APPLICATION OF INTELLECTUAL TECHNOLOGIES IN PROFIT CONTROL

Inna Khmelnitskaya, Oleg Kosenkov

Abstract: The modern level of IT-progress achievements predetermines transformation of management methods in the enterprise. Profit, being one of the main signs of the efficiency of the company's activity, accumulates its effectiveness. The authors (of this article) suggest «a matrix of profit control» in which approaches of economic analysis ABC and XYZ are used together with one of the most effective instruments in decision making on the basis of information technologies – Data Mining – in order to solve the problem production conjugacy with the use of association task. Having divided the whole production into profitability groups, and then, according to the degree of risk for the company's financial condition, into matrix points, it is possible to reveal the existing problems and their causes in terms of both types of products and places of their origin. In the present article we suggest a solution of this task by means of intellectual analysis, which will allow the company to achieve considerable competitive advantages.

Keywords: ABC and XYZ analysis, matrix of profit control, association, algorithms of associative rules searching, rule of anti-monotony, tree of solutions.

Introduction

Profit is the driving force of the market and it determines three crucial interrelated aspects of the company's activity: what, how and for whom to produce. The reward for entrepreneurial activity, evinced as the profit, will depend on how correctly the company managers will answer the above questions. The world has numerous tools of management of various economic objects, such as, for example, ABC and XYZ analysis. The ABC-analysis allows to determine the contribution of certain goods, a group pf goods or clients into the company sales and profit, and to classify them according to the importance degree and the necessity of control, depending on the volume and the sum of inflows. XYZ (frequency) analysis is intended for the evaluation of the company sales stability, for the detection of those clients who are not included into the number of clients making high finances or those not very expensive goods, the purchases of which are stably not too big in quantity but still are very frequent. Using the results of the XYZ-analysis the company will be able to differentiate the goods nomenclature depending in the stability of the profit got and the forecast sales. In our opinion, the optimal and a frequently used means of decision making is the ABC-XYZ-analysis, which was used for the creation of the "matrix of profit control", represented in table 1 [1, 2].

Risk	Profit		
	A (≥ 75%)	B (≥ 20%)	C (≥ 5%)
X (up to 33%)	AX	BX	CX
Y (from 33% to 66%)	AY	BY	CY
Z (over 66%)	AZ	BZ	CZ

Table 1 - Matrix of profit control

The following groups A, B and C are shown as the columns:

Group A – production bringing about 75% of the total profit to the company,

Group B – production bringing about 20% of the total profit to the company,

Group C – production bringing about 5% of the total profit to the company,

Groups of risk X, Y and Z: are represented as rows.

X – group of production with minimal risk, i.e. the volume of sales dynamics is positive, the final economic indexes are good.

Y – group of production with middle risk, i.e. the volume of sales dynamics is not stable or is absent; the final economic indexes are satisfactory.

Z – group of production with maximum risk, i.e. the dynamics is negative, the indexes are marginal.

In this case the risk is calculated from the value of the profit variation index. In accordance with the integrated ABC-XYZ-analysis and the applied bases of production rules knowledge, shown in table 2, the non- structured data are classified into nine groups with different characteristics, which to a big extent simplifies work with them [2].

Table 2 – Original bases of knowledge of production rules of distributing of the production into the cells of the profit control matrix.

Rules	Formulation
1	If the profit is ≥ 75%, then if the risk is ≤ 33%, then «AX» Otherwise if the risk is $\in$ (33%; 66%], then «AY» , otherwise if the risk is > 66%, then «AZ»
2	If the profit is ≥ 20%, then if the risk is ≤ 33%, then «BX» Otherwise if the risk is $\in$ (33%; 66%], then «BY» , otherwise if the risk is > 66%, then «BZ»
3	If the profit is ≥ 5%, then if the risk is ≤ 33%, then «CX» Otherwise if the risk is $\in$ (33%; 66%], then «CY» , otherwise if the risk is > 66%, then «CZ»

