OPPORTUNITIES TO IMPROVE THE MANAGEMENT METHODS OF ROMANIAN ORGANIZATIONS

Gheorghe Gabriela
Faculty of Economics and Business Administration, Dunărea de Jos University, Galaţi, România
gabriela.gheorghe@ugal.ro

Abstract: *Information Technology has responded to the crisis by change and innovation. Competitive organizations are permanently prepared to identify the cutting-edge technologies which will change the future in business by using the concept of Business Intelligence. Information Technology is developing rapidly and providing decision makers with large amounts of information that require processing and analysis. Business Intelligence are critical decisions support applications, their role being to provide to management, based on information from the company, a landscape of its situation at some time, but also predictions on the medium and long term. This paper describes the architecture of the Business Intelligence Platform and highlights the benefits of using decision support systems for business applications. Nowadays, company executives require relevant and accurate business information in real-time to take strategic decisions. It is also essential to have freedom to access this information anywhere and anytime. There is a real need to extend this functionality beyond the office and on the fingertips of the decision makers. After a presentation of the concept of Data mining with examples from a case study, the paper focuses on the potential of the IT&C decisional sector, which embeds simulation and assistance tools for managerial decisions. Developing new methods for predictive modelling and application of existing techniques in many areas will be a permanent concern for both researchers and companies that are interested to gain competitive advantages.*

Keywords: *decision support system, business intelligence, knowledge discovery in databases, data mining technologies.*

JEL classification: *L86, M15, O33*

1. Introduction

The IT researchers' concerns for decision support applications have appeared for over 50 years. Many simulation systems have been developed, making it possible to better exploit the mathematical and statistical programming methods. This formalization effort together with the evolution of new generations of computers and operating systems, have led to the development of the first generation of decision-making tools. The first applications of group decision support systems based on models released in the late '60s. These applications were developed in the '70s and in the early '80s appeared financial planning systems and group decision support systems. Then, starting with the '90s, emerged Executive Information Systems (EIS), OLAP (On Line Analytical Processing) technology and Business Intelligence (BI) applications, continuing with the development of Decision Support Systems (DSS) based on web technologies in the mid '90s, and more recently, the concepts of data mining, text mining, web mining, respectively.

One of the earliest definitions proposed for computer based decision support is “a model-based set of procedures for processing data and judgments to assist a manager in his decision making” (Little, 1970).
In 1992, Kanter said in the paper entitled *Managing with Information* that "DSS is used in less structured problems, where the art of management meets science.”
DSS are "those management information systems that rely on the use of analytical models, specialized databases, judgment and intuition of the decision maker and an interactive computer modelling process, that supports semi-structured or unstructured decision making by managers” (O'Brien, 1999).
The concept of BI appears into IBM Journal in 1958 as "automated system used to disseminate information in various sectors of organizations" (Berta, 2012).
BI are high-level decisions support applications, that provide to management, based on information from the company, a picture of its situation at a moment, but also on the medium and long term, facilitating competent decision making. The BI concepts have ceased to be restrictively considered software applications, these tools are advanced management methods that ensure both organization’s data management and their use as management tools for modelling and forecasting.

2. Data Mining and Knowledge discovery in databases (KDD) definitions:
“We analyze Knowledge Discovery and define it as the nontrivial extraction of implicit, previously unknown, and potentially useful information from data“ (Frawley, Piatetsky-Shapiro and Matheus, 1992).
“KDD is the process of finding hidden information in a database. Data Mining is the use of algorithms to extract information and patterns derived by the KDD process” (Dunham, 2003).
DM is “the extraction of hidden predictive information from large databases” with “sophisticated statistical analysis and modelling techniques to uncover patterns and relationships hidden in organizational databases” (Wang, 2003).
Through the knowledge-discovery process, implicit and potentially useful information is extracted from the raw data (Roiger and Geatz, 2003).
Data Mining is “the process of discovering patterns in data. The process must be automatic or (more usually) semiautomatic. The patterns discovered must be meaningful” (Witten and Frank, 2005).
"The automated or convenient extraction of patterns representing knowledge implicitly stored or captured in large databases, data warehouses, the Web or data streams" (Han and Kamber, 2006).
Analyzing accurately and timely of this large amounts of data is a difficult task, in general, not possible with traditional methods. Ability to analyze and use the massive amounts of data remained far behind the possibilities of storing them. This raises new challenges for businessmen and researchers for the purposes of extracting useful information (Wang, Liao & Rees, 2002).
Increasing accuracy for predictions can save significant amount of money for a company and is a major motivation for using the methods of forecasting and systematic investigation for new models and techniques to improve the current results (Zhang, 2003).
According to market surveys conducted by Gartner Group, at present, the world's leading BI market, with important revenues from the sale of these solutions, are SAP,

