Об'єктом дослідження є процес управління виробництвом продукції на багатономенклатурному харчовому підприємстві. Предметом дослідження є застосування методів інтелектуального аналізу даних для підтримки процесу коригування асортименту продукції. У результаті дослідження розроблено інформаційну технологію, яка включає в себе адаптовані методи інтелектуального аналізу даних. Використання цієї технології дозволить підвищити ефективність господарської діяльності підприємств. Це досягається за рахунок зниження витрат на собівартість продукції. Запропонована технологія може бути використана як у харчовій, так і в інших галузях виробництва

Ключові слова: багатономенклатурне харчове підприємство, інтелектуальний аналіз даних

Объектом исследования является процесс управления производством продукции на многономенклатурном пищевом предприятии. Предметом исследования является применение методов интеллектуального анализа данных для поддержки процесса корректировки ассортимента продукции. В результате исследования разработана информационная технология, которая включает в себя адаптированные методы интеллектуального анализа данных. Использование этой технологии позволит повысить эффективность хозяйственной деятельности предприятия. Это достигается за счёт снижения расходов на себестоимость продукции. Предложенная технология может быть использована как в пищевой, так и в других отраслях производства

Ключевые слова: многономенклатурное пищевое предприятие, интеллектуальный анализ данных

UDC 004.65:005.51:664
DOI: 10.15587/1729-4061.2018.123383

DEVELOPMENT OF INFORMATION TECHNOLOGY FOR SUPPORTING THE PROCESS OF ADJUSTMENT OF THE FOOD ENTERPRISE ASSORTMENT

O. Kharkianen
PhD, Associate Professor*
E-mail: helen@nuft.edu.ua
O. Myakshylo
PhD, Associate Professor*
E-mail: mem2004@ukr.net
S. Hrybkov
PhD, Associate Professor*
E-mail: sergio_nuft@nuft.edu.ua
M. Kostikov
PhD, Senior Lecturer*
E-mail: Kostikov@nuft.edu.ua
*Department of Information Systems
National University of Food Technologies
Volodymyrska str., 68, Kyiv, Ukraine, 01601

1. Introduction

The main task of multiproduct food companies is to gain profit from business activities, continuous satisfaction of consumers' demands and expansion of the seller's market segment. Characteristic feature of a multiproduct food enterprise is a rather complicated production process as well as a large number of varied ingredients with their individual storage conditions and shelf life. In addition, it should be noted that different food products have different effective lives and different types of packing.

A very complicated problem of a multiproduct food business is planning and analyzing the finished product cost price. Cost of raw materials for production and auxiliary materials for packaging is constantly fluctuating because of external economic conditions that are difficult or impossible to predict and control [1, 2].

Effective planning of the assortment of finished products is impossible without the use of modern information technologies for data processing and analysis. The use of information technologies is important for processing large volumes of information with minimal errors and high accuracy. A minor deviation from the required value can lead to significant losses.

Each food business is taking steps to increase its competitiveness. Today, it is not enough to produce high quality products in certain quantities. It is necessary to look for the ways to meet consumer demands, offer a broad spectrum of products simultaneously cutting their cost price. The established information technology makes it possible to prepare numerous solutions for adjusting the assortment while decreasing overall cost price thru processing the accumulated business information. This technology is relevant as it supports decision-making in the managerial process of planning the finished product assortment.

2. Literature review and problem statement

Food business management requires prompt making of many important decisions to optimize the business processes. Specificity of activities of the multiproduct food companies consists in production of a large assortment by processing various raw materials, the use of peculiar formulations

in manufacture and various materials for pre-packaging and packaging. Calculation, planning and product cost-price analysis is a very laborious problem [1, 2] and its solution in the information space is impossible without the use of modern information technologies of data analysis.

Most of the works devoted to improving the management process of multi-product business emphasize implementation of a variety of strategies for managing and controlling organization of production processes. Choice of a hybrid strategy for managing batch production is substantiated in [3]. However, the authors point out that the larger number of product types the longer time of application of the proposed hybrid strategy. Optimization of the plan of multiproduct manufacture using the bee algorithm is given in [4] but it does not take into account the possibility of the assortment adjustment. Work [5] is aimed at modifying the strategy of management of the multiproduct manufacture taking into account solutions of works [6, 7] which, in their turn, are aimed at optimizing management of multi-stage production with a batch workflow. But the common disadvantage of papers [3–7] is the failure to take into account properties and features of the assortment and limitation of their use in various branches of the food industry.

Study [8] is devoted to the methodical approach to solving the multi-step dynamic problem of a proper formation of the optimum assortment structure for a long period. This approach applies only to non-food businesses.

A number of papers are dedicated to solving the problem of processing large volumes of data for the purpose of making optimal decisions in the production management. Creation and use of predictive time models using methods of data mining were proposed in paper [9]. Also, the authors identified the most significant business functions that arise in production guidance loops but the proposed solution just partially relates to the task of planning the product cost price.

Forecast of raw material supplies for food businesses is considered in paper [10]. Conceptual essentials for introduction of an innovative and simplified supply forecast system are presented there. The proposed approach is based on data mining but the impact of such measures on the change of the product cost price is not studied.

Work [11] proposed support for planning food product cost price based on the analytical cube models OLAP (online analytical processing) constructed by the authors. However, this work bears a more general nature of application and is more dedicated to the effective formation of appropriate information structures for data storage.

All methods of intelligent and multidimensional data analysis but without their practical application are most fully represented in [12, 13]. Also, these works consider OLAP and BI (business intelligence) tools but all examples are of educational nature. The author of [14] examines in great detail all aspects of cluster analysis and its application for solving various analytical tasks of management but the study is confined to general simple examples as well. All aspects of application and basics of factor, discriminant and cluster analysis are presented in [15]. It should be noted that the works [12–15] do not address adaptation of the methods of multidimensional and intelligent analysis of data for the needs of the problem of analysis and planning of production in food businesses.

A subsystem of information support for managing risks which may arise in planning the product cost price at a multiproduct enterprise was developed on the basis of OLAP technologies in [16].

