problem, in the considered case connected with estimation. Blind application of data mining methods can lead to the discovery of meaningless and invalid patterns[19]. The proposed method for project duration estimation includes two data mining techniques: neural networks and fuzzy sets.

3. Method of project duration estimation with the use of intelligent systems

New product development requires planning and scheduling of project activities to ensure their controlling, and finally to meet the project requirements. Planning and scheduling of projects uses the estimators of the duration of project activities, their sequence, amount of required resources, etc. The quality of these estimators depends on the type of new product as well as the techniques used for forecasting. A high quality of forecasts is rather difficult to obtain for products that are technological breakthroughs, i.e. that are new to the world and will create their own market. In turn, for modified products that are new to the company but not to the market, there is the possibility to use data from previous product development projects. Modifications of existing products are very often used in industrial companies and the development of this type of new product is further considered.

The presented method is dedicated to industrial enterprises that use an ERP system to support their business processes, including development of new products. The phases of product development depend on the characteristics of the product and the company in which it is designed. However, some common phases can be distinguished, for example planning, concept development, system-level design, detail design, testing and refinement, and production ramp-up[6]. These phases can also be considered in the context of concept initiation, program approval, prototype, pilot, and launch. To obtain a project schedule, the specification of resource availability and duration of project phase is required[26,27]. These parameters can be defined by experts or estimated with the use of an ERP database[28,29]. The first approach is suitable for projects that have a unique form, e.g. innovations and construction projects. In turn, if a company develops modifications of previous products, then it is possible to recognize the patterns from an ERP database and to use them for the improvement of estimation quality of project planning and scheduling. The procedure of the proposed method is presented in Fig. 1.

The presented four stages refer to the steps of the KDD process such as data selection, preprocessing, data mining, and use of the discovered patterns. In order to identify the impact of dimension reduction and preprocessing on estimation quality among different data mining techniques, the four cases have been distinguished: the data before variable selection and feature selection (I), before variable selection and after feature selection (II), after variable selection and before feature selection (III), and after variable and feature selection (IV). Variable selection has been conducted according to two criteria: AIC and BIC. In turn, feature selection has been performed with the use of principal component analysis (PCA) that is the best, in the mean-square error sense, linear dimension reduction technique[23].

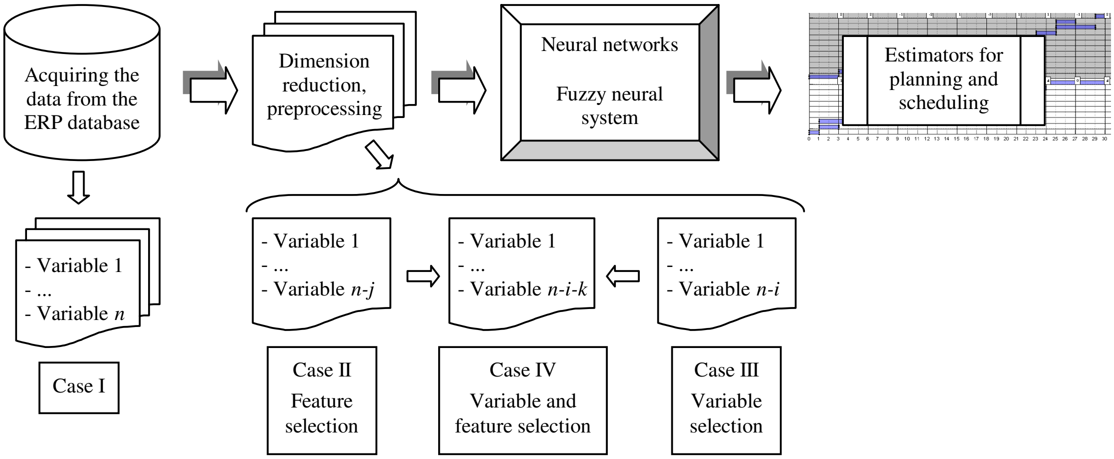

Fig. 1. Stages of the proposed method

An ERP system database comprises an enormous number of attributes that can be considered as potential variables for identifying the duration of developed products. Fuzzy logic and artificial neural networks are complementary technologies and powerful design techniques that can be used in the identification of patterns from within a large database and noisy data.