Oracle, Microsoft, IBM SAS Microstrategy, Actuate, Panorama Software, TIBCO Software etc. (Pettey and van der Meulen, 2012).

3. Confidence in the Data is Key to Business Success

3.1. The emerging need of Business Intelligence instruments

In the decision-computer era, competitive organizations are permanently prepared to identify the cutting-edge technologies which will change the future in business by using the concept of Business Intelligence, proposing as objectives:

- Identifying products and services offer in order to retain customer loyalty
- Best possible risk management and activities optimization
- Intelligent exploitation of stored data for being more informed than competitors
- Anticipate opportunities and threats

Most companies and organizations collect data on a large scale. This huge amount of data can be used to extract knowledge that can represent a real advantage to business managers.

The “Business Intelligence” term is being used in the modern business world to fulfil the emerging need of analyzing data from large and variable data sources and this need brought an evolution and a new line of competition between vendors of different enterprise systems. The dramatic trendy evolution in IT field from the last decade and the increased demand of data driven knowledge, large companies have enabled themselves to store an enormous amount of data. However, the request to derive information quickly and to facilitate the decision making process of business organizations, a new set of tools was required, which leaded to Business Intelligence. BI means the idea of collecting, storing, analyzing and providing access to data and it consists of a set of applications and technologies that support decision making, querying and reporting, online analytical processing, statistical analysis, forecasting and data mining.

This requirement for Business Intelligence applications and technologies has led to the development of BI software. Most major software vendors have launched their own Business Intelligence modules that can be packaged with their software products. Nowadays organizations using enterprise systems are also using the services of available BI software on the market. The functionality of this software is independent of the rest of the modules, but also provides means to seamlessly integrate it with enterprise systems.

In order to identify the BI components, we first need to determine the types of problems are the BI tools designed to solve:

- Structuring data for analysis;
- Analyzing data;
- Creating reports;
- Viewing and sharing reports.

As described above, some components are required to prepare data for analysis, while others are designed for analysis, reporting and sharing.

3.2. DSS requires Big Data

A decision making system requires an overall vision of the Operational Information System. The information it processes is based on the raw operational data, which are subject to processing, that is to the calculations, to which are added different notions of time, location, area, and so on, which can have different meanings depending on the activity field concerned. A common reference system groups elements that can come from several applications, their overlapping representing the memory of the organization.

Enterprises today are beginning to realize the important role Big Data plays in achieving business goals. Concepts like factors that influence a customer to make a purchase, behaviour patterns that point to fraud or misuse, inefficiencies slowing down business processes, now can be understood and addressed by collecting and analyzing Big Data. The insight gained from such analysis helps organizations improve operations and identify new product and service opportunities that they may have otherwise missed. What enterprises need are tools to help them easily and effectively understand and analyze Big Data.

3.3. Business Intelligence Platform

Modern Business Intelligence platform should provide an end-to-end infrastructure, solutions and technologies that support following issues, as seen in the figure 1:

- information integration
- master data management
- data warehousing
- BI tools
- repository of best practices and business models

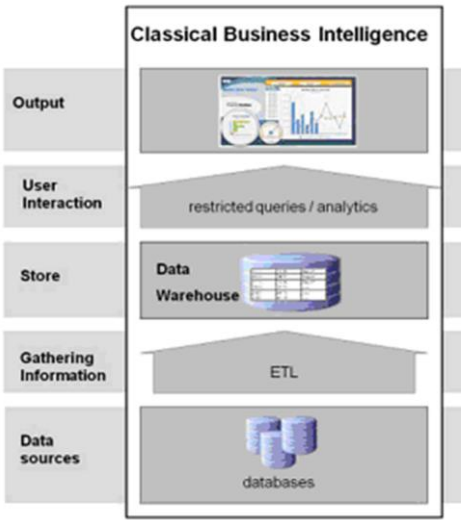

Figure 1: The 5 levels of the traditional BI architecture
Source: BI cubist-project, *http://cordis.europa.eu/projects/rcn/95904_en.html*

3.3.1. Data sources

The data themselves are of two types:

- Operational data from the dedicated transactional processes activities;
- Operational management data consisting of historical data from the activities carried out in a longer period of time, dedicated to strategic management of an organization.