Authors of work [17] proposed the concept of production management using the Six Sigma methodology. They propose to search for the causes of defects in manufacturing processes and finished products based on processing of statistical information. Paper [18] presents results of an effective use of the methods for analyzing time series of intelligent data processing in the study and forecast of production processes. Authors of work [19] proposed the use of cloud technologies in combination with the methods of intelligent analysis in electronic commerce. Authors of paper [20] present the results of effective use of the data mining methods in the process of planning the use of production facilities. This is done on the basis of information accumulated for previous years.

In present-day systems of food business management of various levels, various means and tools of OLAP and BI are used. Analysis of the aforementioned published data [3 20] gives grounds for asserting their limited application. This is especially true for adjusting the finished product assortment in order to raise profitability of the enterprise by reducing the product cost price. No clear algorithm of applying means and tools of OLAP and BI for adjusting the finished product assortment in operational production management was demonstrated.

3. The aim and objectives of the study

This study objective was development of an information technology to support the process of making managerial decisions in operational production management at multiproduct food enterprises. These decisions concern the search for reserves of cutting the total product cost price and obtaining additional profits thru assortment adjustment.

To achieve this objective, the following tasks were solved:

– form the main stages of the information technology in elaboration of proposals for adjusting the assortment structure;

– adapt the algorithms of intelligent analysis of the data obtained in processing statistical business information to support stages of the information technology in adjusting the assortment structure.

4. Adaptation of methods of multidimensional and intelligent data analysis for solving the problem of assortment adjustment

Evolution and implementation of information technologies have brought about introduction and use of various information systems at the food enterprises. The main importance of existing information systems is the large amount of accumulated relevant business information. Its analysis and static processing can provide solution of the management and control problems. Analysis requires solving both standard and non-trivial problems. These problems require submission of proposals to the decision maker (DM) promptly, in a visually convenient form with no resort to specialists in information technologies. To provide support to managerial decision making, the authors have worked out a hybrid data warehouse (DW) based on previously accumulated data.

This relational data warehouse takes into account all data required for analysis and planning of the product cost

price. Its use will unload the enterprise OLTP system, reduce the response time of sampling the data in requests, will form a single logical representation of information for analysis and making managerial decisions.

The data warehouse has a denormalized structure based on the "star" scheme. This scheme is focused on high performance when executing analytical queries. Design of the dimensional data warehouse model structure typical for multiproduct food enterprises was accomplished using the CA AllFusion ERwin Data Modeler CASE tool. This tool ensures creation of the DW structure in the DBMS chosen by enterprise [21, 22]. The data warehouse structure was designed in such a way that it enables construction of hypercubes containing the data necessary for the cost price analysis [11].

The use of the data mining technology was determined by the need for a more subtle, informal analysis of costs and the search for possible ways to their cutting. Classic statistical methods are based on the use of averaged indicators and neglect atypical observations. Application of these methods makes it possible to seek answers to pre-formulated hypotheses. However, such approach does not always reliably characterize the investigated process [9, 10, 12, 13].

Correction of the assortment structure is proposed to be solved with the use of the information technology consisting of 5 main stages. The scheme of implementation of this information technology is shown in Fig. 1

Application of the data mining algorithms forms the basis of the proposed information technology.

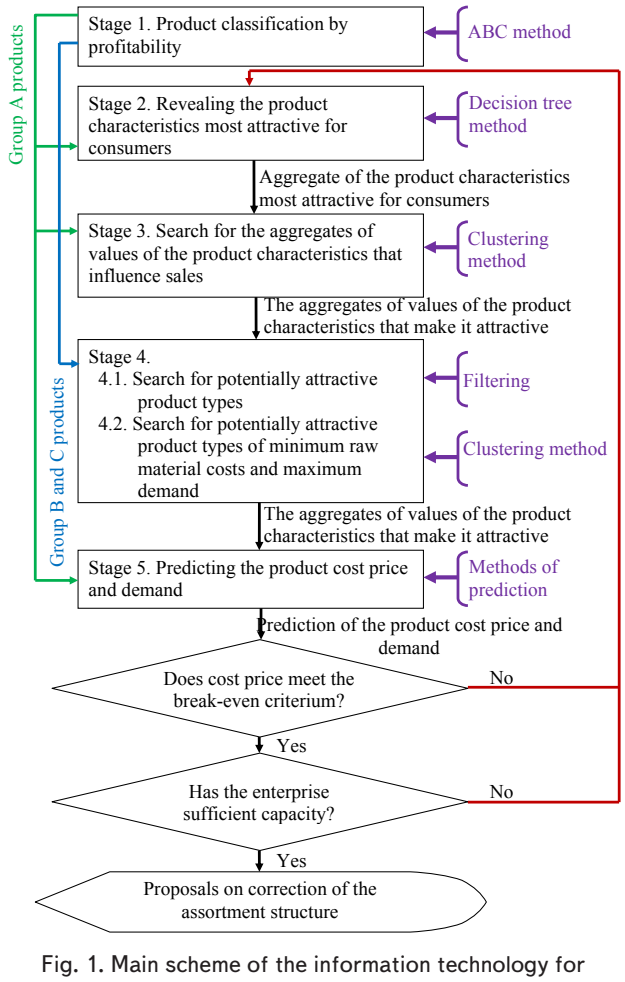


Fig. 1. Main scheme of the information technology for adjusting the assortment structure by means of data mining

Application of the proposed technology provides the DM with information support for assortment adjustment. This is done thru introduction of new types of products, the use of original and popular ingredients in formulations, buildup of production volumes.

To understand the proposed technology, let us consider its main stages in formation of proposals for adjusting the assortment structure for a specific planning period.

The source of information for solving the problem of forming proposals concerning assortment adjustment is the OLAP Costs and Product Sales cube sent to the OLAP client.

Stage 1. Classification of products by the ABC profitability method.

At the first stage, products are classified according to their profitability by the ABC method based on the Pareto principle.

The additive function of profit from production will take the form:

$$P(J_1, J_2, \ldots J_n) = P(J_1) + P(J_2) \ldots P(J_n), \quad (1)$$

where $J_1, J_2, \ldots, J_n$ is the list of products of one assortment group.

