META-KNOWLEDGE MANAGEMENT IN MULTISTRATEGY PROCESS-ORIENTED KNOWLEDGE DISCOVERY SYSTEMS

Mykola Pechenizkiy[*], Alexey Tsymbal[**] and Seppo Puuronen[*]

1. INTRODUCTION

Current electronic data repositories are growing quickly and contain big amount of data from commercial, scientific, and other domain areas. The capabilities for collecting and storing all kinds of data exceed the abilities to analyze, summarize, and extract knowledge from this data. Knowledge discovery systems (KDSs) use achievements from many technical areas, including databases, Data Mining (DM), statistics, AI, machine learning, pattern recognition, high performance computing, management information systems (MIS), decision support systems, and knowledge-based systems. Knowledge discovery is an innovative approach to information management and is associated commonly with the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns and relations in large databases (Fayyad, 1996). Numerous data mining techniques have recently been developed to extract knowledge from large databases.

Since present-day KDSs are armed with a number of available techniques to process data; and, potentially, there are many possible combinations of these techniques to construct a DM strategy for mining a current problem. In a real problem-solving situation it is not computationally feasible to apply every DM strategy. Therefore, dynamic selection of data mining methods in knowledge discovery systems has been under active study (see, for example, (Tsymbal, 2002)). However, at least two contexts of dynamic selection can be distinguished. First, the so-called multi-classifier systems that apply different ensemble techniques (Dietterich, 1997). Their general idea is usually to select one classifier on the dynamic basis taking into account the local performance (e.g. generalisation accuracy) in the instance space. Second, multistrategy learning that applies a strategy selection approach which takes into account the classification problem- related characteristics (meta-data). We are interested in the second context in this study.

Selection of the most appropriate DM technique or a group of the most appropriate techniques is usually not straightforward. Many empirical studies are aimed

[*] Department of Computer Science and Information Systems, University of Jyväskylä, Jyväskylä, Finland

[**] Department of Computer Science, Trinity College Dublin, Dublin, Ireland

to show that one learning strategy can perform significantly better than another strategy on a group of problems that are characterised by some properties (Kiang, 2003).

Certainly, in one way or another, some knowledge is required for making a decision about appropriate techniques' selection and DM strategy construction for a problem at hand. In order to distinguish between the knowledge extracted from data that represents the problem to be mined by means of applying a DM technique and the higher-level knowledge (from the KDS perspective) required for managing techniques' selection, combination and application we will refer to the latter as meta-knowledge.

Meta-learning or "learning to learn" (also known as bias learning) is the effort to automatically induce dependencies between learning tasks and appropriate learning strategies. The meta-learning paradigm is based on the assumptions that it is possible (1) to evaluate and compare learning strategies, (2) to measure the benefits of early learning on subsequent learning, and (3) to use such evaluations to reason about learning strategies and to select useful ones and disregard the useless or misleading strategies (Schmidhuber et al., 1996). In meta-learning in the context of classifier ensembles, where only the data itself is used to make decisions about method selection, rather good practical results are shown in experiments supported by theoretical studies as well (Tsymbal, 2002). Unfortunately, this is not the case with meta-learning for dynamic integration of DM strategies for a data set at hand. This area is less studied and hardly ever applied in practice. There are works on a multistrategy approach based on the ideas of constructive induction and conceptual clustering for conceptual data exploration, i.e. the derivation of high-level concepts and descriptions from data through symbolic reasoning involving both data and background knowledge (Michalski, 1997). Recently, several studies on automatic classifier selection via meta-learning have been reported, see e.g. Kalousis (2002) for an overview. However, the experimental results of presented approaches are not so promising.

In this paper we propose a knowledge-driven approach to enhance the dynamic integration of DM strategies in knowledge discovery systems. Our focus here is on meta-knowledge management aimed to organise a systematic process of meta-knowledge capture and refinement over time. We consider the basic knowledge management processes of knowledge creation and identification, representation, collection and organization, sharing, adaptation, and application with respect to the introduced concept of meta-knowledge.

The rest of the paper is organized as follows. In Section 2 we consider the basics of multistrategy process-oriented knowledge discovery. A meta-knowledge-driven approach for dynamic integration of data-mining techniques is introduced in Section 3. In Section 4 we consider meta-learning approaches for automatic technique selection and discuss their potential and limitations. We conclude with a brief summary and assessment of further research directions in Section 5.

