

Discovering Interesting Association Rules from Legal Databases

SASHA IVKOVIC

School of Information Technology and Mathematical Sciences, University of Ballarat, Ballarat, Victoria, Australia. E-mail: s.ivkovic@ballarat.edu.au

JOHN YEARWOOD

School of Information Technology and Mathematical Sciences, University of Ballarat, Ballarat, Victoria, Australia. E-mail: jly@ballarat.edu.au

ANDREW STRANIERI

Victorian Partnership for Advanced Computing and Department of Computer Science and Computer Engineering, La Trobe University, Bundoora, Victoria, Australia. E-mail: stranier@cs.latrobe.edu.au

ABSTRACT *The Knowledge Discovery from Databases (KDD) technique called ‘association rules’ is applied to a large data set representing applicants for government-funded legal aid. Results indicate that KDD can be an invaluable tool for legal analysts. Association rules discovered identify associations between variables that are present in the data set though are not necessarily causal. Interesting rules can prompt analysts to formulate hypotheses for further investigation. The identification of interesting rules is typically performed using an objective measure of ‘interesting’ although this measure is often not sufficiently accurate to eliminate all uninteresting rules. In this article, a subjective measure of interestingness is adopted in conjunction with the objective measures. This leads to the ability to focus more accurately on those rules that surprise the analyst and are therefore more likely to be interesting. In general, KDD techniques have not been applied to law despite possible benefits because data is often stored in narrative form rather than in structured databases. However, the impending introduction of data warehouses that collect data from a number of organizations across a legal system presents invaluable opportunities for analysts using KDD.*

1. Introduction

The field of Knowledge Discovery from Databases (KDD) integrates techniques from artificial intelligence, mathematics and statistics for the discovery of interesting, previously unknown and potentially useful information from large structured data sets. The benefits of KDD have been realized in numerous commercial and scientific fields where very large databases are the norm. Frawley *et al.* (1991) provide an overview of KDD and its applications.

Data in the legal sphere is not currently stored in structured databases to the same extent as it is in other fields. However, with the increasing use of information systems for case management and litigation support, integrated across many agencies in the legal system, huge volumes of data in warehouse systems will soon become commonplace. By automatically discovering knowledge from data warehouses, KDD techniques can be indispensable tools for ensuring the efficient delivery of legal services and just outcomes in law.

According to Fayyad *et al.* (1996), KDD techniques in general can be grouped into four categories:

- *Classification.* The aim of classification techniques is to group data into predefined categories such as 'pro-plaintiff' or 'pro-defendant'. KDD classification techniques have been applied by Wilkins & Pillaipakkamnat (1997) to predict time to disposition; Zeleznikow *et al.* (1994) to automatically generate rules for dynamic case-based retrieval; Hobson & Slee (1993) to predict theft outcomes; Bench-Capon (1993) to predict social security outcomes; and Stranieri *et al.* (1999) to model judicial discretion in Australian family law.
- *Clustering.* The aim of clustering techniques is to analyze data in order to group the data meaningfully. Clustering techniques have been applied to law by Schweighofer (1999), Pannu (1995), Hayes (1992), Thompson (2001), Bruninghaus & Ashley (2001), Yearwood (2000), Moens *et al.* (1999) and Daniels & Rissland (1997).
- *Series Analysis.* The aim of series analysis techniques is to discover sequences within temporal data. Very few studies have been performed that analyze sequences of data in law. A notable exception is the study by Rissland & Friedman (1995) that aimed to detect significant change in the concept of bankruptcy in US courts.
- *Association.* The objective of association techniques is to discover ways in which data elements are associated with other data elements. For example, an association between the gender of litigants and the outcome of their cases may surprise analysts and stimulate hypotheses to explain the phenomena. Association techniques have not been applied to law to the same extent as classification and clustering techniques.

Stranieri *et al.* (2000) have illustrated that association rule generators can highlight interesting associations in a small data set in family law. In their study, the *Apriori* algorithm advanced by Agrawal *et al.* (1993) was applied to a small family law data set to suggest hypotheses for future investigation. Governatori & Stranieri (2001) applied association rules for the semi-automated discovery of defeasible rules.

The present study illustrates that KDD, using association rule generation on a very large data set (over 380,000 records), can lead to useful information systems for the discovery of knowledge. However, for this to occur, techniques must be developed that filter out interesting rules from the many thousands of potentially interesting rules typically discovered. This has been achieved with the introduction of a new subjective measure of interestingness.

The next section provides an overview of association rules and interestingness. Section 3 discusses issues to do with data selection, pre-processing and transformation. Section 4 presents the web-based association rule generator,

WebAssociator, which uses the interestingness criteria developed here. Section 5 discusses the results obtained.

2. Association rules and interestingness

An association rule identifies a link between two or more data attributes (i.e., variables). A famous, hypothetical example of an association rule that is generated from a supermarket database of purchases is *if nappies, then beer (confidence 80%)*. This is read as ‘beer is purchased together with nappies in the same transaction on 80% of the occasions that nappies are purchased’. The rule is drawn directly from the data. It is not a generalization from the data but merely identifies an association between the purchase of nappies and the purchase of beer that has been observed. The association is not necessarily causal so the rule cannot appropriately be used to predict new purchases.

