An Expert System for Evaluating the 'Knowledge Potential' of Databases

Stephen Potter and Stuart Aitken,
AIAI, University of Edinburgh, Edinburgh, United Kingdom
stephenp@aiai.ed.ac.uk
http://www.aiai.ed.ac.uk/

Abstract

Data Mining techniques can, under favourable conditions, extract valuable knowledge from an organisation's databases. However, the precise nature of these favourable conditions is poorly articulated, and as a result organisations run the risk of instigating costly and time-consuming data mining episodes upon inappropriate or irrelevant data. In response to this problem, the authors have applied expert systems technologies to the task of predicting the extent to which a given database contains knowledge that is both valuable and susceptible to mining. It is intended that this system will form a component of a 'knowledge auditing' method for appraising an organisation's (existing and potential) knowledge resources. This paper describes this application, along with some of the particular issues surrounding its implementation, testing and delivery to prospective users.

1. Introduction

Alongside the increasing capacities and decreasing prices of data storage media there has been a commensurate growth in the number and size of databases that are generated and maintained within organisations. Whilst these undoubtedly provide a useful historical record of the organisation's transactions and entities, there has been a growing recognition that, when considered as a whole, these records can implicitly constitute more general knowledge of the relationships that hold amongst these transactions and entities. This knowledge is potentially of immense value to the organisation, enabling it to better understand and reason about its processes and customers, and then act so as to maximise its profits. Accordingly, the question then becomes one of how this implicit knowledge might be made explicit.

There are several techniques by which information of the sort found in databases might be 'converted' into actionable knowledge. The work described in this paper concerns the application of Data Mining (DM) techniques to databases for the purposes of acquiring knowledge. These techniques exploit algorithms developed through AI research into machine learning and more traditional statistical data analysis algorithms in order to generate useful knowledge inductively, that is, to extract general relationships from a given set of examples of these relationships (provided by the records in the database).

Hence, when successful, a DM episode can convert an organisation's existing database resources into useful knowledge. However, there are many reasons why an application of DM to a particular database might not produce the desired results: the algorithms might be applied inappropriately; the data may be too noisy; or the database records may contain irrelevant or incidental information. Moreover, even if knowledge is extracted, unless it is of practical use to the organisation then the episode cannot be considered a success. Since DM is frequently performed by experts from outside the organisation, and the process of understanding and then applying algorithms to the database can be time-consuming, any failure - regardless of the reason for it - can prove costly.

This factor provided one of the motivations for the work reported here. It was realised that the ability to evaluate a particular database with respect to its 'knowledge potential' (that is, the potential that it has to provide useful knowledge if DM is applied appropriately to it *but without actually applying DM to it*) would be of great value to organisations that possess databases and lack the resources or the inclination to pay to be told that their data are of little worth. Furthermore, if this ability were embodied within an expert system, it would enable it to be effectively transmitted to remote locations where and when it is required. This paper describes such a system, along with issues raised by the use and delivery of its knowledge. It is envisaged that this expert system would form an element in a wider 'knowledge auditing' methodology; since this would be a factor which influenced a number of the choices made in the system, this idea will now be discussed briefly.

1.1 The Knowledge Audit

Since nowadays it is widely accepted that knowledge is a valuable (and in many cases the most valuable) organisational resource, its audit, in order to form a picture of the nature and extent of this resource, would seem as worthwhile to an organisation as would be the audit of any more tangible resource. However, therein lies the difficulty: knowledge is rarely explicit in the manner of these more tangible resources, and is often evinced, if at all, through the processes of the organisation.

Although a browse of the internet reveals a number of knowledge management organisations offering 'knowledge auditing' services, the precise nature of these services is rarely clear (often as an understandable result of their commercial sensitivity), and it was felt to be a worthwhile endeavour, under the umbrella of the *Advanced Knowledge Technologies* initiative[1], drawing on the project members' wide-ranging knowledge research experience, to try to establish a coherent and sound methodology, along with the necessary tools, for performing an audit.

---

[1] The Advanced Knowledge Technologies (AKT) project is a long-term interdisciplinary collaboration between research groups at 5 UK universities which aims to tackle fundamental problems associated with knowledge management.

