Chapter VIII

Knowledge Miner: Assisting in Pattern Discovery and Prediction

Abstract

This chapter addresses the issue of knowledge discovery from data. It reports the results of an empirical investigation of the effectiveness of a specific knowledge-mining tool in discovering and predicting trends from time-series data. The results obtained clearly show that people are capable of enhancing knowledge and subsequent performance using this technology, but they also indicate that there is still room for further improvement. Hence, more research is needed in this area to achieve increased efficiency.
Introduction

The growing amount of data being generated by electronic and traditional transactions between customers, suppliers, and other trading partners represents a potentially valuable source of new knowledge for organisations. Therefore, the discovery of knowledge that is implicit in data is one of the major issues that needs to be addressed in research (Blanning, 2000).

Knowledge discovery has been described as the nontrivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data (Fayyad, Piatetsky-Shapiro, & Smyth, 1996). The term knowledge mining is often used interchangeably with knowledge discovery (Berry & Linoff, 1997). Patterns in data can be extracted by one or more knowledge discovery tasks. A taxonomy suggested by Shaw et al (Shaw, Subramaniam, Tan, & Welge, 1999) includes five categories: dependency analysis, class identification, concept description, deviation detection, and data visualisation. The focus of this study is on deviation detection in combination with visualisation.

It is argued that discovery systems using statistical techniques, such as linear regression, may help forecasters to detect systematic trends (Makridakis, Wheelwright, & McGee, 1983). The parameters of the linear regression equation minimise the mean error values. Thus, a regression line incorporated into the tool design may be a potentially valuable aid to the knowledge discovery process. To explore the knowledge in data more effectively, data visualisation can be used in association with statistical techniques.

Keim (1996) has provided an elaborate analysis of visualisation techniques and classified them as pixel-oriented, geometric projection, and graph based. Of particular interest in this study are graphs. The basic idea of the graph-based technique is to effectively present a graph using a specific layout algorithm, query languages, and abstraction techniques. It has been shown in past research that graphical presentation enhanced accuracy of novices (Lawrence, Edmundson, & O’Connor, 1985). So, one can assume, for example, that a line graph representing a series of historic sales data points may allow a forecaster to view the underlying pattern in product sales over time and help extrapolate that pattern in an estimate of future sales.

The purpose of this chapter is to address this issue of knowledge discovery by describing and testing a specific knowledge discovery tool in the sales forecasting context. In particular, the study reported here examined: (a) whether and how effective was our knowledge discovery tool in enhancing people’s ability to recognise trends in time-series data, and (b) what impact the extracted knowledge had on people’s subsequent forecasting performance.
A PKM–Based Decision-Making Training Program for Personal Healthcare: An Action Learning Approach
www.igi-global.com/article/a-pkm-based-decision-making-training-program-for-personal-healthcare/210689?camid=4v1a

Knowledge Communication
www.igi-global.com/chapter/knowledge-communication/16967?camid=4v1a