Agrawal *et al.* (1993) were the first to describe an algorithm called ‘*Apriori*’ for discovering association rules from databases. In the supermarket example given above, we simply counted the number of times nappies were purchased and the number of times beer was purchased together with nappies and express the result as an association rule with a level of confidence. The difficulty arises in counting all combinations of attributes in order to arrive at all possible association rules. For example, if a database has only 3 Boolean attributes—A, B and C—then there are 24 possible association rules (order in the antecedent is not important) with a single variable consequent:

$A \Rightarrow B; A \Rightarrow C; A \Rightarrow A; B \Rightarrow A; B \Rightarrow C; B \Rightarrow B; C \Rightarrow A; C \Rightarrow B; C \Rightarrow C; A,B \Rightarrow A; A,B \Rightarrow B; A,B \Rightarrow C; A,B \Rightarrow A,B; A,C \Rightarrow A; A,C \Rightarrow B; A,C \Rightarrow C; A,C \Rightarrow A,C; B,C \Rightarrow A; B,C \Rightarrow B; B,C \Rightarrow C; B,C \Rightarrow B,C; A,B,C \Rightarrow A; A,B,C \Rightarrow B; A,B,C \Rightarrow C$

Determining the confidence of each rule by examining every possible association rule is called a ‘brute force’ approach and is feasible only with small databases, though recent advances by Coenen (2000) suggest that brute force may even be feasible with large sets.

Agrawal *et al.* (1993) developed the *Apriori* algorithm in order to discover association rules confidence levels in a more efficient way than counting the combinations of attributes. The mechanics of the algorithm is beyond the scope of this article but it operates by minimizing the number of times each attribute is counted.

The *Apriori* algorithm prompts a user to supply a minimum threshold of confidence. Rules with a confidence level that is very low are usually not very interesting because the association between attributes is low. However, not all association rules with a high confidence are interesting. For example, the ‘*if nappies then beer (80%)*’ rule would probably not be very interesting if there were only a few transactions that involved both nappies and beer in a database of thousands of records. This is measured as the ‘support’ for a rule.

The support for an association rule indicates the proportion of records covered by the set of attributes in the association rule. If there were 1,000 records in total and only 10 of them involved both beer and nappies, then support for the association rule ‘*if nappies then beer*’ is 10/1000 or 1%. A minimum threshold support and confidence can be set in order to define rules that are *interesting*. For example, we may wish only to look at rules that have a support

of 40% or more, and a confidence level of 80% or more. Other measures of interestingness combine support and confidence in other ways.

According to Frawley *et al.* (1991) 'interestingness' refers to the degree to which a discovered pattern is of interest to the user of the system and is driven by factors such as novelty, utility, relevance and usefulness. Bayardo & Agrawal (1999) reviewed several methods proposed for finding the most interesting rules using a variety of metrics. However, it seems unlikely that a single metric will be discovered that quantifies the 'interestingness' or 'goodness' of a rule across all data sets.

Klemettinen *et al.* (1994) demonstrated that templates can be used to elicit the form of interesting rules (and uninteresting) rules from users. Piatetsky-Shapiro & Simoudis (1996) also claim that objective factors alone (such as confidence and support) are deficient and that knowledge-based factors such as the user's domain knowledge also have to be included. Following this line of thought, Sahar (1999) incorporated the user's domain knowledge into his system. Rather than trying to find what is interesting he eliminated a large family of rules that are not interesting. This was done by asking the user to classify a few possible rules so that their elimination can bring about the automatic elimination of many other rules.

Liu *et al.* (2000) divided interestingness measures into two categories: objective or subjective. Objective interestingness measures something about a rule's structure while a subjective measure of interestingness is dependent upon the specific need and prior knowledge of the user. A measure of subjective interestingness was advanced by Piatetsky-Shapiro & Simoudis (1996) in analyzing the difference in support and confidence that a user expects to see in a rule with the values actually observed. Piatetsky-Shapiro & Simoudis claim that deviations are capable of identifying rules that differ from our expectations. Since those rules differ from what we expect, they are by definition interesting. For example, an association rule like *if (Title = Mrs), then (Gender = female) (confidence 0.98)* (read as '98% of the population with title Mrs are female') is not interesting because this corresponds too closely to prior knowledge or expectations.

In the present study a web-based application has been developed to enable users to indicate the level of confidence they expect from rules before they are generated. The confidence values drawn from actual association rules are compared with the values expected in order to calculate a deviation value. A deviation is a discrepancy between the user's confidence prediction and the actual confidence for a rule. Three different categories of deviation are defined: Zero Level of Surprise (ZLS), Positive Level of Surprise (PLS) and Negative Level of Surprise (NLS).