At the operational level, the data is not normalized, there is no conceptual unification. The operational information system can be used to control current activities, but does not allow for long-term decisions, because it does not manage a history of the resulted information. Master data is key business information that may include data about customers, products, employees, materials, suppliers etc. Analysis and reporting is greatly dependent on an organization's master data.

3.3.2. ETL (extract, transform, load) tasks

An Integrated data system is a data collection that gathers consistent data elements from a range of work activities along the continuum of an operational process, stores the data in a coherent way, and is open access by all work activities that require it.
Information Aggregation means that from more information with the same semantic side but from different fields is obtained new information, with the same semantic side and a field formed by joining the fields of the processed information.

3.3.3. Data storage

In traditional BI, data is stored in the Data Warehouses, which can be broken down by activity sector in Data mart. The features of the information are stored in metadata dictionaries that describe the semantics, format and rules of extraction.
A data warehouse in Inmon's (1992) concept is a collection of integrated databases, subject-oriented and designed to provide necessary information for decision making. The elements useful for the analysis are stored as tables (relational databases) or data cubes (multidimensional databases).

3.3.4. Data analysis

The search for these valuable information, yet hidden, patterns and relationships within the data is known as data mining.
Business Intelligence provides to organization support to optimally manage information on their so as organization to able to adopt the best decisions for new competitive advantages. In the same time, BI provides users vital information about techniques used in business and thinking of the competitors, their culture, their intentions and ability to implement them.

3.3.5. BI Tools for Data Mining

The BI tools perform the metadata analysis (data about data) to ensure that meaningful data is collected in order to assure a single version of the truth by consistently performing calculations and reporting against the same data.
The combination of analytics and data visualization should be an integrated component of any BI initiative to enable users to explore data, interact with it and apply analytics to understand it. The analysis and presentation tools, the components visible to the end-user, link the production information, the mid-level decision support information and those needed by the enterprise management.

There are other tools that manage the aggregated and normalized data in a specific decision-oriented database, enabling a multidimensional visualization of information, allowing for decision making at all hierarchical levels of the organization: operational control, management control at department or project level and also strategic planning in the organization.
The reporting tool meets the requirements of operational, planning and regulatory reporting, allowing electronic exchange of data between process and business systems to support effective operations management and business process integration.

4. Case study of a Romanian BI solution

Wizrom Software has released WizWhy application that targets the services sector companies affected by the economic crisis.
The companies that are to change their business partners have a behaviour that can be determined statistically and on which one can create profiles. The proposed solution highlights certain features of the business partner who is about to stop the collaboration, based on the history of the relationship shown in the database. This is just one aspect of the applicability of the data mining solutions for the benefit of the customers. Meanwhile, WizWhy has the ability to also highlight what goes well, so that the management be able to analyze the reasons and act to extend the recipe for success to the whole business.
The operating principle is simple: every company has a database of customers and their business activity. The WizWhy application, the tool we will use as research support for this case study, accesses the database, generating a set of logical rules in which, in addition to quantitative information, qualitative data are also analyzed. The found rules explain with a high percentage of probability why the analyzed results were obtained. Based on these rules, can be highlighted certain profiles used to make predictions. For that matter, the financial services sector, the marketing, market research and insurance are among the main beneficiaries of this software platform.
In addition to the concept of “what if analysis” of the traditional Business Intelligence applications, WizWhy finds on its own those variables that have a significant impact on the results, avoiding subjectivity and can thus detect unusual behaviour of a group of customers through database analysis in order to alert the management. Finding patterns of high-risk companies within financial data enables one to predict to what extent a new customer is risky.
Enterprise Warehouse can reveal interesting phenomena in the data, mainly:

- issuing a summary of the data, reflecting the number of entries, column totals, number of records that meet a set of conditions;
- cases that are exceptions to the rule;
- issuing predictions for new cases starting from the identified rules;
- for auditing purposes: the records that deviate from the discovered patterns in the data might be cases of data entry error or fraud.