According to the Pareto principle, there is such a number $0 < a' < 0.5$, that divides the product list into two groups, *GR1* and *GR2*. It is made in such a way that the size of the group is determined as $GR1 = a' * n$. In this case, the result

$$P(GR1) = (1 - a') * P(GR1, GR2),$$

where $(1 - a')$, is 20 % from the total result of all objects. Present-day researchers are inclined to choose the a' value within [0.2 ... 0.4] which corresponds to [20...40 %]. Thus, by continuing the Pareto analysis in the interval $(0, a')$, we obtain next Pareto points of the required order [23, 24].

In the practice of classical ABC analysis, assortment is divided into 3 groups:

A: products that bring the main income;

B: less demanded types of products;

C: the least demanded types of products.

As a result of the assortment ranking, products of the group A containing the most profitable products are selected from the total volume.

At the end of the first stage, demand forecast for products of the group A is made. In a case of demand growth for products of group A, the plan of production of products of this type is adjusted in the direction of increase within the limits of allowable reserves of production facilities. Otherwise, the production plan for the products of group A is left unchanged.

Stage 2. Identifying characteristics of the products most attractive to consumers.

At the second stage, characteristics of the products most attractive to the consumer are identified.

To formulate recommendations for adjusting the assortment structure, influence of the product characteristics is analyzed taking into account demands of the trade market.

The initial information for implementation of stage 2 is the initial set of data filtered according to the list of products of group A.

To generalize the product characteristics, classification is proposed according to common features:

– quantitative characteristics (the product price/volume ratio, caloric and nutritional value of the products, shelf life);

– qualitative characteristics (presence of specific and high-value raw materials, food additives, natural raw material substitutes, presence of 'No GMO' marking in formulations);

– formulation characteristics (formulation composition and manufacturing standards).

The proposed classification is not unchangeable. It can be expanded depending on the product type. The importance of using these or other characteristics during analysis is determined on the basis of expert opinions of the enterprise specialists.

To evaluate attractiveness of each product type, the data mining algorithm, the "decision tree" is used. The decision tree algorithm allows one to construct nonparametric models which do not contain initial information about the relationship between the data. It would be possible to evaluate attractiveness of each product by classical statistical methods but they are parametric and require knowledge of the model type and available clear hypothesis of the relationship between the data. Statistical methods work only with numerical data and to solve the above problem, it is necessary to classify both numeric and categorical data. Application of the decision tree algorithm provides identification of the cause-and-effect hierarchy of conditions for determining characteristics of the most attractive products. Also, this algorithm allows one to limit the general list of characteristics by determining the most significant of them. For a deeper definition of dependencies between the main characteristics of products and their sales, re-employment of the decision tree algorithm is made.

Implementation of the second stage results in a possibility of getting the following important solutions:

– taking into account the relationship between the product characteristics that are essential for development of new product types of a particular group;

– explore the possibility of applying the revealed characteristics to improve sales of the existing types of products.

Stage 3. Search for the aggregates of values of the product characteristics that affect sales.

At the third stage, search for the sets of values of the product characteristics that affect sales is made. To find the most popular aggregates of characteristics, a model to support decision making by clustering method is constructed [14, 15]. The model is constructed according to the characteristics of the group A products found in stage 2 and the sales volume in quantitative terms.

Application of the clustering analysis makes it possible to find the data structures containing value sets of the characteristics with which the products will be more attractive to consumers.

The initial information for clustering is the initial set of data filtered according to the list of the group A products.

Let us denote the set of product types of group A through $J=\{j_1, j_2, \dots j_i \dots j_n\}$, j_i is a certain type of products.

The task that appears is constructing the cluster set K and mapping E of the set J to the set K, that is $E: J \to K$.

The mapping E sets the data model that is the solution to the problem. Quality of the problem solution is determined by the quantity of correctly qualified data.

Each product is defined by the set of own characteristics and the sales. Let us present the product description as follows (2):

$$j_i=\{z_1, z_2, \dots z_h \dots z_m\}, \quad (2)$$

where z_h is the product characteristic that may take value from a certain set of values of the product characteristics; h is the characteristic number; m is the maximum quantitative value of characteristics for the current product type;

The problem of clustering consists in making the set described by formula (3):

$$K=\{k_1, k_2, \dots k_l \dots k_g\}, \quad (3)$$

where l is the cluster number; g is the quantity of clusters; k_l is the cluster containing similar product types from the set J.

A large number of clustering algorithms are known but the k-means algorithm and its varieties were the most widely used in practice. Application of this method enables uniting of the product types similar by their characteristics and sales into separate clusters. The degree of proximity in the entry of products into a cluster is determined based on the Euclidean distance. Quantity of clusters is set at the stage of algorithm training.

This method is used in the practical implementation of the clustering algorithm by means of Microsoft analysis services.

The study of cluster structures provides a possibility of defining the set of values of the characteristics that make the product attractive. Analysis of clusters makes it possible to distinguish a combination of popular ingredients and other characteristics in preparing formulations for new product types. Identification of the most popular ingredients makes it possible to prepare recommendations for their replacement or introduction into existing formulations.

Stage 4. Search for potentially attractive types of products according to the above value sets of characteristics.

At the fourth stage, search is made for potentially attractive types of products (in groups B and C) based on the above value sets of characteristics most attractive to consumers.

Let us filter the set of products of groups B and C and select only those types of products that contain the value sets of the characteristics found in the third stage. The resulting list of products potentially attractive to the consumer is used to find reserves for building up their production volumes.

Further, products are sought according to the above-mentioned value sets of characteristics with minimal raw material costs and maximum demands.

Taking into account the previously found characteristics, products are classified according to the use of raw materials (especially expensive ones) on the basis of sales of the finished products.

The initial information for clustering is the initial data set filtered according to the product groups B and C.

The study of cluster structures enables identification of a list of products potentially attractive for consumers with minimum raw material costs and maximum sales in quantitative terms in groups B and C.

After that, it is possible to clearly distinguish the types of products that need:

– additional promotional activities to increase sales;

– buildup of production volumes.

Stage 5. Forecasting cost price and demand for products.

At the fifth stage, forecast of the cost price and demand for each type of products determined in Stages 1–4 is made.

This ensures definition of profitability of their production for future periods.

The source of information for forecasting the cost price of individual types of products is the Product Cost Price hypercube. To forecast costs by the costing items, Cost Price by Costing Items and Cost Price by the Expense Objects hypercubes were used. The forecast level of the cost price must meet the criterion of minimum production costs.

