International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-9 Issue-5, January 2021

From Prediction of the Improvement of the Quality towards an Equitable Sharing of the Cost of the Improvement between Business Processes

Jaouad Maqboul, Bouchaib Bounabat

Abstract*: In this work we have developed a quality approach for the quality assessment of data related to the business process for quality projects, this approach uses the cost of the implementation of quality combined with the impact of quality broken down into the benefit and efficiency of data, shapley value helps us choose the business processes that will collaborate to reduce the cost of improvement, Deep learning helps us calculate the quality values for any dimension based on history of previous improvements. To reach our goal, we used the cost-benefit approach (ACB) and the cost-effective approach (ACE) to extract the impact and cost factors then using a multi-optimization algorithm. -objective we will minimize the cost and maximize the impact for each business process and the deep learning introduced will complement our approach to learn from the previous improvements after validation of the processes which will be chosen as well as the values calculated after improvement. The importance of this research lies in the use of impact factors and the cost of the quality evaluation which represent the basis of any improvement, our approach uses generic multi-objective optimization algorithms which will help choose the minimum value of each business process before the improvement, adding a layer of predicting and estimating the quality value of the data generated by the business process before the improvement even, while the value of shapley has aim to minimize the cost of quality projects during fission and merger of companies and even within a company composed of several services and departments to have the lowest possible total cost to help companies manage the portfolios of quality..
Keywords: Artificial neural network, data quality assessment, data quality improvement, deep learning, prediction of improvement in data completeness shapley value.

I. INTRODUCTION

The real capital of companies today is data, information, knowledge bases without forgetting budgetary capital, this capital exhausted by the poor quality of data of which, according to IBM, the cost of poor quality is valued at over three trillion dollars for the US economy, and over 600 billion dollars is wasted on poorly targeted mailings per year. [1], big data produced during the existence of companies represents a heritage that represents the company and its value in the

market, more than 90% of the data produced So far, comes from the last two years [2], the Big data analytics was valued at 37.34 billion US dollars in 2018, in 2027 it will reach 105.08 billion US dollars, in the market 58% of companies adopt Big Data Analytics, in 2019 the revenue is more than 187 billion US dollars [30]; according to the characteristics of Big Data which are (Variety, Variability, Complexity, Value) [31], a large number of erroneous, incomplete, inaccurate data will be added to the existing one, so it is the duty of companies to ensure and correct these imperfections, because the cost of poor quality is enormous and could damage the company's image with partners; to this end that scientists, identify quality as an uplifting and strategic element that will create added value, and satisfy humanity and these financial objectives [3]. The organization of this document is treated as follows: section II deals with the comparative table of research carried out on the impact and cost of the quality of data and processes, section III begins with quality, these dimensions and classification of these dimensions to the dimension chosen for our use case; Section IV our work on the positive impact and cost of improving completeness; Section V describes the use of the artificial neural network to predict whether there will be an improvement in completeness, Section VI is about contribution of the value of shapley for an equitable sharing between the business processes, the conclusions and future work is summarized.

II. RELATED WORK

Manuscript received on January 21, 2021.
Revised Manuscript received on January 27, 2021.
Manuscript published on January 30, 2021.
* Correspondence Author
Jaouad Maqboul, Head of Software Engineering Department, Avenue Mohamed Ben Abdellah Regragui, Rabat, Morocco
Bouchaib Bounabat, Head of Software Engineering Department, Avenue Mohamed Ben Abdellah Regragui, Rabat, Morocco

Table- I: similar work for quality assessment

works	Data Quality for the Informati on Age (Redman, 1996)	Enterprise Knowledge Manageme nt: The Data Quality Approach (Loshin, 2001)	Towards Quantifyi ng Data Quality Costs (Kim & Choi, 2003)	A Classificati on and Analysis of Data Quality Costs (Eppler & Helfert, 2004)	Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Informati on (TM) (McGilvra y, 2008)	Informati on Quality Applied (English, 2009)	DAMA Book of Knowled ge (DAMA, 2009)	mprovin g Data Warehou se and Business Informati on Quality: Methods for Reducing Costs and Increasin g Profits (English, 2011	The Costs of Poor Data Quali ty (Hau g, 2011)	Understand ing the Financial Value of Data Quality Improveme nt (Knowledge Integrity, 2011)	Measuri ng the Business Value of Data Quality (Gartner , 2011)	How to Monetize , Manage, and Measure Informati on as an Asset for Competiti ve X. (Laney, 2017)
data	√	√	√	√	√	√	√	√	√	√	√	√
process								√				
cost of poor data	√	√	√	√	√	√	√	√	√	√		
cost of improvement			√	√								
finalcial impact					√					√	√	√
non finanacial impact												
year	1996	2001	2003	2004	2008	2009	2009	2011	2011	2001	2011	2017

