# Explanation-Oriented Association Mining Using a Combination of Unsupervised and Supervised Learning Algorithms

Y.Y. Yao, Y. Zhao, R.B. Maguire

Department of Computer Science, University of Regina
Regina, Saskatchewan, Canada S4S 0A2
E-mail: {yyao, yanzhao, rbm}@cs.uregina.ca

**Abstract.** We propose a new framework of explanation-oriented data mining by adding an explanation construction and evaluation phase to the data mining process. While traditional approaches concentrate on mining algorithms, we focus on explaining mined results. The mining task can be viewed as unsupervised learning that searches for interesting patterns. The construction and evaluation of mined patterns can be formulated as supervised learning that builds explanations. The proposed framework is therefore a simple combination of unsupervised and supervised learning. The basic ideas are illustrated using association mining. The notion of conditional association is used to represent plausible explanations of an association. The condition in a conditional association explicitly expresses the plausible explanations of an association.

## 1 Introduction

Data mining is a discipline concerning theories, methodologies, and in particular, computer systems for exploring and analyzing a large amount of data. A data mining system is designed with an objective to automatically discover, or to assist a human expert to discover, knowledge embedded in data [2, 6, 21]. Results, experiences and lessons from artificial intelligence, and particularly intelligent information systems, are immediately applicable to the study of data mining.

By putting data mining systems in the wide context of intelligent information systems, one can easily identify certain limitations of current data mining studies. In this paper, we focus on the explanation facility of intelligent systems, which has not received much attention in data mining community. We present a new explanation-oriented framework for data mining by combining unsupervised and supervised learning.

For clarity, we use association mining to demonstrate the basic ideas. The notion of conditional association is used to explicitly state the conditions under which the association occurs. An algorithm is suggested. Conceptually, it consists of two parts and uses two data tables. A transaction data table is used to learn an association in the first step. An explanation table is used to construct an explanation of the association in the second step.

## 2 Motivations

In the development of many branches of science such as mathematics, physics, chemistry, and biology, the discovery of a natural phenomenon is only the first step. The important subsequent tasks for scientists are to build a theory accounting for the phenomenon and to provide justifications, interpretations, and explanations of the theory. The interpretations and explanations enhance our understanding of the phenomenon and guide us to make rational decisions [22].

Explanation plays an important role in learning and is an important functionality of many intelligent information systems [5, 8, 9, 11, 15]. Dhaliwal and Benbasat argue that the role of constructing explanation is to clarify, teach, and convince [5]. Human experts are often asked to explain their views, recommendations, decisions or actions. Users would not accept recommendations that emerge from reasoning that they do not understand [9].

In an expert system, an explanation facility serves several purposes [17]. It makes the system more intelligible to the user, helps an expert to uncover shortcomings of the system, and help a user to feel more assured about the recommendations and actions of the system. Typically, the system provides two basic types of explanations: the *why* and the *how*. A why type question is normally posed by a user when the system asks the user to provide some information. A how type question is posed by a user if the user wants to know how a certain conclusion is reached. Wick and Slagle [19] proposed a journalistic explanation facility which include the six elements *who*, *what*, *where*, *when*, *why*, and *how*.

A data mining system may be viewed as an intermediate system between a database or data warehouse and an application, whose main purpose is to change data into usable knowledge [21]. To achieve this goal, the data mining system should provide necessary explanations of mined knowledge. A piece of discovered knowledge is meaningful and trustful only if we have an explanation. An association does not immediately offer an explanation. One needs to find explanations regarding *when*, *where*, and *why* an association occurs. If a data mining system is an interactive system, it must also provide explanations for its recommendations and actions. For a knowledge-based data mining systems, explanation of the use of knowledge is also necessary to make the mining process more understandable by a user. The observations and results regarding explanations in expert systems are applicable to data mining systems. In order to make data mining a well-accepted technology, more attention must be paid to the needs and wishes for explanations from its end users. Without the explanation functionality, the effectiveness of data mining systems is limited.

On the other hand, studies in data mining have been focused on the preparation, process and analysis of data. Little attention is paid to the task of explaining discovered results. There is clearly a need for the incorporation of an explanation facility into a data mining process.