Foe example, the goods bringing the company the maximum profit and most stable in the volume of sales will go to group AX, and the low-profit goods with unstable sales will go to group CZ. There is also a possibility to sort and filter any table column, which gives significant convenience in work with the table to the analyst. That means, pressing “Enter” on any cell will result in showing the contents of the corresponding group. The goods here are also ordered according to their rating (ABC) and variation index (XYZ), and the matrix in detail, for example in cell AX, shown in figure 1, will look as follows:

	A 1	A 2	A 3
X1	Production of group A1X1 Profit in group A1X1	Production of group A2X1 Profit in group A2X1	Production of group A3X1 Profit in group A3X1
X2	Production of group A1X2 Profit in group A2X1	Production of group A2X2 Profit in group A2X2	Production of group A3X2 Profit in group A3X2
X3	Production of group A1X3 Profit in group A1X3	Production of group A2X2 Profit in group A2X3	Production of group A3X3 Profit in group A3X3

Figure 1 – Matrix of profit control in detail.

From the results of the analyses represented in the matrix, one can determine and regularly review the rules and norms of promotional work and finding out the narrow points of certain goods, work with clients, market segments and other nomenclature. This is an irreplaceable tool to improve the efficiency of the system of goods circulation, increase the profit and, consequently, to improve the financial state of the company.

Nevertheless, creation of the matrix will not either divide the goods into "good" and "bad", or reveal the goods, which should be immediately withdrawn from the list of sales. An additional analysis is always necessary. For example, there is a problem of conjugation of different types of the production sold, i.e. in category BY one can often find associated goods, just like ski- boots for the skies sold. They bring lower profit and are purchased not so regularly by the clients. In category CZ one can often find ski bindings, which bring low profit and are nor purchased regularly. But these goods must be available, otherwise the client, having not found them may not come again to this shop.

The authors solved this problem by means of the task of associations. The application of this task was also caused by a number of objective factors. Modern databases have huge volumes reaching giga- and terabytes, and still there is a tendency for them to increase constantly. This fact causes a stable growth of interest to the methods of finding out certain knowledge in databases. The volumes of modern databases or information depositories being very impressive created stable demand for the scaled algorithms of the data analyses. One of the popular methods of information finding is algorithms of associative rules searching [3, 4].

The first algorithm of associative rules searching called AIS was worked out in 1993 by the staff of IBM Almaden Research Center [5]. Starting from this pioneer work, the interest to association rules increased; in the middle of the 90-s of the previous century there was the peak of the researches in this field, and since then a few algorithms appear every year. Algorithm revelation of association rules appeared as one of the variants to find out typical purchasing patterns, made in supermarkets. Our task is very correlative (practically the same) with that one, to solve which the principal of revelation of associative dependencies in big databases was worked out. The rules, generated by this algorithm, can be presented graphically and in general can be shown as follows:

- If the customer bought goods A_1, goods A_2 … and goods A_n, he is very likely to buy goods B_1, goods B_2 … and goods B_m.…
- If the customer bought goods A_1, goods A_2 … and goods A_n, he is very likely to buy goods C_1, goods C_2 … and goods C_k.…
- If the customer bought goods D_1, goods D_2 … and goods D_t, he is very likely to buy goods E_1, goods E_2 … and goods E_p.

Here (A_1, A_2, …, A_n), …, (D_1, D_2, …, D_t) – are those goods, that the customer has already bought in the current transaction, and (B_1, B_2, … B_m), …, (C_1, C_2, …, C_k) , …, (E_1, E_2, …, E_p) – are the goods, that the customer is inclined to buy during the same visit.

Let there is a database consisting of buyers' transactions. Each transaction is the totality of goods bought by the buyer during one visit. This kind of transaction is also called a market basket.