2. MULTISTRATEGY PROCESS-ORIENTED KNOWLEDGE DISCOVERY

Numerous process-oriented knowledge discovery systems have recently been developed. At the beginning of the millennium there exist about 200 tools that could perform several tasks each (such as clustering, classification, regression, and visualization) for specialized applications (Piatetsky-Shapiro, 2000). This growing trend towards integrating DM tools with specialized applications has been associated with the development of KDSs that are often called “vertical solutions” (Fayyad and Uthurusamy, 2002).

However, adaptability to variations in data characteristics and dynamics of business scenarios becomes increasingly important for data processing systems, as they become an integral part of an organizational decision support system (Kiang, 2003). KDSs should be able to discover knowledge by combining several available techniques, and provide a more automatic environment, or an application envelope, surrounding a highly sophisticated DM system (Fayyad and Uthurusamy, 2002). A similar trend was reported by Michalski (1997) with respect to the orientation of machine-learning systems from single-strategy to multistrategy task-adaptive learning.

Let us consider briefly the basics of the knowledge discovery process according to Reinartz (1999) presented in Figure 1. The life cycle of an idealized knowledge discovery project consists of seven sequential phases from business understanding to deployment. Generally, these phases are not so strictly sequential, and moving back and forth between different phases, caused by the outcome of each phase, is rather natural.

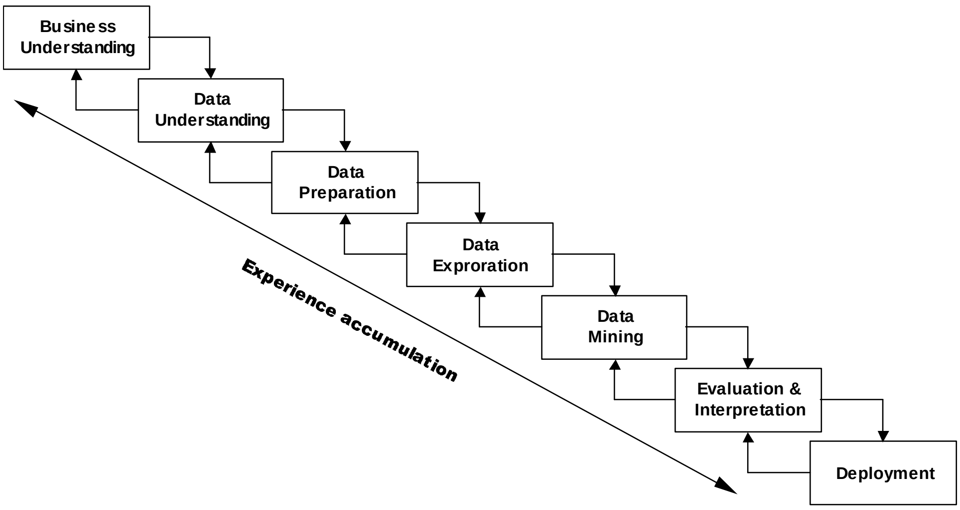

Figure 1. Knowledge discovery process: from problem understanding to deployment (adapted from Reinartz, 1999)

The business-understanding phase is aimed to formulate business questions and translate them into DM goals. The data-understanding phase aims to analyse and document available data and knowledge sources in the business according to the formulated DM goals and provide initial characterization of data. The data preparation phase starts from target data selection that is often related to the problem of building and maintaining useful data warehouses. After selection, the target data is preprocessed in order to reduce the level of noise, preprocess the missing information, reduce data, and remove obviously redundant features. The data exploration phase aims to provide the first insight into the data, evaluate the initial hypotheses, usually, by means of descriptive statistics and visualization techniques. The DM phase covers selection and application of DM techniques, initialization and further calibration of their parameters to optimal values. The discovered patterns that may include a summary of a subset of the data, statistical or predictive models of the data, and relationships among parts of the data are locally evaluated. We would like to clarify that according to this scheme, and some other research literature such as Fayyad (1996), DM is commonly referred to as a particular phase of the entire process of turning raw data into valuable knowledge, and covers the application of modeling and discovery algorithms. In industry, however, both knowledge discovery and data mining terms are often used as synonyms to the entire process of getting valuable knowledge. The evaluation and interpretation phase aims to analyse the discovered patterns, to determine the patterns that can be considered as new knowledge, and to draw conclusions about the whole discovery process as well. The deployment phase aims to transfer DM results that meet the success criteria into the business (Reinartz, 1999).