Databases, when viewed as repositories of (potential) knowledge, constitute resources that fall within the remit of the audit and which should be included within the overall assessment of the organisation's knowledge. Although the wider knowledge auditing research is still very much in its infancy, it was felt that an auditing tool for assessing databases could be constructed in a relatively short period of time and would provide a useful experimental test bed for exploring the various technologies that might play a role in performing the knowledge audit. As such, two principle factors have been influential in its construction:

- First, there is a desire to exploit the information- and knowledge-sharing facilities afforded by the World-Wide Web: the ability to perform a knowledge audit over the internet would, it is hoped, lead to a more efficient (and less costly to implement) procedure.
- Secondly, since commerce is being conducted at increasing rates, and organisations are becoming increasingly dynamic in terms of information (and hence knowledge) flow, it is clear that there is a real need for any auditing process to be quick and efficient to be of any use to these organisations.

Hence, this database auditing tool would be available to the organisation via the medium of the internet, and it should be able to judge a database's worth rapidly, especially when one considers that the organisation might have many databases that it wished to audit. (Although conceived as a component of a knowledge audit methodology, it was also thought desirable that, if possible, the tool should be able to function as a 'stand-alone' application, able to offer useful advice regardless of whether or not it was being used as part of an audit of the organisation in question.)

1.2 Structure of this Paper

In the following section, the Data Mining task, and the processes involved, will be described in some detail. The third section describes the sort of information and knowledge required for an expert system to audit a database. Subsequent sections describe the implementation of this system and some of the particular issues that are raised by the implementation and the testing of this and similar knowledge-based applications. Finally, some conclusions are drawn about the work.

2. Data Mining

According to Berry and Linoff [1] Data Mining can be defined as:

> *"…the process of exploration and analysis, by automatic or semi-automatic means, of large quantities of data in order to discover meaningful patterns and rules."*

They go on to remark that some people have been misled into thinking that DM is "a product that can be bought rather than a discipline that must be mastered." This comment highlights the fact that DM relies as much on the expertise and experience of its practitioners as it does on the availability of particular algorithms – and, consequently, why expert systems might have a role to play in this process. This section is devoted to describing the DM task in greater detail, so providing the background necessary for understanding both the motivation for an expert system and the system itself. To begin with, however, it will be advantageous to discuss briefly databases, and the accompanying terminology used throughout this paper.

2.1 Databases

To simplify matters, a database is considered here to be a collection of one or more *records*. Each record describes some event or entity, in terms of the values that it possesses for each of a number of pre-defined attributes or *fields*, which are common to all records. There are a number of characteristics of databases that can cause particular problems for DM or else influence the decisions that are made. These include the extent to which values are missing in the data, the amount of noise in the values, and the number of records contained in the database and the degree to which these records form a representative sample of the whole. Inasmuch as they have a bearing on the work presented here, these characteristics will be discussed in greater detail in a later section.

2.2 The Data Mining Process

The Data Mining task is one of extracting useful knowledge from a (usually large, but not necessarily so) database. This is an *inductive* process: each record in the data represents some event, and the DM task is to produce some generalised description of these events that is in some way useful to the organisation that owns the database. There is a range of tools for performing this induction; these include algorithms originating from the AI discipline of machine learning (typically, these algorithms will represent attempts to mimic some aspect of human learning capabilities) and from statistical analysis. Limitations of space preclude a detailed discussion of the various algorithms here.

At the highest level, this process is represented by Berry and Linoff's 'virtuous circle' of DM (Figure 1). Here DM is viewed as a continuous cycle, with earlier results being used to modify the business processes, generating more (and perhaps different types) of data, which, in their turn, can then be mined. The work presented here is concerned with the very earliest stages of DM, before such a cycle has been initiated, at the time when the organisation's existing database resources are to be assessed so as to decide whether they contain data that might be transformed into knowledge that is both actionable and useful.

2.2.1 Step 1: Identify Business Problems – and DM Solutions

This initial step involves establishing the goals or the purpose of a DM exercise. This is a poorly defined task, which usually takes as its starting point some particular database, and which involves devising some hypothesis about what might be usefully learned from the database. This notion of 'useful' knowledge is obviously not a uniform one; it depends as much upon the organisation's wider goals and resources, upon the ML algorithms that are available, and upon the DM practitioner's experience and creativity, as it does upon the nature and content of the database itself. Hence, the task at this stage is more one of communication and understanding than one of data analysis: the DM practitioner must gain an understanding of the meaning of the data and its place within the context of the organisation. Only then can notions emerge of what might represent useful knowledge for the organisation.