Let $I = \{i_1, i_2, i_3, \ldots i_n\}$ is the multitude (range) of goods called elements. Let D is the multitude of transactions where each transaction T is a set of elements from I: $T \subseteq I$. Each transaction is a binary vector, where t[k] = 1, if the element i_k is present in the transaction, otherwise t[k] = 0. We say that transaction T includes X, a certain set of elements from I, if $X \subseteq T$. Associative rule is the implication $X \Rightarrow Y$, where $X \subset I$, $Y \subset I$ and $X \cap Y = \varnothing$. The rule $X \Rightarrow Y$ has the support – s, if s % of transactions from D contain $X \cup Y$, $supp(X \Rightarrow Y) = supp(X \cup Y)$. The confidence of the rule shows to what extent it is possible that Y follows from X. The rule $X \Rightarrow Y$ is true with the confidence c, if c % of transaction from D containing X, contains Y too:

$$conf(X \Rightarrow Y) = supp(X \cup Y) / supp(X).$$

In other words, the analysis objective is ascertainment of the following dependences: if in the transaction there is a range of elements X, on these grounds it is possible to conclude that the other range of elements Y must appear in this transaction too. The ascertainment of this kind of dependences allows us to find very simple and intuitively clear rules.

Algorithms of search for associative rules are assigned to find all the rules $X \Rightarrow Y$, for all this support and confidence of these rules must be higher than some thresholds defined beforehand, which are called minsupport and minconfidence respectively.

The search of associative rules is not a trivial task at all, as it may seem from the first glance. One of the problems is algorithmic difficulty at finding frequently met sets of elements, because with the growth of the number of elements in I (| I |), there is an exponential growth of the number of potential sets of elements.

The number of columns in the table is equal to the number of elements, present in the set of transactions D. Each note corresponds to a transaction, where in the respective column there is 1, if the element is present in the transaction, and 0 if otherwise. The necessary condition is the fact that the data should be transformed to their normal look, otherwise the algorithm is not applicable.

Nevertheless, to reveal frequently met sets of elements is an operation, demanding a lot of calculation resources and, consequently, time. The primitive approach to solving this task is a prime enumeration of all the possible sets of elements. This will demand $O(2^{|I|})$ operations, where, |I| is a number of elements. Thus appears the necessity to use one of the support characteristics, saying: support of any set of elements cannot exceed the minimal support of any of its subsets.

This characteristic is called anti-monotony and serves to lower the size of the search space. If we did not have this characteristic, finding multi-element sets would be practically impossible because of exponential growth of calculations.

The characteristic of anti-monotony can be formulated in another way: with the growth of size of set of element, the support decreases or remains the same. From all the above-said follows, that any k- elementary set will be frequently met only when all its (k-1) - elementary subsets are frequently met.

All the possible sets of elements from I can be represented as an array, starting with empty set, then on the 1-st level there are 1-element sets, on the 2-nd – 2-element sets and so on. On k-level k-element sets are represented, and they are connected with their (k-1)-element subsets.

Let us consider figure 2, illustrating the set of elements I – {A, B, C, D, E}. Let us suppose that the set of elements {A, B} has the support lower than the given threshold, and, consequently they are not frequently met. Then, according to the anti-monotony characteristic, all its supersets are not frequently met, either, consequently, they are truncated. All this branch, starting from {A, B}, is marked with grey background. Application of this heuristics allows us to reduce to a great extend the search space.

At the first step of the algorithm 1-element frequently met sets are counted. For this purpose it is necessary to pass through all the set of data and calculate support for them, i.e. how many times they are met in the base.

The following steps will consist of two parts: generation of potentially frequently met sets of elements (they are called candidates) and calculation of the support for the candidates.

After all the frequently met sets of elements have been found, one can start the direct generation of rules.

In order to extract a rule from the frequently met set F it is necessary to find all its nonempty subsets. And for each subset s we can formulate the rule s ⇒(F – s), if the confidence of the rule conf(s ⇒(F – s)) = supp(F)/supp(s) is not less than the threshold minconf.