Using this BI platform we’ll explore the pattern of the first set of data, referred to as training file, and issue predictions for the second data set, the test file.
Sequence of operations:

- First, it opens the data base to be analyzed;
- Second, the working set of data is loaded;

- Once the data set is open, the parameters to be used in the analysis must be defined;
- Analyse the results, trends and reports displayed;
- Generate predictions.

In order to analyse a data set one of the fields should be defined as the dependent variable, while the other fields are considered the independent variables.
A continuous analysis will calculate the specific percent of the dependent variable as a function of the other fields.
The rules that relate between the dependent variable and the other fields are formulated as “if-then” and “if-and-only-if” sentences. On the basis of the discovered rules can be establish the main patterns in the data in order to use them for issuing predictions for new cases.
“Error probability” or “Confidence Level” indicates the degree to which the rule can be relied upon as a basis for predictions (WizWhy Whitepaper, 2011).
Assuming that the data under analysis is a representative sample of an infinite population, the error probability quantifies the chances that the rule does not hold in the entire population and exists accidentally in the file under analysis.
Numeric fields, such as Amount Purchased, are automatically segmented into intervals, and these intervals are the values in the if-then rules. Revealing all the if-then rules is known as the “association rules” method.
Association rule mining, one of the most important and techniques of data mining, introduced in (Agrawal, Imielinski and Swami, 1993) are widely used in various areas such as telecommunication networks, market research studies and risk management, inventory control, web traffic analysis, security on internet surfing, medical research etc.
Various association mining techniques and algorithms are used in order to extract interesting correlations, frequent patterns, associations or casual structures among sets of items in the transaction databases or other data repositories. Association rule mining is to find out association rules that satisfy the predefined minimum support and confidence from a given database.
One of the main challenges of such a method is to validate each possible relationship for a massive data, in a reasonable time-span.
For instance, as we can see in the figure 2, the data we used contain:

- 1,000 records
- 22 fields in each record
- An average of 10 possible values for each field

View Data

Record: 1000 Total: 1000 records

Cash	Company name	Cost o...	Current assets	Current liabi...	Dividend	Equity...	Goodwill ...	Industry
10,80	ArcelorMittal Roman	1430,40	1267,70	973,40	0,00	380,20	280,80	0121 - Iron & Steel
494,30	UPC Romania	1572,70	2830,60	982,10	0,00	2112,60	358,80	0915 - Communications Services
0,00	Ecosal Ltd.	10,90	3,10	6,80	0,00	-3,30	0,10	0975 - Waste Management Services
36,20	WMC Ltd.	880,70	755,40	658,30	0,31	2781,20	0,00	0124 - Metal Mining
0,30	ING Bank				6,00	8,00	0,00	0730 - S&Ls/Savings Banks
0,10	Still Co	6,30	2,60	3,70	0,00	3,40	4,60	1006 - Computer Hardware
0,60	Hiperdia		0,80	1,50	0,00	0,20	0,00	0803 - Biotechnology & Drugs
7,40	Technautic	110,20	56,90	32,20	0,00	24,80	0,00	1024 - Electronic Instruments & Co...
144,50	Viata Libera	994,10	1374,20	1699,60	0,08	424,20	353,90	0903 - Advertising
2,50	MediTest, Inc.	18,80	14,40	5,00	0,00	52,10	32,00	0812 - Medical Equipment & Supplies
3,40	Nexial.ro, Inc.	0,00	3,50	0,30	0,00	3,50	0,00	0963 - Retail (Specialty Non-Apparel)
12,50	Hobby Tour	91,80	59,30	49,20	0,00	76,80	49,20	0430 - Recreational Products

Figure 2: Basic data loading: Customers.mdb
Source: own compilation

When a data set is opened, it is displayed in the WizWhy work area, as follows:

WizWhy Analyzer

File Edit View Issue Settings Window Help

Issue Trends | Issue Rules | Predict to file | Predict online...

