Рис. 1. Главная форма программы

Во многих магазинах это приложение поможет продавцу определять есть ли в магазине то или иной диск, добавлять диски при поступлении, редактировать при нужде, удалять, а также рассказать покупателю об исполнителе.

Список литературы

1. Климова Л. М. Delphi 7. Основы программирования. Решение типовых задач. Самоучитель. М. : КУДИЦ-Образ, 2017. 480 с.

3. Культин Н. Delphi 6. Программирование на Object Pascal. М. : БХВ-Петербург, 2012. 528 с.

7. Культин Н. Основы программирования в Delphi 8 для Microsoft.NET Framework. Самоучитель (+ CD-ROM). М. : БХВ-Петербург, 2013. 400 с.

10. Осипов Д. Delphi. Программирование для Windows, OS X, iOS и Android. М. : БХВ-Петербург. 2014. 464 с.

13. Ревич Ю. Нестандартные приемы программирования на Delphi. М. : БХВ-Петербург. 2016. 560 с.

THEORETICAL ASPECTS OF DATA MINING METHODS

A. Tsitsina
Karaganda Economic Universiti of Kazpotrebsyuz
(Republic of Kazakhstan)

The article considers the goal of data mining, data mining methods, the processes of development in Kazakhstan.

Keywords: intellectual analysis of data, Data Mining, information technologies, intellectual technologies.

В статье рассматривается цель интеллектуального анализа данных, методов интеллектуального анализа данных, процессов развития в Казахстане.

Ключевые слова: интеллектуальный анализ данных, интеллектуальный анализ данных, информационные технологии, интеллектуальные технологии.

The entry of humanity into the information age is associated primarily with the tremendous changes in the field of information activity. Today, our life is almost unthinkable without a computer, the Internet and other information technologies, which every day are becoming more friendly and convenient thanks to the introduction of the latest technological innovations in them, in particular, elements of artificial intelligence. Intellectualization, becoming an imperative for the development of modern means of communication, information retrieval, computing, data processing and analysis, significantly increases the availability of information technology for users with different levels of computer training.

The goal of an intelligent decision analyzer is to determine the correct proposed solution or not; finding what is specifically wrong or incomplete in the answer; and, perhaps, determining what missing or incorrect knowledge might be responsible for the mistake. Intelligent analyzers can provide far-reaching feedback and update the model. Intelligent analysis deals with finite answers to problems.

Changes in the ways of processing and analyzing data arrays, and, as a result, the acquisition of new knowledge about the phenomena under study is associated with the introduction of new methods and tools for analysts who have emerged (and continue to emerge) in the process of developing data mining (IAD). Therefore, the system faces new challenges, due to the need to introduce a program that will allow future professionals to get acquainted with the latest achievements in the field of data processing and analysis. This raises the question of the relevance of such innovations against the background of existing problems in mathematical education.

In our rapidly developing time, information technologies occupy a significant place. Any technology is a key link in any subject area. A distinctive feature of the technology of data mining methods (IAD) is that it is a tool for professionals working in any subject area.

The basis for data analysis is modeling. Building models is a universal way to explore the world. Building models allows you to detect dependencies, extract new knowledge, predict, manage and solve many other problems. Models and modeling are closely related to such a basic concept as a system.

System is a central concept in systems theory and system analysis. The system is commonly understood as a set of objects, components or elements of an arbitrary nature, forming a certain integrity in a particular context. Each system carries the principle of emergence of the system - the system has new properties that do not have its constituent elements.

There are many different methods of data mining, query modeling, processing and collecting information.

The purpose of data mining (eng. Datamining - "data mining", "data mining") is the detection of implicit patterns in data sets. As a scientific area, it began to actively develop in the 90s of the 20th century, which was caused by the wide-

spread use of automated information processing technologies and the accumulation of large amounts of data in computer systems [1, 2, p. 24]. And although existing technologies allowed, for example, to quickly find the necessary information in a database, in many cases this was not enough. There was a need to search for relationships between individual events among large volumes of data, for which we needed methods of mathematical statistics, database theory, artificial intelligence theory, and a number of other fields.

The classic definition is considered to be given by one of the founders of the direction, Grigory Pyatetsky-Shapiro [2, p.42]: DataMining - research and detection by "machine" (algorithms, artificial intelligence tools) in raw data of hidden knowledge that were not previously known, non-trivial, practically useful, available for interpretation.

In essence, data mining is the processing of information and the identification of patterns and trends in it that help make decisions. The principles of data mining have been known for many years, but with the advent of big data, they are even more widely used.

Big data has led to an explosive growth in the popularity of broader data mining methods, in part because there is much more information, and by its very nature and content it becomes more diverse and extensive. When working with large data sets, relatively simple and straightforward statistics are no longer enough. With 30 or 40 million detailed purchase records, it's not enough to know that two million of them are made in the same place. To better meet the needs of customers, it is necessary to understand whether these two million belong to a certain age group and know their average earnings.