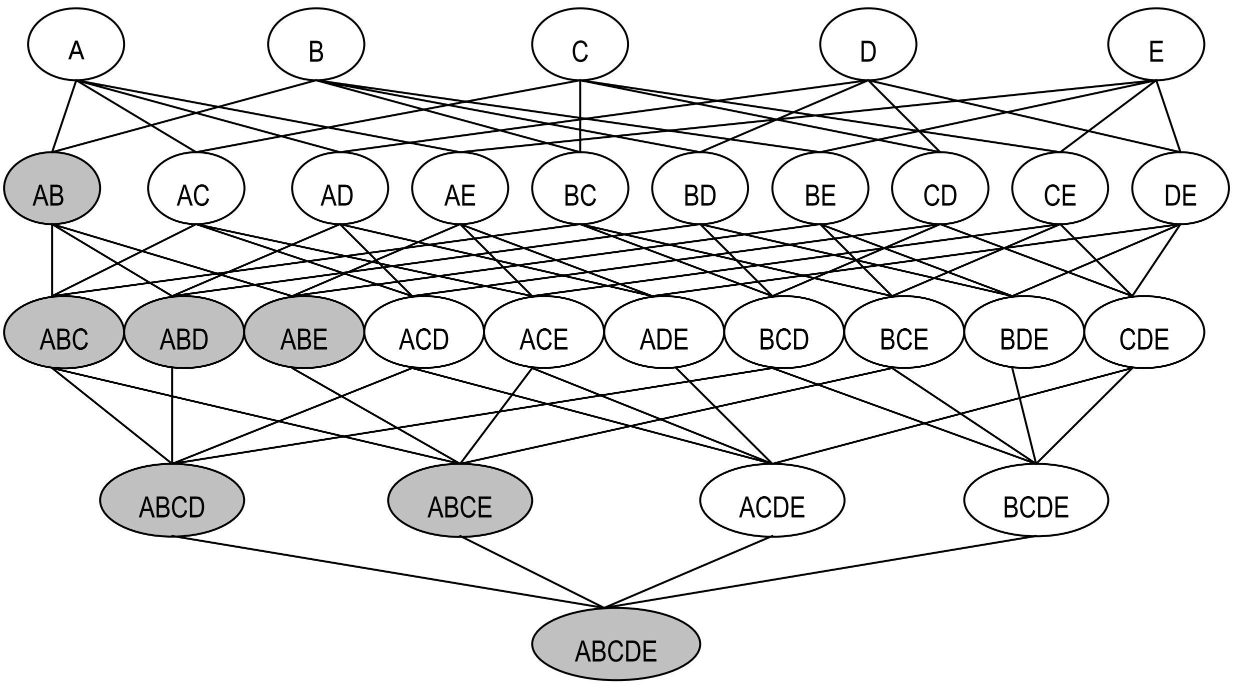

Figure 2 - Graph of application of the rule of anti-monotony at the formation of sets of conjugated production.

Let us comment that the numerator remains constant. The confidence has the minimal value if the denominator has the maximum value, and this happens if in the terms of the rule there is a set, consisting of one element only.

All the supersets of the given set have smaller or equal support, and respectively, the higher value of confidence. This property can be used at the rules extraction. If we start to extract the rules, regarding at first only one element in the rule terms, and this rule has the necessary support, then the rules in the terms of which there are supersets of this element, also have the value of confidence higher than the given threshold. For example, if the rule A $\Rightarrow$ BCDE satisfies the minimal threshold of confidence - minconf, then AB $\Rightarrow$ CDE also satisfies it. To extract all the rules recursive procedure is applied. An important comment: any rule, composed of frequently met set, must contain all the elements of the set. For example, if the set consists of the elements {A, B, C}, then the rule A $\Rightarrow$ B should not be considered.

In this way, applied to our task, it is necessary to establish the level of minimal confidence (minconfidence), the level of which will define the minimal threshold at the selection of the groups of goods (rules), which further on will be considered as one whole at the distribution in matrix cells; in its turn this will allow to take into account the conjugation (mutual dependence) of different types of goods. Still, a problem remains: to define the level of minimal confidence (minconfidence), i.e. if to establish it too high, we may mot take into consideration a part of mutually dependant goods, and consequently, not solve the given task, and if we establish a lower level of minimal confidence, we risk to form groups of goods of little dependence on one another, and, consequently, hide from visual representation the situation with some low-profit goods (having joined them to groups of goods with higher profit). To solve this problem it is possible to use the expert method. As a result of solving a set of tasks (about 20 experiments) it was found out, that the optimal one is the level within 90 up to 100%, depending on the initial quantity of goods. The more it is, the higher should be the level of minimal confidence.

