Hindawi
Scientific Programming
Volume 2019, Article ID 5235706, 4 pages
https://doi.org/10.1155/2019/5235706

Editorial

Data-Driven Computational Intelligence for Scientific Programming

Álvaro Rubio-Largo,[1] **Juan Carlos Preciado**,[2] **and Luis Iribarne**[3]

[1]*Universidade Nova de Lisboa, Lisboa 1099-085, Portugal*
[2]*University of Extremadura, Cáceres 10003, Spain*
[3]*University of Almería, Almería 04120, Spain*

Correspondence should be addressed to Álvaro Rubio-Largo; arl@unex.es

Received 22 July 2019; Accepted 22 July 2019; Published 19 August 2019

Copyright © 2019 Álvaro Rubio-Largo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, big data and its potential to shed valued insights into enhanced decision-making processes has attracted an increasing interest from both academia and industry. Throughout history, there have been eras that have marked turning points in society. Currently, we are faced with the beginning of one of these turning points, a new leap in evolution which arises thanks to the advantages of technology, but which has recently been revolutionizing technical concepts of development and programming. We are in a new age, the era of Data [1].

Currently, the volume of data generated each day is very high, coming from different multiple sources and in diverse formats, usually designed with different goals, methods, profiles, and production and consumption rhythms. The amount of data generated by businesses, public administrations, and numerous industrial and scientific research facilities has increased immeasurably in the past years, turning traditional systems into complex or supercomplex systems. These data may be structured, semistructured, and/or unstructured, extracted from sources as different as Natural Language Processing [2] (chatbots, comments, and social media), multimedia content (videos, images, and audio), geographic information systems (GIS), or sensors (Internet of Things/Everything) on a wide variety of platforms or environments (e.g., machine-to-machine communications, social media sites, sensors networks, or natural interaction) [3].

These are highly relevant data, so much so that they must be taken into account when designing and developing the solutions of the future (or redesigning the present ones). The proper evolution of data concept is of vital importance, due to the fact that they have a huge impact on the economic and social development of the institutions and enterprises they belong to.

In this new cycle of massive data, our administrations, enterprises, and citizens generate colossal amounts of data, data which on their own do not help us in our everyday life or in making decisions, and hence the importance of treating these data in order to turn them into information that is useful and relevant for them [4].

The evolution of computational intelligence and big data practices stimulates an increasing interest in Analytics solution [5]. Accurate prediction, precise identification of trends, and discovering behaviour patterns can optimize the resource usage or consumption, generating new knowledge in science and research, and enabling faster and better decisions in politics, weather, science, research, real estate, sports, or healthcare, among other social application domains [2].

To date, the solutions proposed have been disjointed, and actions put into practice as a result of fashions or through the implantation of a certain known technology, but without a perspective of the specific need of the concept of amount of data. The challenge is no longer the evolution and development of the technology, the current issue is not related with the production and acquisition of data, and now we are treating with new challenges and problems regarding the processing of these data to build information [6]. Based on these concepts and the engineering of complex massive data systems with high availability, we see

that a computational intelligence environment is configured like a system, a smart system. In this context, methodology supporting the design of this system would firstly ask the computational intelligence system's designer, to identify the requirements for the existing data required to continue with the other phases.

The great data acquisition obliges us to have innovative massive data storage systems on a scale not yet contemplated. This mass of data is growing exponentially, and 90% of data on a worldwide scale has been created only in the last two years [7]. The management and processing of this data defines a new knowledge management paradigm which can only be addressed by means of constant methodological and technological innovation regarding the field of computational intelligence applied to scientific programming [8].

The different data sources and formats result in a complexity which makes decision-making expensive, and in cases where the implementation of Big Data, Business Intelligence, and Business Analytics is not enough, users need more and better ways of understanding data. In this scenario, techniques like Data Visualization come into the scene, which enable the visual simplification of all information built from the collected data. This information visualization supports decision-making and the advanced analysis of the data collected, by means of report and graphical representation techniques, with data visualization capacities which enable actors to build the necessary information flexibly on the same set of data, based on this data and placing it or other data at the service of third parties, hence providing an environment of rich collaboration and innovation. One of the trends given clear importance by Computation Intelligence Systems is the visualization of information in a simple, agile, and powerful way, i.e., Computation Intelligence Systems makes possible the process data and display information quickly and in a reasonable time frame to assimilate all possible information by the people for whom it is intended. Definitively, the advance in technology and scientific programming we are experiencing plays an essential role in the evolution of computational intelligence environment [9].