The historical information on the enterprise activities accumulated in the DW was the basis for forecasting.

Solution to this problem was proposed using the methods of mathematical statistics (trend extrapolation and correlation-regression analysis) and the data mining methods for their further comparison.

In solving the problem by the data mining methods, the time series algorithm (Microsoft) which is an aggregate of the ARTXP and ARIMA algorithms was used. The ARTXP algorithm uses the tree model with autoregression and is optimized to forecast the next value of the row. This algorithm is used for short-term forecasting. The ARIMA algorithm uses integrated sliding average autoregressions and is designed for long-term forecasting. The algorithms support automatic detection of seasonality or periodicity in data based on the fast Fourier transform. They enable manual setting the value of the seasonality parameter.

When making a decision on the final level of the product cost price, it is necessary to take into account emergence of possible risks. Risks may be related to market conditions during the planned period. The cost price of finished products in the food industry is highly dependent on a number of factors. One of them is the current market price for raw materials, electricity, and other resources. Deviation of the actual costs from the forecast ones may result in losses not expected by the enterprise. Thus, those companies that will be able to dynamically adapt to the changes in the competitive environment and risk management will successfully operate in that environment. Application of the risk management system enables timely detection, assessment, localization and control of the risk [16].

Having the product cost price forecast and the risks of adopting its level evaluated, let us forecast the demand for products by the data mining method of time series given that this indicator is more predictable and largely determined by consumer habits. Estimate the model used in forecasting on the basis of retrospective data for the previous period and the forecasting risks according to the risk management system [16].

Estimate the managerial decision based on the results obtained by the break-even point method. An increase in production of popular types of products leads to the redistribution of conditionally fixed costs for the entire production volume. This, in the end, will contribute to cutting down the product cost price.

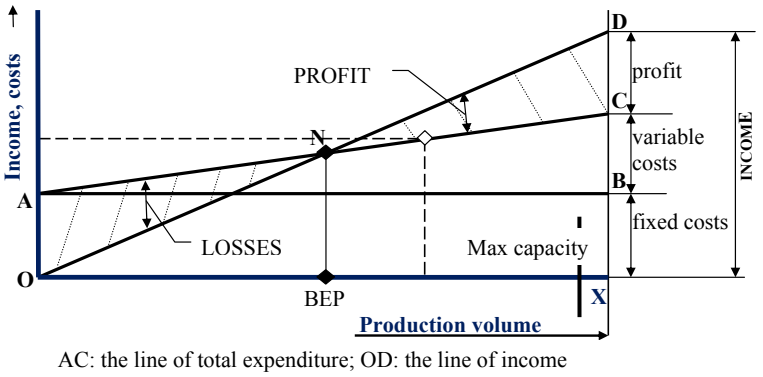
The forecast level of demand for products should satisfy the possibilities of production output at the enterprise facilities and the break-even criterion which is set by the formula (4):

$$X_{BEP} \le X_j^{\text{forecast}} \le A_i, \tag{4}$$

where A_i is maximum capacity of the i-th type of equipment for manufacturing the j-th product type; X_{BEP} is the volume of products at the break-even point.

The break-even calculation method is used to estimate the level of forecast cost price and production volume. Break-even analysis or analysis of costs, profits and output is the analytical approach to studying the relationship between expenses and income at various production levels [16]. Analysis of the break-even point in combination with other methods makes it possible to control the critical volume of production and provide sufficiently substantiated information for making managerial decisions in planning.

Graphical representation of the break-even method (Fig. 2) makes it possible to estimate the relationship between the volume of production X_j, the cost price C_j and the income at various levels of production.



AC: the line of total expenditure; OD: the line of income

Fig. 2. Graphical representation of the method of break-even analysis

The break-even point of the j-th type of products is calculated by (5).

$$BEP = \frac{VP}{G_j - VZ_j}, \tag{5}$$

where VP is the sum of constant costs; G_j is the price of j-th type of products; VZ_j are variable costs for the j-th type of products.

For a multiproduct enterprise, the break-even point is calculated by (6).

$$BEP = \frac{VP}{\sum_{i=j}^{n} G_j - VZ \times \gamma}, \tag{6}$$

where VZ is the sum of variable costs; γ is the percentage of each j-type product in the total volume of sales.

The break-even graph and the analytical calculation of the break-even point for each product on the basis of the obtained forecast values makes it possible to visually estimate whether the cost price is in the zone of profit and whether the forecast product volume can be produced at the production facilities of the enterprise.

The final verification of the solutions resulting from application of the information technology is proposed to be carried out based on optimization of the production program using the simplex method. The criterion of the production program is obtaining of the maximum profit from the assortment offered for the enterprise production line. The quantitative and qualitative results obtained in the previous stages are used to establish and correct constraints of the production program proceeding from the cost price and demand for products.

5. Preparation of managerial decisions on the example of baking industry

For testing the proposed information technology, business information accumulated at the bakery enterprise in recent years has been used. The following questions arise when planning production for future periods:

– which products will be most in demand;

– is there a relationship between the product formulation, other characteristics and sales.

In order to find ways to reduce costs, let us use the above technology applying the data mining methods. With the help of this technology, proposals will be made for adjusting the assortment of bakery and confectionery products.

Search for hidden regularities is carried out on the basis of the OLAP Costs and Sales of Products cube passed to the of MS Excel with The Client of the Intelligent Data Analysis add-on of the OLAP client.

The ABC method has allowed us to distinguish from the assortment a product of group A which brings the highest profit. Having the demand forecast and capacity reserves estimated, the DM may plan measures to increase the production volumes.

Using Group A, we shall find the product characteristics most attractive to the consumer applying the decision tree algorithm. The results are shown in Fig. 3.

