



PROCESS-ORIENTED BUSINESS ARCHITECTURE FOR A CONSCIOUS DECISION MAKING BASED ON USERS' INTERESTS

¹RHIZLANE SELTANI, ²NOURA AKNIN, ³SOUAD AMJAD, ⁴KAMAL EDDINE EL KADIRI

^{1,2,3} Information Technology and Modeling Systems Research Unit (TIMS), Faculty of Science, Abdelmalek Essaadi University, 93030, Tetuan, Morocco

^{1,2,3,4} Operational Research and Applied Statistics Laboratory (LIROSA), Faculty of Science, Abdelmalek Essaadi University, 93030, Tetuan, Morocco

E-mail: ¹sel.rhizlane@gmail.com, ²aknin@ieee.org, ³amjad_souad@uae.ma, ⁴elkadiri@uae.ma

ABSTRACT

Since the appearance of the web 2.0, several new concepts have emerged like social networks which generate big data, characterized by a difficult treatment using traditional administration tools. Though, they represent a rich resource of information that we can use as the basis for decisions to help managers which are always complaining about the time taken to get answers to questions in order to make decisions, and also about the quality of these decisions. In the decision-making task, the most important besides treating the data, especially big data, is to analyze and interpret the results of the treatment in a way to maximize the profit in terms of business management logic. So, we propose in this paper a process-oriented business architecture, founded on users' interests, by analyzing tweets as an example of big data, for decision-making purposes, dealing with the practical functions of text mining. The business orientation of the analysis helps us to get conscious decisions which are more significant and closely adequate to the needs of the enterprise.

Keywords: *Decision Making, Big Data, Data Analysis, Process-oriented, Correlation*

1. INTRODUCTION

Within an enterprise, the decision making constitutes a very crucial task which affects indirectly the overall performance of the company. Seeking for the best ways to implement and operate this process is a necessary mission. On the other hand, a massive mass of a variety of data is produced each day due to the appearance of a large range of applications that accompanied the emergence of the web 2.0 and which can constitute an important additional basis to supply a decision support system. This wealth of information represents a precious resource that we want to exploit in order to support decision-making processes, to generate more meaningful and relevant decisions and to resolve the problem that the most decision support systems are introspective, which leads to a very limited strategic scope. A decision support system which is based on additional and more significant data produces richer analysis and hence, leads to better decisions.

Knowing what the others think is very important for a decision-making system; it is an axis of

analysis which constitutes a crucial importance to assess the performance of a structure. However, the important question is how to deal with the results of these data analysis. The way we treat these results constitutes an essential factor in the success or the failure of the decision-making process. Consequently, it affects severely the quality and the relevance of decisions made by the organization. In this paper, we propose a method which aims to interpret the results of the treatment of tweets using the usual functions of text mining. The architecture we propose is process-oriented according to the structure and the functions of the enterprise to ensure more conformity between needs and analysis, and as a result, provide decisions which are more effective and more suitable to the specific needs of each enterprise department.

2. MATERIAL AND METHODS

2.1 Big Data

Big data refer to data sets which are exchanged by linked objects in the web, and whose the volume, the variety, and the velocity are increased [1]. Consequently, the processing of these data



constitutes a hard mission via traditional tools [2]. Despite the problems the big data raise on different levels, their analysis is still very useful in various areas. Theirs principal features are data volume, data velocity, data variety, and data veracity. These elements can be considered as the characterizing pillars of big data that make their processing and their analysis, a particular challenge.

The characteristics of big data, especially the variety and the veracity, can impede the processing operation, due in particular to the problem of the semi-structured or unstructured nature of big data, and also due to the data complexity which is reflected by the interconnectivity related to the diverse relationships between the different data structures. Consequently, a small modification in these structures can affect the entire meaning of information we can retrieve from them. Therefore, big data processing constitutes one of the most challenging missions which continue to raise more and more of research questions.