The comparison criteria namely: the costs of poor quality, the impact of quality, the cost of quality either for processes or data.

According to the literature in the impact assessment and the cost of quality, the analysis of the effectiveness of a project to improve quality can be broken down as follows:

- The negative value associated with the data, ie the cost of poor quality.
- The financial or business value of improving quality. The costs of improving quality.

However, these works do not offer metrics that allow us to analyze the cost and the tangible and intangible impact, the work done in this direction was mainly based on the economic side, which complicates their applications in any field. or organization.

III. DATA QUALITY DIMENSIONS CLASSIFICATION, DATA COMPLETENESS CHOSEN DIMENSSION

A. Data Quality

Several researchers have proposed definitions of data quality, among these definitions, quality is the degree to which a set of inherent characteristics meets the requirements [4]. According to Philip Crosby, quality is linked to the customer and to the conformity of his requirements [5]. whereas W. Edwards Deming sees quality as a predictable degree of reliability and consistency, suitable for the low-cost market [6]. For Joseph Juran quality is fitness for use [7]. according to article [8], quality is "the manner in which information meets customer expectations and requirements [9, 10].

B. Classification of Quality Dimension

When we talk about data quality, we have to describe the facets of it, and it is the dimension s that play this role, there are three approaches for classification:

- semantic-oriented approach: it is based only on the meaning of the criteria; it is an approach where the criteria are separated from any information framework.
- treatment-oriented approach: it classifies the criteria according to their deployment in the different phases of information processing.
- objective-oriented approach: it is characterized by a ranking of the criteria according to the defined objectives.

The first approach is the most intuitive where the criteria are separated from any information framework, the researchers made a classification from three to four categories of data quality by Wang and Strong: intrinsic, contextual, representational and accessibility, in this study we choose completeness from contextual.

- Intrinsic: indicates that the data is related to the quality that the data has by itself [11]. This aspect of quality is separated of the context and the user's perspective.
- Contextual: Emphasizes the need to view data quality in the context of the task at hand.
- Representational: includes aspects related to the meaning of the data and the format of the data [35], so that they are easy to understand and interpret.

Accessibility: highlights the important role of systems; it is important that the system must be accessible and secure as well the data.

C. Data Completeness

Critical dimensions that are identified by researchers such as accuracy, consistency, completeness, timeliness and updating, this study will examine completeness, considered a fundamental dimension [10], the lack of data can block the entire business process as well as other critical dimensions.

The most widely used definitions of completeness is the extent to which data is not lacking and is of sufficient breadth and depth for the task at hand [11], this means all necessary information available.

The next chapter we will talk about the impact of completeness and cost of its improvement; since it is difficult to quantify the cost, we will take the complexity in its place; evidence proves that the cost of improvement has a linear relationship with complexity [12, 13], since business processes are well defined and will not change often over time, our survey will focus on the data produced and manipulated by these processes.

IV. THE COMPLEXITY, THE IMPACT OF IMPROVING COMPLETENES

A. Impact of data completeness

- Will reduce the cost?

It costs to have good data, it is time and money in it, the cost is divided as cost of low data, cost of prevention, cost of detection, cost of repair and cost of maintenance and improvement; Figure 1 shows that the first three classes of cost are quantifiable, on the other hand the cost of improvement is difficult to calculate [14].

- Will it increase efficiency and profitability?

According to the General Accounting Office 1991 a report estimates that total quality management practices increase the profitability of the company and the efficiency of its employees [15]. A working environment; where the emphasis is on teamwork and where quality is the objective of stakeholder support; the work will be performant and efficient than that of which the quality comes after reclamation of the incident.

- Will increase end customer satisfaction?