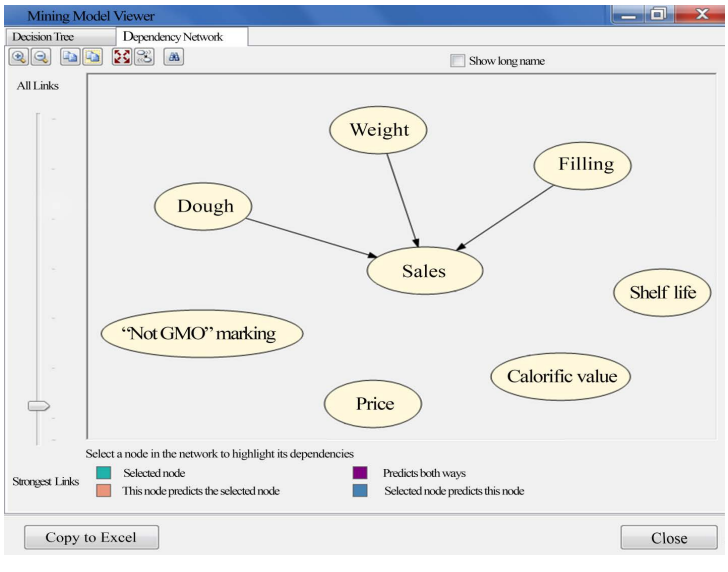


Fig. 3. Characteristics of the most attractive products

To analyze influence of characteristics of the product on its sales, the following characteristics were highlighted: the product weight, shelf life, price, presence of 'No GMO' marking, filling type, dough type. Such characteristics as weight, filling type, dough type have the most significant influence on the product attractiveness (Fig. 3). The value of each characteristic is displayed by the arrow color saturation. In the absence of connection, it is considered that the parameter is less important.

Application of the decision tree algorithm with the software implementation of Microsoft Decision Trees has allowed us to determine values of the characteristics of the most attractive products. A fragment of the constructed decision tree for analysis of influence of characteristics on the product sales is shown in Fig. 4.

By analyzing the constructed decision tree (Fig. 4), we can conclude that the most attractive products are those biscuit dough with a butter cream weighing more than 515 g. This dependence can be seen by the naked eye since namely these nodes are most clearly highlighted in the tree. Products of white dough with butter cream weighing ≥515 g, products of sheet dough with butter cream weighing ≥515 g, etc. are sold to a lesser extent.

Detailed analysis of certain nodes and values of certain characteristics can be done in the built-in model. A more detailed analysis of the model has made it possible to find a relationship between the "filling" characteristic and the sales. The detailed analysis of this fragment (Fig. 5) has revealed that the most attractive products are those with butter filling. Detailed analysis of other options clearly shows that the products without filling ("buns") are more attractive to the consumer.

Thus, the decision tree analysis results in a list of characteristics of product attractiveness. These characteristics are recommended to be taken into account by the DM when planning new types of products. As a result, it is possible to distinguish the main values of the "filling" characteristic in a decreasing order: butter, bun, fruit, etc.

Using the predefined characteristics, we shall build a model of decision support applying the clustering method. This will make it possible to determine the most popular combination by an aggregate of characteristics (weight, filling, dough) and sales. The results of applying the clustering method are shown in Fig. 6.

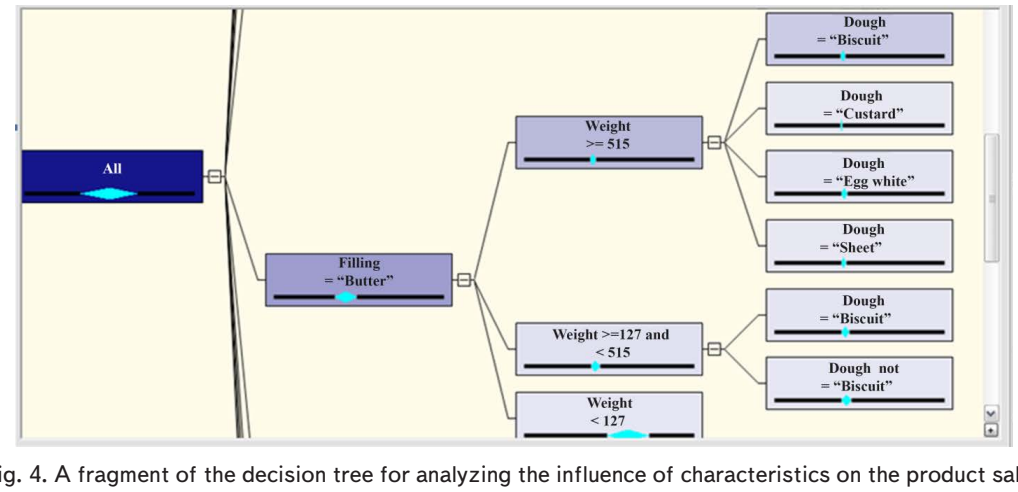


Fig. 4. A fragment of the decision tree for analyzing the influence of characteristics on the product sales

Fig. 6 shows the general distribution of attributes in a cluster. Each line in the color sequence displayed in the cells corresponds to the behavior of customers in the sales channels of the cluster. Different colors correspond to different products. For example, the dark blue color corresponds to a product of a biscuit dough with a butter cream filling. This set of characteristics is present in many clusters in the first place. This makes it possible to conclude that the products with such a combination of characteristics will most often be demanded by various sales channels first of all.

Thus, the found aggregates of characteristics are recommended for consideration by the DM in development of new types of products and if replacement of raw materials is necessary for production.

Applying the previously found sets of characteristics filter out the products of groups B and C to identify products in these groups potentially attractive to consumers.

Since the found products are potentially attractive but do not belong to the product group A, it is advisable to recommend working out the advertizing measures aimed at the product promotion to the markets.

When developing measures to buildup production volumes, it is necessary to take into account the potentially attractive types of products found in groups B and C.

The aggregates of values of characteristics found above were also used to find potentially attractive types of products with minimal costs for *expensive raw materials* and of the highest demand.

To determine the list of products of groups B and C with a butter cream filling which were recommended to produce by the enterprise with minimal expenses for raw materials and with the maximum volume of sales, the clustering model has been applied. The result of work of the clustering model is shown in Fig. 7.

Fig. 7 shows total distribution of attributes in the cluster: the cost of raw materials in the product, the name of the product, the name of the raw material. After analyzing all the data for each cluster, it was determined that the products in cluster 7 with the butter cream filling (Kazka cake, Suvenirny cake, Napoleon cake, Lisovy cake, Carmen pastry and Suvenirne pastry) had the lowest raw material costs (flour, agar, peanuts) and the largest sales. The list of products included in cluster 7 is the most beneficial of groups B and C. This should be taken into account in the planning process.