Computational intelligence techniques form a set of nature-inspired computational methodologies and techniques which have been developed to face the aforementioned complex scientific programming, for which traditional models are unable to work due to high complexity, uncertainty, and the stochastic nature of processes. These techniques typically include parallel/distributed pattern-recognition techniques, genetic programming, fuzzy systems, or evolutionary computation. The overall aim of this special issue is to collect state-of-the-art research findings on the latest developments, as well as up-to-date issues and challenges in the field of computational intelligence applied to scientific programming.

This special issue is configured like a collection of papers on a hot topic of increasing interest within of scientific programming, presenting and describing new research or applications in this field. Several call for papers were sent and distributed among the main mailing lists of the field for relevant researchers to submit their research to this special issue. As an example of the current interest in this field, it is worth mentioning that, for this special issue, we have managed a total of 15 high-quality submissions from different countries: Spain, China, India, Greece, Australia, United Kingdom, and Vietnam. These researches have been managed according to the terms and guidelines of this journal. All the papers included in this special issue were reviewed by at least two expert reviewers. Furthermore, all the papers in the special issue received a minimum of two review rounds. Finally, five papers of high quality in emerging research areas were accepted for inclusion in the special issue (acceptance rate = 5/15 = 33.33%). In summary, we think these papers bring us an international sampling of significant work.

In general, the papers included in this special issue cover detailed scientific aspects related mainly with the fields of recommendation systems, machine learning, pattern recognition, spatial data interaction, and complex events. Research advances are applied to different social domains such as smart cities, scientific digital libraries, urban spatial data, and public health.

The title of our first paper is "Study of Urban System Spatial Interaction Based on Microblog Data: A Case of Huaihe River Basin, China," by Y. Fan, J. Yao, Z. He, B. He, and M. Li. This paper obtains the user microblog information through the sina microblog open platform and studies the urban spatial pattern and urban interaction by means of statistical analysis and spatial analysis. This paper takes the Huaihe basin as the case area to verify the proposal presented on it.

The main research conclusions from our first paper are as follows: (1) the data interface provided by the microblog platform can study the urban spatial pattern. The user trajectory of microblog data can explore the spatial relationship of regional cities, and data acquisition and data quality evaluation can meet the research requirements, and (2) based on microblog data, the spatial and temporal characteristics of the urban system spatial pattern in the Huaihe River Basin are analyzed from network connectivity and urban interaction. The study found that the urban spatial relation in the Huaihe River Basin has the following characteristics: the spatial difference of urban size distribution is obvious; urban layout presents a stratified aggregation phenomenon; and the high-grade cities lead the city's interaction.

As for the application of microblog data in urban research, the current data mainly focus on information text, social relations, and other aspects. The research is mainly about event detection and hot spot exploration. The combination of big data thinking and data mining technology will have more research findings in the study of urban problems.

The second paper is "Practical Experiences in the Use of Pattern-Recognition Strategies to Transform Software Project Plans into Software Business Processes of Information Technology Companies" by C. Arevalo, I. Ramos, J. Gutiérrez, and M. Cruz. The authors' proposal provides a framework to generate software business processes that would otherwise be hidden or wasted in

databases of non-process-aware information systems (non-PAISs). This hidden knowledge can be used to implement the business process management approach in information technology companies (ITCs) that will help them to become more competitive and reduce costs. Compared to other business process discovery methods used with non-PAISs, their results are more adjusted to the reality of processes since they focus on transformations among artifacts that are close to executed processes that exist at different levels of abstraction (i.e., platform level and software expert level). Furthermore, business processes may be enriched with data regarding resources and costs that may also be bound to projects in project management systems. This way, new data will be available to set metrics and study key performance indicators of software business processes.

