Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 34 (2008) 280–290
Expert Systems with Applications
www.elsevier.com/locate/eswa

Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry

Chen-Fu Chien [a,*], Li-Fei Chen [a,b]

[a] Department of Industrial Engineering and Engineering Management, National Tsing Hua University, 101 Section 2 Kuang Fu Road, Hsinchu 300, Taiwan, ROC
[b] Department of Industrial Engineering and Management, Tahua Institute of Technology, 1 Ta-Hwa Road, Chung-Lin, Hsinchu 307, Taiwan, ROC

Abstract

The quality of human capital is crucial for high-tech companies to maintain competitive advantages in knowledge economy era. However, high-technology companies suffering from high turnover rates often find it hard to recruit the right talents. In addition to conventional human resource management approaches, there is an urgent need to develop effective personnel selection mechanism to find the talents who are the most suitable to their own organizations. This study aims to fill the gap by developing a data mining framework based on decision tree and association rules to generate useful rules for personnel selection. The results can provide decision rules relating personnel information with work performance and retention. An empirical study was conducted in a semiconductor company to support their hiring decision for indirect labors including engineers and managers with different job functions. The results demonstrated the practical viability of this approach. Moreover, based on discussions among domain experts and data miner, specific recruitment and human resource management strategies were created from the results.

Keywords: Personnel selection; Human capital; Data mining; Decision tree; Semiconductor industry

1. Introduction

Human capital is one of the core competences for high-tech companies to maintain their competitive advantages in the knowledge economy. Personnel recruitment and selection directly affect the quality of employees. Hence, various studies have been conducted on resumes, interviews, assessment centers, job knowledge tests, work sample tests, cognitive tests, and personality tests in human resource management to help organizations make better personnel selection decisions. Indeed, the existing selection approaches focus on work and job analysis that are defined via specific tasks and duties based on their static properties.

However, owing to the changing nature of knowledge workers in high-tech industry, jobs cannot be easily delineated especially for jobs in the management level. As globalization and technology advance, cross-functional tasks are also increased while new jobs are also constantly created. The requirements of personnel quality in high-technology companies are increasingly strict, while the work processes in these companies are becoming diversified and complicated. Thus, the conventional personnel selection approaches that are developed on the basis of static job characteristics will no longer suffice (Lievens, Van Dam, & Anderson, 2002). In order to find the right people to do the right things for the right jobs, developing effective selection approaches is very critical.

A high-tech industry such as semiconductor industry has many unique or unusual characteristics including complex and highly uncertain manufacturing processes, short product life cycles, low yield problems, and difficulties in

* Corresponding author. Tel.: +886 3 5742648; fax: +886 3 5722685.
E-mail address: cfchien@mx.nthu.edu.tw (C.-F. Chien).

0957-4174/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2006.09.003

acquiring human capital (Chien & Wu, 2003; Sattler & Sohoni, 1999). Thus, the quality of their human resource is very crucial in increasing their competitiveness. In addition, Appleyard and Brown (2001) analyzed the firm-level data from semiconductor manufacturers in the United States, Asia, and Europe and found that engineers play important and growing roles in creating high-performance semiconductor factories. Nevertheless, semiconductor companies, as well as other high-technology companies, often suffer from high turnover rates and difficulties in recruiting the right talents. In order to attract good applicants, companies provide attractive compensation and welfare benefits. However, despite the willingness of many companies to do all that they can to recruit the best people, they usually have difficulties at the selection stage in predicting which applicants would have better work performance and would have longer service time after they are hired. Therefore, selecting the right engineers who can demonstrate the best performance and who will stay with the company for a long time is of great urgency for every high-technology company.

Recently, owing to the advancements in information technology, researchers have developed decision support systems and expert systems to improve the outcomes of human resource management. In particular, data mining is recognized as one of the most salient topics. Data mining refers to the extraction of useful patterns or rules from a large database through an automatic or semi-automatic exploration and analysis of data (Berry & Linoff, 1997; Chen, Han, & Yu, 1996). With the help of data mining techniques, computers are no longer limited to passively storing or collecting data. They can also help the users to actively excerpt the key points from huge amounts of data, and make use of analysis or prediction. Data mining techniques have been widely applied in many fields and have exhibited outstanding results. However, the applications of data mining in the semiconductor industry are mostly related to engineering data analysis and yield enhancement (Braha & Shmilovici, 2002; Kusiak, 2001; Chien, Hsiao, & Wang, 2004; Chien, Wang, & Cheng, 2007). Little research has been done in human resource management.

This study aims to develop a data mining framework for personnel selection to explore the association rules between personnel characteristics and work behaviors, including work performance and retention. An empirical study for indirect labor (IDL) including engineers with different job functions in one of the world largest semiconductor foundry company located in the Hsinchu Science Park in Taiwan is studied to demonstrate the validity of this approach. In particular, we employ decision tree analysis to discover latent knowledge and extract the rules to assist in personnel selection decisions. Furthermore, using the information gathered, domain experts from this company can also generate recruiting and human resource management strategies. Some of the findings have been implemented in this company and the results have shown the practical viability of this approach.

2. Fundamentals

2.1. Personnel selection

Personnel selection plays a decisive role in human resource management in which it will determine the input quality of personnel. Researchers (Borman, Hanson, & Hedge, 1997; Robertson & Smith, 2001) reviewed the personnel selection studies and found that the important issues including change in organizations, change in work, change in personnel, change in the society, change of laws, and change in marketing have influenced personnel selection and recruiting. Hough and Oswald (2000) also reviewed personnel selection studies from 1995 through 1999 and concluded that the nature and analysis of work behavior are changing and hence affecting personnel selection practices. Lievens et al. (2002) identified challenges in personnel selection including labor market shortages, technological developments, applicant perception of selection procedures, and construct-driven approaches.

Meanwhile, advancements in information technology are also affecting personal selection as well as human resource management (Beckers & Bsat, 2002; Kovach & Cathcart, 1999; Liao, 2003). The applications of expert systems or decision support systems on personnel selection and recruitment are increasing (e.g., Hooper, Galvin, Kilmer, & Liebowitz, 1998; Nussbaum et al., 1999). However, little research has employed data mining approaches for personnel selection as the present study does.

2.2. Data mining

Date mining methodologies have been developed for exploration and analysis, by automatic or semi-automatic means, of large quantities of data to discover meaningful patterns and rules. Indeed, such data including personnel data can provide a rich resource for knowledge discovery and decision support. Therefore, data mining is discovery-driven not assumption-driven. Data mining involves various techniques including statistics, neural networks, decision tree, genetic algorithm, and visualization techniques that have been developed over the years.

Data mining problems are generally categorized as association, clustering, classification, and prediction (Fayyad, Piatesky-Shapiro, & Smyth, 1996; Fu, 1997; Han & Kamber, 2001). Association is the discovery of association rules showing attribute-value conditions that occur frequently together in a given dataset. Clustering is the process of dividing a dataset into several clusters in which the intra-class similarity is maximized while the inter-class similarity is minimized. Classification derives a function or model that identifies the categorical class of an object based on its attributes. Prediction is a model that predicts a continuous value or future data trends.