Any business must continually improve its services to its customers, be competitive in the market [16]. The quality of the data serves customers and employees better without investing in marketing which costs money, the other stolen, it is the independence of the company vis-à-vis these employees to serve them well. clients on time even if there are departures from these skills.

- Will improve Cross-Functional Process?

The transverse processes are the real value of the company, they cover several applications, IT systems and organizational areas. The transversal processes are divided into between departments, the improvement of the data managed by these processes will improve the whole work is a big impact about the company.

- Will facilitates decision making and strategy change?

Nowadays, the majority of companies rely on data to make decisions about their business, and have a head start to adjust their strategy for the short term or long term and gain an advantage over these existing competitors. quality information gives a strategic advantage, because this data which is precise, complete and up-to-date gives a head start in the market, make it easy for analysts provide good information to decisionmakers to make good decisions and even make changes to their strategy in the short or long-term.

- Will this have a direct impact in time?

It depends on the context of the company, its position in the market, its image with customers and partners, the completeness or not of the data can lead to callosal facts in the short term such as Facebook, Twitter, Google or long term for small or medium enterprises.

- Will increase security compliance?

Speaking of security, we have to talk about three types of security, hardware security, software security, data security, last stolen is very important in the context of data completeness, we could not have a word of empty pass, otherwise it is a disaster for confidential data and even for everyday life in the company. If the quality of these data is low, the confidentiality is exposed to risks, but the contrary will help to reinforce security, generate savings by refining processes, reduce compliance time by streamlining the reporting process and Reduce the response time.

The complexity of improving data completeness

As mentioned before the cost of quality divides into cost of prevention, detection, correction, improvement and others, we will try to quantify the cost of improving completeness:

- Find data sources to contradict or supplement the data?

There are several ways to complete the data, with an older backup it can prove useful after a failed migration or a system crash even see a cybersecurity attack, knowledge bases are an asset capitalizing on years of experience of the he company, we find there algorithms, instructions for the collaborators, in order to provide a faster service and to satisfy its customers; the fact that experts contribute to this knowledge base will improve archived, documented, shared knowledge and correct other knowledge.

- The nature and type of data processed?

Most businesses rely on huge amounts of data about their vendors, employees, process data, videos, audio, PDFs, knowledge and more; business processes deal with several types of data ranging from golden data to unstructured data, of course the complexity increases along this curve, the data included in tables like employees, customers, products or suppliers are key business entities that support transactions in operational systems [17], therefore unstructured data is expensive, in addition it does not contain information about the data model and is not organized; because these data entered manually may be incomplete due to human intervention, or by machine directly or the processing follows strict rules, and the case is true for processing a faulty algorithm from a knowledge base.

- How the process is executed?

Business Process Automation (BPA) improves organizational efficiency by reducing effort levels and eliminating redundant processes and procedures [18]. The complexity varies depending on how the process is performed manually or automatically, depending on the business situation the complexity varies given the reliability of the information source to algorithms or natural person.

- Identify the attributes of the data that have a very important weight?

Identify the attributes that have very important weight such as the keys of a table which will be consistent between the different data sources, will reduce the complexity of data matching [26], data cleaning [27], data profiling [28], data deduplication [29].

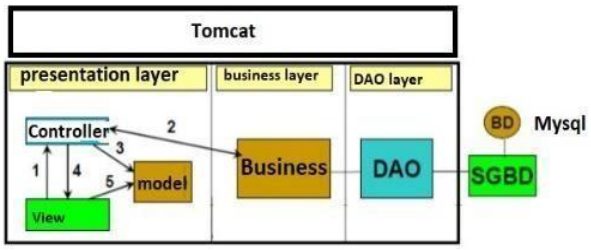
In the next paragraph, we will talk about deep learning of a neural network to classify the completeness prediction and the computed one.

V. FRAMEWORK ARCHITECTURE

A. Technical architecture of the framework

The architecture is based on the MVC framework intended to simplify, accelerate and maintain the development of a web application, allows easy maintainability and scalability [32], it is divided into 3 layers:

- A Model layer contains the data to display.
- A view layer contains the presentation of the graphical interface.
- A control and business layer contains the logic concerning the actions carried out by the user including the business processes.



Technical architecture of the jee application.

B. implementation phase

We present in this section the screenshots of the platform java to quantify the complexity and impact to improve data completeness of business process.