This paper illustrates the AQUA-WS project case study to test the developed MDA-based roadmap. In this case study, they have shown that generated processes are similar to real processes that a business software expert may design. For this reason, they have delivered a semiautomatic proposal to obtain processes of ITCs. As future work, the authors say that they will be able to use further source systems, such as other PMSs, ECMs, ERPs, CRMs, SCMs, or tailor-made software. Besides, they will propose roadmaps to specific SPMLs or GPMLs targets, used by ITCs that work with this type of systems. In this case study, they have considered the following aspects of source systems, target systems, and heuristics to generate business processes.

Our third paper is "A High-Frequency Data-Driven Machine Learning Approach for Demand Forecasting in Smart Cities" by J. C. Preciado, A. E. Prieto, R. Benitez, R. Rodríguez-Echeverría, and J. M. Conejero. This paper presents an approach based on pattern-similarity techniques to forecast water demand, referred to as short-term pattern similarity (STPS). This work faces two important challenges that have been traditionally neglected in previous approaches, namely, a high frequency of predictions (based on measurements in terms of minutes) and the need for external data such as annual seasonality or weather that increases the complexity of the approaches. In that sense, on the one hand, it is based on 1 min steps predictions, and, on the other hand, it does not require estimating annual seasonality since it determines this seasonality by constructing the X and Y patterns in which the series has been normalized.

In order to validate their approach, the study was applied over three different sites of a city in northern Spain. The results obtained provided interesting insights, such as the best predictions obtained in high-density population areas, the difficulties for identifying patterns for Sundays in industrial areas, or the higher random behaviour in low-density areas. Additionally, their pattern-similarity approach (STPS) was also compared to other similar techniques that have been previously used for water forecasting, i.e., $\alpha\beta$ water demand forecast ($\alpha\beta$-WDF) and generalized regression neural network (GRNN). The results obtained evidenced that $\alpha\beta$-WDF was the approach with worst results whilst GRNN and STPS behave similarly. As future work, the authors try to manage some weaknesses already identified in the proposed method. Firstly, predictions success is lower when anomalous days are taken into account. Secondly, with the aim of improving the results for data sites where apparently there is not regularity, such as the low-density population area in our study, other approach like the shorter prediction horizons could be considered, for instance, 4–6 hours. Notwithstanding, this is a subject that remains currently untested. Finally, another interesting line of further work is the application of the presented method for water distribution in different cities with similar water requirements.

A systematic mapping study is addressed in the fourth paper, "Recommendation and Classification Systems: A Systematic Mapping Study," by J. G. Enríquez, L. Morales-Trujillo, F. Calle-Alonso, F. J. Domínguez-Mayo, and J. M. Lucas-Rodríguez. This study has been performed to facilitate researchers and practitioners the task of choosing the most appropriate system, technology, or algorithm to include in the ADAGIO project for satisfying their requirements. In this sense, this paper presents a systematic mapping study (SMS) that analyzes the current state of the art of the recommendation and classification systems and how they work together. Then, from the point of view of the software development life cycle, this review also shows that the work being done in the ML (for classification and recommendation) research and industrial environment is far from earlier stages such as business requirements and analysis. This makes it very difficult to find efficient and effective solutions that support real business needs from an early stage. Then, this paper suggests the development of new ML research lines to facilitate its application in the different domains. As future work, the authors propose a very interesting research line may focus on how to combine these systems to obtain more efficient and effective solutions.

Unlike most SMSs that are focused on the scientific literature, this study has been carried out from two points of view as discussed throughout the paper: the scientific and the industrial scopes. Within the scientific field, the results showed that the most studied technique in recommendation systems is recommendation with the use of collaborative filters, closely followed by those that use content-based filters. Only 14 used hybrid recommendation systems, whereas 31 used collaborative filtering and 29 used content-based methods. This is an interesting suggestion for researchers starting to use recommender systems, to find which of them are more popular and more used in the scientific environment. By conducting market research through systematic industrial mapping, it was found that there are many technologies that offer automatic learning solutions, and most of which are complete systems or libraries. However, the nature of most of them could not be known because the proprietary software did not allow it. Another important issue that must be highlighted is that not only the communities of free software developers are interested in this topic but also there are large companies that are working on it

for commercial purposes. This clearly shows the underlying economic interest, an indicator that it is a branch of long-distance research.