Given that the clustering and decision tree methods have found potentially attractive types of products of groups B and C, it is advisable to forecast their cost prices and sales for future periods. The product cost price was determined by the extrapolation method using correlation-regression models. Data mining (intelligent data analysis) algorithms of time series were also used to compare the obtained forecasts. Let us prognosticate the cost price and sales for the next month for one of the found product types with the butter cream filling, Suvenirne pastry (weight: 200 g).

The results of forecast of the product cost price with the help of extrapolation methods, correlation-regression models and the time series data mining with an expert assessment of the risks of acceptance of the forecast values are given in Table 2.

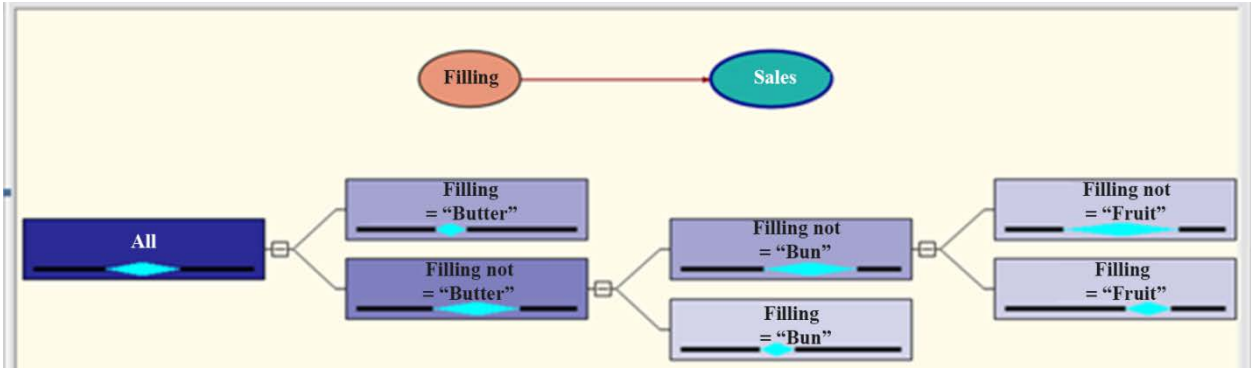


Fig. 5. Decision tree of the product attractiveness depending on the filling

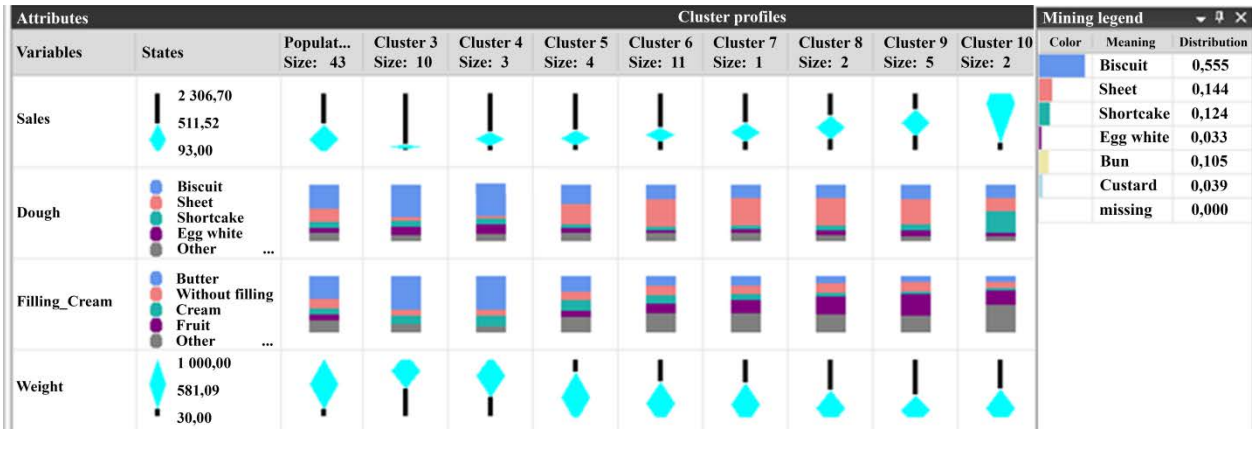


Fig. 6. General distribution of attributes in a cluster

Fig. 7. Cluster profiles

Table 2

Information support for determining the forecast value of cost price

Cost price		Forecast taken into account	Expert formulated probability, %	Expected value, USD	RMS deviation, USD	Coefficient of variation	RMS deviation, USD	Coefficient of variation
					according to the selected forecasts		according to all forecasts	
The forecast obtained by the use of extrapolation methods according to the costing items with further cost price calculation	2.681	+	33.3	2.714	0.048	0.018	0.073	0.265
The forecast obtained by the use of extrapolation methods according to the costing items with further cost price calculation	3.719		0					
Forecast of raw materials, materials and services obtained by the use of correlation-regressive models with further calculation of cost price	3.805		0					
Forecast of total product cost price obtained by the use of the method of correlation-regressive models	2.76	+	33.4					
Forecast obtained with the use of the time series method by the means of data mining	2.768	+	33.3					

Since the values of the four forecasts are very close, they were given the same probability. Instead, two forecasts were rejected by experts. The expected value of the product cost price was USD 2.74 at the RMS deviation of USD 0.0448, that is 1.75 %. Based on this information, the DM having accepted the cost price of the Suvenirne pastry for the plan for the next period at USD 2.74 nearly does not risk because this value can change just by 1.75 %. At the same time, the system warns about a possible change in the cost price by 26.5 % if the decision of rejecting the two forecasts was wrong.

Let us make forecast of the demand for products in the next period by means of data mining. The results are shown in Fig. 8.