Since the big data is a set of a large amount of various data coming from different sources, it is difficult to control the quality of all of them. In big data, the quality can change significantly, due to the veracity characteristic, which can affect the accuracy of the entire analysis.

In addition to all that, many other challenges persist about big data and require elaborating actions related to several and various plans to overcome them [3]. But, a well-developed processing stage can facilitate the matter, to get useful data quickly, efficiency, and in a clear and practical way. All the different applications and channels of the web 2.0 participate significantly in the expansion of the big data. However, social networks are considered as the primary generator of this type of data. Facebook and Twitter are two main examples of successful and widespread social networks which are the most used around the world.

2.2 Text Mining for Enterprise Business Purposes

What people think, constitutes a very important piece of information [4]. So, for a decision-making process, it will be very interesting to get this type of information to enrich decisions. In our context, our approach for the big data analysis is to perform a business-structured analysis, then a process-oriented interpretation to the listening through the channels of the web, here Twitter, to enhance the decision-making task.

To support our treatment process, we use statistical methods through text mining which is considered as one of the most essential and practical strategies for dealing with unstructured data.

Text mining is a specialty of data mining which refers to a set of techniques and methods which aim to exploit documents containing unstructured data in the form of a text, in the different types of files: .txt, .word ... and which often are human writing coded. Therefore, we can extract the knowledge from big data. The operation of text mining is based on the linguistic analysis techniques and can be applied in different fields of application, it is a domain which is very practical and used in the business intelligence field [5].

The text mining is considered as one of the most important and convenient strategies for dealing with unstructured data. The text mining can be defined by the different ways depending on the research domain it is related to, the Information Extraction, the Text Data Mining, or the KDD Process [6].

Usually, in the case of treatment of large amounts of text, text mining is very functional and effective regarding the use of statistical and descriptive practical techniques [7] and which we can deploy and adapt to our case to support our treatment process.

2.3 Decision Making in the Business Context

The enterprise is a for-profit organism that conducts one or more activities in order to generate incomes, contributing to the development of a specific area. The structure and the hierarchy of the company may differ depending on the type and the sector of activity in which it operates. However, most of the companies opt for a typical structure based on a departmental distribution which can change in terms of the number and the type of departments according to the specifications of the company:

- Accounting
- Administration
- Human resources
- Sales
- Production
- Purchasing
- IT
- Marketing



- Research and Development
- Legal Department
- Commercial Department
- Stock

To ensure the good functioning of all these activities, managers are led to take decisions every day. Decision making refers to the process founded on knowledge which aims to make a logical choice or act. The importance of this process becomes more crucial in the managerial context [8]. According to DR. Anderson & al., "Get the decisions right and the organization continues to succeed. Get the decisions wrong and the organization may fail and disappear" [9]. In the business context, decision making remains the pillar of the enterprise's development. It provides the road map of the marketing strategy and the different actions, to maintain and manage the competitiveness of the company, to improve the loyalty of the customers, and to optimize the processes and the costs. Decision making is a process which can contain different steps and varies depending on the treated situation, so we can find many approaches which deal with this subject. Recently, different studies and works are continuously focusing on this problematic, dealing with the strategic and operational levels of the enterprise, by treating the elements of enterprise mathematical models to provide a semantic representation, based on the use of an ontology which is manipulated to capture the significance and the meaning of the mathematical language and which can relate the mathematical elements to the semantic representation of the company [10]. Many techniques of management science can be used in an effective way in the field of decision-making such as Network Models, Inventory Models, Queuing Models ... [9]. These works provided some very interesting results; quantitative techniques are very useful to treat the data. Though, an additional layer of analysis, based on the enterprise's specifications, and according to a process orientation analysis, can be very useful to provide decisions which are very sharp, consistent and relevant, especially if we generate a process with a reduced time lapse.

