



HAL
open science

Editorial: Seventh special issue on Knowledge Discovery and Business Intelligence

Paulo Cortez, Albert Bifet

► **To cite this version:**

Paulo Cortez, Albert Bifet. Editorial: Seventh special issue on Knowledge Discovery and Business Intelligence. *Expert Systems*, 2023, 40 (10), pp.e13466/1-2. 10.1111/EXSY.13466 . hal-04468367

HAL Id: hal-04468367

<https://hal.science/hal-04468367>

Submitted on 4 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Editorial: Seventh special issue on Knowledge Discovery and Business Intelligence

Paulo Cortez¹ Albert Bifet²

¹ ALGORITMI Research Centre/LASI, Department of Information
Systems, University of Minho, 4800-058 Guimarães, Portugal

Email: pcortez@dsi.uminho.pt

ORCID: 0000-0002-7991-2090

² AI Institute - University of Waikato, Hamilton 3240, New Zealand /
LTCI, Télécom Paris, Institute Polytechnic de Paris, 19 place

Marguerite Perey, 91120 Palaiseau, France

Email: albert.bifet@telecom-paris.fr

ORCID: 0000-0002-8339-7773

1 Introduction

Currently, there is a growing interest in the field of Artificial Intelligence (AI) [Russell and Norvig, 2022]. Indeed, there are several successful AI applications that are impacting in diverse real-world domains, such as Industry 4.0 [Silva et al. 2021], Recommendation Systems (e.g., Amazon, Netflix) [Deepjyoti and Mala, 2022] and Virtual Assistants (e.g., Siri, Alexa, ChatGPT) [Grudin 2023]. AI encompasses several approaches to provide machine intelligence, including: Expert Systems (ES) and Decision Support Systems (DSS) [Artnott and Pervan, 2014]; Machine Learning (ML) and Deep Learning (DL) [Alpaydin21]; and Metaheuristics or Modern Optimization [Cortez, 2021]. ES were a popular AI tool in the 1970s and 1980s, assuming an explicit knowledge representation (e.g., via symbolic rules) that was extracted from decision-makers and then processed by inference systems in order to support real-world tasks [Cortez et al., 2018]. After the 1990s, following the growth of big data and computational power, there was a shift towards data-driven AI [Darwiche, 2018]. This shift gave the rise to several data processing terms that often overlap and that aim to extract value from raw data, such as: Knowledge Discovery (KD) and Data Mining (DM) [Fayyad et al., 1996]; Business Intelligence (BI) and Business Analytics (BA) [Artnott and Pervan, 2014]; and Big Data and Data Science [Provost and Fawcett, 2013].

The first Knowledge Discovery and Business Intelligence (KDBI) track of the EPIA conference on Artificial Intelligence was held in 2009, in Aveiro, Portugal. The aim of the track was to promote a research interaction between the KD and BI areas. Following the first edition success, the KDBI track become regularly associated with

the EPIA conference, resulting in nine editions from 2009 to 2023. In 2011, a special KDBI issue was published in Wiley’s Expert Systems journal (EXSY), including extended versions of the best track papers [Cortez and Santos, 2013]. Since then, all KDBI tracks have produced an EXSY special issue (e.g., [Cortez and Bifet, 2020]). This ‘Seventh special issue on Knowledge Discovery and Business Intelligence’ includes extended versions of selected papers presented at the 8th KDBI thematic track of EPIA 2022, held on Lisbon, Portugal. The track received 11 paper submissions and 9 of these papers were presented at the KDBI track of EPIA 2022. This EXSY special issue includes four extended papers that were accepted after two rounds of reviews.

The special issue accepted papers describe novel KDBI methods, systems and challenging real-world applications. For instance, the repair process of digital devices requires time and costs. A critical repairing issue is the proper selection of spare parts. Furthermore, analysing coastal upwelling patterns is essential for studying ocean dynamics and climate models. It is also highly relevant to support the management of coastal resources. Moreover, BI systems often require the development of an Extract-Transform-Load (ETL) system, which can benefit from assuming a service-oriented approach. Finally, valuable marketing trends can be identified by performing a Competitive Intelligence analysis over freely available text resources. These challenges are addressed by the four papers published here. In the next section, we summarize the main contributions of these papers.

2 Contents of the special issue

In the first paper ‘An Empirical Study of Learning-to-rank for Spare Parts Consumption in the Repair Process’, Duarte et al. (2023) perform an empirical study of multi-label classification ML algorithms in order to support an automatic identification of spare parts to be consumed during a digital device repairing process. Using millions of data records, hundreds of ML models were trained under different train set sizes and intervals. Competitive results were obtained by two ML algorithms, a Logistic Regression and an Extreme Gradient Boosting (XGBboost).

In ‘Piece-Wise Constant Cluster Modelling of Dynamics of Upwelling Patterns’, Nascimento et al. (2023) propose a three-stage clustering algorithm to identify season’s coastal upwelling patterns. Using a least-squares clustering criterion, the proposed approach was successfully applied to sea surface temperature data, considering sixteen years of records and two regions, the Portuguese and Morocco coasts.

The third paper, entitled ‘Overcoming Traditional ETL Systems Architectural Problems Using a Service-Oriented Approach’, by Oliveira et al. (2023), presents a service-oriented implementation for ETL design and development. The proposed approach assumes the division of typical ETL processes into self-contained components and it was demonstrated using the Wide World Importers (WWI) Microsoft ETL example.