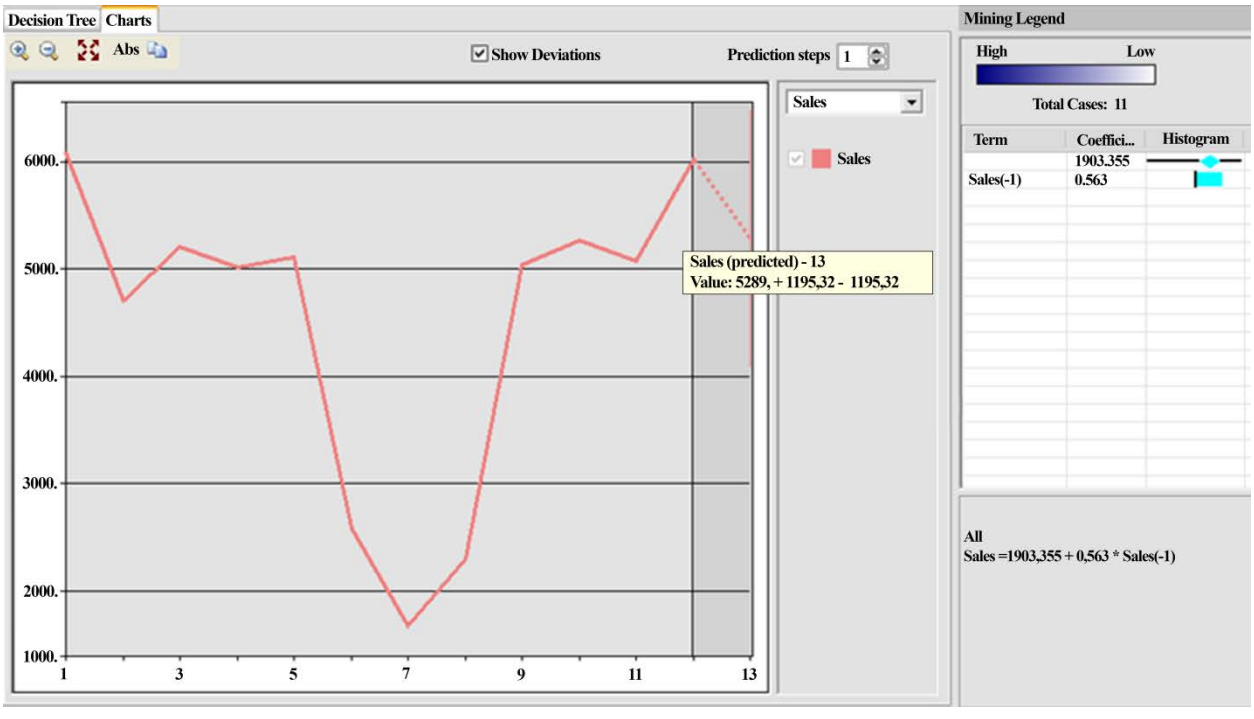


Fig. 8. Forecast of products sales by means of data mining

The curve in Fig. 8 reflects the previous year sales and shows the forecast for the next month. According to the analysis results, the best sales of products with a butter cream filling were in 1 and 12 months and the worst in 6, 7, 8 months. This is explained by the fact that production of products with a butter cream filling was limited precisely in these months when air temperature was above 24 °C. The forecast for the sales of Suvenirne pastry with a butter cream filling for the 13th period will be 5289 pcs.

To draw up recommendations on the adoption of the cost price and sales of the Suvenirne pastry, a break-even graph was used. The graph has shown that the product cost price was in the profit zone and its corresponding output was 2,100 pcs at forecast sales of 5,289 pcs. Taking into account that this volume of production can be realized at the enterprise facilities, the DM can accept the forecast of demand for this pastry at a level of 2,100 to 5,289 pcs.

After determining the products potentially attractive to consumers and forecasting the cost price and demand, the final decision on adjusting the product assortment is taken on the basis of optimization of the production program taking into account the limitations adjusted according to the found solutions.

6. Discussion of the results of development and application of information technology

The peculiarity of the developed information technology of adjusting the product assortment structure for the multiproduct food enterprise is expansion of the ways of using valuable business information accumulated by the enterprise in the course of economic activity due to adaptation of methods of multidimensional and intelligent data analysis.

Use of a corporate data warehouse as a body of data for analysis raises the value and demand of the food enterprise information sources.

The proposed technology adapts the data mining methods to provide information support in planning the assortment of the multiproduct food enterprise. Application of data mining methods makes it possible to take measures of searching for the ways to reduce production costs by analyzing the cost price, sales and characteristics of the products that are in greatest demand.

The proposed information technology combines the use of various approaches and methods, namely:

– the method of expert evaluations for estimating the risks of adopting the forecast values of cost price;

– the break-even method for determining admissible volumes of output of the proposed products;

– the method of optimizing the production program to verify the resulting decisions.

Such an approach reduces likelihood of incorrect management decisions and, accordingly, protects the food business from unnecessary costs.

The proposed information technology of supporting the process of adjusting the multiproduct assortment of finished products along with production scheduling is of a theoretical and practical value for the specialists in the field of management and information technologies.

The proposed information technology is applicable in various fields of production. It has a flexible application capacity. According to the authors, the proposed information technology may not take into account certain features of non-food businesses, nor is it useful for the enterprises providing services. One of its advantages is that it enables processing of large data bodies.

The main drawbacks of the solution include the fact that all models have to be regularly reset with appearance of the need for analysis of the data obtained over the last period of time. It should also be borne in mind that the complex of algorithms for solving the problem of adjusting the product assortment as the initial data necessarily requires large volumes of consolidated corporate data.

Further development of this study and the resulting solution is aimed at building of a web-oriented system that would fully implement all stages of the created information technology. When creating such a web-based system, it is expedient to use public web services or create own modules for data mining.

7. Conclusions

1. The proposed information technology of supporting the process of adjusting structure of the product assortment of the multiproduct food business is based on the integrated use of the data mining methods. The technology features the algorithm of phased joint use of such methods of data mining as clustering, decision tree, forecast for preparation of managerial decisions regarding the food business planning.

2. Evaluation of the obtained results is made by the method of expert evaluations for estimating the risks of adopting the forecast values of cost price, the break-even method for determining the admissible output volumes of the proposed products, the method of optimization of the production program for making final decision on the possibility of manufacturing the proposed assortment at the enterprise facilities.

3. The described approach expands the ways of using the economic data accumulated by the food business and thus makes higher their value. The new knowledge obtained allows the company, without attracting additional funds for marketing studies, to search for profitable and potentially popular types of products, to update their assortment, to work out products with the most demanded components of formulations and other characteristics.