In recent years, all areas were revolutionized after the appearance of the web 2.0, giving organizations original methods to exploit data in order to increase the overall performance of the enterprise. An original way appears: take data from back-end business applications and combine them with unstructured sites data across the web

information, to enable a simple and easy access to essential information, an absolute flexibility of configuration tools and freedom of trade and communication [11], [12]. With large flows of data available across all social networks and the web in general, it is important to isolate the right channel, destroy unnecessary noise and give meaning to the social information.

One of the most innovative concepts that the web 2.0 brings is the ability to switch from a simple spectator to a creator of the web content. So, we are based on one of the most innovative aspects of web 2.0, which is the collaboration to exploit the collective intelligence reflected by the various cooperative applications, such as blogs, social networks, ... and which is used in several disciplines; we are interested in the business area, where it refers to diverse communities of people interacting to produce clever outcomes [13].

Web 2.0 is a very significant resource of information. Users have the possibility to interact continuously with the web content through the collaborative applications, such as blogs, social networks ... which causes the increase in the number of actors on the web, and consequently, the rate of information which circulates on its channels increases as well. This massive data flow generates the phenomenon of the big data. The user is primarily a passive actor, but can become in any moment an active actor, by transmitting opinions which help to ensure the mission of decision making. These views can take, for example, the form of a solution to a particular problem. The collaborative aspect of the decision-making process contributes to solve many business problems efficiently and quickly.

The reviews, including opinions of experts about an issue, may be of great use to make decisions to solve a given problem. Comments can also take the form of feedback to a given subject. Any feedback contains in itself a notice which can be used to extract useful information which enriches the process of the decision making or the form of a proposal for improvement; in any field, application, or system, we always look for ways of improvement, especially in the business case.

Opinions of clients and in particular those which are the most affected by the service proposed by the enterprise, constitute a vital resource of inspiration to make the right decision of improvement. Also, we have the proposals for improvement and complaints about a process, a product, or a service which can also lead to the generation of significant

decisions. So, it will be interesting to search for meaningful data to discover the different rates of associations relating these features to other ones. A framework is elaborated to present the general process starting with the creation of the data on the web and ending with the decision-making task (Figure 1).

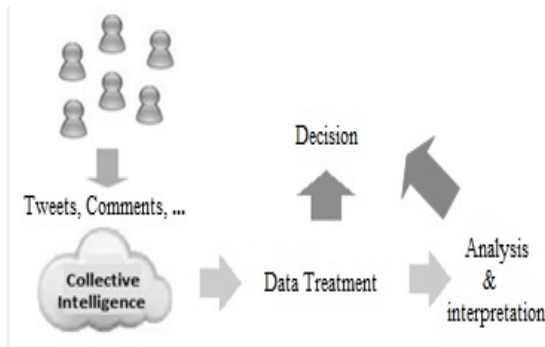


Figure 1 : The Process of the Generation of the Decision 2.0 based on the Big Data Analysis

3. PROPOSED APPROACH

3.1 Adapted Enterprise Business Structure

Taking into consideration our aim of decision making, we adapt the typical structure of the enterprise to get the distribution shown in Figure 2.



Figure 2 : Adapted Enterprise Business Structure

- Production Service

The production department is responsible for everything related to the production process within the enterprise. Usually, this service is the generator of products marketed by the company. It may include production units, factories ... as an example; the production line in a company from the textile industry.

- Marketing & CRM Service

The development of the marketing strategy related to the company, the definition of products, and the treatment of

issues that can negatively affect the company's image, managing contacts with the clients taking into account their expectations and their complaints, are some vital tasks of this department. The same service is responsible for all the tasks related to advertisements, seeking the promotion the company's products and therefore increasing the sales rate.

- Research and Development Service

This service constitutes one of the pillars which contribute to the expansion of the enterprise, exploiting research by doing studies and experiences in order to ensure the development of the company's activity.

- IT Service

The IT service is becoming increasingly essential in all types of business. It uses IT tools to support the company's business. The activity of this department can vary depending on the needs: computer modeling, software development, computer maintenance, networks ...