As one can see from Figure 1, the knowledge discovery process is iterative and consists of various phases and tasks, although the core of the process is DM. We need to admit that current research in knowledge discovery concentrates mostly on the technical details of DM algorithms, whereas the relations between techniques from different fields and how they would fit into the overall knowledge discovery process is often not so clear. It seems, unfortunately, that today still there does not exist any unified methodology that would help practitioners to manage the knowledge discovery process.

An end user of a present-day KDS needs to be a DM expert, or he or she should work in close collaboration with professional data miners, since clear and complete understanding of the knowledge discovery process is essential for successful discovery. And even for business analysts and experienced DM engineers it often remains hard to find the best-suited techniques to solve a business problem under consideration.

3. A META-KNOWLEDGE-DRIVEN APPROACH FOR EFFICIENT USE OF DM TECHNIQUES

In this section we propose to treat a KDS as a complex adaptive system that creates, receives, stores, retrieves, transforms, and transmits meta-knowledge to improve the ability of the KDS to adapt to the environment and to utilize available DM techniques more efficiently and effectively.

First, in Section 3.1 we consider different types of knowledge. Then, in Section 3.2 we discuss the knowledge management process in the context of meta-knowledge. Section 3.3 considers the meta-knowledge repository lifecycle in more detail.

3.1. Different types of knowledge and their transformations

One common definition of knowledge is "justified belief that increases an entity's capacity for effective action" (Nonaka, 1994). An interested reader can find a long history of epistemological debates, and discussion of knowledge from different perspectives in Polanyi (1962). In this section we consider different types of knowledge and their potential in the effective work and performance of a KDS.

Organizational knowledge can be seen as a hierarchical network of rules about specific data or information that has explanatory, predictive, and functional power. These rules are categorized as procedural and declarative. The procedural rules are "know-how" rules and the declarative rules are "know-what" rules. "Knowing where" and "knowing when" represent spatial and temporal contexts of knowledge validity respectively. "Knowing why" provides a KDS with explanatory facilities when it is necessary to argue why a certain DM strategy is recommended or applied.

Beside these technical issues of knowing with respect to knowledge management in KDS, we recognize three basic types of organizational types of knowing. “Knowing what-for” represents DM goals that reflect business goals, and account knowledge of the application domain. “Knowing who” involves information about “who knows what”. As the complexity of the knowledge increases, co-operation between groups (of DM experts, DM practitioners or intelligent knowledge repositories) tends to develop. “Knowing how much” accounts benefits of produced knowledge, resources required, related risks, etc. Although being important the last two knowings are not in the focus of our present study.

In Figure 2 we present the concept of knowledge and its transformations adopted from (Spiegler, 2000). Reality is related to entities whereas data are the attributes from those entities. When the current business problem is formulated as a DM task, data represent those attributes. Information is the result of data processing and information associated with “knowing that and what” type of knowledge. The knowledge term is defined as “knowing how and why” and is the result of information processing. Wisdom is associated with the knowing context of where and when certain knowledge is relevant and valid. All these types of knowing are utilized in every DM technique. In the time dimension, data naturally deals with the past, information is used in the present and knowledge is to be utilized in the future work.

Observing data, hypothesizing on it, and conducting experiments, new knowledge claims can be produced. These claims are validated, placed into the context and become new knowledge. However, what is knowledge for one person or system may be used by another as the initial data for construction of higher-level pieces of knowledge. Therefore, transformations like “data – knowledge – meta-data – meta-knowledge – meta-meta-data – …” are rather natural.