In our platform we record data on business processes, it belongs to which business service which also belongs to a department, during the evaluation of the complexity and the impact of a quality, each question has several choices of answer , this one has a weight that adapts to each situation in which the user of our platform sees that this answer to this question has weight than the other answers, without forgetting the weight attributed to the question itself, it joins the idea of giving a freedom to the user thus to make our framework adaptable to any domain or organization, this data will be dated for a possible use well defined and limited in time.

calculates complexity factor of process improvement

VI. ARTIFICIAL NEURAL NETWORKS (ANN)

A. Principe of ANN - neural network

Deep Learning and Machine Learning are new effective strategies for maximizing corporate profits; neural networks are inherently efficient at finding relationships between data and using them to predict (or classify) data [19]. Artificial neural networks are information processing structures providing the relationship between input data for flexible and powerful machine learning by artificially simulating human brain function [21, 38, 39, 40]. The entry level receives the input information, and the next level receives it. Each successive level receives the output of the previous level, while the last level produces the output of the system that will be the forecast.

B. Predict completeness by an artificial neuron network

ANN uses several hidden layers to create an algorithm that will approximate the complex mathematical function that binds the entry data to calculate or determine completeness. the backpropagation neural network (BPNN) is a learning function employed in an ANN application which uses the backpropagation of the error gradient to decrease the error of nonlinear functions [25]. In this study, a BPNN algorithm was adopted to predict completeness, which will have a favorable impact on the business at less cost.

The BPNN contains three types of layer: input layer, output layer, one or more hidden layers; the first two layers mentioned contains a limited number of neurons depending on the problem, while the intermediate layers can vary, the hidden layers receive inputs from the input layer, they are also input for the output layer, that is a kind of signal converter as for cerebral neurons [22, 24].

Backpropagation is a method to adapt connection weights, to balance and compensate for each error during learning; the aim is to predicted.

VII. COMPLETENESS PREDICTION BY NEURAL NETWORK

A. Principe of ANN - neural network

Deep Learning and Machine Learning are new effective strategies for maximizing corporate profits; neural networks are inherently efficient at finding relationships between data and using them to predict (or classify) data [19]. Artificial neural networks are information processing structures providing the relationship between input data for flexible and powerful machine learning by artificially simulating human brain function [21, 38, 39, 40]. The entry level receives the input information, and the next level receives it. Each successive level receives the output of the previous level, while the last level produces the output of the system that will be the forecast.

B. Predict completeness by an artificial neuron network

ANN uses several hidden layers to create an algorithm that will approximate the complex mathematical function that binds the entry data to calculate or determine completeness.

the backpropagation neural network (BPNN) is a learning function employed in an ANN application which uses the backpropagation of the error gradient to decrease the error of nonlinear functions [25]. In this study, a BPNN algorithm was adopted to predict completeness, which will have a favorable impact on the business at less cost.

The BPNN contains three types of layer: input layer, output layer, one or more hidden layers; the first two layers mentioned contains a limited number of neurons depending on the problem, while the intermediate layers can vary, the hidden layers receive inputs from the input layer, they are also input for the output layer, that is a kind of signal converter as for cerebral neurons [22, 24].

Backpropagation is a method to adapt connection weights, to balance and compensate for each error during learning; the aim is to improve prediction in learning networks; c-thus, the error is calculated as being the difference between the actual output and the calculated output [23, 33]. To decrease the error, we move the weight in the reverse direction of the gradient.

Forward propagation refers to the computation and then to the storage of intermediate variables; in order from input layer to the output layer, an activation donation is required depending on the situation and the value to be predicted.

VIII. COMPLETENESS PREDICTION BY NEURAL NETWORK

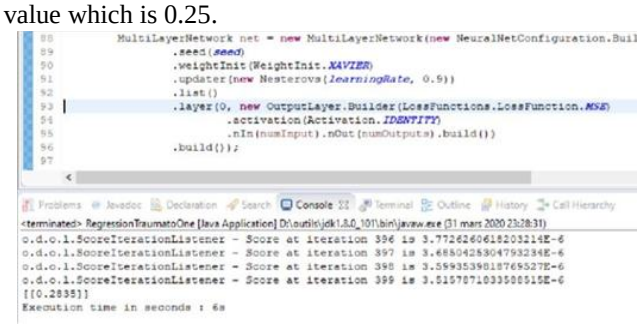
A. Learn from the cost survey responses and the impact of the improvement

Our java application is intended for practitioners and data quality managers, this tool quantified the cost of improving data completeness and its impact on the business; the information accumulated is on business processes, their type (critical, important, normal) in order to have a preference for one over the others, their affiliation to a business service associated with a department; this information will be useful for the company to process data from a single business process by limiting itself in time, will also be useful for a business service including a set of processes or even see relevant to other companies in the same field application [20].

