Chapter XXV
Automation vs. Human Intervention: Is There any Room Left for the Analyst in the Data Mining Process?

Meryem Sevinc
Georgia Southern University, USA

Lawrence Locker
Georgia Southern University, USA

John D. Murray
Georgia Southern University, USA

ABSTRACT

In the contemporary context of knowledge discovery, the amount of information and the process itself has increased in complexity. Relevant to the present chapter is the increased reliance on automaticity in knowledge discovery. Although, there are positive benefits of automation, there is reason to believe that a process that emphasizes greater human participation may produce more meaningful results. Through a description of the human information processing system and its attributes, this chapter discusses why an analyst-centered approach to a knowledge discovery system is a desirable goal. We argue that a perspective based on cognitive psychology can serve as a useful guide in achieving a desirable synergy between automated knowledge discovery tools and the human analyst.
INTRODUCTION

In the present technological age, there is an increasing need in complex organizations for the rapid acquisition, interpretation, and practical application of data. Fifty years ago, it was considered a great success for organizations to be able to answer a question such as “What were their revenue had been over the previous four years.” Today, however, the questions are much more sophisticated, such as “What is the estimated unit sales over the next ten months?” and “What are the reasons behind these projections?” Technological advances such as efficient computer systems and the World Wide Web (WWW), now allow organization analysts to access easily enormous data sets which can, in turn, be analyzed in any number of ways that can be helpful to an organization. For example, a retail company could use available data to gain a better understanding of customer preferences, leading to more effective use of advertising dollars and overall improvement in marketing strategies. Alternatively, companies could use data for information about internal functioning that could lead to a better understanding of employee communication or effective use of technology. Indeed, in an age of competitive global markets, effective acquisition and use of data is not only a benefit, but may actually be necessary for an organization to stay competitive. Such a climate has led to the label “inquiring organizations” which refers to organizations that are involved in the creation of knowledge (e.g., data) that serves their mission to stay current and competitive (e.g., Churchman, 1971; Murray, Case, & Gardiner, 2005). Technological processes that are critical to inquiring organizations are Knowledge Discovery in Databases (KDD) and Data Mining (DM).

KDD refers to the general process of discovering useful information and patterns in datasets. DM is a specific form of KDD involving the use of computational algorithms to extract from large data sets information and statistical patterns that directly point to actionable findings. In this chapter, we focus on the role of the data mining analyst, the individual who applies the computational algorithms to a data set and then interprets the output in light of the organizational goals for strategic change or improvement. Currently, there is a trend in DM towards a greater reliance on automation. That is, once the analyst selects the appropriate algorithms, their execution is largely automated (e.g., Murray et al., 2005). Consequently, the search for meaningful patterns is computer-based, whereas the role of the analyst is centered primarily on interpretation of outcomes. The heavy reliance on automation and relatively low analyst involvement in DM has benefits and liabilities. The benefits of a heavy reliance on automation include speed and efficiency with which data analytic processes can be executed. In contrast, a liability of an automated algorithm-execution stage is that the analyst is unable to flexibly employ and interject into the process valuable background experiences and domain knowledge. For example, in the “data extraction process” low analyst involvement may be associated with missed data or patterns. Stated differently, if pattern detection is left solely to a computer-based algorithm, then it is probable that many patterns will be discarded. Furthermore, some of these (discarded) patterns might, to a human analyst, be judged to be important based on the background knowledge of the analyst (e.g., an insight that might indicate a new approach to the data and consequently new model parameters that might lead to the identification of statistical patterns that might otherwise never be considered). We would argue that although automaticity has an invaluable pragmatic value in its ability to reduce large bodies of data to manageable proportions, it is also important to determine how the typically automated components of DM can be augmented by potentially valuable human (i.e., analyst) involvement given the rich knowledge and inferential abilities that humans bring to any task. In this chapter, we will attempt to articulate why a heavily automated approach to
Related Content

Assessing Knowledge Management System User Acceptance with the Technology Acceptance Model
www.igi-global.com/chapter/assessing-knowledge-management-system-user/25208?camid=4v1a

A Novel Meta-Heuristic Approach for Load Balancing in Cloud Computing
www.igi-global.com/article/a-novel-meta-heuristic-approach-for-load-balancing-in-cloud-computing/190601?camid=4v1a

A Holistic Approach for Understanding Project Management
www.igi-global.com/chapter/holistic-approach-understanding-project-management/66913?camid=4v1a

Effective Stakeholder Knowledge Sharing for Effective Organizational Memory
www.igi-global.com/chapter/effective-stakeholder-knowledge-sharing-effective/5983?camid=4v1a