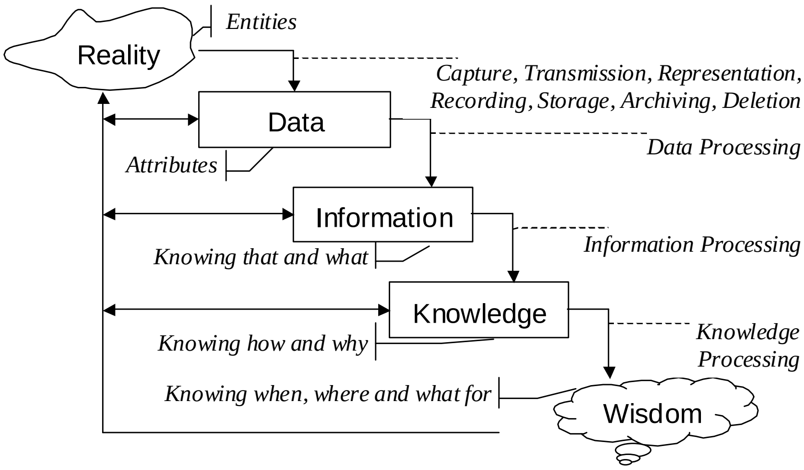

Figure 2. Transformations of data and knowledge concepts (adapted from Spiegler, 2000)

Thus, the knowledge discovery transformation of data into knowledge (Fig. 2) may be applied at any level of knowledge, as the knowledge-data difference is inessential and subjective in our case. Any level may have a meta-level. Replacing data by meta-data, the transformation produces meta-knowledge instead of knowledge, and so on at the next level. Therefore it is often not so easy to determine whether knowing belongs to meta-data or meta-knowledge.

In the next subsection we emphasize the view on knowledge as an entity that can be produced, moved, inspected, rejected, and assessed, just as a widget in a factory. We consider the primary knowledge management processes including knowledge creation, knowledge organization, knowledge distribution, and knowledge application.

3.2. The knowledge management process in the context of meta-knowledge

The goal of meta-knowledge management is to make more effective and efficient use of available DM techniques.

According to King's et al. (2002) analysis of the most important issues in knowledge management, there are 4 groups of such issues: (1) executive/strategic management, (2) operational management, (3) costs, benefits, and risks management, and (4) standards in the KM technology and communication. In this paper we address the issues of the second group that include the identification of available knowledge, seeking ways to capture it in a KM process, and analysing the ability to design an meta-knowledge management system including its tools and applications.

Generally, the problem of knowledge capture, storage, and dissemination is similar to data and information management in ISs, and therefore some executives prefer to view KM as a natural extension to IS functions (Alavi and Leidner, 1999). According to Zack (1999) the most practical way to define KM is to show on the existing IT infrastructure the involvement of: (1) knowledge repositories, (2) best-practices and lessons-learned systems, (3) expert networks [these are DM experts], and (4) communities of practice [these are end-users].

The main idea of the continuous KM process is presented in Figure 3. We separate five key phases of this process. The first phase deals with knowledge acquisition or creation. Having a collection of data sets and a collection of classifiers, we can characterize them producing meta-data as a result of their (algorithms and data sets) characterization. Michalski (1997) proposed a functional approach to knowledge management with regard to both symbolic and numerical data exploration tasks at both the data and knowledge bases levels. Michalski introduced knowledge generating operators (KGOs) that can be used for creating, modifying, combining, deleting, and selecting rules and other structures in the knowledge base. KGOs can be applied (1) to generate and test meta-rules (hypotheses), which are inferred from summarizing the facts or discriminating between the groups of facts about DM strategy performance on problems with certain characteristics; (2) to construct a decision structure from a set of decision rules; and (3) to generate new features/concepts.

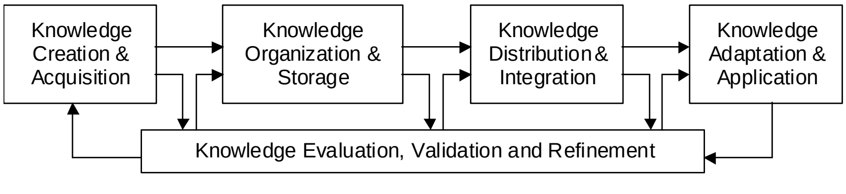

Figure 3. The knowledge management process

When meta-data is available, a machine-learning algorithm can be applied to it. As a result, a meta-learning model that maps dataset characteristics onto classifiers' characteristics with respect to the introduced performance criteria is built. We consider this general idea with respect to knowing a data set and DM techniques, knowing characteristics of the data set and DM techniques, knowing how these properties can be mapped from one set onto another in the context of meta-learning in Section 4.