B. case study

Case 1

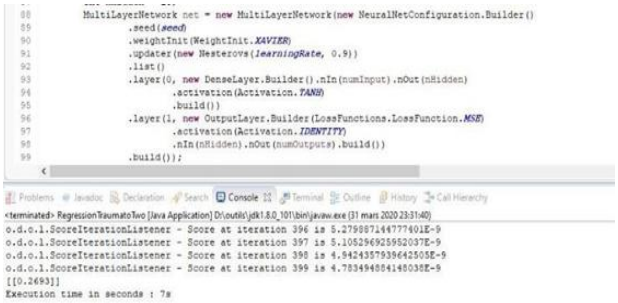
We start with a single hidden layer with 10 neurons in order to predict completeness; we are using the dl4j library, the completeness value of 0.2835 is greater than the desired value which is 0.25.



The predicted value of completeness for one hidden layer

Case 2

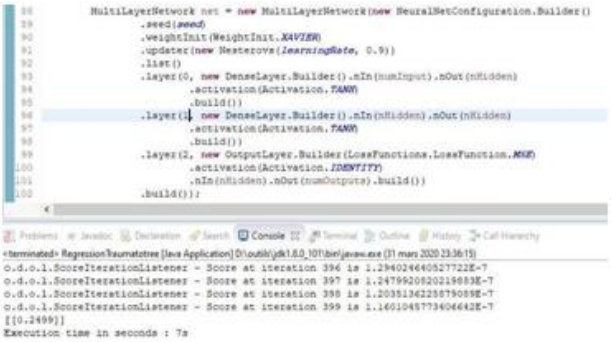
The next step is to add a second hidden layer, so we have two layers, each layer contains 10 neurons, the value calculated after learning is 0.2693, it exceeds the desired value which is 0.25.



The predicted value of completeness for two hidden layers

Case 3

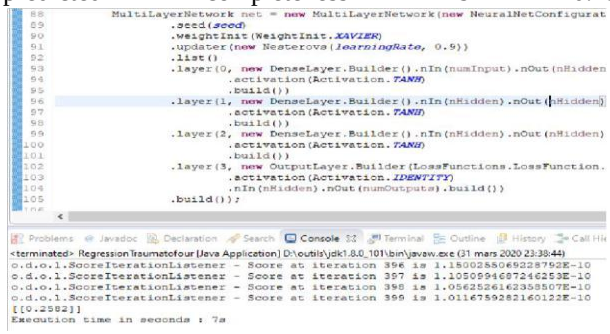
Adding another hidden layer to have three hidden layers, the completeness after using the algorithm is 0.2499 approaching the desired value.



The predicted value of completeness using three hidden layers

Case 4

We add another hidden layer to accurately predict completeness, so be aware that we had 4 hidden layers, the predicted completeness is 0.2583.



The predicted value of completeness using four hidden layers

According to the results found, it turned out that three hidden layers are adapted to our problematic and predicted with great precision the value after improvement, see table 2 is summarizing the four cases.

Table-II: comparison between four cases

	calculated completeness	the desired completeness	calculation time
Case 1	**0.2835**	**0.25**	**6s**
Case 2	**0.2693**	**0.25**	**7s**
Case 3	**0.2499**	**0.25**	**7s**
Case 4	**0.2582**	**0.25**	**7s**

IX. SHAPLEY VALUE AS A DECISION SUPPORT SYSTEM IN THE CHOICE OF BUSINESS PROCESSES AND SERVICES TO IMPROVE

A. Shapley value

In game theory, more specifically in a cooperative game, the value of Shapley gives players a fair distribution of winnings. It is named in honor of Lloyd Shapley who introduced the concept in 1953[36,37].