It is commonly accepted that a data mining process consists of the following steps: data selection, data preprocessing, data transformation, pattern discovery, and pattern evaluation [6]. Several variations have been studied by many authors [7, 10, 16]. By adding an extra step, explanation construction and eval-

uation, we can obtain a framework of explanation-oriented data mining. This leads to a significant step from detecting the *existence* of a pattern to searching for the underlying *reasons* that explain the existence of the pattern.

## 3 Explanation-oriented association mining

Association mining was first introduced using transaction databases and deals with purchasing patterns of customers [1]. A set of items are associated if they are bought together by many customers. Some authors extended the original associations to negative associations [20].

### 3.1 Conditional associations and explanation evaluation

The reasons for the occurrence of an association can not be provided by the association itself. One needs to construct and represent explanations using other information. More specifically, if one can identify some conditions under which the occurrence of the association is more pronounced, the condition may provide some explanation. By adding time, place, customer features (profiles), and item features as conditions, we may identify *when*, *where* and *why* an association occurs, respectively.

The notion of conditional associations has been discussed by many authors in different contexts [4, 14, 18]. Typically, conditions in conditional associations mining are used as constraints to restrict a portion of the database to mine useful associations. For explanation-oriented association mining, we take a reverse process. We first mine association and then search for conditions.

We can profile transactions by customers, places, and time ranges. Domain specific knowledge is used to select a set of profiles and to form an explanation table. Different explanation tables can be constructed, which lead to different explanations. Each explanation table may or may not be able to provide a satisfactory explanation. It may also happen that each table may be able to explain only some aspects of the association.

Let $\phi\psi$ denote an association discovered in a transaction table. Let $\chi$ denote a condition expressible in the explanation table. A conditional association is written by $\phi\psi \mid \chi$. Suppose $s$ is a measure that quantifies the strength of the association. An example of such measures is the *support* measure used in association mining [1]. Plausible explanations may be obtained by comparing the values $s(\phi\psi)$ and $s(\phi\psi \mid \chi)$. If $s(\phi\psi) > s(\phi\psi \mid \chi)$, namely, the association $\phi\psi$ is more pronounced under the condition $\chi$, we say that $\chi$ provides a plausible explanation for $\phi\psi$, otherwise, $\chi$ does not. We may also introduce another measure $g$ to quantify the quality of conditions [22]. Explanations are evaluated jointly by the two measures.

### 3.2 Explanation construction

Construction of explanations is equivalent to finding conditions in conditional associations from an explanation table.

Suppose $\phi\psi$ is an association of interest. We can classify transactions into two classes, those that satisfy the association, and those that do not satisfy the association. With this transformation, searching for conditions in conditional associations can be stated as learning of classification rules in the explanation table. Any supervised learning algorithm, such as ID3 [12], its later version C4.5 [13], or PRISM [3], may be used to perform this task.

### 3.3 An algorithm for explanation-oriented association mining

Explanation-oriented associating mining consists of two steps. In the first step, an unsupervised learning algorithm, such Apriori [1] or a clustering algorithm, is used to discover an association. In the second step, an association of interest is used to create a label in the explanation table. Any supervised learning algorithm, such as ID3 [12] or PRISM [3], is used to learn classification rules, which are in fact conditional associations.

The framework of explanation-oriented association mining is thus a simple combination of existing unsupervised and supervised learning algorithms. As an illustration, the combined Apriori-ID3 algorithm is described below:

**Input:** A transaction table and explanation profiles.
**Output:** Conditional associations (explanations).

**1** Use the Apriori algorithm to generate a set of frequent itemsets in the transaction table. For each $\phi\psi$ in the set, $support(\phi\psi) \geq minsup$.
**2** **If** $\phi\psi$ is interesting
  **2.a** Introduce a binary attribute named *Decision*. Given a transaction $x \in U$, its value on *Decision* is "+" if it satisfies $\phi\psi$ in the transaction table. Otherwise, its value is "-".
  **2.b** Construct an information table by using the attribute *Decision* and explanation profiles. The new table is called an *explanation table*.
  **2.c** By treating *Decision* as the target class, we can apply the ID3 Algorithm to derive classification rules of the form: $\chi \Rightarrow Decision = \text{"}+\text{"}$, which corresponds to the conditional association $\phi\psi \mid \chi$. The condition $\chi$ is a formula in the explanation table, which states the condition $\chi$ under which the association $\phi\psi$ occurs.
  **2.d** Evaluate conditional associations based on statistical measures.

