Dynamic Programming Approach for Construction of Association Rule Systems

Fawaz Alsolami[1], Talha Amin[1], Igor Chikalov[1], Mikhail Moshkov[1], and Beata Zielosko[2]

[1] Computer, Electrical and Mathematical Sciences and Engineering Division
King Abdullah University of Science and Technology
Thuwal 23955-6900, Saudi Arabia
{fawaz.alsolami, talha.amin, mikhail.moshkov}@kaust.edu.sa
igor.chikalov@gmail.com
[2] Institute of Computer Science, University of Silesia
39, Bedzinska St., 41-200 Sosnowiec, Poland
beata.zielosko@us.edu.pl

Abstract. In the paper, an application of dynamic programming approach for optimization of association rules from the point of view of knowledge representation is considered. Experimental results present cardinality of the set of association rules constructed for information system and lower bound on minimum possible cardinality of rule set based on the information obtained during algorithm work.

Key words: association rules, decision rules, dynamic programming, set cover problem, rough sets.

1 Introduction

Association rules are popular form of knowledge representation. They are used in various areas such as business field for decision making and effective marketing, sequence-pattern in bioinformatics, medical diagnosis, etc. One of the most popular application of association rules is market basket analysis that finds associations between different items that customers place in their shopping baskets.

There are many approaches for mining association rules. The most popular, is Apriori algorithm based on frequent itemsets [1]. During years, many new algorithms were designed which are based on, e.g., transaction reduction [2], sampling the data [13], and others [7, 9].

The most popular measures for association rules are support and confidence, however in the literature many other measures have been proposed [8, 9]. In this paper, we are interested in the construction of rules which cover many objects. Maximization of the coverage allows us to discover major patterns in the data, and it is important from the point of view of knowledge representation. Unfortunately, the problem of construction of rules with maximum coverage is NP-hard [6]. The most part of approaches, with the exception of brute-force and Apriori algorithm, cannot guarantee the construction of rules with the maximum coverage. The proposed dynamic programming approach allows one to construct such rules.

Application of rough sets theory to the construction of rules for knowledge representation or classification tasks are usually connected with the usage of decision table [12] as a form of input data representation. In such a table, one attribute is distinguished as a decision attribute and it relates to a rule consequence. However, in the last years, associative mechanism of rule construction, where all attributes can occur as premises or consequences of particular rules, is popular. Association rules can be defined in many ways. In the paper, a special kind of association rules is studied, i.e., they relate to decision rules. Similar approach was considered in [10, 11], where a greedy algorithm for minimization of length of association rules was studied. In [15], a dynamic programming approach to optimization of association rules relative to coverage was investigated.

When association rules for information systems are studied and each attribute is sequentially considered as the decision one, inconsistent tables are often obtained, i.e., tables containing equal rows with different decisions. In the paper, two possibilities of removing inconsistency of decision tables are considered. If in some tables there are equal rows with, possibly, different decisions, then (i) each group of identical rows is replaced with a single row from the group with the most common decision for this group, (ii) each group of identical rows is replaced with a single row from the group with the set of decisions attached to rows from the considered group. In the first case, usual decision tables are obtained (decision tables with single-valued decisions) and, for a given row, we should find decision attached to this row. In the second case, decision tables with many-valued decisions are obtained and, for a given row, we should find an arbitrary decision from the set of decisions attached to this row.

For each decision table obtained from the information system, we construct a system of exact rules in the following way: during each step, we choose a rule which covers the maximum number of previously uncovered rows. We stop the construction when all rows of the table are covered. If the obtained system of rules is short enough, then it can be considered as an intelligible representation of the knowledge extracted from the decision table. Otherwise, we can consider approximate rules, and stop the construction of the rule system when the most part of the rows (for example 90% of the rows) are covered.