- Commercial & Sales Service

Commercial tasks comprise prospecting new clients, elaborating offers, taking the necessary actions to meet the defined objectives, evaluating, controlling, and managing sales ...

- Human Resources Service

The principal mission of the human resources service is to recruit new people. But, it is also responsible for the management of the company staff: organize training; manage holidays and promotions of the employees.

3.2 Process-Oriented Architecture for Decision Making based on the Clustering Principle

The architecture is elaborated following the company's structure previously defined, the principle of the clustering is adopted to categorize and split the results into six main preliminary areas, in a way that the elements which belong to the same area have much more similarities than with those belonging to other zones. These areas represent the business processes related to the different services. The pillars of the structure of the company constitute the axes of action plans to do by the company. Secondary zones can be set

according to vertical or horizontal directions in the case of mutual interests between the services.

In the business context, a usual marketing tradition aims to inspect the products which are the most sold together to put them side by side in the store in order to improve sales rates. Always, it is about studying correlations. The correlation which exists between two or more random variables refers to how much the association that exists between these variables is strong. The quantitative measure which corresponds to this type of statistical relationship is named a correlation coefficient (CC).

For the company, a great need emerges, is about to detect what are the most associated elements to their products when they are discussed by the customers or the web community. We analyze correlation coefficients, calculated using usual functions of text mining, to know how strongly our studied feature (product, brand, company's name...) is associated with other ones.

The results of the data treatment undergo a process-oriented interpretation, we use the correlation coefficient as an indicator to define the elements to which the studied feature is linked and subsequently, we can determine the degree of strength of this connection. After that, each of the correlation coefficients determines the rate in terms of time and energy to devote to the corresponding characteristic in the area. The general model is represented in Figure 3.

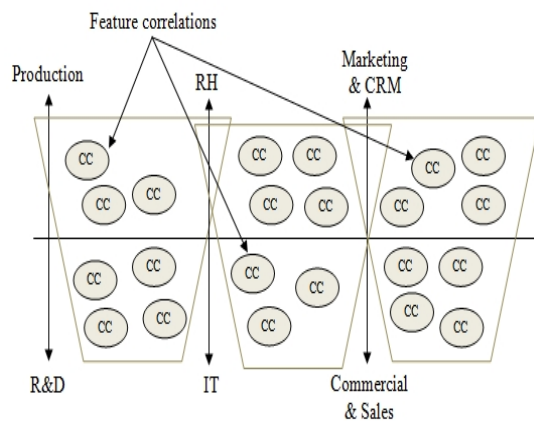


Figure 3 : Axes of Measures and Action Plan with Correlation Coefficients

Each area generates a set of decisions to make by the enterprise according to the different characteristics. These decisions are reflected by the answers to the following questions:

- What are the actions to take?

- What are the issues to solve?
- What are the challenges to overcome?
- What must be improved? Developed?

Here some examples of actions to do with reference to decisions made, in a different way every time, according to the obtained results:

On the Human Resources Plan:

- The intervention of experts: experts can be intervened in order to assist the enterprise staff to resolve a problematic or to improve a product.
- Organization of seminars: can be very useful to create a sharing environment with external and potential collaborators.
- Ongoing training and external courses: to improve the level of knowledge or the professional skills of the company's staff regarding a domain, a subject or a specific characteristic.

On the Marketing / CRM Plan:

- Common advertising campaigns: if the feature studied (the name of the company, one of its products ...) is connected with a name of another company, one of its products..., common advertising campaigns can be organized to increase the sales revenues of the two companies.
- Actions to involve clients in the process must be taken, by presenting the current issues to them to get their opinions about the solution or their propositions of improvement of a specific aspect.

On the Production Plan:

- Some production collaborations can be established when the correlation coefficient proves the existence of connection with other products or companies.

Once the different characteristics are distributed according to the corresponding areas, we classify them in a way to establish a treatment priority plan at each area, based on the correlation coefficients. These priorities determine how the treatment of the associated problematic is crucial or urgent relative to other problems in the same area.