These business requirements led from simple search and statistical data analysis to more sophisticated data mining. To solve business problems, such data analysis is required, which allows you to build a model for describing information and ultimately leads to the creation of a resulting report. This process is illustrated in Figure 1.
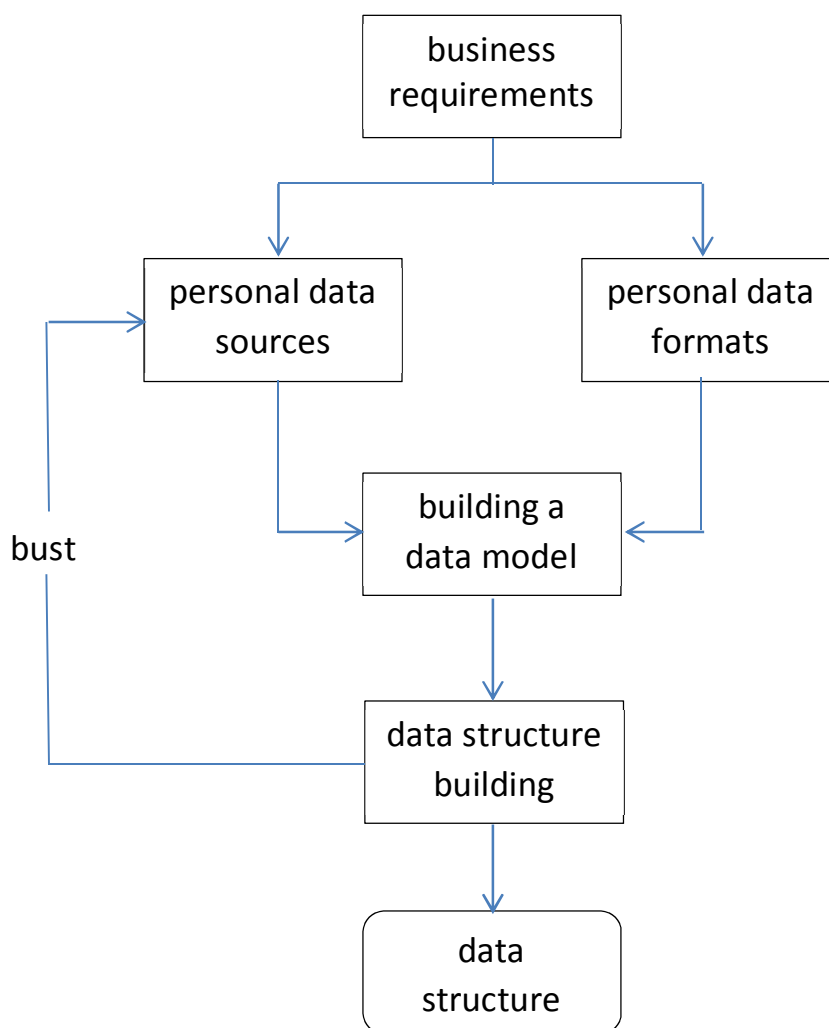
```
                    ┌──────────────┐
                    │   business   │
                    │ requirements │
                    └──────────────┘
             ┌─────────────┴─────────────┐
             ▼                           ▼
    ┌──────────────┐            ┌──────────────┐
    │ personal data│            │ personal data│
    │   sources    │            │   formats    │
    └──────────────┘            └──────────────┘
             │                           │
             │      ┌──────────────┐     │
   bust      └─────▶│  building a  │◀────┘
                    │  data model  │
                    └──────────────┘
                            │
                            ▼
                    ┌──────────────┐
                    │data structure│
                    │   building   │
                    └──────────────┘
                            │
                            ▼
                    ╭──────────────╮
                    │     data     │
                    │  structure   │
                    ╰──────────────╯
```

Fig. 1. Diagram of the process

The process of analyzing data, finding and building a model is often iterative, as it is necessary to find and identify various information that can be extracted. It is also necessary to understand how to link, convert and combine them with other data to get the result. After the discovery of new elements and aspects of data, the approach to identifying sources and formats of data with the subsequent comparison of this information with a given result may change.

The specificity of modern requirements for information processing (a huge amount of data and their heterogeneous nature) makes both statistical and expert approaches powerless in many practical areas, including economic ones. Therefore, to analyze information accumulated in modern databases, methods must be effective, i.e. easy to use, have a significant level of scalability and a certain automatism. This concept is the basis of two modern technologies Data Mining and KDD - Knowledge Discovery in Databases.

The classic definition of the technology of "data mining" (Data Mining) reads as follows: it is the detection in the initial ("raw") data of previously unknown, non-trivial, practically useful and accessible interpretation of knowledge. That is, the information found in the process of applying Data Mining methods should be non-trivial and previously unknown, for example, average sales are not.