Finally, the fourth paper ‘MapIntel: A Visual Analytics Platform for Competitive Intelligence’, by Silva and Bação (2023), describes MapIntel, a system for acquiring Competitive Intelligence from large collections of text data. The system can handle complex natural language queries, also performing a visual exploration

of the analysed corpus, aiming to provide meaningful insights for the decision-maker.

Acknowledgments

We wish to thank the other KDBI 2022 track (of EPIA) co-organizers, namely João Gama, Luís Cavique, Manuel Santos and Nuno Marques. Also, we would like to thank the authors, who contributed with their papers, and the reviewers (from the KDBI 2022 program committee and the EXSY journal). The work of P. Cortez was supported by FCT - *Fundação para a Ciência e Tecnologia* within the R&D Units Project Scope: UIDB/00319/2020.

References

- [Alpaydin21] Alpaydin, E. (2021). Machine learning. MIT Press.
- [Artnott and Pervan, 2014] Artnott, D. and Pervan, G. (2014). A Critical Analysis of Decision Support Systems Research Revisited: The Rise of Design Science. *Journal of Information Technology*, 29(4): 269-293.
- [Cortez, 2021] Cortez, P. (2021). Modern Optimization with R. 2nd edition, Springer.
- [Cortez and Bifet, 2020] Cortez, P. and A. Bifet (2020). Fifth Special Issue on Knowledge Discovery and Business Intelligence. *Expert Systems*, 36(6): e12628. doi: 10.1111/exsy.12628.
- [Cortez and Santos, 2013] Cortez, P. and Santos, M. F. (2018). Knowledge Discovery and Business Intelligence. *Expert Systems*, 30:4: 283-284. doi: 10.1111/exsy.12042.
- [Cortez et al., 2018] Cortez, P., Moro, S., Rita, P., King, D. and Hall, J. (2018). Insights from a text mining survey on Expert Systems research from 2000 to 2016. *Expert Systems*, 35(3): e12280. doi: 10.1111/exsy.12280.
- [Darwiche, 2018] Darwiche, A. (2018). Human-level intelligence or animal-like abilities?. *Communications of the ACM*, 61(10): 56-67.
- [Deepjyoti and Mala, 2022] Deepjyoti, R. and Mala, D (2022). A systematic review and research perspective on recommender systems. *Journal of Big Data*, 9(1): 59.
- [Duarte et al., 2023] Duarte, E., Moraes, D. and Padula, L. L. (2023). An Empirical Study of Learning-to-rank for Spare Parts Consumption in the Repair Process. *Expert Systems*, pages –.
- [Fayyad et al., 1996] Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). *Advances in Knowledge Discovery and Data Mining*. MIT Press.
- [Grudin 2023] Grudin, J. (2023). ChatGPT and Chat History: Challenges for the New Wave. *Computer*, 56(5): 94-100.

- [Nascimento et al., 2023] Nascimento, S., Martins, A., Relvas, P., Luís, J. F. and Mirkin, B. (2023). Piece-Wise Constant Cluster Modelling of Dynamics of Upwelling Patterns. *Expert Systems*, pages –.
- [Oliveira et al., 2023] Oliveira, B., Oliveira, Ó and Belo, O. (2023). Overcoming Traditional ETL Systems Architectural Problems Using a Service-Oriented Approach. *Expert Systems*, pages –.
- [Provost and Fawcett, 2013] Provost, F. and Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1): 51-59.
- [Russell and Norvig, 2022] Russell, S. and Norvig, R. (2022). Artificial Intelligence: A Modern Approach. 4th edition, Pearson, Harlow, UK.
- [Silva and Bação., 2023] Silva, D. and Bação. (2023). MapIntel: A Visual Analytics Platform for Competitive Intelligence. *Expert Systems*, pages –.
- [Silva et al. 2021] Silva, A., Cortez, P., Pereira, C. and Pilastrri, A. (2021). Business analytics in Industry 4.0: A systematic review. *Expert Systems*, 38(7).

The editors

Paulo Cortez

Paulo Cortez is a Full Professor at the Department of Information Systems, University of Minho, Portugal. He is also assistant director of the ALGORITMI R&D center. His research interests include: Business Analytics (Decision Support, Data Science and Forecasting); and Artificial Intelligence (Machine Learning, Evolutionary Computation and Applications). Currently, he is associate editor of the *Expert Systems* and *Decision Support Systems* journals. He is co-author of more than 201 indexed (e.g., Scopus) publications in international journals (e.g., Decision Support Systems, Information Processing Management) and conferences (e.g., IEEE IJCNN, ACM GECCO). Web-page: <https://pcortez.dsi.uminho.pt/>

Albert Bifet

Albert Bifet is a Professor of Artificial Intelligence (AI) and Director of the *Te Ipu o te Mahara* AI Institute of at University of Waikato, New Zealand. He is also a Professor of Big Data at LTCI, Telecom ParisTech, France. He is the author of a book on Adaptive Stream Mining and Pattern Learning and Mining from Evolving Data Streams. He is one of the leaders of MOA and Apache SAMOA software environments for implementing algorithms and running experiments for online learning from evolving data streams.