The data treatment is realized using common text mining functions. Once we retrieve the useful information from the raw data by recuperating correlations, we implement our architecture, by splitting the corresponding features into process sectors according to the clustering principle, grouping features correlations which have elements generating decision-making processes which follow the same business reasoning and process purpose. Action plans are established providing rates in terms of time and energy to devote to the different issues or treatment to make decisions. Also, coordination and collaboration zones between the different services of the enterprise can be defined. A conscious decision-making system is elaborated focusing on the exploitation of the collective intelligence and giving a business interpretation to it, in order to get decisions which are more significant and relevant.

4. RESULTS AND DISCUSSION

Twitter is a microblogging tool which counts today more than hundreds of millions of users. It constitutes a main pillar of the social networks space, and it is characterized by several features like the shortness of tweets, so messages are limited to 140 characters, which reduces the time of the creation and provides a high reactivity [14], this is why it constitutes an ideal candidate for our study.

There are many datasets which deal with social networks content, especially Twitter. One of the datasets the most used in the literature, STS-Gold, a

dataset containing tweets, human annotated and representative of the real-world [15]. Given that our finality is the decision making, we were inspired from it, to elaborate our dataset of tweets, without annotation, for our final test and which fits more the decision-making context. First of all and to be able to realize our text analysis correctly; our dataset must undergo some treatments according to our context and our purpose. A preprocessing stage is crucial; the text must undergo several textual transformations which aim to make the text more suitable for further treatments and to remove objects which are not relevant for our analysis. This process contains some usual techniques of transformation and adaptation to prepare the textual content of the big data source, here Twitter, and which are necessary to be able to treat data. The dataset, containing thousands of tweets, is properly managed to extract the appropriate information for our purpose.

In particular, our studied product is a much known brand of phones which we note 'X Phone'. To be able to make our study, we first calculate the correlations between it and other characteristics in the data set, to reveal terms to which, it is the most associated. We opt for a correlation limit of 0.1 because we need to fly over the maximum possible of associations so we can get a better view of features connected to X Phone. Results are represented in Figure 4.

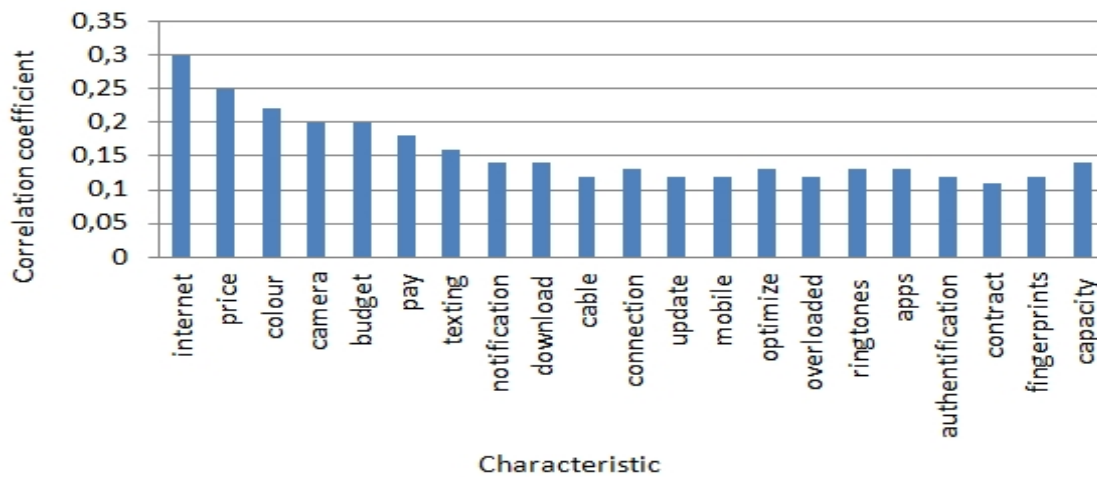


Figure 4 : Correlation Coefficients of 'X Phone' with a Threshold of 0.1

After determining the various links and their associated correlation coefficients, each characteristic is assigned to its corresponding area. Many specified zones can be created in different directions, as required, such as coordination zones and collaboration ones.