Artificial neural networks (ANNs) are an important class of tools for quantitative modeling. Today, neural networks are treated as a standard data mining tool and used for many data mining tasks such as pattern recognition, time series analysis, prediction, and clustering. Neural networks are computing models for information processing and are particularly useful for identifying the fundamental relationship among a set of variables or patterns in the data. Several important characteristics of neural networks make them suitable and valuable for data mining. For instance, ANNs do not require several unrealistic *a priori* assumptions about the underlying data generating process and specific model structures, the mathematical property of the neural network in accurately approximating or representing various complex relationships has been well established, ANNs are nonlinear models, and ANNs are able to solve problems that have imprecise patterns or data containing incomplete and noisy information with a large number of variables. ANNs with their nonlinear and nonparametric nature are more effective for modeling complex data mining problems[15].

The fuzzy neural system has the advantages of both neural networks (e.g. learning abilities, optimization abilities and connectionist structures) and fuzzy systems (simplicity of incorporating expert knowledge). As a result, it is possible to bring the low-level learning and computational power of neural networks into fuzzy systems and also high-level human-like IF-THEN thinking and reasoning of fuzzy systems into neural networks. The fuzzy neural method is rather a way to create a fuzzy model from data by some kind of learning method that is motivated by learning procedures used in neural networks. This substantially reduces development time and cost while improving the accuracy of the resulting fuzzy model. Being able to utilize a neural learning algorithm implies that a fuzzy system with linguistic information in its rule base can be updated or adapted using numerical information to gain an even greater advantage over a neural network that cannot make use of linguistic information and behaves as a black box[30]. The combination of fuzzy systems and neural networks has recently become a popular approach in engineering fields for solving problems in control, identification, prediction, pattern recognition, etc[31,32]. One well-known structure is the adaptive neuro-fuzzy inference system (ANFIS), which has been used in the study.

4. Example of project duration assessment with the use of intelligent systems

The example refers to four steps of knowledge discovery: data selection, preprocessing data, data mining, and interpretation. The example also includes the comparison of results for project duration assessment in two perspectives: before and after data selection, as well as before and after preprocessing. The input variables of the prototyping phase in the product development project are presented in Table 2. These variables derive from an enterprise's internal database (ERP system) and concern departments such as purchasing, materials management, production, research and development (R&D) and project management. The development of a product prototype includes many processes that require, for instance, the purchase of materials from the suppliers, storage of materials, and usage of materials in production.

Table 2. Input variables for product prototype phase

Purchasing	Materials Management	Production	R&D	Project Management
Value of material purchase	Number of materials in warehouses	Productive capacity (actual / maximal)	Number of R&D employees leaving the company / total employees in R&D	Number of standard tasks in the project phase
Number of suppliers selling required materials	Number of supplier withdrawal notices	Number of resource overloads	Number of day absence in R&D department / total working days	Number of unique tasks in the project phase
Number of subcontractors	Number of warehouse transfers	Time of machine inspection	Average number of distinct products per design platform	Number of changes in the project phase specification
Delivery duration	Changes of price list	Number of machines	Percentage of existing parts used in new products	Number of materials needed in the project phase
Delay of delivery		Number of work orders	Number of new products introduced during the last year / total products	Number of team members

The first step of the knowledge discovery process concerns data selection and has been divided into two phases. First, the expert chooses the variables according to his/her experience from the enterprise's internal databases. In the considered example these variables are presented in Table 2. Then, the number of variables can be reduced with the use of Akaike information criterion and forward inclusion method. The data set has been reduced from 24 input variables to 6 (number of subcontractors – NoS, delivery duration – DD, number of work orders – NoWO, percentage of existing parts used in new products – EPinNP, number of unique tasks in the project phase – NoUT, number of team members – NoTM).

Data preprocessing is the second step of the knowledge discovery process and helps artificial neural networks and a fuzzy neural system obtain more accurate results. The considered example includes data from 33 previous product development projects (P1-P33). The collected data has a numerical format, but sometimes different units, for example, purchasing is in monetary unit, delivery duration in days, and existing parts used in new products in percent. As a result, the data requires transformation to make the use of data mining techniques more effective. In the presented example, the principal component analysis has been used for preprocessing.