In [4], the presented algorithm was proposed as application for multi-stage optimization of decision rules for decision tables. We extend it to association rules. The presented algorithm can be considered as a simulation of a greedy algorithm for construction of partial covers. So we can use lower bound on the minimum cardinality for partial cover based on the information about greedy algorithm work which was obtained in [10].

The paper consists of five sections. Section 2 contains main notions. In Sect. 3, algorithm for construction of system of association rule systems is presented. Section 4 contains experimental results for decision tables from UCI Machine Learning Repository, and Section 5 – short conclusions.

2 Main Notions

An *information system* I is a rectangular table with $n+1$ columns labeled with attributes $f_1, \ldots, f_{n+1}$. Rows of this table are filled by nonnegative integers which are interpreted as values of attributes.

An association rule for I is a rule of the kind

$$(f_{i_1} = a_1) \wedge \ldots \wedge (f_{i_m} = a_m) \rightarrow f_j = a,$$

where $f_j \in \{f_1, \ldots, f_{n+1}\}$, $f_{i_1}, \ldots, f_{i_m} \in \{f_1, \ldots, f_{n+1}\} \setminus \{f_j\}$, and $a,a_1,\ldots,a_m$ are nonnegative integers.

The notion of an association rule for I is based on the notions of a decision table and decision rule. We consider two kinds of decision tables: with many-valued decisions and with single-valued decisions.

A *decision table with many-valued decisions* T is a rectangular table with n columns labeled with (conditional) attributes $f_1, \ldots, f_n$. Rows of this table are pairwise different and are filled by nonnegative integers which are interpreted as values of conditional attributes. Each row r is labeled with a finite nonempty set $D(r)$ of nonnegative integers which are interpreted as decisions (values of a decision attribute). For a given row r of T, it is necessary to find a decision from the set $D(r)$.

A *decision table with single-valued decisions* T is a rectangular table with n columns labeled with (conditional) attributes $f_1, \ldots, f_n$. Rows of this table are pairwise different and are filled by nonnegative integers which are interpreted as values of conditional attributes. Each row r is labeled with a nonnegative integer $d(r)$ which is interpreted as a decision (value of a decision attribute). For a given row r of T, it is necessary to find the decision $d(r)$. Decision tables with single-valued decisions can be considered as a special kind of decision tables with many-valued decisions in which $D(r) = \{d(r)\}$ for each row r.

For each attribute $f_i \in \{f_1, \ldots, f_{n+1}\}$, the information system I is transformed into a table I_{f_i}. The column f_i is removed from I and a table with n columns labeled with attributes $f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_{n+1}$ is obtained. Values of the attribute f_i are attached to the rows of the obtained table I_{f_i} as decisions.

The table I_{f_i} can contain equal rows. We transform this table into two decision tables – with many-valued and single-valued decisions. A decision table $I_{f_i}^{m-v}$ with many-valued decisions is obtained from the table I_{f_i} by replacing each group of equal rows with a single row from the group with the set of decisions attached to all rows from the group. A decision table $I_{f_i}^{s-v}$ with single-valued decisions is obtained from the table I_{f_i} by replacing each group of equal rows with a single row from the group with the most common decision for this group.

The set $\{I_{f_1}^{m-v}, \ldots, I_{f_{n+1}}^{m-v}\}$ of decision tables with many-valued decisions obtained from the information system I is denoted by $\Phi^{m-v}(I)$. We denote by $\Phi^{s-v}(I)$ the set $\{I_{f_1}^{s-v}, \ldots, I_{f_{n+1}}^{s-v}\}$ of decision tables with single-valued decisions obtained from the information system I. Since decision tables with single-valued decisions are a special case of decision tables with many-valued decisions, we consider the notion of decision rule for tables with many-valued decisions.

Let $T \in \Phi^{m-v}(I)$. For simplicity, let $T = I^{m-v}_{f_{n+1}}$. The attribute f_{n+1} will be considered as a decision attribute of the table T. We denote by $Row(T)$ the set of rows of T. Let $D(T) = \bigcup_{r \in Row(T)} D(r)$.