In principle, these zones determine areas which combine decision features belonging to more than two zones and which can require the intervention of other services. The corresponding decision plan is defined by the process-oriented business architecture represented in Figure 5.

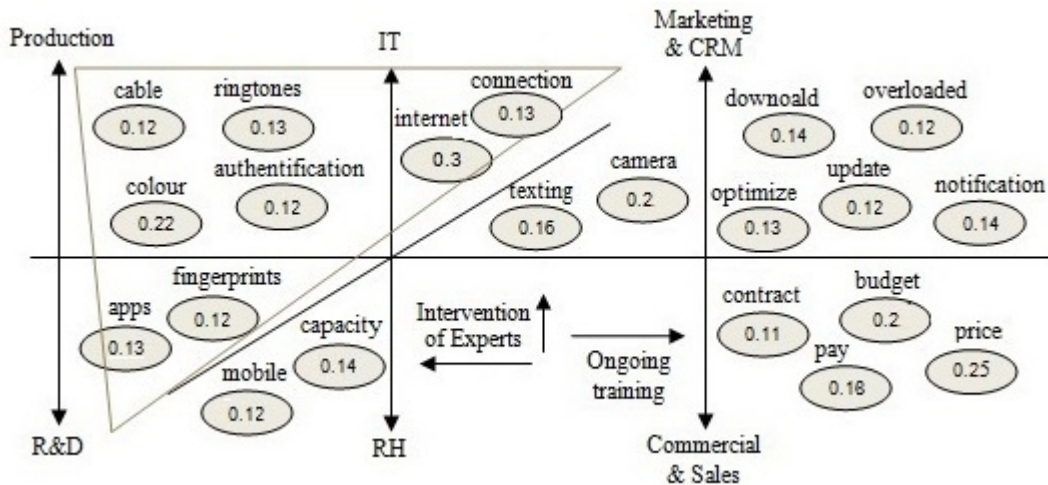


Figure 5 : Process-oriented Business Architecture for Decision Making regarding 'X Phone'

Decision making is a process that requires time and energy to ensure its good functioning. Energy means how much effort is required to ensure a task. It refers as well to how many people belonging to the service will be in charge of the mission, which leads to a better management of the resources of the company, so we can conserve more time and money.

Of course, enterprise's services have sometimes needs to work together to make a decision. In our architecture, potential collaborations between employees from different services can be established if needed; in this case, the distribution is adapted to the new area created.

In order to organize the distribution according to the two pillars of time and energy, the correlation coefficients are used as an indicator of the need in terms of time and energy required to effectively identify issues and solve problems as part of a conscious and efficient decision making.

Some examples of adequacy representations are shown in Figure 6.