WizWhy1

Basic Data | Rule Parameters | Error Costs | Rule Report | Manual Select | Data Format | Prediction Input

Open Data of Type: OLE DB View Data...

Data Source: C:\Users\Public\WizSoft\WizWhy\Customers.mdb/

Field Grid:

	Field Name	Field Type	Analyze if Empty	Ignore Field	Dependent Variable
2	Company name	Category			
3	Cost of goods sold	Number			
4	Current assets	Number			
5	Current liabilities	Number			
6	Dividend	Number			
7	Equity (common)	Number			
8	Goodwill and intangibl	Number			
9	Industry	Category			
10	Liabilities and equity	Number			
11	Long-term debt	Number			

☑ Analyze the Dependent Variable as Boolean

Figure 3: WizWhy Analyzer work area displays the tabbed dialog boxes

Basic Data | Rule Parameters | Error Costs | Rule Report | Manual Select | Data Format | Prediction Input

Dependent Variable: Sales

The total range of the Field's values is: 0,10 - 111.630,00

Average: 1.355,39 Standard Deviation: 6.491,88 Median: 73,40

Select the range of values to be analyzed in the rules (Boolean analysis):

Predicted Value: ☑ More than: 1000 ☐ Less or equal than: 111630 Frequency: 15,83 %

Minimum probability of if-then rules: 22 %

Minimum probability of if-then-NOT rules: 91 %

Minimum number of cases in a rule: 100 10.0%

Maximum number of conditions in a rule: 3

☐ Re-display Rule Parameters after reading the data.

☑ Search for Unexpected Rules

Figure 4: Settings the rule parameters: Sales>1000 and cases in a rule>=100
Source: own compilation

For the prediction: *'Sales is more than 1.000'* when the Issue Rules button is pressed the following summary report is generated:
Summary Report:
C:\Customers.mdb
Total number of records: 1000
Minimum number of cases in a rule: 100
Dependent Variable: Sales
Predicted Value (analyzed as Boolean): *more than* 1.000,00
The total number of case explained by the set of conditions: 928.
The total number of cases in the data: 935.
Average probability of the predicted value is 0,158.
When performing an analysis on a database of clients in order to discover the rules of loyalty winning, the first step is to identify the basic rules and trends. Parameters are calculated, representing the "importance" of each field to identify customers who are likely to head for the competitors.
If-then rules represent sufficient conditions (the "if" condition is a sufficient condition for the result). If-and-only-if rules go one step further: they represent necessary and sufficient conditions.
Obviously such a relation cannot be accidental, and therefore might be relied upon when issuing predictions. Indeed, when WizWhy reveals if-and-only-if rules it takes them into account when issuing predictions for new cases.
"If-and-only-if" type of relationship offers a higher degree of safety and is considered in order to develop predictions, as we can see in figure 5:

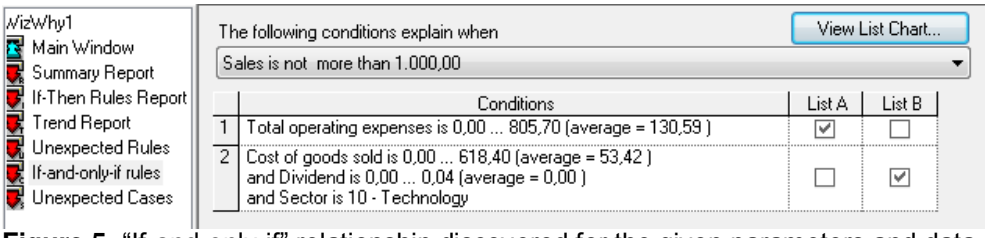

Figure 5: "If-and-only-if" relationship discovered for the given parameters and data
Source: analysis performed with WizWhy

In the previous example, the two conditions, (1) and (2), are necessary and sufficient conditions for the dependent variable 'Sales' being less than 1.000. If at least one of them holds there is a high probability that the Boolean analysis 'Sales grater than 1.000' is false. If the two of them do not hold, there is high probability, respectively 0,961 (147 out of 153 cases) that the condition 'Sales is more than 1.000' is true. In the second if-and-only-if rule above, the first condition refers to the case where the value in the Cost of goods sold field is the interval between 0 and 618,40. The platform employs a unique algorithm for the optimal segmentation of numeric (continuous) fields.