A decision table is called *empty* if it has no rows. The table T is called *degenerate* if it is empty or has a *common* decision, i.e., $\bigcap_{r \in Row(T)} D(r) \neq \emptyset$. We denote by $N(T)$ the number of rows in the table T and, for any $t \in \omega$, we denote by $N_t(T)$ the number of rows r of T such that $t \in D(r)$. By $mcd(T)$ we denote the *most common decision* for T which is the minimum decision t_0 from $D(T)$ such that $N_{t_0}(T) = \max\{N_t(T) : t \in D(T)\}$. If T is empty then $mcd(T) = 0$.

For any conditional attribute $f_i \in \{f_1, \ldots, f_n\}$, we denote by $E(T, f_i)$ the set of values of the attribute f_i in the table T. We denote by $E(T)$ the set of conditional attributes for which $|E(T, f_i)| \geq 2$.

Let T be a nonempty decision table. A *subtable* of T is a table obtained from T by the removal of some rows. Let $f_{i_1}, \ldots, f_{i_m} \in \{f_1, \ldots, f_n\}$ and $a_1, \ldots, a_m \in \omega$ where ω is the set of nonnegative integers. We denote by $T(f_{i_1}, a_1) \ldots (f_{i_m}, a_m)$ the subtable of the table T containing the rows from T which at the intersection with the columns $f_{i_1}, \ldots, f_{i_m}$ have numbers $a_1, \ldots, a_m$, respectively.

As an *uncertainty measure* for nonempty decision tables we consider *relative misclassification error* $rme(T) = (N(T) - N_{mcd(T)}(T))/N(T)$ where $N_{mcd(T)}(T)$ is the number of rows r in T containing the most common decision for T in $D(r)$.

A *decision rule over* T is an expression of the kind

$$(f_{i_1} = a_1) \wedge \ldots \wedge (f_{i_m} = a_m) \rightarrow f_{n+1} = t \quad (1)$$

where $f_{i_1}, \ldots, f_{i_m} \in \{f_1, \ldots, f_n\}$, and $a_1, \ldots, a_m, t$ are numbers from ω. It is possible that $m = 0$. For the considered rule, we denote $T^0 = T$, and if $m > 0$ we denote $T^j = T(f_{i_1}, a_1) \ldots (f_{i_j}, a_j)$ for $j = 1, \ldots, m$. We will say that the decision rule (1) *covers* the row $r = (b_1, \ldots, b_n)$ of T if r belongs to T^m, i.e., $b_{i_1} = a_1, \ldots, b_{i_m} = a_m$.

A decision rule (1) over T is called a *decision rule for* T if $t = mcd(T^m)$, and if $m > 0$, then T^{j-1} is not degenerate for $j = 1, \ldots, m$, and $f_{i_j} \in E(T^{j-1})$. We denote by $DR(T)$ the set of decision rules for T.

Let ρ be a decision rule for T which is equal to (1). The value $rme(T, \rho) = rme(T^m)$ is called the *uncertainty* of the rule ρ. Let α be a real number such that $0 \leq \alpha \leq 1$. We will say that a decision rule ρ for T is an *α-decision rule for* T if $rme(T, \rho) \leq \alpha$. If $\alpha = 0$ (in this case, for each row r covered by ρ, the set $D(r)$ contains the decision on the right-hand side of ρ) then we will say that ρ is an *exact* rule. We denote by $DR_\alpha(T)$ the set of α-decision rules for T.

3 Algorithm for Construction of Association Rule System

α-Decision rules for tables from $\Phi^{m-v}(I)$ can be considered as α-association rules (modification for many-valued decision model) for the information system I. α-Decision rules for decision tables from $\Phi^{s-v}(I)$ can be considered as α-association rules (modification for single-valued decision model) for the information system I. In this section, we consider an algorithm for construction of an association rule system for

I in the frameworks of both many-valued decision model and single-valued decision model. Since decision tables with single-valued decisions are a special kind of decision tables with many-valued decisions, we will discuss mainly many-valued decision model.