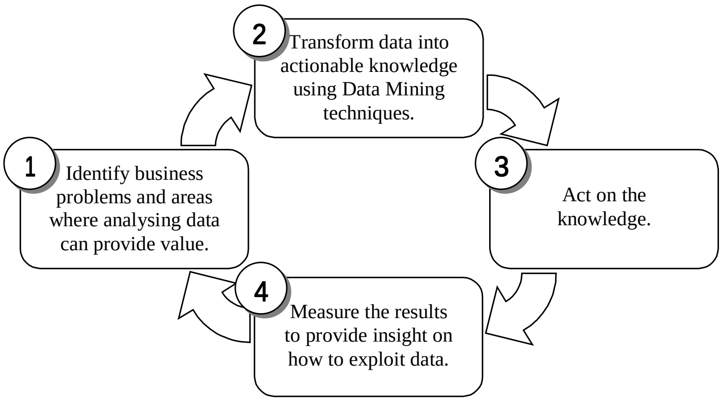

Figure 1. The ‘virtuous circle’ of Data Mining (after Berry and Linoff [1]).

If these (potentially conflicting) factors can be reconciled, the outcome of this stage will be one or more models or descriptions of the knowledge that might be learned from the data: the in the next stage the DM task becomes one of learning these models. The quality of the analysis and choices made during this step will have a major bearing on the efficiency and success of the endeavour.

2.2.2 Step 2: Transform Data into Actionable Knowledge

This step involves a number of sub-tasks, applied for each knowledge model in turn. First, a particular DM algorithm is selected as being suitable for learning the knowledge (in practice this will have had a bearing on the previous step, influencing the construction of the knowledge descriptions, and the choice may well have been (implicitly) made already). With the algorithm selected, the next sub-task is to construct the appropriate datasets from the database. Depending on the algorithm chosen, three datasets may be needed:

- The *training set* –contains the data to be used to try to learn the knowledge.
- The *validation set* – this has the same form as the training set, but consists of distinct sets of values. It is used when applying certain algorithms to set parameters or guide the acquisition process (not all algorithms require this set).
- The *test set* – a third disjoint set drawn from the database that shares the form of the previous two. This set is used to assess the quality of the learned knowledge once training has concluded.

Together these sets may include all the records in the relevant subset of the database, or else some sample of them: a representative smaller sample may provide the desired knowledge more efficiently. Ensuring that each sample/subset is representative of the whole is extremely difficult, if not impossible. (It is sometimes appropriate to manipulate the datasets so that they are no longer representative, so as to emphasise certain rare – but crucial – events in the data.)

Before these datasets can be constructed, however, it will usually be necessary to preprocess the data. This preprocessing can be a difficult task, and requires a certain amount of expertise to do well – as it must be done if the DM exercise is to succeed. Han and Kamber [2] identify four different forms of data preprocessing:

- *Data cleaning* – remove (or reduce) noise and correct inconsistencies and incompleteness.
- *Data integration* – merge data from a number of sources.
- *Data transformation* – manipulate data into forms more suited to the DM algorithm in question.
- *Data reduction* – reduce numbers of data to more manageable levels, and in such a way as to preserve as far as possible its potential for DM purposes.

The next step is to apply the algorithm to the suitably processed datasets. In order to do this, values must be chosen for the particular learning parameters of the algorithm. Generally, there is little assistance available for selecting these values, and DM practitioners must rely on their experience or intuition; nevertheless, there will usually be some degree of trial-and-error and 'tweaking' of values involved, until either apposite settings are found and the knowledge is successfully learned, or else the attempt is abandoned in the wake of repeated failure. During the actual learning process little human interaction will, in general, be needed. If learning terminates successfully, and evaluation using the test data suggests that the resulting knowledge conforms to the description of the desired knowledge, then the task may be considered to have concluded satisfactorily - the knowledge has been 'mined' successfully from the database. If not, then an alternative choice must be made at one of the previous stages in the DM process, and the task repeated accordingly. Once the knowledge has been validated, it will be necessary to analyse it as a precursor to acting upon the knowledge. In order to facilitate this, it may be necessary to 'visualise' the knowledge, that is, to make it more immediately comprehensible through the use of graphical representations, etc.