The Zero Level of Surprise (ZLS) category of deviations represents rules that were previously known to the user because the degree of association observed was similar to that expected. Positive Level of Surprise (PLS) deviations are those where the user was surprised because they expected a low association between two attributes but a high association was observed. In this case, the user underestimated the association. A user perplexed by an association demonstrates positive surprise because the association is presumably unexpected. A Negative Level of Surprise (NLS) occurs when the user overestimates the association in that a high level of association is expected but a low level is observed. PLS and NLS deviations are associated with rules that were previously unknown. In this study, these are considered interesting.

The effectiveness of this new subjective measure of interestingness is demonstrated with a web-based data analysis tool that generates association rules from a large data set drawn from over 380,000 applications for legal aid. Applications are made to a semi-government legal aid organization called Victoria Legal Aid (VLA). VLA aim to use the KDD techniques illustrated here to further their objective of providing legal aid in the most effective, economic and efficient manner to those in the community with the greatest need.

The data set and web-based application, WebAssociator, is described in the next section.

3. Data selection, pre-processing and transformation of the VLA data set

Data selection is the process of determining which data will be used in data mining. This involves selecting a target set of variables for consideration and a set of records from which to gather data for the selected variables. The VLA data set contains over 380,000 records collected from 11 Victorian regional offices during the three-year period 1997 to 1999.

Victoria Legal Aid collects data on over 300 variables related to applications for legal aid. An association rule generation mining exercise using all available variables requires extensive resources and probably could not be done in real time on existing serial processing machines. Some selection of variables for consideration was required.

Following the observation by Yoon *et al.* (1999) that domain knowledge is useful for restricting the knowledge discovery process, VLA experts were involved in the selection of variables for this study. Experts selected seven variables that were important and interesting for rule generation. These were: sex (3 possible values 'M', 'F', 'Both'); age ('date of birth'); occupation (7 numbered categories); reason for refusal (pre-specified VLA refusal codes); law type (e.g., 'civil law', 'family law'); decision ('granted' or 'not granted'); and dealing type ('advice', 'court appearance', etc.).

Data pre-processing is a necessary step in dealing with large data sets that contain noisy, missing or irrelevant data. However, the VLA data set is presumed to contain a minimum of noise. 'Noise' is generally regarded as values that are recorded incorrectly because of data entry processes. The database management system in use by VLA performs integrity checks on all input data so absurd values cannot be entered.

Similarly, minimal data transformation was required. The attribute 'date of birth' is too finely grained for meaningful processing in that a rule that indicates *if date of birth = 26 September 1959, then decision = grant aid (with a support x%, confidence y%)* is unlikely to be useful. To meaningfully mine such quantitative (numerical) variables, the values had to be partitioned into intervals. Domain experts were again approached to define the intervals. Partitions identified as useful were: 'under 18', '21-25', '26-28', '29-30', '31-35', '36-40', '41-50' and 'over 50'.

The data selection, pre-processing and transformation phases of the KDD exercise for the VLA data set were performed relatively easily. This may not be the case for all large data sets. The data mining phase described in the next section involves the use of the web-based tool and subjective measure of interestingness.

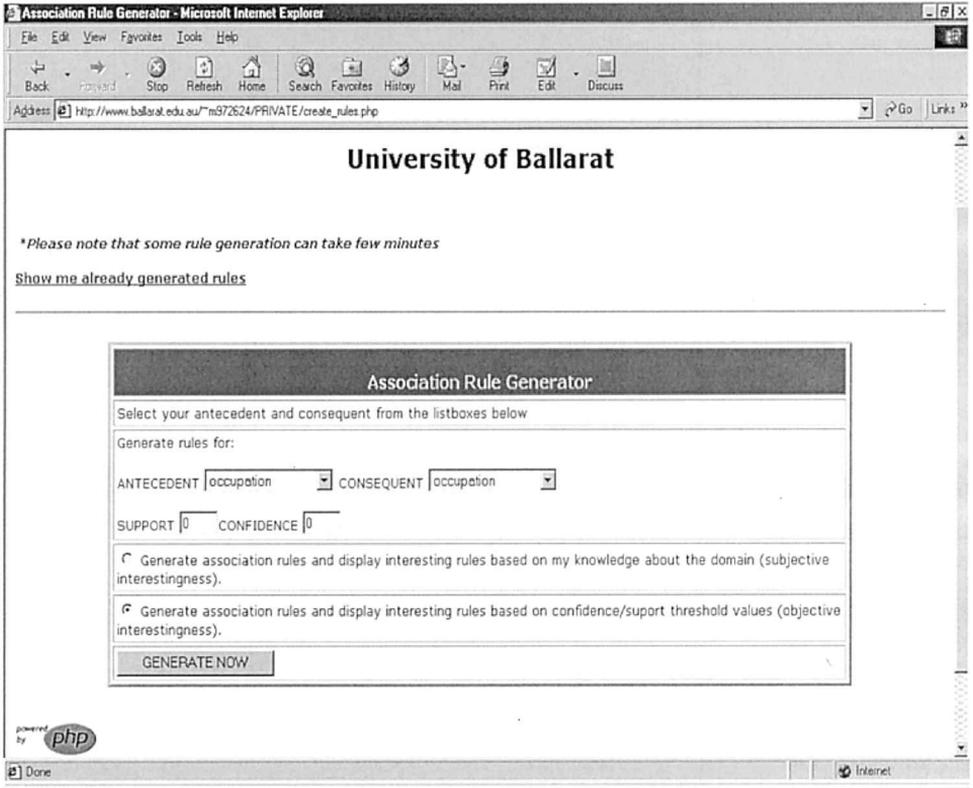


Figure 1. WebAssociate form for the user selection of an antecedent and consequent.