Let $T = I^{m-v}_{f_{n+1}}$. Let S be a nonempty finite set of α-decision rules for T (*system* of α-decision rules for T), and β be a real number such that $0 \leq \beta \leq 1$. We say that S is a *β-system of α-decision rules for T* if rules from S cover at least $(1-\beta)N(T)$ rows of T.

We describe an algorithm α-β-Rules which, for a decision table T, and real numbers α and β, $0 \leq \alpha \leq 1$, and $0 \leq \beta \leq 1$, constructs a β-system of α-decision rules for T. During each step, we choose (based on a dynamic programming algorithm [4]) a decision rule which covers maximum number of uncovered previously rows. We stop when the constructed rules cover at least $(1-\beta)N(T)$ rows of T. We denote by $Rule_{\alpha,\beta}(T)$ the constructed system of rules.

We denote by $C(T, \alpha, \beta)$ the minimum cardinality of a β-system of α-decision rules for T. It is clear that $C(T, \alpha, \beta) \leq |Rule_{\alpha,\beta}(T)|$. Using information based on the work of algorithm α-β-Rules, we can obtain lower bound on the parameter $C(T, \alpha, \beta)$. During the construction of β-system of α-decision rules for T, let the algorithm α-β-Rules selects consequently rules $\rho_1, \ldots, \rho_t$. Let $B_1, \ldots, B_t$ be sets of rows of T covered by rules $\rho_1, \ldots, \rho_t$, respectively. Set $B_0 = \emptyset$, $\delta_0 = 0$ and, for $i = 1, \ldots, t$, set $\delta_i = |B_i \setminus (B_0 \cup \ldots \cup B_{i-1})|$. The information derived from the algorithm's work consists of the tuple $(\delta_1, \ldots, \delta_t)$ and the numbers $N(T)$ and β.

From the results obtained in [10] regarding a greedy algorithm for the set cover problem it follows that $C(T, \alpha, \beta) \geq l(T, \alpha, \beta)$ where

$$l(T, \alpha, \beta)) = \max \left\{ \left\lceil \frac{\lceil (1-\beta)N(T) \rceil - (\delta_0 + \ldots + \delta_i)}{\delta_{i+1}} \right\rceil : i = 0, \ldots, t-1 \right\}.$$

Using algorithm α-β-Rules, for each decision table $T \in \Phi^{m-v}(I)$, we construct the set of rules $Rule_{\alpha,\beta}(T)$. As a result, we obtain the system of rules (α-association rules for the information system I – modification for many-valued decision model) $Rule^{m-v}_{\alpha,\beta}(I) = \bigcup_{T \in \Phi^{m-v}(I)} Rule_{\alpha,\beta}(T)$. This system contains, for each $T \in \Phi^{m-v}(I)$, a subsystem which is a β-system of α-decision rules for T. We denote by $C^{m-v}(I, \alpha, \beta)$ the minimum cardinality of such system. One can show that

$$L^{m-v}(I, \alpha, \beta) \leq C^{m-v}(I, \alpha, \beta) \leq U^{m-v}(I, \alpha, \beta),$$

$L^{m-v}(I, \alpha, \beta) = \sum_{T \in \Phi^{m-v}(I)} l(T, \alpha, \beta)$ and $U^{m-v}(I, \alpha, \beta) = \left| Rule^{m-v}_{\alpha,\beta}(I) \right|$.