The third step of the knowledge discovery process regards the use of data mining techniques. Among the available intelligence systems two have been chosen: artificial neural networks and the adaptive neuro-fuzzy inference system (ANFIS). In order to eliminate the overtraining of ANNs and ANFIS (too strict function adjustment to data) and to increase the estimation quality, the data set has been divided into learning (P1-P25) and testing sets (P26-P33). The results have been obtained in the Matlab® software and presented in Table 3 as the root mean square errors (RMSE) for the testing set. The results for the intelligence systems are compared with the average and linear model. The performance comparison includes four cases that refer to the data before and after the use of variable selection and feature selection (see Fig. 1).

The results presented in Table 3 indicate that the least error in the testing set for the duration of project prototype phase was generated with the use of ANNs trained according to the Levenberg-Marquardt algorithm (LM). Surprisingly, the least RMSE is for the entire set of data, i.e. before data selection, and after data preprocessing (case II), which can indicate existing complex and hidden patterns in the entire data set. The ANNs trained according to gradient descent momentum and an adaptive learning rate (GDX) algorithm generated worse results than ANFIS. This demonstrated the importance of proper choice of learning algorithm for ANNs and the need of result comparison for different intelligent systems that have ambiguous learning procedures. It is noteworthy that RMSE generated with the use of intelligent systems are smaller than RMSE for the average and linear models. The comparison of different forecasting models is especially recommended in the case of significant variance for the dependant variable (in the considered case for the duration of a project phase).

Table 3. Comparison of RMSE for different models

Model	Case I	Case II	Case III	Case IV
ANN - LM	23.005	1.537	13.009	2.607
ANN - GDX	30.505	3.093	30.388	6.460
ANFIS - grid partition	–	11.217	5.468	4.284
ANFIS - subtractive clustering	16.777	7.570	3.491	3.086
Linear model	43.480	16.425	10.886	7.515
Average		30.596		

In studies, a multilayer feed-forward neural network has been applied by the use of the back-propagation algorithm, and optimisation weights according to the LM and GDX algorithms. The neural network structure has been determined in an experimental way, by the comparison of learning and testing sets for the different number of layers and hidden neurons. RMSE have been calculated as the average of 20 iterations for each structure of neural network with a number to the extent of 20 hidden neurons. Figure 2 presents the comparison of RMSE for different numbers of hidden neurons for networks trained according to the LM algorithm, and in the context of learning and testing sets for four cases respectively. For instance, the minimal value of testing set RMSE has been obtained in case I for 13 hidden neurons in one hidden layer (the top left side of Fig. 2).

