Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 38 (2019) 84–91

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

29th International Conference on Flexible Automation and Intelligent Manufacturing: (FAIM2019), June 24-28, 2019, Limerick, Ireland

Machine Learning and Process Mining applied to Process Optimization: Bibliometric and Systemic Analysis

Ederson Carvalhar Fernandes[a*], Barry Fitzgerald[b], Liam Brown[b], Milton Borsato[a]

[a]*Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba, Brazil*
[b]*Limerick Institute of Technology (LIT), Limerick, Ireland*

Abstract

The highly competitive business environment has been increasing with the advent of Industry 4.0, since the fast-changing market requirements need rapid decision-making to improve productivity. Hence, the smart factory has been highlighted as a digitized and connected production facility, which can use and combine data analytics and artificial intelligence algorithms and techniques to manage and eliminate failures in advance by accurate prediction. Thus, the purpose of this study is to identify the unfilled gaps and the opportunities regarding machine learning and process mining applied to process optimization, through a literature review based on the last five years of study. In order to accomplish these goals, the current study was based on the Knowledge Development Process – Constructivist (ProKnow-C) methodology. Firstly, a bibliographic portfolio was created through Articles Selection and Filters Application. This found that, from 3562 published articles across five databases between 2014 and 2018, only 32 articles relating to the topic were relevant. Secondly, the bibliometric analysis allowed the interpretation and the evaluation of the bibliographic portfolio regarding its impact factor, the scientific recognition of the articles, the publishing year and the highlighted authors. Thirdly, the systemic analysis carried out thorough reading of all selected articles to identify the main researched problems, the proposed goals and resources, the unfilled gaps and the opportunities.

* Corresponding author. Tel.: +55-41-98487-4449.
E-mail address: edersonfernandes@alunos.utfpr.edu.br

2351-9789 © 2019 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the scientific committee of the Flexible Automation and Intelligent Manufacturing 2019 (FAIM 2019)
10.1016/j.promfg.2020.01.012

Keywords: Machine Learning; Process Mining; Process Optimization; Bibliometric Analysis; Systemic Analysis.

1. Introduction

Industrial competitiveness challenges have been increasing with the advent of the Industry 4.0. The fourth industrial revolution puts great emphasis on the smart factory, which encompasses the development of intelligent production and manufacturing processes, the new capabilities through communication between the physical and virtual environment, and the interoperability among computer systems [1].

Process management is an important component of the smart factory and has been deployed in several ways to help productivity improvement. Unforeseen events can cause deviations from procedures, resulting in excessive downtime, overproduction, defects and inventory issues. In this way, process management methods can provide valuable insights for organizations, such as the identification of bottlenecks and time lags [2].

Process optimization depends on the field engineer's knowledge and expertise. Yet, many companies still have been managing processes for one single process or workflow at a time, comparing results through regular meetings, mappings and manual documents [3]. However, the advent of Artificial Intelligence (AI) has enabled computer science to invent efficient and new solutions to human problem-solving tasks [4]. One of the main branches of AI, that gives computers the ability to learn without being explicitly programmed, is called Machine Learning (ML).

Along with ML, the use of Process Mining (PM) has increasingly grown in the manufacturing industry. [5] declare that PM is a family of techniques used to discover and improve real business processes by extracting knowledge from event logs available in process-aware information systems. Event logs are basic resources that help provide information about network traffic, usage and other conditions.

The focus of this paper, therefore, is to present a literature review for the methods of ML and PM applied to process optimization. The main goal of this research is the identification of the trends and proposed solutions from the last five years relating to the different approaches and strategies developed for process optimization, as well as finding unfilled gaps.

The sequence and explanations of how the activities carried out will be shown more detailed in section 3. In section 2 the methodological aspects will be shown. Section 4 includes the final considerations.

2. Methodological Aspects

According to [6], the research classification can be goals-based and technical procedures-based. The goals-based classification contains three main groups: exploratory, descriptive and explanatory. Technical procedures-based classification splits the research into the following categories: bibliographic, documentary, experimental, ex-post facto, survey and case-study. This study was classified as exploratory and bibliographic, respectively.

Exploratory research is research conducted for a problem that has not been studied more clearly and helps to determine the best research design, selection of subjects and data collection. Bibliographic research is based on published material, basically constituting books and scientific articles.

The goal of the presented work was obtained using the Knowledge Development Process – Constructivist (ProKnow-C). This methodology was proposed by [7] and [8] and developed to help the researchers to find the most relevant content for the research in an easier way than they were used to.