2.2.3 Step 3: Act upon the Mined Knowledge

Armed with this new knowledge, the organisation must put it into practice. How this is done obviously depends on both the organisation and the nature and content of the knowledge. The knowledge may merely serve to confirm that the organisation's existing beliefs and operations are well-founded and so increase the confidence with which they are applied, or it may lead to the modification of existing operations so as to move into more profitable regions, or even to the introduction of entirely new operations to exploit hitherto unrecognised areas into which the organisation can expand.

2.2.4 Step 4: Measure Results

If the value of the knowledge – and of DM, in general – to the organisation is to be ascertained, the results of its use must be monitored. Analysis of these results may provide interesting insights, and if the organisation is convinced of the worth of DM and committed to its application, they may suggest additional areas in which data may be mined, and, indeed, the sort of data that should be collected to optimise the returns from future DM episodes.

3. Data Mining and the Knowledge Audit

Under the right circumstances and with the right resources available, then, the discipline of Data Mining can transform data into knowledge that can improve the operations and activities of an organisation. For this reason, it was thought that as yet unmined databases should be included within the remit of a knowledge audit of an organisation, with some measure of their potential value as repositories of knowledge being incorporated into this assessment. However, a DM episode on a single database might be a lengthy process, requiring DM expertise and resources - and even then, a failure to extract useful knowledge might not necessary mean that the database is worthless as a store of knowledge: poor decisions may have undermined the attempt where better ones would have resulted in success. When, in addition, one considers that an organisation might have a great number of databases, it is clear that this audit must be done, if it is to be done at all, without recourse to actually performing DM on the database.

The authors, drawing on their practical experience of DM applications, feel that it is possible to perform an appraisal of a database in this manner, and, moreover, that the knowledge for performing this audit might usefully be embodied within an expert system, providing a means by which the audit might be automated - hence the work reported in this paper. This section gives an overview of the knowledge used to build this system; the sources of this knowledge were the authors themselves, DM textbooks (chiefly [1] and [2]) and Microsoft's Object Linking and Embedding (OLE) description of DM [3].

The assumption was made that the user of the system would be an employee of the organisation that is undergoing the knowledge audit, and that the system will be executed once for each database that is included within the scope of the audit. This user is expected to be familiar with the database and its contents, and with its place in the wider context of the organisation, and able to answer questions about these: the system would ask a sequence of relevant questions, and make an evaluation of the database on the basis of the answers supplied by the user. Since the system would be part of a general-purpose auditing methodology, the questions would necessarily refer to organisations and databases in the abstract, since it would be infeasible to incorporate specific knowledge of each organisation and its databases.

Furthermore, since this system will be operating as a component in a wider knowledge audit of the organisation, and there may be a number of databases that require auditing, there exists an added practical consideration: the system should be able to make its evaluation without the need for drawn out, time-consuming interaction with the user. In other words, there is a need to find some compromise between, on the one hand, gathering enough information to make an accurate assessment of the database and, on the other, asking as few questions as is possible.

Finally, although this expert system would form but a component in a knowledge auditing methodology, it would, presumably, be of great practical advantage if the system were able to draw conclusions and make recommendations that in some way 'stand alone'; that is, if the system were able to offer general advice on the suitability of databases for DM. With these considerations and assumptions in mind, and taking into consideration the foregoing discussion of the DM process, the relevant information required from the user can be thought of as falling into three general categories:

- The relationship that exists between the database, knowledge and the organisation;
- The availability and nature of background knowledge of the domain that may assist DM, and;
- The nature and content of the database itself.

Each of these will now be considered in turn.

