

PREDICTIVE ANALYSIS IN THE BIG DATA ERA

Drd., Ramona MARGE, *Universitatea din Oradea, ROMÂNIA*
Dr., Stefan IOVAN, *Universitatea de Vest, Timișoara, ROMÂNIA*
Ing., Alina-Anabela IOVAN, *Informatica Feroviară SA, București, ROMÂNIA*

ABSTRACT: *In the big data era, companies in a multitude of industries increasingly feel the need to collect the most relevant information about their business. They want to find a way to make decisions based on correct information at the right time. To achieve this, the creation of systems that can transform the data collected insights to generate actions that bring business benefits. In a world that is increasingly connected and interdependent, it often happens that corporate decision-makers feel that they are literally attacked by data, but they can not rely on their true potential. The right insights are the basis of operational excellence, providing a significant competitive advantage and leading to business success. New Business Intelligence (BI) trends and innovations overcome complex analytics solutions of large data sets to provide users with other solutions and services with transformation potential.*

KEY WORDS: *predictive analysis, big data, retrospective analysis, consumer behavior, relevant information.*

1. INTRODUCTION

In a world that is increasingly connected and interdependent, it often happens that corporate decision-makers feel they are literally attacked with data, but they fail to capitalize on their true potential. The right insights are the basis of operational excellence, providing a significant competitive advantage and leading to business success.

They want to find a way to make decisions based on correct information at the right time. To achieve this, the creation of systems that can transform the data collected insights to generate actions that bring business benefits. Some of these benefits could be:

► *Identifying growth opportunities* – internal and external data analysis can help shape and predict business outcomes by identifying the most profitable growth opportunities as well as business differentiators.

► *Improving business performance* – data analysis facilitates agile planning, more accurate forecasting, improved budget planning and is an important tool in decision making.

► *Better risk management and regulatory requirements* – data analysis allows for

improved reporting procedures, identification of risk areas such as compliance breaches, fraud or damage to reputation.

► *Use of emerging technologies* – new opportunities can be identified to obtain relevant business management information based on new technologies.

1.1. Predictive analytics transforms businesses

Very few companies use the full potential of predictive analytics [1]. On the other hand, this approach often conflicts with trying to keep control and lower IT costs. Therefore, quantifying and capitalizing on available information and identifying sources of information that can support the generation of new opportunities have become the main challenge.

In the digital age in which we operate, the volume of data generated is steadily increasing. Every minute of every day, over 200 million emails are sent globally, and Google receives over 2 million search requests. It is estimated that by 2020, about 450 billion online transactions will take place every day. Given this context, organizations view data as a fourth factor of production, in

addition to capital and human and material resources.

Effective integration of predictive analytics into business management has a measurable impact on performance as it allows for better planning, clearer forecasts and more informed decisions, leading to profit growth, risk reduction and business agility [2].

The use of predictive analysis is useful for organizations to ensure that all relevant functions are involved in the process so as to obtain an overview and minimize information leaks [1 - 2].

Consumer information is a typical example: sales have billing and transaction logging data, marketing has information from consumer feedback, and the logistics department has details of actual deliveries. All of this information may double or may vary from one department to another.

A consistent analysis of all these data can be a challenge, but a fair and consolidated analysis can generate added value for business. Companies that monitor and estimate how consumer behavior will evolve and its preferences, without exceeding confidentiality limits, can gain important benefits [2].

Predictive analytics helps companies to move from a decision-making process to a proactive, information-driven one. Based on this approach, companies can build models to make better predictions about realistic scenarios, and provide opportunities and associated issues [1].

In a steadily evolving digital world, with increasing volumes of data generated, only those companies that will rely on information will be able to increase their competitiveness. Business performance will depend on the organization's ability to access and exploit the right information.

Those organizations that will be able to understand and filter the relevant information, who will know how to discover patterns and act, based on the results thus obtained, will become top performing businesses.

2. DEFINING TRENDS FOR BUSINESS INTELLIGENCE

In a world that is increasingly connected and interdependent, it often happens that

corporate decision-makers feel they are literally attacked with data, but they fail to capitalize on their true potential [3]. The right insights are the basis of operational excellence, providing a significant competitive advantage and leading to business success.

New *Business Intelligence* (BI) trends and innovations overcome complex analytics solutions of large data sets to provide users with other solutions and services with transformation potential. We present a list of the most important trends in Business Intelligence [4].

The Natural Analytics approach redefines the BI domain. We, people, are now working in a world of extremely complex information with an overload of data from which we have to learn how to choose the essentials. But when it comes to data analysis, it is a process that we naturally do. People have extraordinary abilities to recognize patterns and a natural inclination to associative thinking and to find the meaning of things [3]. Natural Analytics is a BI design and BI approach designed to enhance and expand these natural abilities, helping users gain insight into data and make decisions from the simplest to the most complex. Natural Analytics is a unique QlikView approach and will continue to be a key factor in the next generation of products to be launched.