The second phase deals with knowledge organization and storage. In our context these processes are related mainly to knowledge representation issues. Minsky (1991) discusses pro and cons of connectivist and structural approaches to knowledge representation, concluding that their combination would be natural, since usually at the lower levels of abstraction it tends to have net architecture and tends to organize clusters and hierarchical structures at the higher levels of abstraction.

The third phase is related to knowledge distribution and knowledge integration processes. Generally, we have four potential sources of knowledge that has to be integrated in the repository of KDS system: (1) meta-knowledge from an expert in data-mining, knowledge discovery, statistics and related fields; (2) meta-knowledge from a data-mining practitioner; (3) meta-knowledge from laboratory experiments on synthetic data sets; and, finally, (4) meta-knowledge from field experiments on real-world problems. Beside this, research and business communities, and similar knowledge discovery systems themselves can organize different so-called trusted networks, where participant are motivated to share their knowledge.

The fourth phase deals with knowledge adaptation and application processes. It is often impossible to apply certain knowledge directly. Therefore knowledge adaptation process needs to be undertaken. Multi-Case-based reasoning (Leake and Sooriamurthi, 2002) might be one approach to perform such adaptation.

The fifths phase deals with the knowledge evaluation, validation and refinement processes. In order to keep the meta-knowledge updated there is a need to have a monitoring process to control whether the discovered meta-knowledge remains valid and a technique for continuous enhancement of knowledge. We consider these issues in the next subsection.

3.3. Meta-knowledge repository lifecycle

Since the repository is created it tends to grow and at some point it naturally begins to collapse under its own weight, requiring major reorganization (Zack, 1999). Therefore, the repository needs to be continuously updated, and some content needs to be deleted (if misleading), deactivated or archived (if it is potentially useful). Content may become less fragmented and redundant if similar contributions are combined, generalized and restructured.

The process of filtering knowledge claims into accepted or suppressed is commonly applied in KM. This is even more important in meta-knowledge management since a plenty of claims are produced automatically (and therefore usually need to be filtered automatically).

In Section 3.1 we mentioned the “knowing when” and “knowing where” contexts. The basic idea here is that when the environment changes (that in general may happen all the time), all of the general rules without specifying the context could become invalid. Therefore, it is highly desirable to make the knowledge repository adaptive, i.e.

some knowledge should exist that would guide an organization to change the repository when the environment calls for it.

Some knowledge claims are naturally in constant competition with the other claims. Disagreements within the knowledge repository need to be resolved by means of generalization of some parts and contextualization of the others. In order to increase the quality and validity of knowledge, it needs to be continually tested, improved or removed (deactivated). Refinement leads to formulating a new knowledge claim, which requires a new process of testing and validation.

Some basic principles of triggers can be introduced in the knowledge repository. Thus, for example, when some knowledge is falsified, the deductively inferred claims from the claims to be deleted should be deleted as well.

We would like to clarify the notions of knowledge validity and knowledge quality with respect to the knowledge refinement process.

The contexts “knowing when” and “knowing where” can be discovered before it appears in a real situation. So-called zooming in and zooming out procedures can be used to find a context where theory can be falsified or supported. The goal of such procedures is in search for balance between generality, compactness, interpretability, and understandability and sensitiveness to the context, exactness, precision, and adequacy of meta-knowledge.

The quality of knowledge can be estimated by its ability to help a KDS produce solutions faster and more effectively. To determine the relative quality of a validated knowledge claim, its value needs to be compared to the values of the other claims according to the existing criteria. In any case knowledge claims have both a degree of utility and a degree of satisfaction. However, the quality of knowledge is often context-dependent. Therefore where and when context conditions may be important in many situations not only for knowledge validation but also for quality estimation.

The quality of a knowledge claim is further dependent on the accuracy of the criteria used to evaluate it. Such criteria as complexity, usefulness, and predictive power are well formalised and easy to estimate. On the contrary, such criteria as understandability, reliability of source, explanatory power are rather subjective and therefore inaccurate.

4 AUTOMATIC DATA-MINING TECHNIQUE SELECTION VIA META-LEARNING

In this section we introduce the basic idea of meta-learning in the context of efficient use of available DM techniques. In Section 4.1 we show how meta-knowledge can be produced by meta-learning. Section 4.2 briefly overviews what meta-data can be produced and used in the meta-learning process. In Section 4.3 we discuss the limitations of the meta-learning approach and suggest some directions how to proceed with the research on adaptive selection of a DM technique.