B. Use the Shapley value to choose candidate business processes to improve

A business service represents the added value that a company delivers to its environment. we can distinguish between internal and external services: Internal services is about the added value that is being delivered within the domain that the service belongs to. External services are about the added value that is being delivered to other domains or to the environment like customers.

A business service supports one or more business process and is offered to business customers. we will use the collaborative games represented by the value of shapley for the minimization of the cost of the improvement of the business service.

1) Shapley value to share the cost of business service improvement

As mentioned before the complexity for improving business process is:

$$CF(P_x) = \frac{\sum_{n=1}^{N}(w_c(n)*\sum_{h=1}^{T} w_c(h)*V_c(k))}{\Sigma w_c(n)} \quad (1)$$

Where $w_c(n)$ is weight of question

$w_c(h)$ is weight of h value of question

$V_c(h)$ the value of h element

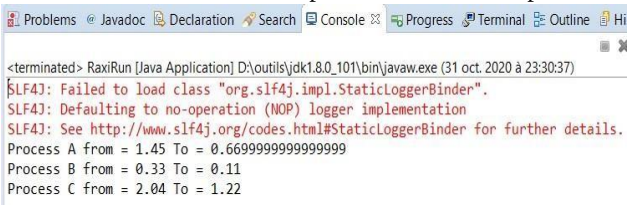
The cost is proportional to complexity, so the equation of the cost improvement of business process is:

$$\text{Cost}(P_x)) = \frac{CF(P_x)*((Compl_D - Compl_i) + |Compl_D - Compl_i|)}{2(1 - Compl_D)} (2)$$

Where $Compl_D$ is target completeness and $Compl_i$ initial completeness.

2) Use Case

In our case we choose one business service from our organization, publication of book, this service contains many business processes like, Arbitration, the sale of book, correction and other, for these three processes we use our application java to calculate cost of each process and use implementation of shapley value in java to calculate shared cost between all three-process chosen for this improvement of business service, the figure 8 talks about the shareable cost between the processes. we see a decrease in cost of improvement since the process belong to the same business service. in our example we have chosen to make a coalition between the three processes which constitutes the business service, if we have a cost constraint, not to exceed it, the coalitions can be between two processes than three processes.



The shared cost by shapley value

X. RESULTS AND DISCUSSION

The results of this case study show that a neural network driven by three hidden layers was better suited and more accurate for predicting the value of completeness of data used by a business process and when adding a fourth. layer, the completeness has been decreased, the results found for cooperation between business processes, experience has shown that the processes of the same business departments must collaborate, which will lead to a decrease in the cost of evaluation and quality improvement and the process will share the cost equitably.

XI. CONCLUSION AND FUTURE WORK

Our framework aims to find a way to judge quality projects according to the impact on the business and how much it would cost to improve data quality, our use case based on completeness given the importance of this dimension quality; the use of neural networks brings a deep analysis of the data accumulated to predict the improvement, this cost and impact evaluation system adapts to any domain based on factors with a weight of each question and a weight for each chosen value; each company will build its own model and can be generalized to another company or service in the same field.

After each improvement, it is necessary to compare the expected value with the actual value; update the value and recalculate if the prediction will be complete and have completeness and prediction accuracy over time. the value of shapley is used as a decision-making system to share the cost for a post-service for a department and even between companies during mergers or fissions.

Some future work: (i) the generalization of our approach for more dimensions of data quality (ii) application of our approach by dimension category (iii) use the shapley value between data quality dimensions.