The data reveals its secrets. It is obvious that those organizations that use their day-to-day data to make decisions are more successful, while those organizations that do not, see their position on the market threatened. If, traditionally, data analysis enters expert assignments as an "analyst", today it is increasingly among the skills of an ordinary business user [3]. The objective of today's most successful Business Intelligence solutions is to offer to any user, regardless of their technological abilities, the power to harness valuable hidden information in the vast amounts of data they have available to make important business decisions.

Visual Data Discovery solutions are growing in popularity. Starting in 2013, worldwide, visual data discovery solutions have become a hot topic of discussion. Specialty vendors such as QlikView have grown rapidly,

encouraged by their more agile and intuitive platforms, allowing users to create their own data query methods without having to turn to specialized IT departments. Visual data discovery tools have made reporting, querying, and creating dashboards easier, more attractive and more flexible than other BI modules.

Organizations choose mobility. Moving from inflexible solutions that emit static reports to user-oriented BI solutions will accelerate as more organizations recognize the importance of obtaining data that each user can base their decisions on. Therefore, BI solutions adapt the way people work, giving them access to information wherever they are at any given time. For both large and small organizations, "business world speed" means, business intelligence solutions are commonly used, not just occasionally. Business users want to be able to access information in the context of their day's natural flow, not just when they're physically in office. PC tablet devices, smart phones, for example, are ideal for mobile BI applications support and encourage collaborative decision-making process.

Data from social media becomes useful tools for business. Increasingly, social media becomes an essential component of the presence of a market organization and contributes to more direct and authentic communication with its audience. At the same time, however, beyond the number of likes or followers that an organization earns on different social networking platforms, social media becomes a relevant source of information that can measure its reputation in the market or reveal unexpected insights about the composition of its audience.

Switching to cloud. For some years, the cloud has been a topic of interest in almost all segments of software except Business Intelligence. 2014 was probably the year when the cloud became an important element of BI solutions. The change was driven, on the one hand, by the maturing of the international BI market and, on the other hand, by organizations that realized that the cost of maintaining their own infrastructure is high compared to a cloud solution. Because BI resources are limited, organizations prefer their people to spend time exploring data,

testing innovative ideas, and highlighting business insights instead of being blocked with infrastructure maintenance and software upgrades.

Big Data at user-controllable sizes. The truth is that organizations need help in capitalizing on their huge amounts of data, regardless of the platforms on which they are stored [5]. Most BI solutions focus on processing Big Data rather than distributing them in the format and size appropriate for business users. In the years to come, successful organizations will be able to offer users the technology tools that will make them feel they can control very large amounts of data with the same ease they would handle with lower volumes [6].

3. BIG DATA EXAMPLIFICATIONS

3.1. European response to the Big Data challenge

The fact that Big Data is a concept with revolutionary potential, a so-called technological disrupter for the evolution of human society [5 - 6], is well illustrated by a European initiative that ended in early May 2015.

In essence, this is a new project called Strategic Research and Innovation Agenda (SRIA) to discover the value of the Big Data for Europe concept taken as a whole. The objective of the SRIA is to describe the main challenges and needs related to the implementation of the Big Data concept at European level over the next five to ten years. The project was developed by NESSI, the European Technology Platform for Software and Services, along with partners from the Big Data project within the European FP7 project axis. According to the data presented on the website www.bigdatavalue.eu, the vision of this new project is to establish a public-private partnership in the Big Data area at European level.

The authors of the project have gone from a few fundamental findings. One of these is that by 2020, smart IT & C applications such as smart grids, smart logistics or smart cities will be deployed widely across Europe.

Another finding, based mainly on the predictions of the major market research companies, is that expanding access to broadband Internet, intensifying the use of mobile technologies, and generalizing the use of M2M communications (Machine-to-Machine) will contribute decisively to the explosion of generated data volumes, which by the end of this decade will reach a total of 40.000 Exabyte's.

Intelligent extraction and use of this huge volume of data will generate real revolutions, the project's authors say, in the decision-making process in business, science and society. And this will inevitably lead to the emergence of new industries, jobs and services [5].

The main areas of impact in terms of using Big Data at European level are considered to be public health, transport and logistics, energy, public sector, the media and media content, production of goods, trade and environment.

All of the above are elements that lead the project's authors to conclude that Big Data will become, as a concept, one of the main engines of European economic growth [7].

Europe must aim high and to provide a future in which Europe will be taking the lead in exploiting Big Data and application of this concept in some of the most critical sectors for the entire continent of socio-economic prospects.

The value extracted from the use of Big Data is a crucial tool to overcome foreseeable obstacles in the future related to resource congestion, more efficient use of energy, but also to support a better public health system across Europe. It remains to be seen, within a not too distant horizon, which will be Romania's response to the challenge launched for the entire continent we belong to.

3.2. Big Data applied in agriculture

As agriculture and the agricultural industry is one of the strategic engines for the functioning of the Romanian economy in the future, a case study on the use of Big Data in an American agricultural area could awaken interest in such solutions to make production more efficient in Romania [7].

Specifically, this is a project currently running in Flint Valley River, an eminently agricultural land in Southwestern Georgia. At the end of this project, farmers in this region will be able to get highly accurate weather forecasts three days in advance, each farmer for the farm he cares for.