4. Generating association rules from the VLA data set

There are over 26,000 possible association rules that could be generated from the VLA data set. As discussed above, the use of an objective measure of interestingness, such as a ratio of confidence to support, could considerably reduce the number to be assessed by users. However, the objective of this work was to devise a new measure of interestingness that was subjective in that it involved domain knowledge from experts.

WebAssociate restricts the generation of rules to those that involve only one antecedent and one consequent. Users select the antecedent and consequent of interest using the web form illustrated in Figure 1.

Following selection of the antecedent and consequent of interest to the user, the form illustrated in Figure 2 is presented to the user. This prompts the user to provide a degree of correlation (confidence) between 0 and 1 for each association rule possible involving the antecedent and consequent selected. For example, a confidence of 0 on the rule *if sex = F, then law type = criminal law* indicates that the user believes there to be no association between female gender and criminal law type.

The deviation between the user's confidence level and the confidence actually observed is displayed in graphical form in Figure 4. The horizontal axis represents the user's expected level of confidence. The vertical axis represents

C:\AAA-Research\May-2001-Thesis\myRules\VLA_RULES.html - Microsoft Internet Explorer

File Edit View Go Favorites Help

Your name:

Id	Rule	Correlation
1	IF sex = F THAN law type = CIVIL LAW	<input type="text"/>
2	IF sex = F THAN law type = CRIMINAL LAW	<input type="text"/>
3	IF sex = F THAN law type = FAMILY LAW	<input type="text"/>
4	IF sex = M THAN law type = CIVIL LAW	<input type="text"/>
5	IF sex = M THAN law type = CRIMINAL LAW	<input type="text"/>
6	IF sex = M THAN law type = FAMILY LAW	<input type="text"/>
7	IF sex = F THAN reason for refusal =	<input type="text"/>
8	IF sex = F THAN reason for refusal = GUIDELINES	<input type="text"/>
9	IF sex = F THAN reason for refusal = INSUFFICIENT FUNDS	<input type="text"/>
10	IF sex = F THAN reason for refusal = MERIT	<input type="text"/>
11	IF sex = M THAN reason for refusal =	<input type="text"/>
12	IF sex = M THAN reason for refusal = GUIDELINES	<input type="text"/>
13	IF sex = M THAN reason for refusal = INSUFFICIENT FUNDS	<input type="text"/>
14	IF sex = M THAN reason for refusal = MERIT	<input type="text"/>

Figure 2. WebAssociate form for the elicitation of user association values.

the actual level of confidence. Each point on the graph represents a rule. For example, the point uppermost in the right hand corner corresponds to a rule where the user expected a very high (1.0) association and the observed value was also high (0.82). This point is close to the 45 degree ($x = y$) line which is the line of perfect agreement.

Points near the line of perfect agreement represent rules that have not surprised the user. These rules are not interesting. Points that fall in the triangle in the upper left hand corner represent rules where a positive level of surprise (i.e., rules where the user has underestimated the association value) has occurred. Points in the triangle in the lower right hand corner represent rules where a negative level of surprise has occurred.

The boundary between Positive Level of Surprise (PLS) and Zero Level of Surprise (ZLS), and that between Negative Level of Surprise (NLS) and ZLS, is determined as follows:

- Set the ZLS threshold to 0.1 and -0.1. This is because points within 0.1 of the line of perfect agreement can be regarded as sufficiently close to zero.
- Set the PLS line to 0.45 so that only those points above that line are considered interesting. However, if the majority of points fall between the 0.45 line and the ZLS line then too few points will be considered interesting. This can occur if the mean of a user's values are less than 0.45. That is, the values are bunched near the line of perfect agreement. In order to guard against this, the mean of the positive deviation values between 0.1 and 1.0 is calculated first. If this mean is less than 0.45, then the PLS line is set at the mean. If it is greater than 0.45, then the PLS line stays at 0.45. The PLS line in Figure 4 remains set at 0.45.