REFERENCES

1. A. Haug, F. Zachariassen, D. van. Liempd, (2011, 04). The costs of poor data quality. JIEM. Vol 4(No 2), pp. 168-193. Available: https://www.jiem.org/index.php/jiem/issue/view/11
2. I. Yaqoo, I. A. T. Hashem, A. Gani, S. Mokhtar, E. Ahmeda,N. B. Anuar, A. V. Vasilakos. (2016, 12), Big data: From beginning to future.
3. IJIM. Vol 36(issue6), pp 1231-1247, Available: https://www.sciencedirect.com/science/article/abs/pii/S0268401216304753
4. R. Kantsperger, W. H. Kunz. (2010, 01), Consumer Trust in Service Companies:a Multiple Mediating Analysis. Managing Service Quality: An International Journal. Vol 20(issue1). Available: https://www.emerald.com/insight/content/doi/10.1108/09604521011011603/full/html
5. F. Sidi, P. H. S. Panahy, L. S. Affendey, M. A. Jabar, H Ibrahim, A. Mustapha, " Data Quality:A Survey of Data Quality Dimensions," in International Conference on Information Retrieval & Knowledge Management., Kuala Lumpur, Malaysia, 2012.
6. P.B. Crosby, "Quality is free: The Art of Making Quality Certain", New York: McGraw-Hill,1979, p 39.
7. E.W.Deming, (1982, juin 01). Quality, Productivity and Competitive Position. (2nd ed.). Cambridge, MA.
8. Juran on Leadership for Quality, 1998, pp. 15.
9. Maqboul J., Bounabat B. (2018) "Towards a Completeness Prediction Based on the Complexity and Impact," presented at the 1st int. ITCS, khouribgha, MR, Mar. 28-29, 2017.

10. JURAN'S QUALITY HANDBOOK, 5th ed., McGraw-Hill, 1979, pp.21.
11. R. Y. Wang and M. Diane. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. Management Information Systems. Vol 12, pp. 5-33, Available: http://www.jstor.org/stable/40398176
12. L. L. Pipino, W. Lee. Yang, and R. Y. Wang. (2002, April,). Data quality assessment. Commun. Vol 45, pp, 211–218, Available: https://doi.org/10.1145/505248.506010
13. S. Arestie, R. Harris, N. Ricci, M.A. Schaffner, M.M. Shaw, C. Whitlock, D. Miller, "A bottom-up modeling approach for the profit analysis of cellularized spacecraft architectures." In IAC, Beijing, China, Sept. 23-27, 2013
14. A. Hvenegaard, J.N. Arendt, A. et al. Street. (2011), Exploring the relationship between costs and quality: Does the joint evaluation of costs and quality alter the ranking of Danish hospital departments? Eur J Health Econ. Vol 12, pp. 541–551, Available: https://doi.org/10.1007/s10198-010-0268-9
15. Anders Haug, Frederik Zachariassen, Dennis van Liempd, The costs of poor data quality,2011.
16. Juran on Leadership for Quality, 1998, pp. 231.
17. L. Theresia, and R. Bangun. (2017), "Service quality that improves customer satisfaction in a university: a case study in Institut Teknologi Indonesi," at 10th in International Seminar on Industrial Engineering and Management "Sustainable Development In Industry and Management", Tanjung Pandan - Belitung, Indonesia, Nov 7-9, 2017.
18. S. E. Zoder, MIPP, "Improving Enterprise Master Data Quality What Is The ROI?." presented at the 8th in AMCIS.,Savannah, Georgia,Aug. 710, 2015.
19. S. A. Sirohey, A. I. Hunjra and B. Khalid, (2012). Impact of Business Process Automation on Employees' Efficiency. Bulletin of Business and Economics, Vol 1(1), pp. 1-12, Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3228559
20. D. Fister, J. Mun, V. Jagric and T. Jagric, (2019). deep learning for stock market trading: a superior trading strategy. NNW Vol 29, pp. 151-171.
21. T. Karatas, A. Oskoui, A. Hirsa. (2019), Supervised Deep Neural Networks (DNNs) for Pricing/Calibration of Vanilla/Exotic Options Under VariousDifferent Processes.
22. Md. Haidar Sharif, Osman Gursoy, (2018, Feb). Parallel Computing for Artificial Neural Network Training using Java Native Socket Programming. PENS. Vol 6(No 1), pp. 1-10, Available: http://pen.ius.edu.ba
23. M. Buscema. (1998), Back Propagation Neural Networks, in Substance Use & Misuse. SUM Vol33(2), Available: https://www.tandfonline.com/doi/abs/10.3109/10826089809115863.
24. S. H. Oh, (1997, May), Improving the error backpropagation algorithm with a modified error function. IEEE TNN, Vol 8(3), pp. 799803, Available: https://ieeexplore.ieee.org/abstract/document/572117.
25. A. Semenov, V. Boginski, E. L. Pasiliao. (2019)," Neural Networks with Multidimensional Cross-Entropy Loss Functions," presented at the 8th International Conference, CSoNet 2019, Ho Chi Minh City, Vietnam, Nov 18–20 2019.
26. J. V. N. Lakshmi, "Stochastic Gradient Descent using Linear Regression with Python",2016.
27. P. Christen," Data Matching Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection", Canberra, ACT, Australia, 2012, pp. 8.
28. O. Azeroual, G. Saake, M. Abuosba. (2018, 2), Data Quality Measures and Data Cleansing for Research Information System. 12 Journal of Digital Information Management0, Vol 16(1).
29. W. Dai1, I. Wardlaw, Y. C. Kashif Mehdi, L. Yanyan, J. Long. (2016, Mar 29). Data Profiling Technology of Data Governance Regarding Big Data: Review and Rethinking", in (eds) Information Technology: New Generations. Advances in Intelligent Systems and Computing, vol 448. Springer. Avcailable: https://doi.org/10.1007/978-3-319-32467-8_39
30. N. Kuma,Shobha,S. C. Jain. (2018, Jul 10), Efficient Data Deduplication for Big Data Storage Systems, Progress in Advanced Computing and Intelligent Engineering. AISC. Available: https://doi.org/10.1007/978981-13-0224-4_32.
31. E. Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. International Journal of Information Management, Vol 38(1), pp. 187-195.
32. K. Avita, M. Wazid, R. H. Goudar. (2017), "Big Data: Issues, Challenges, Tools and Good Practices," presented at the 6th int.International Conference on Contemporary Computing, NOIDA, India, Aug. 8-10, 2013.
46. H. I. Sarker, K. I. Z. Apu, (2014, 09). MVC Architecture Driven Design and Implementation of Java Framework for Developing Desktop 45. Application. IJIT, vol 7(5), pp.317-322, Available: DOI:10.14257/ijhit.2014.7.5.29.
47. D. E. rumelhart, G. E. hinton, and R. J. Williams. (2014), "Learning Internal Representations by Error Propagation," in Readings in Cognitive Science, 1989, pp. 399-421
48. R. Majumdar, (2014), Business decision making, production technology and process efficiency, IJEM, Vol. 9(1), pp. 79-97, Available: https://doi.org/10.1108/IJoEM-03-2011-0023.
49. P. H. S. Panahy, F. Sidi, L. S. Affendey, M. A. Jabar, H. Ibrahim, A. Mustaph, (2013, 01), A Framework to Construct Data Quality Dimensions Relationships, IJST, Vol 6(5), pp. 4422-4431.
50. L. S. Shapley, "A Value for n-Person Games", RAND Corporation, Santa Monica, Calif, (295), 1952. As of October 20, 2020: https://www.rand.org/pubs/papers/P0295.html.
51. S. Hart, "The New Palgrave: Game Theory,", Norton, 1989, pp. 210–216.
52. H. Bandy, (1997). Developing a Neural Network System. AAII Working Paper. S. Haykin, "Neural Networks: a comprehensive foundatio, pp. 32, 2008.

