Creating a Decision Making Model Using Association Rules

Višnja Istrat
Faculty of organizational sciences, University of Belgrade, Serbia
Visnja.istrat@gmail.com

Abstract: Being a highly significant and complex function of management, decision making requires methods and techniques which simplify the process of selecting one choice among all available options. Decision making is therefore selection of that particular choice over any of several alternatives. Due to the process complexity, a continuous research and improvement of the methods and techniques modern decision making involves is required. One of many modern business challenges is to discover any possible improvement in the decision making process managers shall use in making the right decision. Any decision made by managers directly impacts the realized profit, business and company’s position on the market. The contribution of the paper is showing the importance of association rules in modern decision making.

Key words: business intelligence, association rules, management, decision making.

1. INTRODUCTION

Three dimensions determining complete development of this discipline need to be highlighted: qualitative, quantitative and information-communication aspect. These three aspects of decision making completely satisfy all the concepts of modern decision making development, both theoretically and practically.

The fact is that mankind faces the decision making problem in each phase of its social development, which has resulted in increased need for learning more about it. In this work both the significance and application of association rules will be analyzed on an example of car sales business. The research was conducted on a sample of 1728 transactions in order to recognize and establish the association rules and then determine their impact on the sales and profit. For the purpose of this research, a large car sales database was used as a source of information, which is also described in this work. Once these association rules were established, they were then used to create a better and more complete market supply.

Quantitative approach in modern decision making defines the basic formalism of general decision making problem [1]. According to [2], the decision making problem is a five item problem \((A, X, F, \Theta, \succ)\) where:

- \(A\): represents a definite set of available alternatives (actions) ranked by a session participant in order to select the most acceptable one;
- \(X\): represents the set of possible outcomes as a consequence of selecting an alternative;
- \(\Theta\): represents a set of world states, and depends on the unknown state \(\theta \in \Theta\) because the consequences of selecting alternative \(a \in A\) may differ;
- \(F: A \times \Theta \rightarrow X\), for each world state \(\sigma\) and for each alternative \(a\), determines the resulting consequence \(x = F(a, \sigma)\);
- \(\succ\): weak order relation on \(X\), i.e. a binary relation that satisfies the two following criteria:
  - Completeness: either \(x \succ y\) or \(y \succ x\), \(\forall x, y \in X\);
• Transitivity: if \( x \succ y \) and \( y \succ z \) then \( x \succ z \). \( \forall x, y, z \in X \).

Relation \( \succ \) features the decision maker and is called a preference relation. Two other important relations can be derived from the preference relation. The first is the strict preference relation where \( x \succ y \) if and only if both \( x \succ y \) and not \( y \succ x \). The second is the indifference relation where \( x \sim y \) if and only if \( x \succ y \) and \( y \succ x \). The most often way of solving decision making problems is the transformation of weak order \( \succ \) on \( X \) into normal order \( \succeq \) over the field of real numbers by the means of utility functions.

As the perfect solution rarely exists, a decision making process will be deemed successful if it produces the most acceptable decision for given problem. As stated in [1], the moment of taking such a decision is unquestionably both the most creative and most critical moment in the complete process of decision making.

II THEORETICAL FRAMEWORK

According to [1], Agrawal was first to establish association rules back in 1993 in order to perform the analysis of market basket. Cornerstone literature for fundamental basis of association rules use is [12]. Association rules are data mining technique where the goal is to find the data structure, regardless of the variables value. The task of the association rules is to determine which items go together. Typical example is grouping the items that can be bought together in shopping in supermarket – market basket analysis.

Supermarket chains use association rules to plan the timetable of the items on the shelves or in the catalogue so that the items that can be bought together are most often together on the shelves. They are used to identify chances for cross-selling and designing of attractive packaging or grouping the products or services. It is simply approach on how to create the rules from large databases. If two items, i.e. computer and web camera are often bought together, we can produce two association rules:

- Buyers that buy computer can also buy web camera with probability of P1
- Buyers that buy web camera can also buy computer with probability of P2

The market basket method has predominantly been used in sales analyses since, although its significance has been proved in other fields such as the analysis of credit card sales, identification of insurance companies committing a fraud or analysis of telecommunication services. However, this method cannot be explicitly used in case of simultaneous events, but only in case of successive events, which may be very useful in marketing for instance.