Let us take as an example of one of Belarusian companies to examine the solution of the task considered above. As a result of the analysis of more than 120 kinds of products, the task of association was solved, and owing to this some goods were substituted for the goods-set. That means the distribution of the production into groups A, D and C was made not only for the production which gives a certain share of profit, but also for the acquired with the probability over 90%. The normalized value of mean-square of profit deviation (σ) was used as the risk indicator. At σ value of not more than 0,33 the production was placed to subgroup X, at σ value from 0,33, to 0,66 the production was placed to subgroup Y, at σ value over 0,66 the production was placed to group Z. Thus, the matrix was build, in each cell of which one can find the sum of profit and the group of production placed in this cell (figure 3).

Before the association carrying out

	A	B	C
X	A, C; 7133	E, G; 627	N, S, W, U; 215
Y	B; 868	F, Q, H; 484	X, K, L, T, Z; 143
Z		P, R, Y; 412	D, I, O, M, V; 113

After the association carrying out

A	B	C	
C; 3060	GN, E; 667	S, W, U; 175	X
AFD; 4247	QI, H; 348	X, K, T, Z; 104	Y
BO; 901	PRL, Y; 451	M, V; 42	Z

Figure 3 – Profit control matrixes of one of Belarusian companies before and after the association carrying out

Then the problem of the choice of the analysis order for the production in the cells arose, i.e. production of which cell should be analyzed first of all, which next, etc. For this purpose quadric matrix of pair alternatives comparison was created, and it was filled with estimates from the priorities scale of T. Saati, based on the expert considerations on the following rule. If one alternative of the pair (K1) is more important than the other one (K2) on both criteria of the profit sum and risk, then the cell on the crossing of row K1 and column K2 is filled with the number of expert's evaluation, and the cell of the matrix situated on the crossing of row K2 and column K1- with the reciprocal number. Then, applying one of the possible methods - geometric middling of the line - the vector of priorities is created; in this vector to each row of the matrix corresponds to the meaning of the share of geometric middling value, calculated from the values of each line in the total middling geometric value [8]. The cell with the maximum value will be analyzed first of all and so on according to decrease of values of priority vectors. As a result of practical application within the frameworks of the task solution, one of Belarusian companies got the following vector of priorities (table 3):

Table 3 – Vector of priorities of the analysis of cells of profit control matrix.

Cell	AX	AY	AZ	BX	BY	BZ	CX	CY	CZ
Priority vector	0,391	0,391	0,391	0,391	0,391	0,391	0,391	0,391	0,391
№ of cells analysis	1	2	9	5	4	3	8	6	7

After the order of the analysis is defined, it is necessary to reveal the problems for each group and type of the production, for this purpose one of the methods of Data Mining – the tree of solutions is used [2]. Economic characteristics influencing the profit are shown as nodes, the sign of the branch represents the stimulating direction to profit, thus the volume of sales should be increased (+), and costs should be lowered (–). The tree of factors of profit dynamics formation is represented as a cognitive card in figure 4.