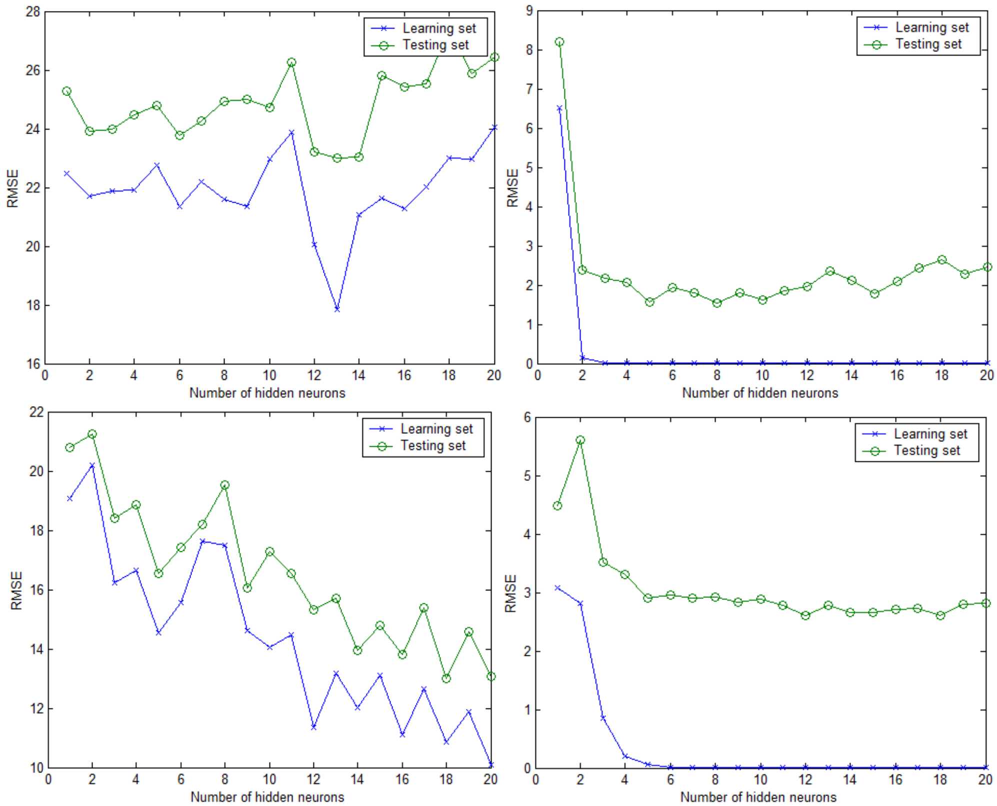

Fig. 2. RMSE and number of hidden neurons for the considered cases (I-IV)

Also the use of a fuzzy neural system requires the declaration of a few parameters concerning the learning phase e.g. a membership function type of fuzzy sets (e.g. triangular, Gaussian function), method of defuzzification and weights optimisation, and stop criterion (e.g. error value or the number of iteration). After the declaration of fuzzy neural structure, the system is learnt according to, for example, the back-propagation algorithm, and consequently, the shape of membership function is determined. The rules are generated according to the shape of membership functions, and can be presented for the decision maker in descriptive form. After learning phase, the testing data is led to input to the system to compare the error between different models (see Table 3). The ANFIS trained according to grid partition method for the entire set of variables (case I) had to be terminated because of generating enormous number of rules. ANFIS trained with the use of subtractive clustering generated 21 rules for case I. In turn the number of rules for case II-IV equal 729 and 14 for grid partition method and subtractive clustering, respectively.

The membership functions and rules are a basis for evaluating the duration of an actual product development project. Let us assume that for the actual project the following values are considered: number of subcontractors – 5, average delivery duration – 21 days, number of work orders – 4, number of team members – 5, number of unique tasks in the project phase – 4, and percentage of existing parts used in new products – 0.205. Figure 3 presents the memberships functions for 12 rules that determine the planned duration of project phase at 120 days.

