# Exception Rules in Data Mining

**Olena Daly**
*Monash University, Australia*

**David Taniar**
*Monash University, Australia*

## INTRODUCTION

Data mining is a process of discovering new, unexpected, valuable patterns from existing databases (Frawley, Piatetsky-Shapiro, & Matheus, 1991). Though data mining is the evolution of a field with a long history, the term itself was only introduced relatively recently, in the 1990s. Data mining is best described as the union of historical and recent developments in statistics, artificial intelligence, and machine learning. These techniques are then used together to study data and find previously hidden trends or patterns within.

Data mining is finding increasing acceptance in science and business areas that need to analyze large amounts of data to discover trends, in which they could not otherwise find. Different applications may require different data mining techniques. The main kinds of knowledge that could be discovered from a database are categorized into association rules mining, sequential patterns mining, classification and clustering.

In this article, we concentrate on exception rules mining.

## BACKGROUND

Exception rules mining has attracted a lot of research interest (Déjean, 2002; Grosof & Poon, 2002, 2003; Hellerstein, Ma, & Perng, 2002; Hussain, Liu, Suzuki, & Lu, 2000; Keogh, Lonardi, & Chiu, 2002; Liu, Hsu, Mun, & Lee, 1999; Padmanabhan & Tuzhilin, 2000; Suzuki, 2002a, 2002b; Yamada & Suzuki, 2002; Zhang, Zhang, Yan, & Qin, 2002). Exception rules have been defined as rules with low support and high confidence (Hussain et al., 2000). A traditional example of exception rules is the rule Champagne=>Caviar. The rule may not have high support but it has high confidence. The items are expensive so they are not frequent in the database, but they are always brought together so the rule has high confidence. Exception rules provide valuable knowledge about database patterns.

Exception rules discovery can be classified as either directed or undirected. A directed search obtains a set of exception rules each of which contradicts to a user-specified belief (Liu et al., 1999; Padmanabhan, 2000). An undirected search obtains a set of pairs of an exception rule and a general rule (Hussain et al., 2000; Suzuki, 2002a, 2002b; Yamada & Suzuki, 2002).

Directed search of exception rules will be described next. User-specified beliefs are obtained first. Each of the discovered exception rules contradicts to user-supplied beliefs.

In Liu et al. (1999), post-analysis of the discovered database patterns is performed to identify the most interesting patterns. The technique is characterized by asking the user to specify a set of patterns according to his/her previous knowledge or intuitive feelings. This specified set of patterns is then used by a fuzzy matching algorithm to match and rank the discovered patterns. The assumption of this technique is that some amount of domain knowledge and the user's interests are implicitly embedded in his/her specified patterns. In general, the discovered patterns are ranked according to their conformities to the user's knowledge or their unexpectedness, or their actionabilities.

In terms of unexpectedness, patterns are interesting if they are unexpected or previously unknown to users. In terms of actionability, patterns are interesting if users can do something with them to their advantage. With such rankings, a user can simply check the few patterns on the top of the list to confirm his/her intuitions (or previous knowledge), or to find those patterns that are against his/her expectation, or to discover those patterns that are actionable.

Padmanabhan and Tuzhilin (2000) focus on discovering unexpected patterns and propose methods for discovering a minimal set of unexpected patterns that discover orders of magnitude fewer patterns and retain most of the interesting ones. The approach has been experimentally tested using a case study application in a marketing domain.

The rule A=>B is defined in Padmanabhan and Tuzhilin (2000) to be unexpected with respect to the belief X=>Y on the dataset D if B and Y logically contradict each other, the antecedents of the belief and the rule hold on the same statistically large subset of D, and the rule A, X=>B holds.

Now the undirected method of searching exception rules will be explained. Exception rules will be obtained based on general rules or common sense rules.

In Hussain et al. (2000), a method for mining exception rules is presented based on a novel measure which estimates interestingness relative to its corresponding common sense rule and reference rule. Common sense rules are rules with high support and high confidence. Reference rules are rules with low support and low confidence. Exception rules are defined as rules with low support and high confidence.

