

Using Information Technology for Human Resource Management Decisions

MOJCA BERNIK, JOŽE FLORJANČIČ, DUŠAN ČRNIGOJ, IGOR BERNIK

Faculty of Organisational Science

University of Maribor

Kidriceva 55a, SI- 4000 Kranj

SLOVENIA

Abstract: The usage of information technology in human resource management is presented in the article. New information techniques, such as data mining, can be integrated in the human resource management information system. They help us to explain the existing human resource processes, predict or classify the data and thus support the human resource management decisions. The result is integration of information into a whole and the subsequent creation of new human resource knowledge.

Key-Words: Information technology, human resource management, decision making, data mining

1 Introduction

The fast and unstoppable development of information and communications technology that we have witnessed in the past few years influenced not only the quality of an individual's life, but also enabled the individual to discover complex insight on self, the human race and the world, thus increasing their creativity and innovativeness in all levels of life. There have been changes especially on social science disciplines, mostly on those, who are investigating human way of life and work. Human resource discipline as one of them is therefore investigating influence of usage information and communications technology, which allows not only faster acquirement of information, but offers extra help at decision making on human resource field.

As we are now in the knowledge era, the basic element of human resources (HR) research is the acquirement of knowledge. The amount of information currently available is huge and the problem of the HR management is mainly the filtering and integration of information. Numerous theoreticians in the area of HR management that delved in this problem have reached the conclusion that it's not only the acquirement of information, but mostly its management, stemming from the fact that information is one of the basic resources of every organization. Like an organization manages material, human and other resources, it also manages the information that flow either within the company or outside. Because of that, information management has recently become one of the priority tasks of HR management.

The knowledge management that HR management has to face is in a wider respect composed of two elements: the people and the information technology (IT) [1]. The people are the knowledge carriers and the IT enables them to amass the knowledge, organize it, access

it and use it. The HR management uses IT mostly as the HR information system (IS), which is connected to other information systems in the organizations and other external IS [2]. In this situation the point is that management can use the knowledge to create added value by getting the needed information at the right time and at the right place. Therefore the modern HR information systems are being upgraded to include decision support systems (DSS) that help to find different aspects in problem solving. The probably most important benefit of DSS is the transparency of produced solutions that enables close to optimal solutions even in the so-called "fuzzy" areas such as human resources. Some of the DSS are based on the knowledge discovery in databases methods (KDD) that is presented in the remainder of the article in more detail with the example of its use in the analysis of absenteeism - one of the most often analysed unwanted phenomena that the HR management encounters.

2 Absenteeism in the HR management

Many researches in various fields treat the problem of absenteeism. This is the reason for the multitude of definitions of absenteeism, the most basic being any form of temporary absence from work. Most researchers consider absenteeism in a narrow sense of sick leaves, also called "medical absenteeism". The results of medical absenteeism are lost days of work or a time when the employee is temporarily unable to work due to disease or injury. This distinguishes medical absenteeism from e.g. permanent disabilities that lead to permanent absence, even if the permanent disability condition may start as a disease or an injury that leads only to medical absenteeism at the beginning.

In a wider sense, absenteeism includes approved leaves, such as holidays, study leaves, special leaves etc. and the unapproved leaves, such as being late to work, leaving work during work hours without permission, skipping entire days etc. [3]. A separate problem is the so-called hidden absenteeism that is formally classified as medical absenteeism, however the reasons for absence are dysfunctional family relations, poor financial state of the family, lack of organized day-care for the children or elderly, long distance to work, menial labour, poor employee training, poor motivation or stimulation, lack of scheduled lunch breaks etc. These forms of absenteeism are hard to track and control as they're usually undetected. Thus they are absent from records even though their influence on the level absenteeism is severe.

To reduce absenteeism the organizations should have a good overview of the types of absenteeism and not only statistical data on work leaves. They should also employ modern methods for detection of hidden absenteeism. They should discover the quantity of detection, its reasons and suggest measures to decrease it, based on the discovered facts. Data mining based DSS can be of a big help for the HR management.

Most organizations regularly tracks the unapproved daily absence, that is not based on the health of employees but things such as poor relations, job satisfaction, family problems and similar. For example a financial IS requires the data on the duration and type of job absence to calculate wages. If the organization lacks an IT supported HR information system, the files are plain paper files.

The content and tracking of the absenteeism indicators depend on the company size and its activities. In an organization with only a couple of dozen employees, the number of tracked indicators is smaller than in an organization with several hundred or thousand employees. The reason is the smaller amount of HR experts [4] in charge of tracking and management of the problem and the detailed knowledge of every employees work. In small companies, absenteeism is usually not a large problem as the small workforce requires employees to show up regularly, and all absences are quickly noticed, as the employees know each other well.