4.1 Meta-learning

Meta-learning (also known as bias learning) is the effort to automatically induce correlations between tasks and inductive learning strategies (Kalousis et al., 2004). Meta-learning can produce rules concerning the use of the alternative strategies, methodological knowledge, or no more than correct predictions concerning the ranking of strategies for a new task.

Successful meta-learning in the context of (automatic) DM technique(s) selection would be really very important and beneficial in the data mining practice. It is obvious that in a real-world situation it is very unlikely to accomplish the brute-force search comparing all the applicable approaches.

Several meta-learning approaches for automatic DM technique selection have been introduced in the literature. A comprehensive overview can be found in Kalousis (2002). The most popular strategies for meta-learning are characterisation of a dataset in terms of its statistical/information properties or more recently introduced landmarking approach (Pfahringer and Bensusan, 2000) and the characterization of algorithms.

The general idea of meta-learning with respect to the selection of a data mining technique for a data set at hand is presented in Figure 4. Having a collection of datasets and a collection of classifiers, we can characterize them producing meta-data ("knowing that") as a result of their (algorithms and data sets) characterization. When meta-data is available, a machine-learning algorithm can be applied to it. As a result, a meta-learning model ("knowing how") that maps dataset characteristics onto classifiers' characteristics with respect to the introduced performance criteria is built. When a new data set is introduced to the system, necessary dataset's characteristics are estimated so that the meta-model is able to suggest an algorithm or a combination of algorithms according to a performance criterion. Trying to validate obtained "knowing how" in different contexts we will produce new pieces of "knowing where" and "knowing when".

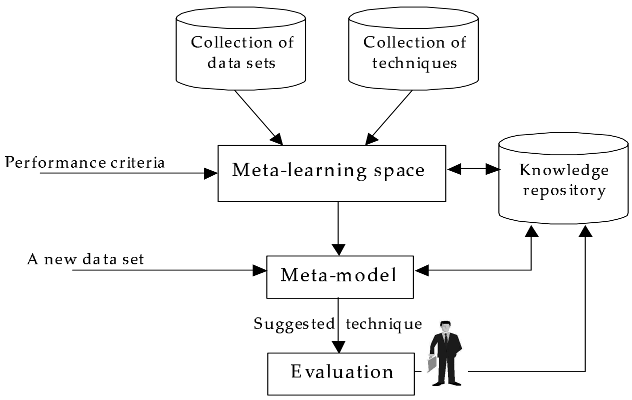

Figure 4. Meta-learning system for suggestion of a technique for a dataset at consideration

4.2 Meta-data used in the meta-learning process

When providing a recommendation to the user, the suggestion may come as a list of applicable algorithms, the best algorithm or a ranking of the algorithms.

The characteristics of a dataset that can be used for meta-learning are commonly divided into those that describe the nature of attributes, attributes themselves, associations between attributes, and associations between attributes and a target variable. To account this, three categories of dataset characteristics can be used: (1) simple characteristics (number of examples, number of attributes, number of classes, number of binary attributes etc.), (2) statistical characteristics (standard deviation ratio, mean absolute correlation of attributes, mean skewness and mean kurtosis of attributes, etc.), (3) information theory characteristics (entropy of class, mean entropy of attributes, noise-signal ratio, mean mutual information of class and attributes, etc.). The exhaustive list of statistical and information measures of a dataset and the data characterization tool (DCT) is proposed by Lindner and Studer (1999).

The second approach to characterize a dataset is landmarking. The idea is to directly characterize a dataset by relating the performance of some learners (landmarkers) to the performance of some other algorithm (Pfahringer and Bensusan, 2000). This idea has reflection in the common practice of data miners. When a new problem is stated, some preliminary exploration is performed usually in order to quickly come up with some few potentially promising approaches.

Naturally, beside characterization of a data set from the data side, some application restrictions or priorities can be introduced. Thus, a user may like to define the most (un)desirable or the most crucial characteristic(s) of an algorithm to be selected for a certain application. The most common characteristics that are taken into account are: algorithm taxonomy, interpretability of the model, importance of results interpretability and algorithm transparency, explanation of decision; training time, testing time, accuracy, and cost handling for misclassification. A good overview of potential algorithms' characteristics is given in Kalousis (2002).