Currently, a team of researchers from the Flint River Soil and Water Conservation District, the American Department of Agriculture, University of Georgia and IBM, uses sophisticated Big Data tools to analyze large volumes of meteorological, geographic, historical and other data. The aim is modeling weather events with high precision and location higher than can be achieved today.

Such weather forecasts will help farmers make more effective decisions about irrigation, sowing, harvesting and soil fertilization, which will enable them to better conserve the water used and to increase in the long run the volumes of harvested crops.

Flint River Valley is an important part of Georgia's agricultural industry. Farms in this 27-county region contribute about \$ 2 billion annually to United States revenue.

About ten years ago, the water conservation authority of this region signed a partnership with the American Department of Agriculture, the University of Georgia and other local, regional and state agencies to promote water conservation practices among farmers the whole area.

The effort has since begun to bring important results. A so-called VRI (*Variable Rate Irrigation*) technology developed by University of Georgia researchers has quickly moved from concept to commercial.

Based on GPS, this technology allows farmers to arrange sprinklers for irrigation in the soil so that they do not spread water over areas that do not need it, and direct their jets to areas that need irrigation.

The project on using Big Data for more efficient weather forecasting is the second largest project between the aforementioned entities and does not have a clear end date.

But it is certain that at the end of the project, farmers will be able to benefit from such improved forecasts by using smart tablets or phones to increase their mobility at the same time.

4. CONCLUSIONS

In a steadily evolving digital world with increasing volumes of generated data, only those companies that will rely on information will be able to increase their competitiveness. Business performance will depend on the organization's ability to access and exploit the right information [7].

Moving from inflexible solutions that emit static reports to user-oriented solutions will accelerate as more organizations recognize the importance of obtaining data that each user can base their decisions. Therefore, predictive analysis solutions are adapted to the way people work, giving them access to information wherever they are at a given time. For both large and small businesses, business speed means that business intelligence solutions are commonly used, not just occasionally [8]. Business users want to be able to access information in the context of their day's natural flow, not just when they're physically in office. PC tablet devices, for example, are ideal for mobile applications of predictive analytics support and encourage collaborative decision-making process.

Those organizations that will be able to understand and filter the relevant information, who will know how to discover patterns and act based on the results thus obtained, will become top performing businesses [7].

No project is really completed until you can measure your results by comparing them with the goals we have set to see how effective it was. Perhaps that is why one of the major benefits analytical solutions provide to companies that regularly use them refers to sustained performance [8].

By analyzing the data we get from daily activity, regardless of the type of industry in which it is present, a company can constantly take those business decisions to ensure its success on the market. Otherwise, if the same company chooses to ignore the data and does not make efforts to properly analyze it, its business potential drops significantly. The question then arises: what do we do to fully harness the potential of data analysis solutions to ensure the best decision support?

The article tries to answer, at least partially, this question.

5. REFERENCES

- [1] Ștefan Iovan, (2017), *Predictive Analytics for Transportation Industry*, București: Editura Universitară, *Journal of Information Systems & Operations Management*, (**JISOM**), Vol. **11**, No. **1**, pag. 58 - 71;
- [2] Ștefan Iovan, Cristian Ivănuș, (2013), *Management for Adaptable and Agile Organizations in Crisis Situations*, Railway PRO Science & Technology, *Club Feroviar Conferences “Concepts and technologies for sustainable transport. A greener railway passenger and freight transport”*, Sibiu, Romania, pag. 32 – 43;
- [3] Ștefan Iovan, (2013), *The Importance and the Definition of e-Skills for Europe*, Iași: Editura PIM, Proceedings of the International Conference: “*Transforming the educational relationship: intergenerational and family learning for the lifelong learning society*”, România, pag. 258 – 269;
- [4] Ștefan Iovan, Cristian Ivănuș, (2014), *Business Intelligence and the Transition to Business Analytics*, Târgu Jiu: “Academica Brâncuși” Publisher, *Annals of the “Constantin Brâncuși” University, Engineering Series*, Issue **4/2014**, (**CONFERENG 2014**), pag. 150 - 156;
- [5] Mihai Dinu, Ștefan Iovan, (2014), *Harnessing Big Data Volumes*, Proc. of 7th Symposium “*Durability and Reliability of Mechanical Systems*”, (**SYMECH 2014**), Polovragi – Gorj, pag. 250 - 256;
- [6] Ștefan Iovan, (2015), *Big Data Security Problems*, Proc. of 16th European Conference (**E_COMM_LINE 2015**), București, România, ISSN: 2392-7240;
- [7] Ștefan Iovan, Cristian Ivănuș, (2016), *Modeling of Management Processes in an Organization*, Târgu - Jiu: “Academica Brâncuși” Publisher, *Annals of the “Constantin Brancuși” University, Fiability & Durability Series*, Issue: Supplement **1/2016**, (**SYMECH 2016**), pag. 213 - 219;
- [8] Ramona Marge, Ștefan Iovan, (2018), *The impact of economic company data theft*, Târgu Jiu: “Academica Brâncuși” Publisher,

*Annals of the “Constantin Brâncuși”
University, Fiability & Durability Series,
Issue: 1(21)/2018, (SYMECH 2018), pag. 275
- 280;*