We can do the same for the set $\Phi^{s-v}(I)$ of decision tables with single-valued decisions. As a result, we obtain the system of rules (α-association rules for the information system I – modification for single-valued decision model) $Rule^{s-v}_{\alpha,\beta}(I) = \bigcup_{T \in \Phi^{s-v}(I)}$ $Rule_{\alpha,\beta}(T)$ which contains, for each $T \in \Phi^{s-v}(I)$, a subsystem which is a β-system of α-decision rules for T. Denote $C^{s-v}(I, \alpha, \beta)$ the minimum cardinality of such system. One can show that

$$L^{s-v}(I, \alpha, \beta) \leq C^{s-v}(I, \alpha, \beta) \leq U^{s-v}(I, \alpha, \beta),$$

Table 1. Total number of rules (upper bound / lower bound)

Information system	Rows	Attr	$\alpha = 0, \beta = 0$ single-val	$\alpha = 0, \beta = 0$ many-val	$\alpha = 0.3, \beta = 0.2$ single-val	$\alpha = 0.3, \beta = 0.2$ many-val
adult-stretch	16	5	11 / 10	9 / 8	11 / 10	9 / 8
balance-scale	625	5	**847** / **709**	**507** / **373**	**390** / **364**	50 / 28
breast-cancer	266	10	**1027** / **570**	**984** / **522**	**401** / 208	**362** / 180
cars	1728	7	**446** / **237**	**358** / 203	62 / 41	51 / 33
hayes-roth-data	69	5	**169** / 117	**161** / 105	115 / 73	105 / 68
lenses	24	5	29 / 18	25 / 16	17 / 12	15 / 12
monks-1-test	432	7	50 / 36	40 / 36	35 / 27	31 / 27
monks-3-test	432	7	36 / 30	22 / 19	25 / 21	16 / 14
shuttle-landing	15	7	47 / 35	42 / 33	30 / 21	27 / 20
teeth	23	9	114 / 86	114 / 86	65 / 51	65 / 51
tic-tac-toe	958	10	**5186** / **4112**	**5186** / **4112**	**1508** / **737**	**1508** / **737**
zoo-data	59	17	159 / 84	143 / 70	51 / 37	49 / 37

$L^{s-v}(I, \alpha, \beta) = \sum_{T \in \Phi^{s-v}(I)} l(T, \alpha, \beta)$ and $U^{s-v}(I, \alpha, \beta) = \left| Rule^{s-v}_{\alpha,\beta}(I) \right|$.

4 Experimental Results

Experiments were made using data sets from UCI Machine Learning Repository [5] and software system Dagger [3]. Some decision tables contain conditional attributes that take unique value for each row. Such attributes were removed. In some tables, there were equal rows with, possibly, different decisions. In this case each group of identical rows was replaced with a single row from the group with the most common decision for this group. In some tables there were missing values. Each such value was replaced with the most common value of the corresponding attribute.

Prepared 12 data sets were considered as information systems. Table 1 contains name (column "Information system"), number of rows (column "Rows"), and number of attributes (column "Attr") for each of the considered information systems. Table 1 presents also upper / lower bounds (see descriptions at the end of the previous section) on $C^{m-v}(I, \alpha, \beta)$ (column "many-val") and on $C^{s-v}(I, \alpha, \beta)$ (column "single-val") for pairs ($\alpha = 0$, $\beta = 0$) and ($\alpha = 0.3$, $\beta = 0.2$).

We can see that, for tables with many-valued decisions, upper and lower bounds on the number of rules are less than or equal to the bounds for decision tables with single-valued decision. We considered a threshold

$$30 \times (number\ of\ attributes)$$

Table 2. Total number of rules for information system balance-scale with 5 attributes

Single-valued decisions

β \ α	0	0.1	0.3
0	847 / 709	773 / 667	670 / 606
0.01	833 / 695	759 / 653	656 / 592
0.05	772 / 634	698 / 592	595 / 531
0.1	693 / 555	619 / 516	516 / 463
0.15	614 / 499	540 / 460	447 / 414
0.2	534 / 444	460 / 404	390 / 364