AUTHORS PROFILE

Jaouad Maqboul phd doctorate; PMP® training (Project Management Professional) 2016; Technical Director since 2013; technical project manager since 2011; Team leader and application manager from 2010 to 2011; technical reference JEE, STruts; Engineer of study and development since 2008; Specialized Master in Engineering & Information Systems at the Faculty of Sciences Semlaliya Marrakech - Cadi Ayyad University Marrakech; Applied Computer Science Degree at FSSM; University Certificate of Scientific Studies (CUES) in Maths-Physics at FSSM 1999-2000; baccalaureate Mathematical Sciences Series at Cadi Ayyad High School in Marrakech; 2nd Prize winner of the 10th edition of the student week in March 2005 at the faculty of Sciences Semlalia Marrakech.

Bouchaib Bounabat Professor in ENSIAS, (National Higher School for Computer Science and System analysis), Rabat, Morocco (since 1994) ; Member of University Mohammed V –Souissi Management Commission (since 2009); Responsible of Doctoral Programs, ENSIAS (since 2006); Head of Software Engineering Department, ENSIAS (2005-2008); Responsible of Qualsadi R&D Team (Quality of Software Architectures, their Development and Integration) (2006-2010); National Consultant to the Ministry of Higher Education, and Scientific Research for the Evaluation of Moroccan Universities Computer Science curricula's (2001 – 2003)