3.1 The Relationship between Database, Knowledge and Organisation

Assuming that the user has some database, from a DM perspective the knowledge audit task is to try to answer the following question: *to what extent do the data in this database contain knowledge that: (a) might be data-mined, and (b) would (potentially) be of value to the organisation?* To answer this question, it is necessary to first consider – at a generic level - about what sorts of knowledge can be acquired through DM. The literature contains several classifications of the types of knowledge that can be mined; there is some degree of consensus to suggest that, at the highest level, this knowledge falls into two categories, which may be termed (following Han and Kamber [2]) *predictive* and *descriptive* knowledge:

- *Predictive knowledge* – this sort of knowledge allows some value to be suggested for one or more attributes (each attribute will correspond directly to a field in the database or else will be derived from some combination of fields by some known operation) of a particular record, based on the values of the other fields of the record. So, this form of knowledge might be used to make predictions about the behaviour of certain attributes, or to otherwise supply values where they are missing for whatever reason.
- *Descriptive knowledge* – this knowledge is in the form of some useful general description or representation of the database or some subset of the data contained therein. This description will make evident general patterns and relationships that occur in the data that were either previously unknown, or else were suspected and are now confirmed.

Any given database might contain both predictive and descriptive knowledge. The first question of the expert system, then, would simply ask the user whether or not, on the basis of the sort of information recorded in the database fields, he/she believes that one or both of these types of knowledge may be present in the database under consideration, and if so, whether this knowledge would be of use to the organisation if it were to be made explicit. Since a negative answer to either the first or second part of this question would indicate that DM could extract nothing of use from the database, the database is adjudged worthless and the session ceases.

Following a positive response to both parts, the interrogation of the user proceeds to the next question, which also concerns the nature of knowledge learnt through DM. Since DM algorithms operate inductively and, accordingly, any knowledge they acquire remains susceptible to invalidation by the occurrence of a single counter-example, this knowledge cannot be guaranteed to be wholly and completely accurate; rather, confidence in its correctness grows with the number of examples – with no counter-examples – seen. Hence, at this stage user should be asked if all the (prescriptive and descriptive) knowledge gained from the database would only be of use if it were wholly correct – if so, the database has no value from a DM perspective, and questioning ceases.

Since, for the purposes of the audit, it is necessary to quantify the potential value of the knowledge in the database, it is necessary to gain some idea of the possible extent of knowledge in the database. Therefore, the next question asks how many 'scraps' (that is, distinct elements) of knowledge of both types are thought potentially to exist in the database, and how many record fields are 'involved' in each scrap of knowledge, on average. This latter question is an attempt to gain some sort of idea of the complexity of the knowledge (and so, the number of data necessary to acquire it) – as a rule of thumb, it is assumed that complexity is proportional to the number of involved fields, though this is not necessarily so.

Now that the some idea has been gained of the 'amount' of knowledge that the database may contain, attention can turn to the question of its worth - any DM exercise can be considered a success only if the knowledge it derives from the data is of value to the organisation. Failure to consider the resources and goals of the organisation can lead to situations in which, for example, valid, seemingly 'good' knowledge is mined, but since it is already well-known within the organisation the DM effort confers no additional value. Hence, the value of the potential knowledge can only be assessed by some agent (in this case, the system user) familiar with the objectives and operations of the organisation. The user is asked to judge the relative worth of the knowledge scraps, if they were made available as a whole, on a scale of 1 ("knowledge of minor usefulness, perhaps confirming existing knowledge and validating current procedures") to 5 ("knowledge of the utmost importance to the future of the organisation"). Once a clearer picture has been gained in this way of the relationship between organisation, database and potential knowledge, attention can turn to the next area of questioning.

3.2 Background Knowledge of the Domain

The existence and availability of additional background domain knowledge can also have a bearing on the data mining exercise. This might be quite generic in nature, or else be specific to the organisation and to the data stored in the database. The next phase in the questioning involves asking about background knowledge. During DM background knowledge may be used for two general purposes:

- When preprocessing of the data, background knowledge can be used to 'clean' or to better represent the data for the chosen algorithm (its most common use).
- During the operation of the algorithm, to guide and suggest possible generalisations that might hold across the data.

The existence of this background knowledge is *not* a necessary condition for successful data mining – useful knowledge can be mined in its absence. However, in a typical DM application, its presence will facilitate the exploitation of the database, will make the process more efficient (and thereby reduce costs) by reducing the amount of analysis of the data necessary to preprocess it, and may increase the quality and value of the acquired knowledge by expressing it in more readily comprehensible terms, suited to both the domain and the organisation.