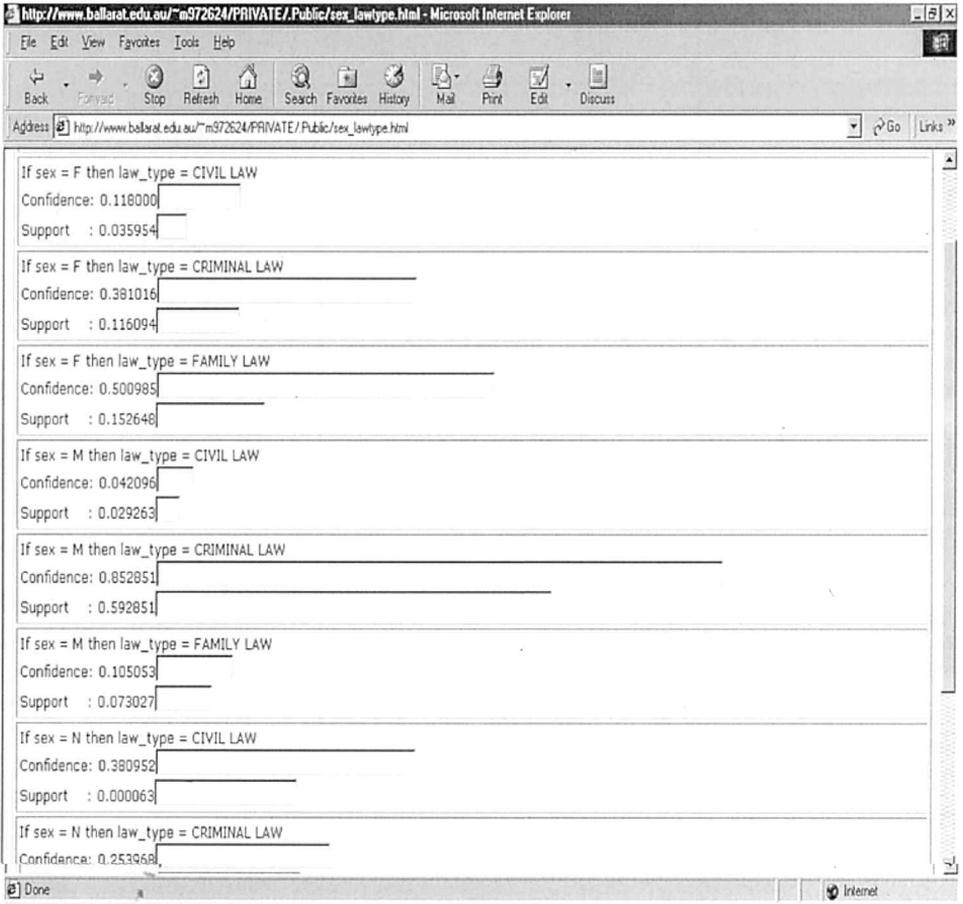


Figure 3. Sample confidence and support for actual association rules.

- Step 2 is repeated for the NLS line using -0.45 rather than 0.45 . If the mean of deviations between -0.1 and -1.0 is greater than -0.45 then the NLS is set at the mean otherwise it is set at -0.45 . The NLS line in Figure 4 has been brought in to -0.25 .

Figure 5 illustrates the plot of actual versus expected confidence as it appears to the expert in WebAssociate.

5. Interpretation

WebAssociate was tested with values from three experts on 144 different single antecedent and single consequent rules. For Expert 1, 27 of the 144 rules (22.9%) were regarded as interesting according to our measure of interestingness. A total of 20 of the interesting rules had a *positive level of surprise*. This means that the expert underestimated the strength of 13.9% of all rules and 75% of the interesting rules. A total of seven of the interesting rules had a *negative level of surprise*. This expert overestimated the strength of 4.8% of all rules and 25.9% of

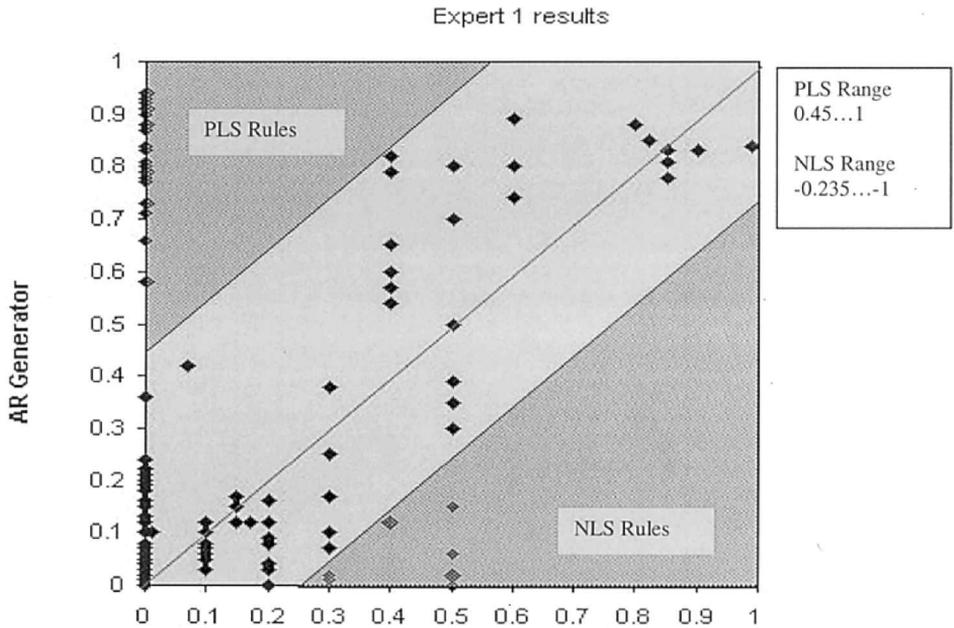


Figure 4. Deviation between expected and actual confidence values for one expert.