The fifth paper is entitled "Facilitating the Quantitative Analysis of Complex Events through a Computational Intelligence Model-Driven Tool" by G. Díaz, H. Maciá, V. Valero, J. Boubeta-Puig, and G. Ortiz. Complex event processing (CEP) is a computational intelligence technology capable of analyzing big data streams for event pattern recognition in real time. In this paper, the authors illustrate the use of the MEdit4CEP–CPN approach for the complex event analysis through a case study based on the sick building syndrome. The event patterns have been graphically modeled with MEdit4CEP–CPN and then automatically transformed into both Event Processing Languages (EPL) and Coloured Petri Nets (CPN) code. Additionally, CPN Tools has been used to make quantitative analysis of events produced for this case study. Given the flexibility provided by MEdit4CEP–CPN, this analysis could be applied to other cutting-edge real-world case studies, such as eHealth, robotic, and mobile edge, and cloud computing applications.

The main advantage of MEdit4CEP-CPN is that supports many functionalities that other approaches do not provide, such as (1) modeling CEP domains and event patterns in a user-friendly way by dragging and dropping elements on a canvas, (2) validating the pattern syntax, (3) automatically transforming the graphical patterns into a CPN model, (4) automatically transforming the CPN model to the XML code executable by CPN Tools and validating the pattern semantics, (5) automatically generating the Esper EPL code and deploying it in a particular event-based system, and (6) providing a quantitative analysis of complex events through the CPN Tools executable model automatically generated by the tool. As future work, the authors plan to add additional features and functionalities to MEdit4CEP-CPN, such as further EPL operators and new transformation techniques.

Data and computational intelligence are outstanding research issues in the field of computer sciences which combined together represent one of the most emerging topics at present. We sincerely hope that you enjoy this special issue. We also have hopes that paper collection as a whole can pleasantly introduce readers to the composite and challenging arena of the application of computational intelligence to the scientific programming field, giving a fresh view of several state-of-the-art solutions from diverse perspectives. All accepted papers are within the scope of the journal and particularly the special issue, and all of them provide relevant and interesting research techniques, models, and work directly applied to the area of scientific programming. Before concluding, we want to express our sincere gratitude to all the authors who submitted their paper to this special issue and the many reviewers whose dedicated efforts made this special issue possible.

Conflicts of Interest

The editors declare that they have no potential conflicts of interest.

Acknowledgments

The editors wish to acknowledge the collaborative funding support from Ministerio de Economía e Innovación (Spain) (RTI2018-098652-B-I00 (MINECO/ERDF, EU) project), Ministry of Economy and Competitiveness (Spain) (CoSmart TIN2017-83964-R project), and Consejería de Economía e Infraestructuras/Junta de Extremadura (Spain)–European Regional Development Fund (ERDF) (GR18112 project and IB16055 project). The editors would also like to thank the reviewers for their generous time in providing detailed comments and suggestions that helped us to improve the quality of this special issue.

Álvaro Rubio-Largo
Juan Carlos Preciado
Luis Iribarne

References

[1] C. L. P. Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: a survey on big data," *Information Sciences*, vol. 275, pp. 314–347, 2014.

[2] S. Sun, C. Luo, and J. Chen, "A review of natural language processing techniques for opinion mining systems," *Information Fusion*, vol. 36, pp. 10–25, 2017.

[3] H. Harb, A. Makhoul, and C. Abou Jaoude, "A real-time massive data processing technique for densely distributed sensor networks," *IEEE Access*, vol. 6, pp. 56551–56561, 2018.

[4] L. Cao, "Data science: a comprehensive overview," *ACM Computing Surveys*, vol. 50, no. 3, pp. 1–42, 2017.