Hence, the availability of background knowledge will, in general, serve to enhance the potential of the database for DM, and so increase the value of the database from the perspective of a knowledge audit. Accordingly, the system should ask its user about the availability of background knowledge. In order to do this, thought needs to be given the to nature of this knowledge. It is considered that background knowledge in DM applications can be classified as one of the following:

- Knowledge of useful or appropriate categories or ranges into which (particularly numerical) values fall.
- Knowledge of potentially useful attributes that can be derived from the fields of the database, and of how such attributes are derived.
- The definitions of any potentially useful (taxonomic, hierarchical, etc) relationships describing the fields of the data.

For the purposes of the system, the user is not required to supply any of this knowledge, but merely to be able to indicate its availability or otherwise.

3.3 The Database and its Contents

Now the focus can turn to the database itself. Obviously, the nature and the content of the database will be a major factor governing the success of the DM enterprise – if the knowledge is simply not present in the data, then no amount of effort can acquire it. The system user has already expressed a belief that the fields of the database are such that useful relationships amongst them potentially exist; here, then, the task is to try to gather more information about the database and the values it contains so as to evaluate the suitability of the database for acquiring this knowledge. To this end, the user is asked for the following information:

- The number of records in the database - knowledge needs to be well-exemplified if it is to be learned, and, in general, the more complex the knowledge, the greater the number of examples that are needed to learn it.
- The number, if any, of 'missing' fields – attributes that the user thinks of as possible factors in the potential knowledge but which – for whatever reason – have not been recorded.
- An indication of the extent to which values are missing from relevant fields – for a number of reasons, values are often go unrecorded; this can serve to limit the usefulness of a record, in the worst case rendering it entirely useless.
- The accuracy of the recorded values – inaccurate values can result in the learning of inaccurate knowledge. Unfortunately, inaccuracy is not apparent from the database itself – rather, some external knowledge of the methods of gathering and recording values is needed. Accuracy can be considered to manifest itself in two related forms, called here *precision* (the degree by which numerical values differ from the corresponding 'actual' values) and *correctness* (the degree to which non-numerical values are 'right').
- The content of the relevant fields – on the whole, current DM algorithms operate on symbols and numbers; any other forms of data – for example,

images, or complete English sentences, either need to be preprocessed extensively using background knowledge or else disregarded for the purposes of DM. The user is asked to indicate whether the fields contain information beyond simple symbolic and numerical values.

Once these questions have been asked the consultation of the user is completed. Many more questions could be asked and more information gathered, but the need for a quick assessment process should once again be emphasised – the authors consider the above questions sufficient to form an evaluation of the database.

3.4 Evaluating the Database

The answers supplied are used to generate an overall score for the database. ‘Positive’ characteristics of the database and its context increase its score, while ‘negative’ characteristics diminish it. Once everything has been considered, the database will have been awarded a score for its ‘knowledge potential’ on a scale of 0 (“The nature and content of the database render it of little or no use to your organisation for DM purposes”) to 5 (“Your answers suggest that the database represents an extremely valuable source of organisational knowledge, and one that would amply repay the costs incurred as a result of a DM episode”). At a later stage, it will almost certainly be necessary to ‘normalise’ this score to bring it into line with the wider knowledge auditing evaluation (perhaps expressing it in more concrete financial terms). This will not be a trivial task, and will require detailed analysis of the return generated by knowledge and the costs of implementing new knowledge within organisations, as well as of the costs of DM itself.

4. Implementation

The knowledge described above can most naturally be described in terms of forward-chaining rules. This, coupled with the desire to investigate the possibility of performing the audit remotely using the internet led the authors to the use of the Jess shell [4] for the implementation of the system. Jess is a rule engine and scripting environment written in Sun System’s Java language. Originally inspired by the CLIPS shell, it provides an inference engine based (like CLIPS) on the Rete algorithm. Being written in Java, it offers programmers the opportunity to embed the shell within an applet, and so, to make their expert systems available across the WWW using standard web browsers. GUI facilities permit remote users to interact with the system, providing it with the requested information, and, later, to be presented with the results of the system’s reasoning.