the interesting rules. Table 1 depicts a sample of the 144 rules and the deviation value for Expert 1. Rule 121 in the top row indicates a deviation of 0.77. Expert 1 expected little association between applicants in the 35 to 40 age group and approved decisions whereas there was a high association observed.

An interesting rule such as *if place of birth = Vietnam, then law type = criminal law* had an actual confidence value of 0.901 which was much higher than that expected by the experts. An interesting rule is useful in that it suggests

Table 1. Deviation values for a sample of rules

Id	Rule	Deviation
121	IF age = 35–40, THEN decision = APPROVED	0.77
124	IF age = 40–50, THEN decision = APPROVED	0.73
131	IF age = 21–25, THEN dealing type = CASE	0.92
133	IF age = 25–28, THEN dealing type = CASE	0.94
135	IF age = 28–30, THEN dealing type = CASE	0.93
137	IF age = 30–35, THEN dealing type = CASE	0.94
25	IF occupation = 8, THEN law type = CRIMINAL LAW	0.58
49	IF occupation = H, THEN decision = APPROVED	0.84
127	IF age = under 18, THEN decision = APPROVED	0.71
129	IF age = 18–21, THEN dealing type = CASE	0.91
9	IF sex = F, THEN reason for refusal = INSUFFICIENT FUNDS	–0.50
10	IF sex = F, THEN reason for refusal = MERIT	–0.29
12	IF sex = M, THEN reason for refusal = GUIDELINES	–0.44
14	IF sex = M, THEN reason for refusal = MERIT	–0.48
23	IF occupation = 7, THEN law type = FAMILY LAW	–0.35
81	IF age = 18–21, THEN law type = CIVIL LAW	–0.28

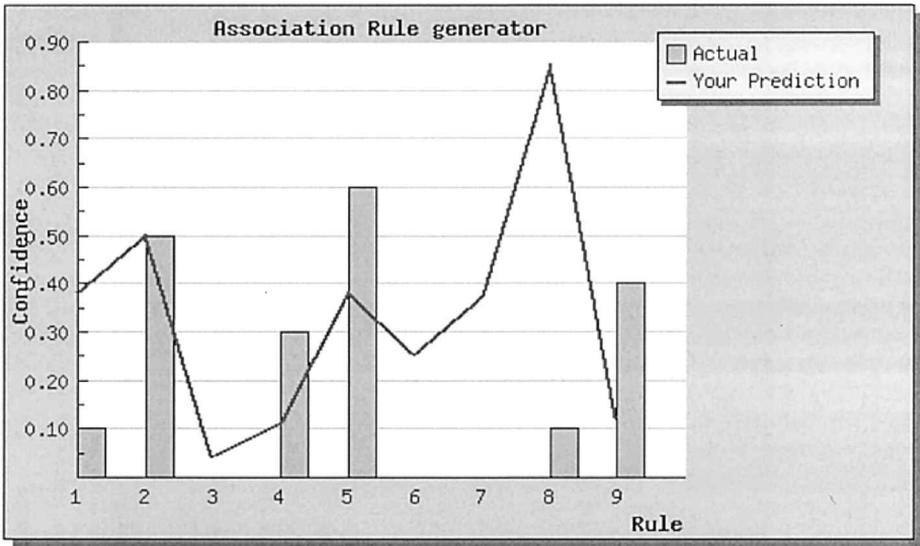


Figure 5. The categorization of all interesting rules for one expert.

hypotheses. In this case, experts in an unstructured discussion advanced the following hypotheses:

- Applicants from Vietnam tended only to apply for support for criminal law matters whereas other groups applied for legal aid for a range of matters including family and civil law.
- Applicants from Vietnam committed more crime than other groups.
- There was a spurious effect. For example, perhaps males from Vietnam are much more likely to apply for aid than females from Vietnam. The link between males and criminal matters is already clearly established.
- VLA promotional material was not readily accessible to persons of Vietnamese origin. This may be because criminal justice groups within the legal system come into direct contact with applicants and suggest they apply for aid.

Each hypothesis direction was explored in informal discussion and the first seemed to carry far more credence than others—very few persons of Vietnamese origin seem to apply for legal aid for family law matters, for instance.