Knowledge should describe new relationships between properties, predict the values of some signs based on others.

There are several methods of data analysis. Among them, the most frequently used are the following: methods based on a predictive rating system, and methods based on Data Mining technology [5, p.38].

The consequence of the above disadvantages is the low quality of the forecast based on the ratings, as well as the large time, labor and financial costs of the preparation of the forecast.

An approach to forecasting based on Data Mining technologies eliminates the disadvantages of traditional approaches.

Approaches based on Data Mining technologies:
  – logical regression;
  – Решений decision trees;
  – neural networks [5, p. 38].

Data mining is a set of approaches united by the idea of computer mathematics and the use of the theory of artificial intelligence.

This group includes the following methods:
  – artificial neural networks (recognition, clustering, forecast);
  – evolutionary programming (including algorithms for the method of group accounting of arguments);
  – genetic algorithms (optimization);
  – associative memory (search for analogs, prototypes);
  – fuzzy logic;
  – decision tree;
  – expert processing systems.

All methods of IBA are divided into two large groups according to the principle of working with initial training data [3, p. 28].

As can be seen from the table, each of the methods has its own strengths and weaknesses. But no method, whatever its assessment from the point of view of its inherent characteristics, can provide a solution to the entire spectrum of Data Mining tasks.

The dynamics of the economic, social and socio-political situation in Kazakhstan imposes new requirements for information and analytical support of management activities in both the public and commercial sectors [9].

The basis of the modern software industry and the decisive factor for success in creating information and anilithic systems is the technology of their creation. Information-analytical systems are a special class of information systems designed for analytical data processing, and not for automating the daily activities of an organization. Information and analytical systems combine, analyze and store as a single whole information extracted from the accounting databases of organizations and from external sources. Data warehouses that are part of information analysis systems provide for the conversion of large volumes of highly detailed data into generalized verified information that is suitable for making informed decisions. Unlike conventional databases, warehouses contain a processed,

streamlined, and understandable for managers. The data warehouse is an assembly line for preparing information in an integrated, consistent, visual form to support management decision making [6, p. 26].

Creation of information and analytical systems that really meet the goals and objectives of organizations is a rather complex process, including the stages of the formation of concepts, design, development, implementation and maintenance. Thus, a general method of creating information and analytical systems is needed, containing the composition and sequence of work and tasks, the composition of role functions and the documents generated.

It is obvious that intellectual technologies reveal new ways to improve the quality of services in the conditions of the modern information society. So an adaptive presentation provides an individual approach, support in solving problems and intelligent analysis of solutions with interactive communication processing can save considerable time, model selection technologies can enhance managerial and communicative aspects.

References
1. The history of the development of data mining methods/ Data Mining. Internet resource. Access mode. URL: http://azfor.ucoz.ru/publ/3-1-0-3.
2. Marina Chapot. Intelligent data analysis in decision support systems. Internet resource. - Access mode. URL: http://www.osp.ru/os/1998/01/179360/
3. Dyuk V. A., Samoylenko A. P. DataMining: training course. SPb. : Peter. 2011.
4. B. de Ville. Microsoft Data Mining. DigitalPress. 2011.
5. Han J., Kamber M. Data Mining: Concepts and Techniques // Morgan Kaufmann, 2010.
6. Mikhail Petrovskiy. Fuzzy Kernel-based Method for Real-time Network Intrusion Detection // Springer-Verlag, Lecture Notes in Computer Science. 2010. Vol. 2887. P. 189–200.
7. Igor Mashechkin, Mikhail Petrovskiy and AndreyRozinkin. Enterprise Anti-spam Solution Based on Machine Learning Approach // Proceedings of 7th International Conference on Enterprise Information Systems. USA. Miami. 2010. Vol. 2. P. 188–193.

УДК 334.78

ОБЗОР РЕШЕНИЙ ОСНОВНЫХ ПРОИЗВОДИТЕЛЕЙ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ РАЗРАБОТКИ И ОБРАБОТКИ БОЛЬШИХ ДАННЫХ

Т. Л. Тен, В. Г. Дрозд, Б. Ж. Спанова
Карагандинский экономический университет Казпотребсоюза
(Республика Казахстан)

В работе рассмотрены особенности использования технологии Microsoft Reporting Services для предоставления аналитической информации конечным пользователям на базе реализации модуля формирования и доступа к выходным документам посредством возможностей служб Reporting Services. В данной публикации дается обзор решений основных производителей программного обеспечения для разработки ХД. При изложении материала используется, по возможности, следующая схема: название проекта компании и его цель; архитектурные решения; СУБД и используемая модель данных; возможности языка обработки данных; степень охвата жизненного цикла (анализ – проектирование – реализация – поддержка); возможные конкурентные преимущества.