4. Practical value of the study consists in elaboration of a standard tool for preparing decisions on assortment planning in the decision making support system of the food enterprise. The tests carried out at several food enterprises with various production pattern confirmed effectiveness of this solution since the profit growth has been achieved due to: conditional annual savings, increase in production volumes, product cost price reduction and structural shifts in the product assortment.

References

1. Tsymbaliuk L. H., Antoshkina L. I., Skryhun N. P. Planuvannia diyalnosti pidpryiemstv kharchovoi promyslovosti: pidruch. / Berdian. un-t menedzhmentu i biznesu. Donetsk: Yuho-Vostok, 2014. 340 p.
2. Tsymbaliuk L. H., Skryhun N. P., Antoshkina L. I. Formuvannia ta upravlinnia vytratamy vyrobnytstva. Donetsk: Yuho-Vostok, 2013. 240 p.
3. A comparison of HK-CONWIP and BK-CONWIP control strategies in a multi-product manufacturing system / Onyeocha C. E., Wang J., Khoury J., Geraghty J. // Operations Research Perspectives. 2015. Vol. 2. P. 137–149. doi: 10.1016/j.orp.2015.07.001
4. Ajorlou S., Shams I. Artificial bee colony algorithm for CONWIP production control system in a multi-product multi-machine manufacturing environment // Journal of Intelligent Manufacturing. 2012. Vol. 24, Issue 6. P. 1145–1156. doi: 10.1007/s10845-012-0646-5
5. Onyeocha C. E., Khoury J., Geraghty J. Evaluation of the Effect of Erratic Demand on a Multi-Product Basestock Kanban-CONWIP Control Strategy // Proc. of the 9th conf. on stochastic models of manufacturing and service operations (SMMSO 2013). Kloster Seeon, 2013. P. 147–155.
6. Baynat B., Buzacott J. A., Dallery Y. Multiproduct Kanban-like Control Systems // International Journal of Production Research. 2002. Vol. 40, Issue 16. P. 4225–4255. doi: 10.1080/00207540210146198
7. Onyeocha C. E., Geraghty J. A Modification of the Hybrid Kanban-CONWIP Production Control Strategy for Multi-Product Manufacturing Systems // Proc. of the 29th International Manufacturing Conference (IMC29). University of Ulster, Belfast (UK), 2012.
8. Shorikov A. F., Rassadina E. S. Dinamicheskaya optimizaciya kompleksnogo programmnogo upravleniya strukturoy tovarnogo assortimenta predpriyatiya // Ekonomika regiona. 2012. Issue 3. P. 261–271.
9. Modeling complex manufacturing process activities using data mining approach / Tseng T.-L., Nadackal A. J., Kwon Y., Zhang J., Yao T. // IIE Annual Conference and Expo 2007 – Industrial Engineering's Critical Role in a Flat World: Conference Proceedings. 2007. P. 560–565.
10. Holimchayachotikul P., Phanruangrong N. A Framework for Modeling Efficient Demand Forecasting Using Data Mining in Supply Chain of Food Products Export Industry // Advances in Intelligent and Soft Computing. 2010. P. 1387–1397. doi: 10.1007/978-3-642-10430-5_106
11. Miakshylo O. M., Kharkianen O. V. Planuvannia sobivartosti produktsiyi kharchovoho pidpryiemstva na osnovi analitychnykh modelei OLAP-kubiv // Kharchova promyslovist. 2011. Issue 10-11. P. 332–337.
12. Metody i modeli analiza dannyh: OLAP i Data Mining: ucheb. pos./ Barsegyan A. A., Kupriyanov M. S., Stepanenko V. V., Holod I. Sankt-Peterburg: BHV-Peterburg, 2004. 336 p.
13. Analiz dannyh i processov: ucheb. pos. / Barsegyan A. A., Kupriyanov M. S., Holod I. I., Tess M. D., Elizarov S. I. Sankt-Peterburg: BHV-Peterburg, 2009. 512 p.
14. Dyuran B., Odell P. Klasterniy analiz. Moscow: Statistika, 1977. 128 p.
15. Faktorniy, diskriminantnyy i klasternyy analiz / Kim Dzh.-O., M'yuller Ch. U., Klekka U. R. et. al.; I. S. Enyukov (Ed.). Moscow: Finansy i statistika, 1989. 215 p.

16. Miakshylo O. M., Kharkianen O. V. Doslidzhennia i rozrobka metodiv upravlinnia ryzykamy v diyalnosti kharchovoho pidpryiemstva // Naukovi pratsi Natsionalnoho universytetu kharchovykh tekhnolohiyi. 2014. Vol. 20, Issue 5. P. 105–114.
17. Hajduová Z., Weiss E., Mixtaj L. Application of Statistical Methods at Copper Wire Manufacturing // Metalurgija. 2009. Vol. 48, Issue 1. P. 55–58.
18. Perzyk M., Krawiec K., Kozłowski J. Application of time-series analysis in foundry production // Archives of Foundry Engineering. 2009. Vol. 9, Issue 3. P. 109–114.
19. Cloud Computing and Big Data as Convergent Technologies for Retail Pricing Strategies of SMEs / Suciu G. et al. // Proc. of the 8th International Conf. „Challenges of the Knowledge Society" (CKS 2014). Bucharest (Romania), 2014. P. 1044–1052.
20. Bubeník P., Horák F. Proactive Approach to Manufacturing Planning // Quality Innovation Prosperity. 2014. Vol. 18, Issue 1. P. 23–32. doi: 10.12776/qip.v18i1.208
21. Tumanov V. E., Maklakov S. V. Proektirovanie relyacionnyh hranilishch dannyh. Moscow: Dialog-MIFI, 2007. 333 p.
22. Maklakov S. V. Modelirovanie biznes-processov s AllFusion PM. Moscow: Dialog-MIFI, 2008. 224 p.
23. Koch R. Living the 80/20 Way: Work Less, Worry Less, Succeed More, Enjoy More. London: Nicholas Brealey Publishing, 2014. 196 p.
24. Davydov A. A. Ubyvayushchie chislovye posledovatel'nosti v sociologii: fakty, ob'yasneniya, prognozy // Sociologicheskie issledovaniya. 2001. Issue 7. P. 113–119.
25. Târnăveanu D. Pentaho Business Analytics: a Business Intelligence Open Source Alternative // Database Systems Journal. 2012. Vol. 3, Issue 3. P. 23–34.