The relatively small number of rules permitted the rapid implementation of a prototype expect system; it was felt desirable that the system should be developed in this manner so as to provide a useful experimental platform for exploring the potential and testing the viability of ‘delivering’ knowledge via the internet – one of the key concepts underlying the notion of the knowledge audit as a whole. The system was developed along fairly typical lines, incorporating rules to ask the various questions outlined above and then to evaluate these answers. Routines for displaying additional information and advice to users was incorporated into the latter rules, explaining, for instance, that the proportion of missing values in the database hamper its worth or that the evaluation of the database as a whole suggests that a limited preliminary DM study might be worthwhile before embarking on a full-blown exercise. In doing this, the intention is to increase the usefulness of the system as a 'stand-alone' system, independent of the goals of the knowledge audit.

This system was constructed into an applet complete with user interface, offering drop-down menus and the like to simplify communication with the user. This was incorporated within a web-page (Figure 2), which also included links to further information and to provide feedback via email, along with a glossary of the technical terms used within the expert system, implemented in such a way (using frames) as to allow a user to look-up a definition in the course of using the system.

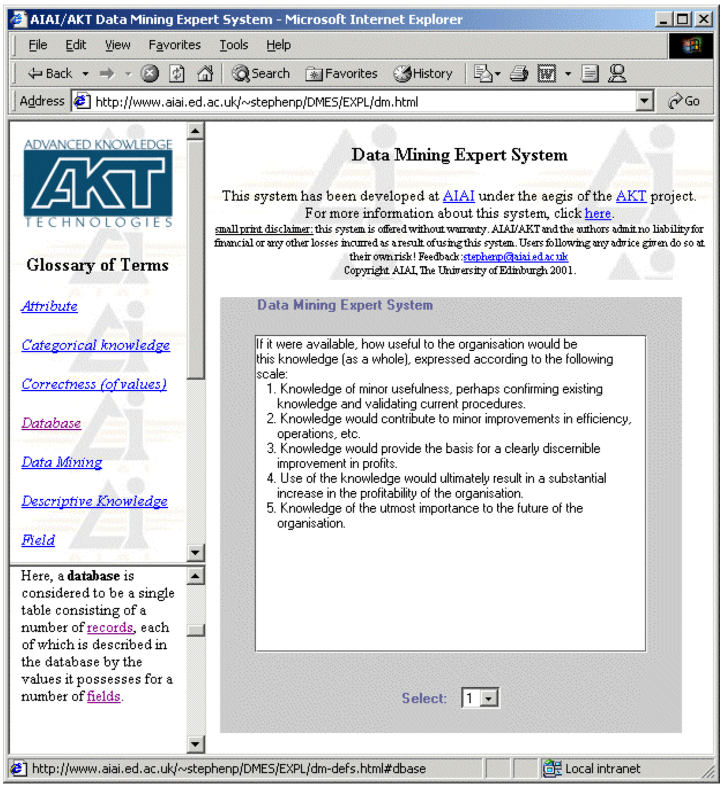

Figure 2. The expert system applet in a browser (in this case Microsoft's Internet Explorer).

5. Testing the Application

The implementation and intended use of this system – in particular, the use of the internet – throw up challenges (and opportunities) for testing beyond those normally encountered in the course of expert system development. Testing of a more conventional expert system can be thought of as a process of confirming that the system can be used as intended to solve real problems. Beyond the detection of bugs in the operation of the various components of the system and interface, this involves, on one hand, validating as far as possible that the knowledge embodied in the system is correct and complete, and is that required to solve the problem in question, and, on the other hand, testing that the system can be (and will be) used effectively by its intended users. Although comprehensive testing of the system described in this paper is still very much in its early stages, each of these testing objectives has interesting implications for applications of this sort.

5.1 Testing the Knowledge

Typical conventional approaches to validating the knowledge in an expert system include comparing the performance of the system on a set of test cases with the expected performance, and asking experts to criticise the knowledge base. The former approach is problematic in this case, since this audit task is, in some respects, a new one – no cases exist of the 'correct' appraisal of databases in the fashion proposed here. Furthermore, it would not be adequate merely to seek out examples of applications (both successful and otherwise) of DM, since, for example, the failure of a particular application does not necessarily mean that the database contains no knowledge that might be mined.