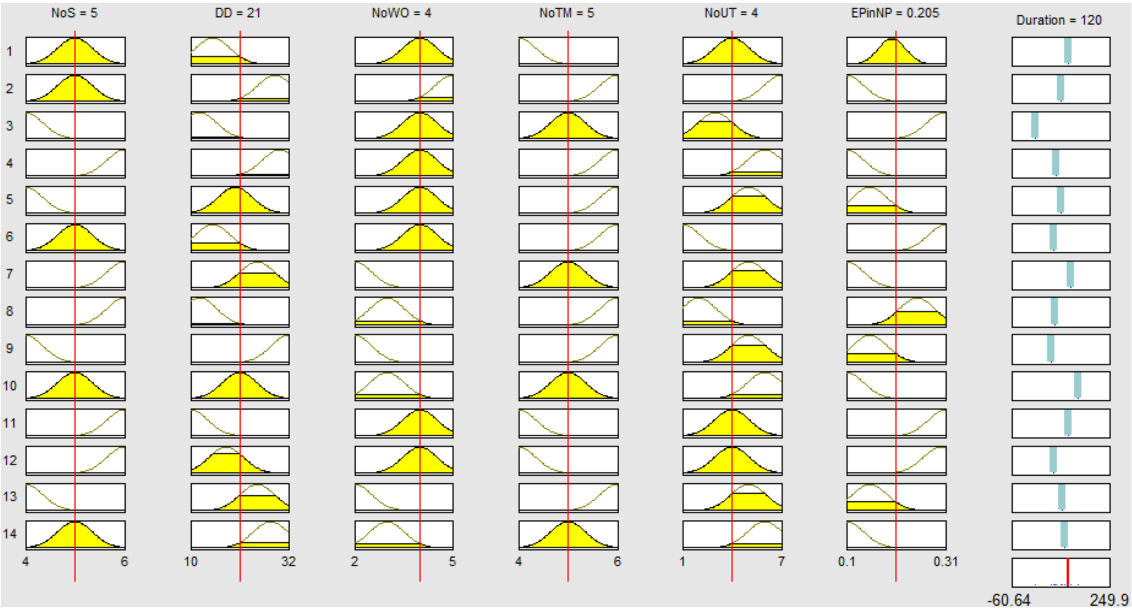

Fig. 3. Estimation of project duration with the use of fuzzy neural system

The presented analysis can be broadened into multidimensional analysis to support the decision-maker in determining optimal values of some parameters. For instance, the decision-maker wants to investigate the impact of the number of team members (from 5 to 6) and the number of subcontractors (from 5 to 6) on the planned duration of a project phase. Table 4 presents the results for these constraints that indicate a decline of the duration of project phase with an increase in the number of subcontractors or/and team members.

Table 4. What-if analysis for project duration (in days)

Number of subcontractors	Number of team members	
	5	6
5	120	98.3
6	95.9	87.5

The above-presented study is conducted for each phase of the product development project and the obtained estimates can be used for further evaluation, for instance, in the project scheduling, cost planning, and cash flow projections.

5. Conclusions

The enormous amount of data and attributes in an ERP database require attention to be paid to the proper choice of variable reduction and data mining methods. In the paper, various data mining and knowledge discovery models and methodologies have been compared and generalized into four steps: data selection, data transformation, data mining, and using patterns. Among the data mining techniques, two have been chosen – artificial neural networks and the fuzzy neural system. These techniques have been applied to estimating relationships between the duration of the previous product development projects and the selected attributes from an ERP database. The results indicate better estimation quality for intelligent systems than statistical models, and that variable selection and preprocessing significantly influences on the obtained results. The experiments also show the importance of careful choice of parameters of intelligent systems, for instance, the structure and learning algorithm of ANN, and the need for result comparison for different forecasting models.

The advantages of the use of intelligence systems include the search of complex and potentially useful relationships into an ERP database and their use in the estimation of project duration. More exact identification of project duration enables more precision of project planning and scheduling, as well as the improvement of cost planning. This approach is especially useful for production enterprises that have a database of previous product development projects. The application of the proposed approach encounters some difficulties, among other things, with the collection of sufficient amounts of data from past similar projects. Moreover, the lack of uniform rules for building the structure of neural networks and fuzzy neural systems may cause an acceptance problem for the decision-makers. However, the presented approach seems to have promising properties for acquiring information from an ERP system. Further research focuses on the development of the proposed method of project duration estimation towards the choice of an optimal set of new products for development.

References

1. Relich M. *Knowledge acquisition for new product development with the use of an ERP database*. In: Proceedings of the Federated Conference on Computer Science and Information Systems; 2013. p. 1285-1290.
2. Chin KS, Tang DW, Yang JB, Wong SY, Wang H. Assessing new product development project risk by Bayesian network with a systematic probability generation methodology. *Expert Systems with Applications* 2009; **36**:9879-9890.
3. Kahraman C, Buyukozkan G, Ates NY. A two phase multi-attribute decision-making approach for new product introduction. *Information Sciences* 2007; **177**(7):1567-1582.
4. McCarthy IP, Tsinopoulos C, Allen P, Rose-Anderssen C. New product development as a complex adaptive system of decisions. *Journal of Product Innovation Management* 2006; **23**(5): 437-456.
5. Cooper LP. A research agenda to reduce risk in new product development through knowledge management: a practitioner perspective. *Journal of Engineering and Technology Management* 2003; **20**(1):117-140.
6. Park S, Kim J, Choi HG. *A risk management system framework for new product development*. In: International Conference on Economics and Finance Research, IPEDR 4, IACSIT Press; 2011. p. 51-56.
7. Zammori FA., Braglia M, Frosolini M. A fuzzy multi-criteria approach for critical path definition. *International Journal of Project Management* 2009; **27**:278-291.
8. Bonnal P, Gaurc K, Lacoste G. Where do we stand with fuzzy project scheduling? *Journal of Construction Engineering and Management* 2004; **130**:114-123.
9. Long LD, Ohsato A. Fuzzy critical chain method for project scheduling under resource constraints and uncertainty. *International Journal of Project Management* 2008; **26**:688-698.
10. Fortin J, Zielinski P, Dubois D, Fargier F. Criticality analysis of activity networks under interval uncertainty. *Journal of Scheduling* 2010; **13**:609-627.
11. Maravas A, Pantouvakis JP. A fuzzy repetitive scheduling method for projects with repeating activities. *Journal of Construction Engineering and Management* 2011; **137**:561-564.
12. Sitek P, Wikarek J. *A hybrid method for modeling and solving constrained search problems*. In: Proceedings of the Federated Conference on Computer Science and Information Systems; 2013. p. 385-392.