The situation in medium and large organizations is different, as the low intensity of interpersonal relations and poor knowledge of other people's work leads to higher absenteeism. Employees don't feel sufficient loyalty and aren't motivated to do their best. Conversely, the size of an organization influences which indicators of absenteeism and in what detail they will be monitored. Therefore in production based organizations the monitoring of absenteeism indicators is much more detailed than in service based organizations as the level of absenteeism is much higher in production based

companies due to physical nature of work and the work conditions in production.

The data on absenteeism in the HR information system mostly refer to the absentee and the time of absence. The disability data is kept separate and are not counted as absenteeism. The category of data that would describe the reasons for absenteeism is not present in most companies as only the times and types of absenteeism are recorded. This is why the hidden absenteeism is hard to detect and reduce.

The HR information system data on absenteeism are mostly built around the identity of the absentee and the time of absence. The disability data is managed separately and is not connected to absenteeism. Most HR information systems do not include a data entity that would describe the reasons for absenteeism, but only times and types of absence. Because of this, hidden absenteeism is sometimes very hard to detect and reduce. One of the ways to discover the reasons of absenteeism and link them together is the use of data mining methodology, which when coupled with the HR IS enables us to:

- Transform data into meaningful information,
- Transform information into knowledge and link it to the process content,
- Exchange knowledge with other participants in the process.

3 A showcase of the data mining methodology use in the research of absenteeism

The HR management uses the HR data that may be located in the central database of the organization or solely inside the HR IS, which is kept separate from the central database. Regardless of the location, the HR database mostly contains data on the employee and the job position that are sorted into these categories [5]:

- basic employee data,
- data on professional, work related and personal development,
- data on work and job performance,
- HR social data,
- work safety data, data on managerial and societal activities of the employees.

These data are connected to absenteeism either directly or indirectly. With the aid of data mining tools (e.g. Weka, Intelligent Miner, etc.) that perform data mining by building decision trees, the mass of data is analysed to find the links and rules in the data. In the nodes of the tree attribute variables are processed, while leaves represent classes. A tree is built on the base of known reasons for absenteeism and enables us to

classify new reasons. The tree construction starts at the root and moves towards the leaves.

In the case of a decision tree meant for the search for reasons for absenteeism we have used a classifying decision tree and the data were processed with the Weka tool [6] that allows us to enter values as either discrete or continuous variables. The basic principle of tree construction is the separation into very “clean” subsets. A subset is clean when all its cases belong to a single class. In such a case branching is halted and a leaf is created.

Due to a large amount of data a decision tree can become very complex. If we have a complex tree with lots of classes, we may doubt the choice of attributes and the reliability of lower branches [7]. In such a case it makes sense to reduce the number of attributes and try to build a smaller tree. The advantage of smaller trees is their transparency and comprehensibility.

In the further part of the paper a case of absenteeism data processing using the Weka tool is shown. Fig. 1 shows the attribute classifier within Weka, which enables us to build trees.