Consequently, the authors' attention has turned to the second form of testing, the employment of experts to criticise the knowledge. One of the chief problems encountered during this form of testing is the difficulty of gaining access to experts; with the various demands placed upon their expertise, finding the time necessary to perform this evaluation can be difficult. Here, however, the use of the internet suggests an alternative approach. The system can easily be made available to a greater number of experts, at geographically remote sites, than is usually available for appraisals, and, furthermore, since the system would be constantly on-line, the appraisal can be performed when at any time of the expert's choosing.

For such an approach to succeed it will probably be necessary, at the very least, to implement an alternative version of the expert system in which the knowledge is made much more explicit, or even to dispense with the system entirely, and simply 'publish' the knowledge in some sort of on-line discussion forum, opening the possibility of experts critiquing the knowledge collaboratively. Such an approach raises interesting practical and theoretical issues – about, for example, how experts might be encouraged to interact in this fashion, and where ownership of collectively developed knowledge would rest – but the authors think that this approach to the knowledge assessment has great potential, and intend to pursue it in the near future.

5.2 Testing the System

An important task during the early stages of conventional expert system development is to obtain a clear idea of the prospective users of the system, in terms of their capabilities, resources and context. This should influence the implementation and, in particular, the interface, of the resulting system. However, in this case this is difficult, since the system is intended for use by a variety of people in a variety of organisational contexts, and, as such, the idea of the user is at most, a vague, generic one. Once again, this makes testing the system problematic. Questions raised range from the technical to the sociological; they include:

- Can the system be accessed (in an acceptable amount of time) and run on a range of different platforms using different browsers? Does the appearance of the interface remain as uniform as possible? Do users understand the interface concepts (links, menus, buttons, etc) with which they are presented?
- Can the system be used by the full range of prospective users, whose context, abilities and knowledge might vary substantially?
- Assuming that the system can be used, is the knowledge it contains delivered in the right form? Can users answer the questions and then act upon the advice they are given? Do users trust knowledge delivered in this manner?

While there has been some work in remote evaluation of software systems (e.g. see [5]), there has, so far, been little reported by way of methodologies for testing the delivery of knowledge-based systems over the internet in this manner. The authors believe that as the internet continues to grow and its exploitation becomes more sophisticated, 'knowledge publishing' ventures, such as the one reported here, will gain in prominence, and, as a consequence, techniques for their evaluation will come to be of great importance. As with the testing of the knowledge, in the near future the authors intend to try to develop methodologies to assist in this task.

6. Conclusions

This paper describes an expert system application for evaluating the extent to which a database contains knowledge that Data Mining algorithms might be used to acquire. While the implementation of this system is quite straightforward and conventional, the desire to use internet technologies and the notion of the system as being a single component in a wider knowledge auditing methodology have imposed certain constraints on this system, influencing many of the decisions made, and have interesting implications for the application and testing of the system.

Acknowledgements

This work is supported under the Advanced Knowledge Technologies (AKT) Interdisciplinary Research Collaboration (IRC), which is sponsored by the UK Engineering and Physical Sciences Research Council under grant number GR/N15764/01. The AKT IRC comprises the Universities of Aberdeen, Edinburgh, Sheffield, Southampton and the Open University. Java, Jess, CLIPS and Microsoft Explorer are registered trademarks.

References

1. Berry, M. J. A. & Linoff, G. Mastering data-mining: the art and science of customer relationship management, Wiley, New York, 2000.
2. Han, J. & Kamber, M. Data mining: concepts and techniques, Morgan Kaufmann, San Francisco, CA, 2001
3. Microsoft, Inc. OLE DB for data mining specification, 2000. Available online at: http://www.microsoft.com/data/oledb/dm.htm.
4. Jess, the Java Expert System Shell, v.5.2, http://herzberg.ca.sandia.gov/jess/. (Jess is available free of charge for academic research purposes.)
5. Hartson, H.R., Castillo, J.C., Kelso, J., Kamler, J., & Neale, W.C. Remote Evaluation: The Network as an Extension of the Usability Laboratory. Procs. CHI'96 Human Factors in Computing Systems, 1996, pp. 228-235.