The formula for the relative interestingness measure RI in Hussain et al. (2000) is derived based on information theory and statistics. The measure has two components, which are interestingness based on the rule's support and interestingness based on the rule's confidence.

Suzuki (2002a) introduces undirected discovery of exception rules, in which a pattern represents a pair of an exception rule and its corresponding strong rule. Proposed scheduled discovery and exception rule discovery guided by a meta-pattern are described and tested on data sets.

Suzuki (2002b) presents an algorithm for discovering exception rules from a data set without domain-specific information. The method is based on sound pruning and probabilistic estimation. The normal approximations of the multinomial distributions are employed as the method for evaluating reliability of a rule pair. The method has been validated using two medical data sets and two benchmark data sets in the machine learning community.

The main contribution of Yamada and Suzuki (2002) is the formalization of spiral discovery for interesting exception rules and a method that employs initial knowledge, MDL-based discretization and reduction of the number of discovered rule pairs. The experimental evaluation was performed on meningitis data set.

## EXCEPTION RULES MINING

A new approach to mine exception rules will be proposed in this section. The approach belongs to the category of directed search. An interconnection between exception rules and strong association rules will be considered. As opposed to the research work described in the previous section, both strong positive and negative association rules are considered.

Based on the knowledge about positive and negative association rules in the database, the candidate exception rules will be generated. A novel exceptionality measure will be proposed to evaluate the candidate exception rules. The candidate exceptions with high exceptionality will form the final set of exception rules.

In order to formulate the proposed approach, a few data mining terms have to be defined. Itemset is a set of database items. For example, itemset XY means a set of two items X and Y. Association rule is an implication of the form X=>Y, where X and Y are database itemsets. An example of an association rule could be supermarket items Chips=>Coke purchased together frequently.

The rule X =>Y has support s, if s% of all transactions contain both X and Y. The rule X =>Y has confidence c, if c% of transactions that contain X, also contain Y. In association rules mining user-specified minimum support (minsup) and minimum confidence (minconf) are given.

Association rules with support greater or equal to minsup and confidence greater or equal to minconf are referred to as strong rules (Agrawal, Imielinski, & Swami, 1993; Agrawal & Srikant, 1994).

Itemsets that have support at least equal to minsup are called frequent itemsets. Negative itemsets are itemsets that contain both items and their negations. For example, consider the negative itemset X~Y. In this itemset ~Y means negation of item Y (absence of item Y in the database record).

Negative association rule is an implication of the form X=>~Y, ~X=>Y, ~X=>~Y, where X and Y are database items, ~X, ~Y are negations of database items. An example of a negative association rule is Coke=>~Pepsi, which means that people do not buy Coke and Pepsi together.

In our approach, the search for exception rules will be based on the knowledge about strong association rules in the database. An example: we discover a strong association rule in the database, for instance, shares of companies X and Y most times go up together X=>Y. Then those cases when shares of the companies X and Y do not go up together, X=>~Y or ~X=>Y, we call *exceptions* when satisfying the proposed *exceptionality* measure explained next. An algorithm for mining exception rules based on the knowledge about association rules will be proposed in as well.

We explain a few proposed definitions first. For exception rules mining instead of minsup we employ *lower* and *upper bounds*, satisfying the conditions: 0<lower bound<upper bound<minsup;

*Low* support belongs to the range [lower bound; upper bound]. *Infrequent itemsets* have low support. Note that the lower bound is always greater than 0, as we are not interested in rules with 0 support or close to 0. Upper bound is lower than minsup. The lower and upper bounds are chosen specifically for each data mining application.

*Exception Rules* are rules with *low* support and high *exceptionality* values. Infrequent itemsets with high exceptionality are called *exceptional* itemsets.

In the proposed exception rules mining the confidence measure is not applicable to evaluate the exception rules.