5. Conclusions

The organizational repository of enterprise wide data across the lines of business contains massive and integrated data and represents the complete organizational view of information needed to run and understand the business.

This unprecedented development of IT technologies has led to different changes in the business environment: the appearance and development of knowledge-based organizations, and of the knowledge-based management within the organizations. Within the total resources of a knowledge-based organization, the informational resources register a spectacular dynamic.

Both the explosive growth in the amount of information inside companies, and the rapid evolution of IT equipments, determine more and more companies to adapt their strategies. Due to the diversification of the type of data to be stored, companies need a new way to manage unitarily both their applications and data files, thus adapting themselves easier to the market dynamics.

The IT resources are now becoming the engine of successful enterprises, representing not only the operational base, but also the means for marketing and communication strategies. At the same time, organisations must review the relevance of knowledge management to operational managers.

The knowledge-based organizations are able to gather, extract, label, organize process and share knowledge to make superior performances possible: better business solutions and decisions, a better collaboration and share of information.

In order to face new challenges, the companies must improve their digital and technological knowledge capacities, even during crises or financial restraints.

New software applications accelerate the decisional process in this big data era. Important IT companies announce the launching of new software applications that offer clients a complex method to decrease the big data volume and to speed up the business processes, by making it easy to obtain useful information.

Companies that have deployed data visualization solutions to help derive value from Big Data are finding that there are behind-the-scenes benefits to using these tools. Not only does the combination of strong analytics and data visualization give users the power to make the right business decisions, it also facilitates the joining of different disciplines within an enterprise to help solve business problems.

References

Agrawal, R., Imielinski, T. and Swami, A. N. (1993) “Mining association rules between sets of items in large databases”, *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, Vol.22, No. 2, pp. 207-216.

Berta, D. A. (2012) *Business Intelligence in education*, Leveraging Technology for Learning, Volume II, Proceedings of The 8th International Conference eLearning and Software for Education, Bucharest, 26-27 April, pp. 62-67.

BI cubist-project (2011) [online], http://cordis.europa.eu/projects/rcn/95904_en.html available: http://www.cubist-project.eu/uploads/media/CUBIST_factsheet_short.pdf

Dunham, M. H. (2003) *Data Mining: Introductory and Advanced Topics*, Prentice Hall.

Frawley, W. J., Piatetsky-Shapiro, G. and Matheus, C. J. (1992) "Knowledge discovery in databases: An overview." *AI magazine* 13.3: 57.

Han, J. and Kamber, M. (2006) *Data Mining: Concepts and Techniques*, 2nd edition, Morgan Kaufmann.

Inmon, W.H. (1992) *Building the Data Warehouse*, QED, Wellesley, Mass.
Kanter, J. (1992) *Managing with Information,* Prentice Hall.
Little, J. D. C. (1970) Models and Managers: The Concept of a Decision Calculus, *Management Science*, Vol. 16(8), pp. 35-43.
O'Brien, J. A. (1999) *Management Information Systems: Managing Information Technology in the Internetworked Enterprise*, Boston: Irwin McGraw-Hill.
Pettey, C. and van der Meulen (2012) Analysts Examine the Future of BI During the Gartner Business Intelligence Summit, April 1-3, Los Angeles, Gartner's CIO survey [Online], available: http://www.gartner.com/newsroom/id/1971516 .
Roiger, R.J. and Geatz, M.W. (2003) *Data Mining - A Tutorial-Based Primer*, Addison Wesley, Pearson Education Inc.
Wang, M., Rees, S.J. and Liao, S.Y. (2002) *Building an online purchasing behavior analytical system with neural network*, DataMining III*.,* WIT Press.
Witten, I. and Eibe, F. (2005) *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd edition, Morgan Kaufmann.
WizWhy Whitepaper, [Online], http://wizsoft.com/default.asp?win=7&winsub=31 .
Zhang, G. P. (2003) *Time series forecasting using a hybrid ARIMA and neural network model*, Neurocomputing, vol. 50, pp. 159-175.