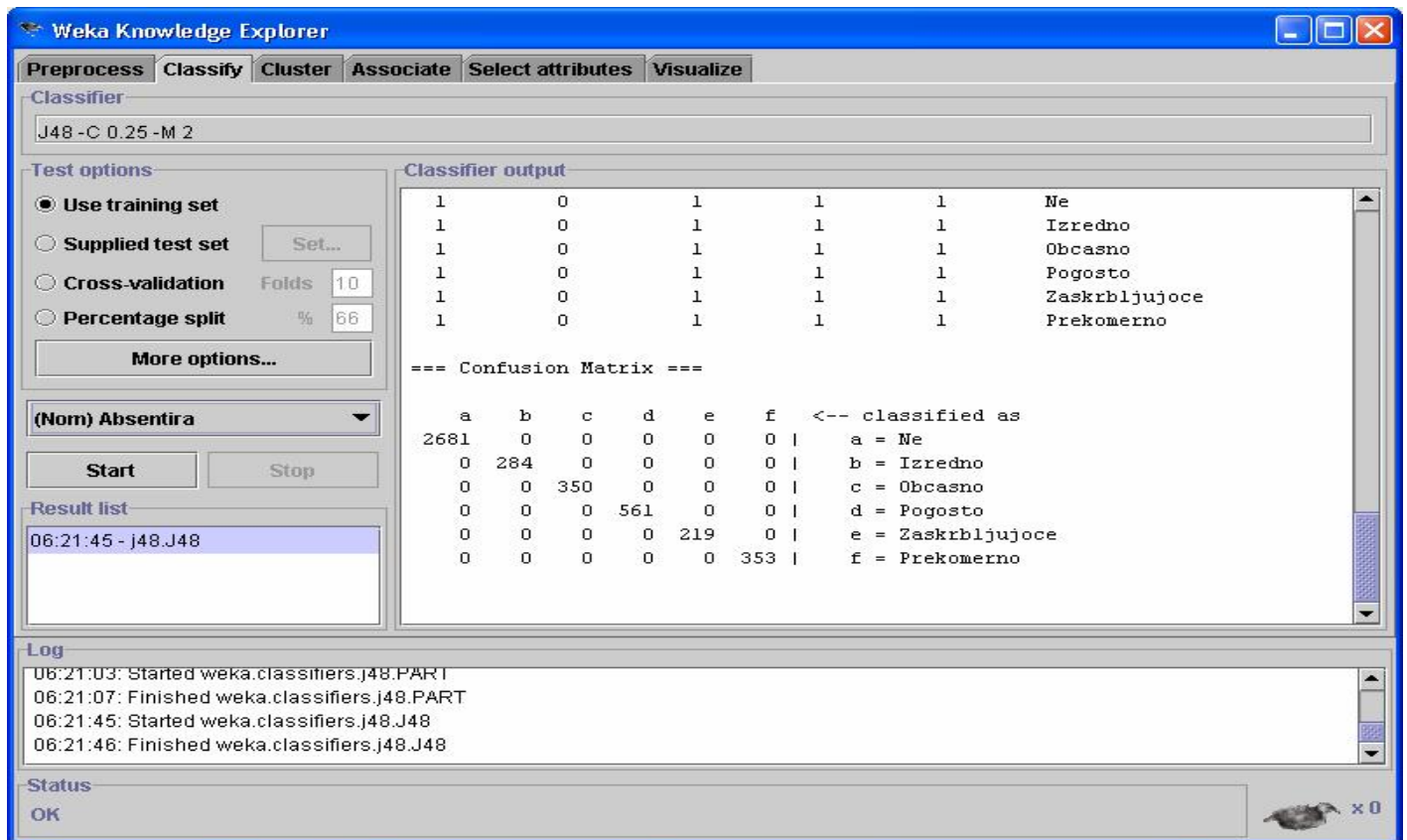


Fig. 1. Weka Knowledge Explorer with classified data

Weka Knowledge Explorer is a graphical user interface which contains all the main tabs of Weka: Preprocess, Classify, Cluster, Associate, Select attributes and Visualize. The Preprocess tab is the starting point for rules discovery. It displays the analysis-ready data, attribute characteristics, lets us select the attribute for classification and tree building and apply an arbitrary combination of filters. The Classify tab lets us configure and use any of the Weka’s classifiers on the given data set. We can select cross checking or a separate data set test. Classification errors and decision trees can also be displayed. Similarly, in the Cluster tab Weka’s data sorting methods are applied. The Associate tab searches for association rules and the Select attributes tab lets us select the most appropriate attributes in a data set by

using any combination of attribute evaluation and search methods. The Visualise tab lets us display the data in one or two-dimensional graphs. For discrete attributes, values are depicted by different discrete colours, while continuous attributes are shown using a spectrum of colours. This tab also lets us visualise forecasts (it can be opened independently of the Classify/Cluster tab). If a class is discrete, the wrong classification points are shown in a frame in a colour of the predicted class. If a class is discrete, the point size depends on the size of the classifier’s error [6]. Figure 2 displays a tree of HR data analysis of medical absenteeism. The tree has a high classification precision, is comprehensible and is sized proportionally to the attributes.