Many-valued decisions

β \ α	0	0.1	0.3
0	507 / 373	433 / 331	330 / 270
0.01	493 / 359	419 / 317	316 / 256
0.05	432 / 298	358 / 256	255 / 195
0.1	353 / 219	279 / 180	176 / 127
0.15	274 / 163	200 / 124	107 / 78
0.2	194 / 108	120 / 68	50 / 28

as a reasonable upper bound on the number of rules if a system of rules is used for knowledge representation. In the case $\alpha = 0$ and $\beta = 0$, the threshold is exceeded for five information systems (see numbers in bold): balance-scale, breast-cancer, cars, hayes-roth-data, and tic-tac-toe. The consideration of approximate rules and partial covers can improve the situation. In the case $\alpha = 0.3$ and $\beta = 0.2$, the threshold is exceeded for three information systems (balance-scale, breast-cancer, and tic-tac-toe) if we consider decision tables with single-valued decisions and for two information systems (breast-cancer and tic-tac-toe) if we consider decision tables with many-valued decisons.

For four information systems (balance-scale, breast-cancer, cars, and hayes-roth-data), upper / lower bounds on $C^{m-v}(I, \alpha, \beta))$ and on $C^{s-v}(I, \alpha, \beta))$ for $\beta \in \{0, 0.01, 0.05, 0.1, 0.15, 0.2\}$ and $\alpha \in \{0, 0.1, 0.3\}$ can be found in Tables 2, 3, 4, and 5.

5 Conclusions

In the paper, an algorithm for construction of association rule system is proposed. It simulates the work of greedy algorithm for set cover problem. Experimental results present cardinality of the set of association rules constructed for information system and lower bound on minimum possible cardinality of such set based on the information about the algorithm work. In the future, the length of constructed association rules will be studied also. We are planning to extend an approach proposed in [14] for decision rules to construction of association rule systems. This approach allows one to construct rules with coverage close to maximum and requires less time than the dynamic programming approach.

Table 3. Total number of rules for information system breast-cancer with 10 attributes

Single-valued decisions

$\beta \backslash \alpha$	0	0.1	0.3
0	1027 / 570	995 / 559	786 / 439
0.01	1007 / 556	975 / 546	766 / 424
0.05	905 / 496	873 / 485	665 / 364
0.1	784 / 426	753 / 416	562 / 296
0.15	684 / 369	660 / 361	478 / 246
0.2	596 / 321	574 / 314	401 / 208

Many-valued decisions

$\beta \backslash \alpha$	0	0.1	0.3
0	984 / 522	952 / 512	746 / 400
0.01	964 / 506	932 / 498	726 / 385
0.05	862 / 446	830 / 438	625 / 326
0.1	741 / 383	710 / 376	522 / 261
0.15	645 / 335	622 / 327	439 / 212
0.2	562 / 295	540 / 288	362 / 180

Table 4. Total number of rules for information system cars with 7 attributes

Single-valued decisions

$\beta \backslash \alpha$	0	0.1	0.3
0	446 / 237	440 / 235	391 / 204
0.01	398 / 213	396 / 212	349 / 181
0.05	264 / 136	263 / 135	221 / 107
0.1	152 / 102	152 / 102	133 / 74
0.15	111 / 72	111 / 72	92 / 56
0.2	74 / 42	74 / 42	62 / 41

Many-valued decisions

$\beta \backslash \alpha$	0	0.1	0.3
0	358 / 203	350 / 203	304 / 170
0.01	320 / 183	317 / 183	272 / 150
0.05	218 / 112	218 / 112	174 / 82
0.1	124 / 85	124 / 85	104 / 57
0.15	93 / 60	93 / 60	73 / 45
0.2	64 / 34	64 / 34	51 / 33

Table 5. Total number of rules for information system hayes-roth-data with 5 attributes

Single-valued decisions

β \ α	0	0.1	0.3
0	169 / 117	169 / 117	169 / 117
0.01	169 / 117	169 / 117	169 / 117
0.05	158 / 108	158 / 108	158 / 108
0.1	143 / 95	143 / 95	143 / 95
0.15	128 / 83	128 / 83	128 / 83
0.2	115 / 73	115 / 73	115 / 73

Many-valued decisions

β \ α	0	0.1	0.3
0	161 / 105	161 / 105	159 / 104
0.01	161 / 105	161 / 105	159 / 104
0.05	150 / 96	150 / 96	148 / 95
0.1	135 / 85	135 / 85	133 / 84
0.15	120 / 75	120 / 75	118 / 75
0.2	107 / 69	107 / 69	105 / 68

Acknowledgements

Research reported in this publication was supported by the King Abdullah University of Science and Technology (KAUST).