13. Relich M. A declarative approach to new product development in the automotive industry. In: *Environmental Issues in Automotive Industry*. Berlin Heidelberg: Springer; 2014. p. 23-45.
14. Bach I, Bocewicz G, Banaszak Z, Muszyński W. Knowledge based and CP-driven approach applied to multi product small-size production flow. *Control & Cybernetics* 2010; **39**(1):69-95.
15. Zhang GP. Neural Networks For Data Mining. In: *Data mining and knowledge discovery handbook*. 2nd ed. Springer; 2010.
16. Han J, Kamber M. *Data mining. Concepts and techniques*. 2nd ed. San Francisco: Morgan Kaufmann Publishers; 2006.
17. Cios KJ., Pedrycz W, Swiniarski RW, Kurgan LA. *Data mining: a knowledge discovery approach*. New York: Springer; 2007.
18. Li T, Ruan D. An extended process model of knowledge discovery in database. *Journal of Enterprise Information Management* 2007; **20**(2):169-177.
19. Fayyad U, Piatetsky-Shapiro G, Smith P. From data mining to knowledge discovery in databases. *American Association for Artificial Intelligence*; Fall 1996. p. 37-54.
20. Cabena P, Hadjinian P, Stadler R, Verhees J, Zanasi A. *Discovering data mining: from concepts to implementation*. Prentice Hall; 1998.
21. Marban O, Mariscal G, Segovia J. A data mining & knowledge discovery process model. In: *Data Mining and Knowledge Discovery in Real Life Applications*, Vienna: I-Tech; 2009.
22. Pedrycz W. Data and dimensionality reduction in data analysis and system modeling. In: *Encyclopedia of Complexity and Systems Science*, Springer; 2009. p. 1775-1789.
23. Chizi B, Maimon O. Dimension reduction and feature selection. In: *Data Mining and Knowledge Discovery Handbook*. 2nd ed. Springer; 2010.
24. Guyon I., Elisseeff A. An introduction to variable and feature selection. *Journal of Machine Learning Research* 2003; **3**:1157-1182.
25. Li T, Ruan D. An extended process model of knowledge discovery in database. *Journal of Enterprise Information Management* 2007; **20**(2):169-177.
26. Bach I, Bocewicz G, Banaszak Z. Constraint programming approach to time-window and multiresource-constrained projects portfolio prototyping. In: *New Frontiers in Applied Artificial Intelligence.* Berlin Heidelberg: Springer; 2008. p. 767-776.
27. Bocewicz G, Nielsen I, Banaszak Z. Automated guided vehicles fleet match-up scheduling with production flow constraints. *Engineering Applications of Artificial Intelligence* 2014; **30**:49-62.
28. Relich M, Jakabova M. *A decision support tool for project portfolio management with imprecise data*. In: Proceedings of the 10th International Conference on Strategic Management and its Support by Information Systems; 2013. p. 164-172.
29. Sitek P, Wikarek J. *A hybrid approach to modeling and optimization for supply chain management with multimodal transport.* In: Proceedings of the 18th International Conference on Methods and Models in Automation and Robotics (MMAR); 2013. p. 777-782.
30. Azar AT. Adaptive neuro-fuzzy systems. In: *Fuzzy systems*. InTech; 2010.
31. Cheng MY, Tsai HC, Sudjono E. Evolutionary fuzzy hybrid neural network for project cash flow control. *Engineering Applications of Artificial Intelligence* 2010; **23**:604-613.
32. Chien SC, Wang TY, Lin SL. Application of neuro-fuzzy networks to forecast innovation performance. *Expert Systems with Applications* 2010; **37**:1086-1095.