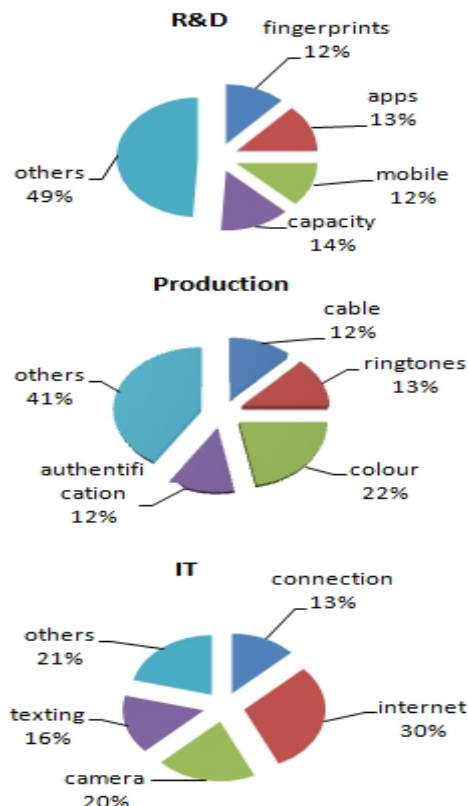


Figure 6 : Time / Energy Distribution for the Decision-making Process



Based on these representations, managers are assisted in their decision-making process. Decisions reflecting the subjects which interest more the clients and the web community are provided. Implementing the distribution according to the business architecture provides more flexibility, an interesting saving of time, and more precision regarding repartition of tasks dealing with making decisions, according to specific coefficients determining how much these tasks need in terms of time and energy, which ensures more organization and efficiency.

5. CONCLUSION

Making good decisions within a business is a decisive factor in the development of the organization. With the growth of collaborative tools on the web, an important collective intelligence circulates on social networks channels, constituting a flow of big data. Using these data in the best ways is of great benefit to the company.

The main research objective behind this work is to find the adequate ways to improve the decision-making process in the enterprise to get more relevant and meaningful decisions. We deal with an important concept which is big data, in order to exploit the wealth of information in terms of volume and variety, and we mainly focus on taking consideration of what people think about the company or its products, to support the system with decisions which are very significant and conscious. So, we proposed a method of decision making in the enterprise context, to exploit the added value contained in the opinions of web users, in order to facilitate the process of decision making, providing more significant decisions through a process-oriented architecture generator of decision plans which fit the needs and the specificities of the company. Consequently, we allow to managers to take the appropriate actions, with the intention of improving the overall performance of their enterprise.

REFERENCES:

- [1] Zikopoulos P, Eaton C. *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill Osborne Media. 2011.
- [2] Letouzé E. Big Data for Development: Challenges & Opportunities. Global Pulse is a United Nations initiative. 2012;(May):47.
- [3] Katal A, Wazid M, Goudar RH. Big data: Issues, challenges, tools and Good practices. In 6th International Conference on Contemporary Computing, IC3 2013. 2013. p. 404–409.
- [4] Pang B, Lee L. Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*. 2008;2(1–2):1–135.
- [5] Gupta P, Narang B. Role of text mining in business intelligence. *Gyan Jyoti E-Journal*. 2012;1(2).
- [6] Hotho A, Nürnberger A, Paaß G. A brief survey of text mining. In *Ldv Forum (Vol. 20, No. 1)*. 2005. p. 19-62.
- [7] Meyer D, Hornik K, Feinerer I. Text mining infrastructure in R. *Journal of statistical software*. 2008;25(5):1–54.
- [8] Bazerman M. H, Moore D.A. *Judgment in managerial decision making*. John Wiley & Sons. 2008.
- [9] Anderson DR, Sweeney DJ, Williams TA, Camm JD, Cochran JJ. *An introduction to management science: quantitative approaches to decision making*. Cengage learning. 2015.
- [10] Muñoz E, Capón García E, Lainez JM, Espuña A, Puigjaner L. Mathematical knowledge management for enterprise decision making. In *Computer Aided Design: Proceedings of the 23rd European Symposium on Computer Aided Chemical Engineering, Lappeeranta Finland*. 2012, p. 637–642.
- [11] Nelson GS. Business Intelligence 2.0: Are we there yet. In *SAS Global Forum*. 2010, p. 50-62.
- [12] Raden N. Business intelligence 2.0: Simpler, more accessible, inevitable. *Intelligent Enterprise*. 2007;1.
- [13] Gill Z. User-driven collaborative intelligence: social networks as crowdsourcing ecosystems. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*. 2012. p. 161–170.
- [14] Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS one*. 2011;6(12):e26752.
- [15] Saif H, Fernandez M, He Y, Alani H. Evaluation datasets for Twitter sentiment analysis a survey and a new dataset, the STS-Gold. In *ESSEM 2013*. P. 9–21.