Numerous examples show how wide the application of the association rules methods in business can be. The researchers have drawn some significant conclusions based on which the right business decision can be taken. The methods proved to be most effective when applied to the market basket analysis, where association rules were used to identify patterns in purchasing products related to the car sales business.

III METHODOLOGY

Data mining is a process of analysing large data sets in order to discover significant patterns and rules. As modern companies are constantly seeking for higher goals, which especially refers to their productivity, it becomes absolutely necessary to improve the functioning of their organization through better understanding of their customers’ needs. Association rules determine which items have been purchased together. According to [12], the term association rules was first introduced by Agrawal. The task is to identify a set of rules which co-exist in some data set. According to [13], knowledge is a theoretical or practical understanding of facts and information, while wisdom is the synthesis of knowledge and experience that deepens our understanding of connections between different entities and uncovers some hidden message in their
existence. If knowledge is considered tools, then wisdom can be considered a set of skills using knowledge as its tools. There is a case of association rules in the field of market basket [10], where these rules classify into groups the items purchased together in supermarkets.

IV RESULTS AND DISCUSSION

The data used for the purpose of this research are all car sales related. A total of 1728 transactions were processed. Attributes were assigned to shopping (very high, high, medium and low rate of transactions), car maintenance (very high, high, medium, low level), number of car doors (two, three, four or five), number of seats (two, three, four), trunk size (small, medium, large) and safety (low, medium, high). All the data were classified as: accurate, inaccurate, good and very good.

In this paper one application of business intelligence techniques on a real business problem has been described. Some modern software architecture has been used for this purpose, namely Orange data mining software, which is a very popular data mining tool among managers, who use it massively to support themselves in decision making. Orange is extremely handy and user-friendly software which contains numerous data mining options for data analysis, such as model integration, testing, data visualization, solution application, etc. One of the most popular and well-proven technologies for data mining is CRISP-DM methodology. In order to discover the association rules in a large database of car sales transactions, the data have been processed using Orange software. The association rules can further help customers select a car to purchase by providing them with relevant sales data based on the information gathered from previous sales.

Figure 1: Desktop appearance of Orange software

Figure 2: Data entry options of Orange

Figure 3: Model of association rules created in Orange's canvas interface
The first step was uploading the data file into the program. The Support and Confidence parameters were defined next. As the research progressed, these parameters were altered in order to determine their impact on the final result of the association rules searching process.

Different results were obtained when the minimal Support was reduced from 30% to 10 %, and the minimal Confidence was increased from 50% to 60%. The number of discovered association rules consequently increased. There were thirty new association rules with the Support ranging from 11% to 33% and Confidence ranging from 62% to 100%. Further analysis of the obtained results revealed the purchasing patterns to be used in forming the optimal set of products in order to increase the sales rate and satisfy the customers to the furthest extent possible. Highlighted is the association rule with the highest Support and Confidence, of 0.333 and 1 respectively. The analysis of each association rule showed precision in characteristics and patterns in the customers’ behaviour.

Car attributes (characteristics) used for the purpose of this research were: price, maintenance costs, number of doors, number of seats, trunk size and safety. The obtained association rules suggested that those customers interested in purchasing a low safety two-seated car would most likely purchase a car classified
as inaccurate. Marketing experts and experts from the sales sector are recommended to work to attract potential customers by evaluating their needs and offering them a suitable car with the highest probability of purchase. Furthermore, the obtained association rules also show that when a car is purchased, the car radio is commonly purchased as well being part of the car’s additional equipment. Therefore any marketing strategy should be developed with respect to this information on order to be successful.

V CONCLUSION
By analyzing the current research work of various domestic and foreign experts in the field of business intelligence, and despite the fact this field is relatively new in some segments of application, a great popularity and potential it carries have been proved. The obtained results have been analyzed using modern scientific methods and several recommendations for the future research course have been made.

For the purpose of this research, a large car sales database has been used for extracting data and discovering association rules. The obtained association rules have been used to create better and more comprehensive market offer. The process of improving the car sales model has been presented having in mind that managers learn about their customers’ purchase habits through association rules. The significance of multidisciplinary approach has been explained – the dependence between decision making, business intelligence, human resource management, knowledge management and modern ICT has been proved to be effective in the case of business system management.

VI REFERENCES