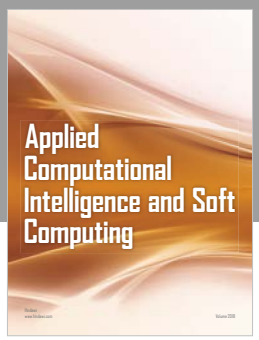
[5] K. Lepenioti, A. Bousdekis, D. Apostolou, and G. Mentzas, "Prescriptive analytics: literature review and research challenges," *International Journal of Information Management*, vol. 50, pp. 57–70, 2020.

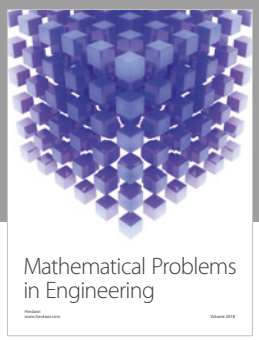
[6] A. L'Heureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz, "Machine learning with big data: challenges and approaches," *IEEE Access*, vol. 5, pp. 7776–7797, 2017.

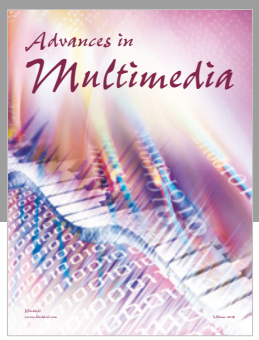
[7] B. P. L. Lau, S. H. Marakkalage, Y. Zhou et al., "A survey of data fusion in smart city applications," *Information Fusion*, vol. 52, pp. 357–374, 2019.

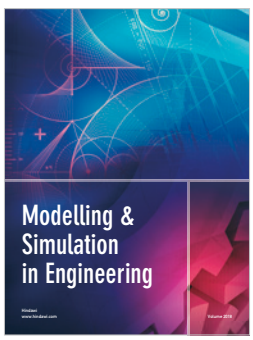
[8] Y. Jin and B. Hammer, "Computational intelligence in big data [guest editorial]," *IEEE Computational Intelligence Magazine*, vol. 9, no. 3, pp. 12-13, 2014.

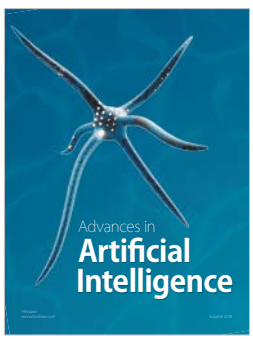
[9] Z. Lv, H. Song, P. Basanta-Val, A. Steed, and M. Jo, "Next-generation big data analytics: state of the art, challenges, and future research topics," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 1891–1899, 2017.

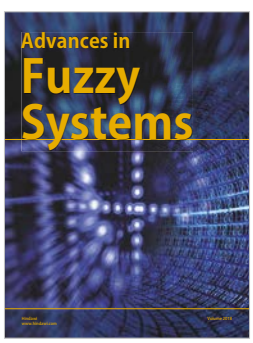


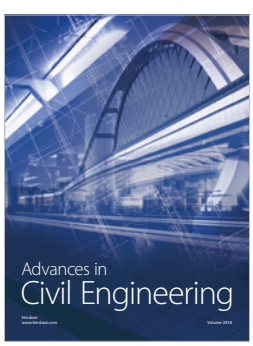


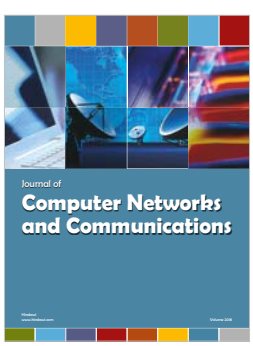


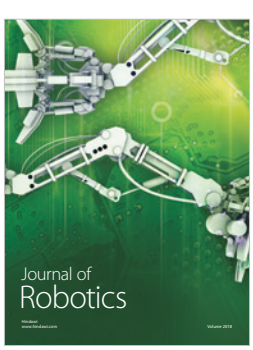


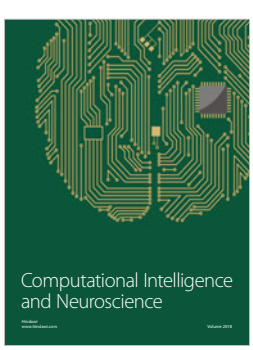












Hindawi

Submit your manuscripts at
www.hindawi.com