Association rule generators are particularly useful for discovering associations that would otherwise not be considered. Surprising associations prompt the creation of hypotheses that would explain the associations. The Vietnamese example above is limited in that the association is discovered only between two items: ‘Vietnam’ as the place of birth and ‘criminal offence’ as the type of offence. A more comprehensive association rule generator discovers associations between multiple item sets. For example, there may well be interesting associations between a particular social background, level of education, proficiency with English, place of birth, type of crime and other sociocultural factors. Association rules with more antecedents would lead to the suggestion of hypotheses that are more complex, though undoubtedly more apt as explanations of social phenomena.

The association rule generator may facilitate the process of forming hypotheses. However, once formed, hypotheses need to be explored using conventional

methods. A causal link cannot be assumed to exist in an association so experimental. Therefore, ethnographic or other types of research methodologies must be applied to investigate hypotheses.

WebAssociate is most appropriately deployed in a judicial support system as a device that supports a judge in forming hypotheses regarding the relationship between decisions made and case features. The hypotheses are unlikely to be invoked to support a decision in any one case but rather are useful to ensure that decision-making consistency is maintained and spurious associations that reflect unjust processes are identified. This is most readily illustrated in common law jurisdictions as support for types of *stare decisis* that Wassestrom (1961) calls 'local' and 'personal' *stare decisis*.

Wassestrom (1961) identified three types of *stare decisis*:

- *Traditional stare decisis*. The same, equivalent or more favorable fact pattern in the same or higher court leads to the same decision. It is this kind of *stare decisis* that Kovacs (1992) claims has not occurred fully in the discretionary domain of family law because the High Court of Australia has failed to lay down specific constraints for trial judges to follow. Higher courts openly publish judgements that provide, to one extent or another, clear principles. Published higher court judgements enhance consistency in decision making in lower courts.
- *Local stare decisis*. The same, equivalent or more favorable fact pattern in the same court leads to the same decision. For example, this manifests itself in family law as a desire for Family Court judges to exercise discretion in a manner that is consistent with other judges of the court. In a legal aid setting, this manifests itself as a desire for a grant officer to be consistent with other grant officers.
- *Personal stare decisis*. The same, equivalent or more favorable fact pattern found by the same judge in different cases leads to the same decision. This manifests itself in the Family Court as the tendency of an individual judge to be consistent with the way he or she exercised discretion in past/similar cases.

The application of the association rule generator described in this work can enhance local and personal *stare decisis*. To enhance local *stare decisis*, associations drawn from decisions made by one judge can be compared with associations from those made by other judges. This provides a mechanism for the identification of associations that reflect undesirable inconsistencies across judges' many decisions. Similarly, an individual judge, intent on analyzing his or her own decision-making processes to enhance personal *stare decisis* can compare associations from decisions in one time period to those from other periods.

Periodic reports generated by a database management system can, in some circumstances, provide the same information as association rule generators. However, reports must largely be pre-specified. An association rule generator, particularly when integrated with a subjective measure of interestingness, can automatically generate rules at a fine level of granularity and identify those that are likely to be interesting.

6. Conclusion

Association rules generators scan a database to generate rules of the form *if A and B, then C* (*confidence = x%, support = y%*). Typically too many rules are

generated for real time processing and manual analysis. The *Apriori* algorithm (Agrawal *et al.*, 1993) generates only those rules that meet an objective level of interestingness specified as a minimum support and minimum confidence threshold value. Although this leads to the generation of rules in real time, manual analysis of discovered rules is still tedious.

In this article a subjective measure of interestingness has been presented. This measure operates by asking a user to select a subset of rules and then to estimate the degree of association they expect between antecedent and consequents of those rules. The expected association is compared with that actually observed. Rules where there is a zero-level deviation are assumed to be of little interest, whereas rules that depict a positive or a negative deviation are taken to be interesting.

The association rule generator and subjective measure of interestingness have been integrated into a web-based program called WebAssociator and tested with a large data set from Victoria Legal Aid, a provider of government-funded legal services in Australia. The tool has been found to be very useful in discovering associations between attributes of persons applying for legal aid that would otherwise not have been noticed.

Future system development aims to extend the current tool by enabling the generation of multiple antecedent, single consequent rules in addition to performance improvements with some data sets. Future research aims to evaluate the effectiveness of data analysis with this tool.