4.3 Limitations of the presented meta-learning approaches

Most meta-learning approaches for automatic algorithm selection (such as meta-decision tree and a meta-instance based learner) assume that the features used to represent meta-instances are sufficiently relevant. However, it was experimentally shown that this assumption often does not hold for many learning problems. Some features may not be directly relevant, and some features may be redundant or irrelevant. Even those meta-learning approaches that apply feature selection techniques, and can eliminate irrelevant features and thus somehow account the problem of high dimensionality, often fail to find good representation for meta-data. This happens because of the fact that many features in their original representation are weakly or indirectly relevant to the problem.

Beside the limited success reported by researchers, someone needs to admit that the meta-learning approach as such has several shortcomings. Lindner and Studer (1999) reported two general problems with meta-models that associate a data set and algorithm characteristics. The first problem is the representativeness of meta-data examples. The possible space of learning problems and thus a meta-learning-space is vast and getting larger with the invention of new algorithms, consideration of new characteristics and parameters. But the size of meta-datasets used in the studies is naturally rather small because of the computational complexity of producing a single meta-example – usually a time-consuming cross-validation process is used to estimate the performance of every algorithm used in the study. The other problem (related especially to the landmarking approach) is the computational complexity (up to $O(n^3)$, where n is the number of examples) of some sophisticated statistical measures. Such measures are not scalable to large datasets and actually such amount of resources can be directly used for applying a sophisticated DM technique instead of doing meta-learning.

We believe that a deeper analyses of a restricted set of data mining techniques (e.g., feature transformation techniques and classifiers) at both theoretical and experimental levels as a more beneficial approach rather than application of the meta-learning approach only to the whole range of machine learning techniques at once. We emphasize the necessity to integrate meta-knowledge produced by DM experts, DM practitioners and meta-learning approaches. We discussed related issues in Section 3.4.

We consider a decision support system-based approach that uses (1) the meta-learning paradigm as a way of knowledge acquisition from the experiments, and (2) a methodology used in expert systems design for representation, accumulation and use of expert knowledge and knowledge acquired through the meta-learning process (Pechenizkiy et al., 2003).

The second point is related to the problem of a meta-dataset size reported by many researchers. Therefore, we stress on a possibility of artificial dataset generation. This possibility allows to significantly increase the repository of datasets and thus to increase the sample size of a meta-dataset. However, an additional and probably more significant benefit of using a synthetic dataset is the possibility to check (or search for) individual hypothesis under the controlled settings, since in such datasets their characteristics can be predefined, one part of which can be fixed, and the other one varied.

In the case of synthetic data sets, however, we have another problem, the problem of knowledge verification. Hypothesis, produced and tested on synthetic data sets need, to be verified on real-world (or benchmark) data sets to make valid conclusions that could be used in practice. Afterward, the verified knowledge can be integrated into the recommendation system.

5. CONCLUSION

Although there are a huge number of data mining techniques – one can hardly find a technique that would be best for all data sets. Our main research goal is to contribute to knowledge in the problem of DM strategy selection for a certain DM problem at hand. Certainly, in one way or another, some knowledge is required for making a decision about appropriate techniques' selection and DM strategy construction for a problem under consideration. Unfortunately, there does not exist canonical knowledge, a perfect mathematical model, or any relevant tool to select the best technique. Instead, a volume of accumulated empirical findings, some trends, and some dependencies have been discovered in a number of various studies. On the other hand there are certain assumptions on performance of DM techniques under certain conditions.

In order to distinguish between the knowledge extracted from data that represents the problem to be mined by means of applying a DM technique and the higher-level knowledge (from the KDS perspective) required for managing techniques' selection, combination and application we introduced the latter one as meta-knowledge.

In this paper we proposed a knowledge-driven approach to enhance the dynamic integration of data-mining strategies in knowledge discovery systems. Our focus was on meta-knowledge management aimed to organise a systematic process of meta-knowledge capture and refinement over time. We considered the basic knowledge management processes of knowledge creation and identification, representation, collection and organization, sharing, adaptation and application with respect to the introduced concept of meta-knowledge.