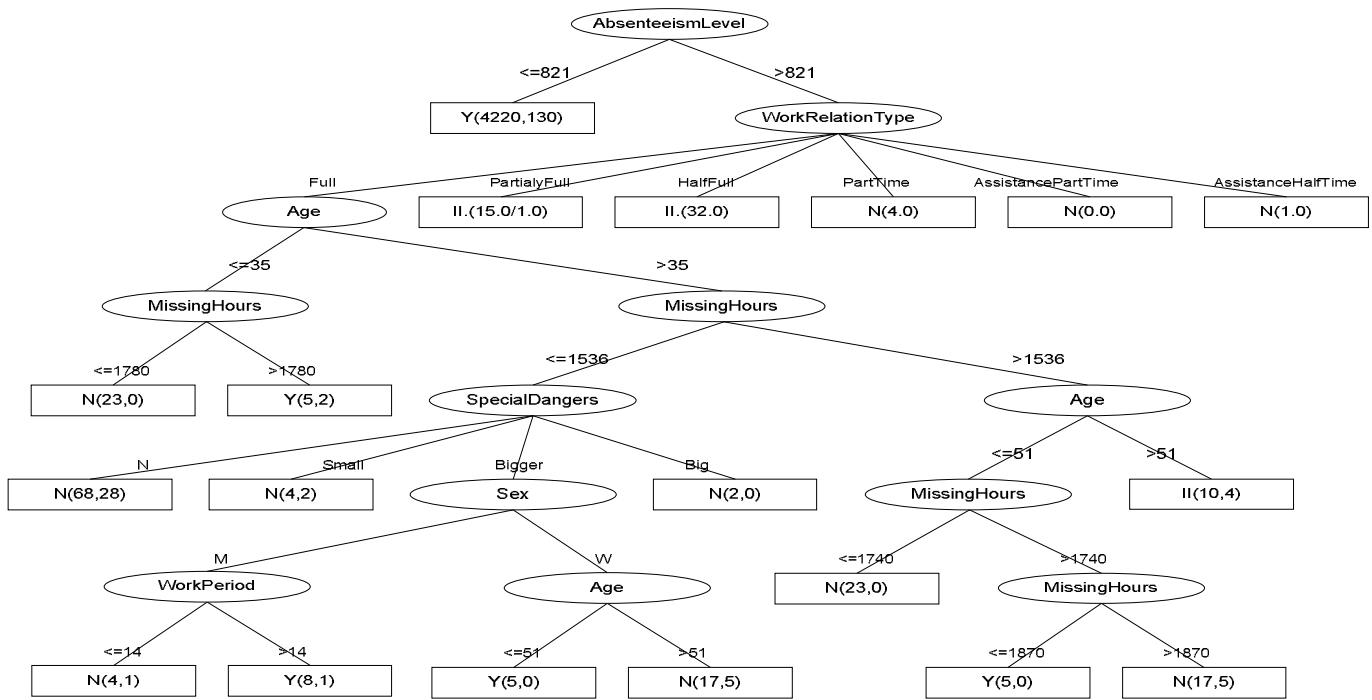


Fig. 2. A decision tree – classification of absenteeism by the absenteeism level

We may conclude from the constructed tree that absenteeism level is a link between the total missing hours, work period and job special dangers, which is evident from the high classification precision and the tree comprehensibility. Apart from the discovery of reasons for absenteeism and related connections, we can use the shown methodology to find other reasons that are linked to absenteeism in any way.

4 Conclusion

The shown case of knowledge discovery in data methods and the explanation of absenteeism data suggest that the described methodology is appropriate to be used in human resources management. It allows us to evaluate human resources data from several aspects, which gives a new dimension and useful value to the data. The methodology of knowledge discovery in data can be simply used within the modern human resources information systems, as the user is not burdened with data analysis and can concentrate on the discovery of links and rules in the data. When coupled with appropriate explanations of the collected data, this leads to better decision making.

The described methodology is not useful only in the human resources, but in all areas where an explanation of data, collected from the transactional parts of the IS, is needed in the decision making. With the explanation based on the discovered rules the data is better understood and that means we can make faster and better decisions.

References:

- [1] Bahra N.: Competitive knowledge management, Palgrave 2001
- [2] Florjančič J., Bernik M., Bernik I.: Human Resource Information System Development, Organization and Management: Selected Topics, ed.: Florjančič J., Paape B., Peter Lang, 2003, pp. 181 – 193
- [3] Dierendonck D., LeBlanc P. M., Breukelen W.: Supervisory behavior, reciprocity and subordinate absenteeism, Leadership & Organization Development Journal, Vol. 23, No. 2, 2002, pp. 84 - 92
- [4] Florjančič J., Bernik M., Bernik I.: Personnel Planning, Organization and Management: Selected Topics, ed.: Florjančič J., Paape B., Peter Lang, 2003, pp. 13 – 45
- [5] McLeod R., DeSanctis G.: Resource Flow Model of the Human Resource Information System, Journal of Information Technology Management, No. 3, 1995, pp. 1-15
- [6] Weka: Weka3, Machine Learning Software in Java, Department of Computer Science University of Waikato, Hamilton, New Zealand, (<http://www.cs.waikato.ac.nz/~ml/weka/>), 2006
- [7] Witten I.H., Frank E.: Data Mining – Practical Machine Learning Tools and Techniques with Java Implementations, Morgan Kaufmann Publishers, 2006