References

- Agrawal, R., Imielinski, T. & Swami, A. (1993) Mining association rules between sets of items in large databases. *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data* (New York, ACM Press), pp. 207–216.
- Bayardo, R. & Agrawal, R. (1999) Mining the most interesting rules. *Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, ACM Press), pp. 145–154.
- Bench-Capon, T.J.M. (1993) Neural networks and open texture. *Proceedings of the 4th International Conference on Artificial Intelligence and Law* (New York, ACM Press), pp. 292–297.
- Bruninghaus, S. & Ashley, K.D. (2001) Improving the representation of legal case texts with information extraction methods. *Proceedings of the 8th International Conference on Artificial Intelligence and Law* (New York, ACM Press), pp. 42–49.
- Coenen, F. (2000) *A Second Brute Force Algorithm for 'Basket Analysis'*. Technical Report, Department of Computer Science, University of Liverpool, UK.
- Daniels, J. & Risland, E. (1997) Finding legally relevant passages in case opinions. *Proceedings of the 6th International Conference on Artificial Intelligence and Law* (New York, ACM Press), pp. 39–46.
- Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P. (1996) The KDD process for extracting useful knowledge from volumes of data. *Commun. ACM* 39(11): pp. 27–34.
- Frawley, W.J., Piatetsky-Shapiro, G. & Matheus, C.J. (1991) Knowledge discovery in databases: An overview. *Knowledge Discovery in Databases* (Cambridge, MA, AAAI/MIT Press), pp. 1–27.
- Governatori, G. & Stranieri, A. (2001) Towards the application of association rules for defeasible rule discovery. *Fourteenth Annual International Conference on Legal Knowledge and Information Systems, Jurix'01* (December) (Amsterdam, IOS Press), pp. 63–75.
- Hayes, P. (1992) Intelligent high volume text processing using shallow, domain specific techniques. In P.S. Jacobs (Ed.) *Text-Based Intelligent Systems: Current Research and Practice in Information Extraction and Retrieval* (Hillsdale, NJ, Lawrence Erlbaum), pp. 227–236.
- Hobson, J.B. & Slee, D. (1993) Rules, cases and networks in a legal domain. *Law, Computers & Artificial Intelligence* 2(2): pp. 119–135.
- Klemettinen, M., Mannila, H., Ronkainen, P., Toivonen, T. & Verkamo, A.I. (1994) Finding interesting rules from large sets of discovered association rules. *Proceedings of the 3rd International Conference on Information and Knowledge Management* (New York, ACM Press), pp. 401–407.

- Kovacs, D. (1992) *Family Property Proceedings in Australia* (Sydney, Butterworths).
- Liu, B., Hsu, W., Chen, S. & Ma, Y. (2000) Analyzing the subjective interestingness of association rules. *IEEE Intelligent Systems* (September/October): pp. 47–55.
- Moens, M.F., Uyttendaele, C. & Dumortier, J. (1999) Abstracting of legal cases: The SALOMON experience. *Proceedings of the 7th International Conference on Artificial Intelligence and Law* (New York, ACM Press), pp. 114–122.
- Pannu, A.S. (1995) Using genetic algorithms to inductively reason with cases in the legal domain. *Proceedings of the 5th International Conference on Artificial Intelligence and Law (ICAIL'95)* (New York, ACM Press), pp. 175–184.
- Piatetsky-Shapiro, G. & Simoudis, E. (1996) Mining business databases. *Communications of the ACM* (New York, ACM Press), pp. 42–48.
- Rissland, E.L. & Friedman, M.T. (1995) Detecting change in legal concepts. *Proceedings of the 5th International Conference on Artificial Intelligence and Law (ICAIL'95)* (New York, ACM Press), pp. 127–136.
- Sahar, S. (1999) Interestingness via what is not interesting. *Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, ACM Press), pp. 332–336.
- Schweighofer, E. (1999) The revolution in legal information retrieval or: The empire strikes back. *Journal of Information, Law and Technology* at <http://elj.warwick.ac.uk/jilt/99-1/schweigh.html>
- Stranieri, A., Zeleznikow, J., Gawler, M. & Lewis, B. (1999) A hybrid rule-neural approach for the automation of legal reasoning in the discretionary domain of family law in Australia. *Artificial Intelligence and Law* 7(2–3): pp. 153–183.
- Stranieri, A., Zeleznikow, J. & Turner, H. (2000) Data mining in law with association rules. *Proceedings of IASTED International Conference on Law and Technology* (Innsbruck, Austria, IASTED), pp. 129–135.
- Thompson, P. (2001) Automatic categorisation of case law. *Proceedings of the 8th International Conference on Artificial Intelligence and Law* (New York, ACM Press), pp. 70–77.
- Wassestrom, R. (1961) *The Judicial Decision: Toward a Theory of Legal Justification* (Stanford, CA, Stanford University Press).
- Wilkins, D. & Pillaipakkamnatt, K. (1997) The effectiveness of machine learning techniques for predicting time to case disposition. *Proceedings of the 6th International Conference on Artificial Intelligence and Law* (New York, ACM Press), pp. 39–46.
- Yearwood, J. (2000) Case-based retrieval of refugee review tribunal text cases. In Oskamp, A., De Mulder, R., Van Noordwijk, C., Grutters, C.A.F.M., Ashley, K. & Gordon, T. *Legal Knowledge-based Systems (JURIX'97)* (Amsterdam, IOS Press), pp. 67–84.
- Yoon, S., Henschen, L., Park, E.K. & Makki, S. (1999) Using domain knowledge in knowledge discovery. *Proceedings of the 8th International Conference on Information Knowledge Management* (New York, ACM Press), pp. 243–250.
- Zeleznikow, J., Vossos, G. & Hunter, D. (1994) The IKBALS project: Multimodal reasoning in legal knowledge based systems. *Artificial Intelligence and Law* 2(3): pp. 169–203.