We considered meta-learning approaches for automatic DM technique selection as means of meta-knowledge creation, evaluation and refinement. We analyzed main limitations of such approaches, and suggested ways of their improvement. We see our further research efforts in the implementation of presented knowledge-driven framework for a KDS that contains a limited number of DM techniques of a certain type (presumably different feature transformation methods) and evaluation of this framework in practice for real-world problems in a distributed environment.

Acknowledgments

This research is partly supported by the COMAS Graduate School of the University of Jyväskylä, Finland and by the Science Foundation Ireland under Grant No. S.F.I.-02IN.1I111.

REFERENCES

Alavi, M., and Leidner, D. 1999, Knowledge management systems: Issues, challenges, and benefits. *Communications of the AIS* **1**(7), 2-36.

Fayyad, U.M. 1996, Data Mining and Knowledge Discovery: Making Sense Out of Data, *IEEE Expert* **11**(5), 20-25.

Fayyad, U.M., and Uthurusamy, R. 2002, Evolving data into mining solutions for insights. *Communications of the ACM* **45**(8), 28-31.

Dietterich, T. G. 1997, Machine Learning Research: Four Current Directions. *AI Magazine* **18** (4), 97-136.

Kalousis, A. 2002, Algorithm Selection via Meta-Learning. University of Geneve, Department of Computer Science. PhD Thesis.

Kalousis, A., Gama, J., and Hilario, M. 2004, On data and algorithms: understanding inductive performance, *Machine Learning*, **54**(3): 275-312, 2004.

Kiang, M. 2003, A comparative assessment of classification methods, *Decision Support Systems* **35**, 441-454.

King, W., Marks, P., and McCoy, S. 2002, The most important issues in knowledge management. *Communications of the ACM* **45**(9), 93-97.

Leake, D.B., and Sooriamurthi, R. 2002, Automatically Selecting Strategies for Multi-Case-Base Reasoning. In: Susan Craw, Alun D. Preece (Eds.): *Advances in Case-Based Reasoning*, 6th European Conference (ECCBR 2002), Proceedings. LNCS, 2416 Springer, 204-233.

Lindner, G., and Studer, R. 1999, AST: Support for algorithm selection with a CBR approach. In *Proc. of the 3rd European Conf. on Principles and Practice of Knowledge Discovery in Databases*, 418-423.

Michalski, R.S. 1997, Seeking Knowledge in the Deluge of Facts, *Fundamenta Informaticae* **30**, 283-297.

Minsky, M. 1991, Logical Versus Analogical or Symbolic Versus Connectionist or Neat Versus Scruffy. AI *Magazine* **12**(2), 34-51.

Nonaka, I. 1994, A dynamic theory of organizational knowledge creation. *Organization Science* **5**(1), 14-37.

Pechenizkiy, M., Puuronen, S., and Tsymbal, A. 2003, Feature Extraction for Classification in the Data Mining Process. *International Journal on Information Theories and Applications* **10**(1), Sofia, FOI-Commerce, 321-329.

Pfahringer, B., Bensusan, H. and Giraud-Carrier, C. 2000, Meta-Learning by Landmarking Various Learning Algorithms. In Langley P. (Ed.) *Proc. of the Seventeenth International Conference on Machine Learning*, San Francisco, Morgan Kaufmann, 743-750.

Piatetsky-Shapiro, G. 2000, Knowledge Discovery in Databases: 10 years after. *SIGKDD Explorations* **1**(2), 59-61.

Polanyi, M. 1962, Personal knowledge. Chicago, University of Chicago Press.

Reinartz, T. 1999, Focusing Solutions for Data Mining. LNAI 1623, Berlin Heidelberg.

Schmidhuber, J., Zhao, J., and Wiering, M. 1996, Simple principles of metalearning. Technical Report IDSIA-69-96, IDSIA.

Spiegler, I. 2000, Knowledge Management: A New Idea or a Recycled Concept? *Communications of the Association for Information Systems*, **3**, Article 14.

Tsymbal, A. 2002, Dynamic Integration of Data Mining Methods in Knowledge Discovery Systems. PhD Thesis. University of Jyväskylä, Finland.

Zack, M. 1999, Managing codified knowledge. *Sloan Management Review* **40**(4), 45-58.