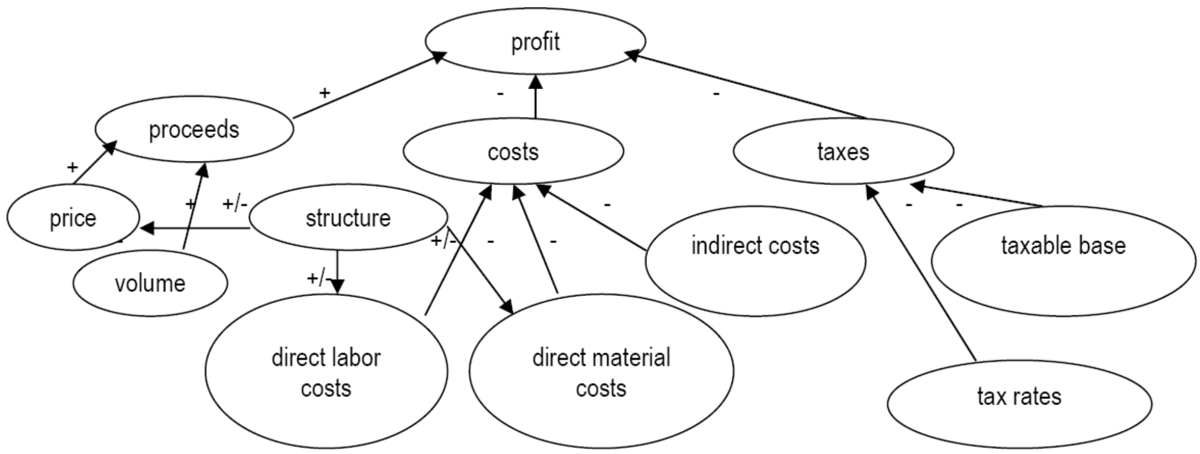

Figure 4 - Cognitive card of interrelation of economic characteristics in their influence on profit

The reason of profit decrease is revealed on the bases of the methodology of marginal analysis application. The critical (maximum permissible value) is calculated for every characteristic (tree node), then the actual value of the characteristic is analyzed in dynamics and compared to the critical value. Thus, for characteristics with negative value, branches of actual value should be lower than critical (maximum permissible) and decrease in dynamics, and for the branches with positive value, the actual value should be higher than the critical (minimal permissible) and increase in dynamics. If the value of the node satisfies the condition of permissibility, then one should move on the following branch. If the condition of permissibility is not observed, the identification of this node takes place- it is identified as the reason of profit decrease and so on and moving on all the branches takes place.

Thus, the matrix of profit control, where such tasks and methods as Data Mining, association, tree of solutions, production rules were realized, will allow the company (depending on the period of time- a day, a month, a year, a few years) to analyze the situation effectively (to reveal the risk degree) both in short-time and long-time periods. This will be realized on the bases of the production classification according to the degree of its meaning for the company financial stability in enlarged groups and in the section of assortment. As a result of the hierarchical structure one can make fragmentation, and the method of the tree of solutions allows to reveal practically all the “narrow” points and, consequently, to take well-grounded management solutions which will allow to greatly improve the actual system of economic analysis on the base of finding out fundamentally new facts and become the basis for considerable competitive advantages.

Bibliography

1. Хмельницкая И.В., Косенков О.А. Новые подходы к анализу и управлению прибылью //Новые информационные технологии в образовании: Доклады и выступления участников седьмой Международной научно-практической конференции «Использование программных продуктов фирмы «1С» в инновационной деятельности учебных заведений» - М.: Федеральное ГОУВПО «Финансовая академия при правительстве Российской Федерации», Фирма «1С» , 2007, С.307-310.
2. Хмельницкая И.В., Косенков О.А Применение Data Mining в экономическом анализе финансово-хозяйственной деятельности предприятия //Инженерный вестник, 2006, №1 (21)/1.- С.192-195.
3. Дюк В., Самойленко А. Data Mining: учебный курс (+CD). –СПб: Питер, 2001. – 368с.:ил.
4. Методы и модели анализа данных OLAR в Data Mining . – СПб., 2004. - 336с.
5. R. Agrawal, T. Imielinski, A. Swami. 1993. Mining Associations between Sets of Items in Massive Databases. In Proc. of the 1993 ACM-SIGMOD Int’l Conf. on Management of Data, 207-216.
6. S. Brin et al., 'Dynamic Itemset Counting and Implication Rules for Market Basket Data', In Proc. ACM SIGMOD Int’l Conf. Management of Data, ACM Press, New York, 1997.
7. J. Hipp, U. Guntzer, and G. Nakaeizadeh. Algorithms for Association Rule Mining – A General Survey and Comparison. In Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2000.
8. Саати Т. Принятие решений. Метод анализа иерархий. - М.: Радио и связь,1989. – 316с.

Authors' Information

Inna Vladimirovna Khmelnitskaya – senior lecturer of the Belarusian-Russian University, the town of Mogilev, the Republic of Belarus; postgraduate student of the Belarusian State University, the city of Minsk, the Republic of Belarus; e-mail: inna21@gmail.com

Oleg Alekseevich Kosenkov – senior lecturer of the Belarusian-Russian University, the town of Mogilev, the Republic of Belarus; undergraduate of the Belarusian State University, the city of Minsk, the Republic of Belarus; e-mail: kos.oleg@gmail.com