The authors wish to express their gratitude to anonymous reviewers for useful comments.

References

1. Agrawal, R., Imieliński, T., Swami, A.: Mining association rules between sets of items in large databases. In: SIGMOD '93, pp. 207–216. ACM (1993)
2. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases. In: Bocca, J.B., Jarke, M., Zaniolo, C. (eds.) VLDB, pp. 487–499. Morgan Kaufmann (1994)
3. Alkhalid, A., Amin, T., Chikalov, I., Hussain, S., Moshkov, M., Zielosko, B.: Dagger: A tool for analysis and optimization of decision trees and rules. In: Computational Informatics, Social Factors and New Information Technologies: Hypermedia Perspectives and Avant-Garde Experiences in the Era of Communicability Expansion, pp. 29–39. Blue Herons (2011)
4. Alsolami, F., Amin, T., Chikalov, I., Moshkov, M.: Multi-stage optimization of decision rules for knowledge discovery and representation. Knowledge and Information Systems (2015), (submitted)
5. Asuncion, A., Newman, D.J.: UCI Machine Learning Repository (http://wwwicsuciedu/~mlearn/, 2007), http://www.ics.uci.edu/~mlearn/
6. Bonates, T., Hammer, P.L., Kogan, A.: Maximum patterns in datasets. Discrete Applied Mathematics 156(6), 846–861 (2008)
7. Borgelt, C.: Simple algorithms for frequent item set mining. In: Koronacki, J., Raś, Z.W., Wierzchoń, S.T., Kacprzyk, J. (eds.) Advances in Machine Learning II, Studies in Computational Intelligence, vol. 263, pp. 351–369. Springer Berlin Heidelberg (2010)
8. Glass, D.H.: Confirmation measures of association rule interestingness. Knowledge-Based Systems 44(0), 65–77 (2013)
9. Han, J., Kamber, M.: Data Mining: Concepts and Techniques. Morgan Kaufmann (2000)
10. Moshkov, M.J., Piliszczuk, M., Zielosko, B.: Greedy algorithm for construction of partial association rules. Fundam. Inform. 92(3), 259–277 (2009)

11. Moshkov, M.J., Piliszczuk, M., Zielosko, B.: On construction of partial association rules. In: Wen, P., Li, Y., Polkowski, L., Yao, Y., Tsumoto, S., Wang, G. (eds.) RSKT, LNCS, vol. 5589, pp. 176–183. Springer (2009)
12. Pawlak, Z., Skowron, A.: Rudiments of rough sets. Inf. Sci. 177(1), 3–27 (2007)
13. Toivonen, H.: Sampling large databases for association rules. In: Vijayaraman, T.M., Buchmann, A.P., Mohan, C., Sarda, N.L. (eds.) VLDB, pp. 134–145. Morgan Kaufmann (1996)
14. Zielosko, B.: Optimization of approximate decision rules relative to coverage. In: Kozielski, S., Mrózek, D., Kasprowski, P., Małysiak-Mrózek, B., Kostrzewa, D. (eds.) BDAS 2014. LNCS, vol. 8537, pp. 170–179. Springer (2014)
15. Zielosko, B.: Global optimization of exact association rules relative to coverage. In: Kryszkiewicz, M., Bandyopadhyay, S., Rybiński, H., Pal, S.K. (eds.) PReMI 2015. LNCS, vol. 9124, pp. 428–437. Springer (2015)