## 4 Conclusion

By drawing results from artificial intelligence in general and intelligent information systems in specific, we demonstrate the needs for explanations of mined results in a data mining process. We show that explanation-oriented association mining can be easily achieved by combining existing unsupervised and supervised learning methods. The main contribution is the introduction of a new point of view to data mining research. An explanation facility may greatly increase the effectiveness of data mining systems.

## References


1. Agrawal, R. and Srikant, R., Fast algorithms for mining association rules in large databases, *Proceedings of VLDB*, 487-499, 1994.
2. Berry, M.J.A. and Linoff, G.S. *Mastering Data Mining: the Art and Science of Customer Relationship Management*, John Wiley & Sons, New York, 2000.
3. Cendrowska, J., PRISM: an algorithm for inducing modular rules, *International Journal of Man-Machine Studies*, **27**, 349-370, 1987.
4. Chen, L., *Discovery of Conditional Association Rules*, Master thesis, Utah State University, 2001.
5. Dhaliwal, J.S. and Benbasat, I., The use and effects of knowledge-based system explanations: theoretical foundations and a framework for empirical evaluation, *Information Systems Research*, **7**, 342-362, 1996.
6. Fayyad, U.M., Piatetsky-Shapiro, G. and Smyth, P. From data mining to knowledge discovery: an overview, in: *Advances in knowledge discovery and data mining*, Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P. and Uthurusamy, R. (Eds.), 1-34, AAAI/MIT Press, Menlo Park, California, 1996.
7. Han, J. and Kamber, M., *Data mining: Concept and Techniques*, Morgan Kaufmann, Palo Alto, CA, 2000.
8. Hasling, D.W., Clancey, W.J. and Rennels, G., Strategic explanations for a diagnostic consultation system, *International Journal of Man-Machine Studies*, **20**, 3-19, 1984.
9. Haynes, S.R., *Explanation in Information Systems: A Design Rationale Approach*, Ph.D. Dissertation, The London School of Economics, University of London, 2001.
10. Mannila, H. Methods and problems in data mining, *Proceedings of International Conference on Database Theory*, 41-55, 1997.
11. Pitt, J., *Theory of Explanation*, Oxford University Press, Oxford, 1988.
12. Quinlan, J.R., Learning efficient classification procedures, in: *Machine Learning: An Artificial Intelligence Approach I*, Michalski, J.S., Carbonell, J.G., and Mirchell, T.M. (Eds.), Morgan Kaufmann, Palo Alto, CA, 463-482, 1983.
13. Quinlan, J.R., *C4.5: programs for machine learning*, Morgan Kaufmann, Palo Alto, CA, 1993.
14. Rauch, J., Association rules and mechanizing hypotheses formation, *Proceedings of ECML workshop proceedings: machine learning as experimental philosophy of science*, 2001.
15. Schank, R. and Kass, A. Explanations, machine learning, and creativity, in: *Machine Learning: An Artificial Intelligence Approach, III*, Kodratoff, Y. and Michalski, R. (Eds.), Morgan Kaufmann, Palo Alto, CA, 31-48, 1990.
16. Simoudis, E. Reality check for data mining. *IEEE Expert*, **11**, 1996.
17. Turban, E. and Aronson, J.E. *Decision Support Systems and Intelligent System*, Prentice Hall, New Jersey, 2001.
18. Wang, K. and He, Y., User-defined association mining, *Proceedings of PAKDD*, 387-399, 2001.
19. Wick, M.R. amd Slagle, J.R. An explanation facility for today's expert systems, *IEEE Expert*, **4**, 1989, 26-36.
20. Wu, X., Zhang, C. and Zhang, S. Mining both positive and negative association rules, *Proceedings of ICML*, 1997.
21. Yao, Y.Y. A step toward foundations of data mining, manuscript, 2003.
22. Yao, Y.Y., Zhao, Y. and Maguire, R.B. Explanation oriented association mining using rough set theory, *Proceedings of International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing*, to appear, 2003.