This methodology is divided into five stages: *(i)* article selection that constitutes the bibliographic portfolio, *(ii)* bibliometric analysis, *(iii)* systemic analysis, *(iv)* research question definition, and *(v)* research goal definition.

The current study was conducted only with the first three stages from the Proknow-C since the goal of this study is the identification of the trends and opportunities relating to the research. Mendeley and Microsoft Excel software were applied to this methodology for helping with the bibliographic management and data tabulation.

While executing this methodology two restrictions were applied to this process. Firstly, it was decided to check only articles from journals, and secondly, a 5-year period was selected to limit the research (i.e., between 2014 and

2018).

3. Research Development

In this study, the applicability and effectiveness of PM and ML methods for the Process Optimization are investigated. Firstly, the bibliographic portfolio was constructed in order to gather publications with relevant content and scientific recognition, and then the bibliometric and systemic analyses were executed and studied to uncover new opportunities and obtain deep knowledge about the topic.

3.1. Bibliographic Portfolio

The bibliographic portfolio must include all published articles related to the research topic, and within the period from 2014 to 2018.

The alignment for this selection was established from three research axes: Process Optimization, Machine Learning and Process Mining. For these three research axes, 17 keywords were associated with them in order to narrow the search and get the best selection for this study.

As a strategy to obtain the research axes connected to the same articles, the Boolean operators "AND" and "OR" were used to associate the keywords with each other. Hence, the searches were carried out with the following combination: ("Process Improvement" OR "Process Mapping" OR "Process Optimization" OR "Value Stream" OR "Waste Elimination") AND ("AI" OR "Artificial Intelligence" OR "Ensemble Learning" OR "Machine Learning" OR "Pattern Classification" OR "Predictive Modeling") AND ("Bottleneck Analysis" OR "Conformance Checking" OR "Petri Net" OR "Process Discovery" OR "Process Mining" OR "Workflow Net").

Thus, searches were performed in five databases: Academic Search Ultimate, Emerald Insight, IEEE, Science Direct and Engineering Village (Compendex), between July 20 and 24, 2018 through associated keywords, period and document type. The databases were selected according to their availability and their alignment with the areas of interest. These searches selected 3562 raw articles for the base.

The software Mendeley was used for helping to manage the selected articles. This tool is capable of not only managing the articles, but once imported, of identifying and eliminating all duplicate articles as well as published documents that have not been published in journals, such as conferences, books, patents, etc. Hence, from the 3562 articles, only 3557 articles were imported by the tool, due to the unavailability of 5 articles. The next step is related to the reading of the remaining article titles to determine the alignment with the research topic. Hence, the title alignment filter resulted in 782 remaining articles.

This quantity of articles was then submitted to the scientific recognition filter. Firstly, they were transferred to the software Microsoft Excel along with their titles, authors name, publication year and journal name. Then, the number of citations for each of them was verified through the Google Scholar [9] between July 24 and 28, 2018. The number of citations was also transferred to Microsoft Excel and all were arranged in descending order. This arrangement was extremely necessary to create a cut-off value and classify them for the most and less cited articles, which was determined according to Pareto's principles [10].

Pareto's principles highlighted that 80% of all citations are represented by 20% of publications. Therefore, the sum of all citations resulted in 9864 citations and 80% of this number is represented by approximately 7891 citations. It was observed the last article within the 80% had 12 citations. Therefore, the cut-off value related to scientific recognition was identified as 12 citations or more. The application of this cut-off value resulted in 193 approved articles for the base. This is illustrated in Figure 1, which shows the most cited articles classified for the bibliographic portfolio. It highlights that the most cited article [11] has 1262 citations.

The most and less cited articles were divided into two repositories called K and P, respectively, for better organization of this process. The 193 articles in K repository were applied to the abstract filter, which meant it was necessary to read all the 193 abstracts and determine which articles were aligned with the research topic. Thus, the application of the abstract filter resulted in 12 articles remaining in the K repository.

The P repository consisted of 589 articles, considered the less cited articles. These went through a different analysis criterion, over which some articles from this repository could still make part of the bibliographic portfolio. This process defined two possible conditions to analyze their alignment:

- A) Published articles from the last 2 years of this analysis (i.e., 2017 and 2018) mainly in order to not eliminate good works which could be relevant, but haven't had the proper acknowledgement because they were published too recently;
- B) Published articles more than 2 years from this analysis (i.e., 2014 to 2016) for which their authors are the same authors of the articles belonging to K repository.

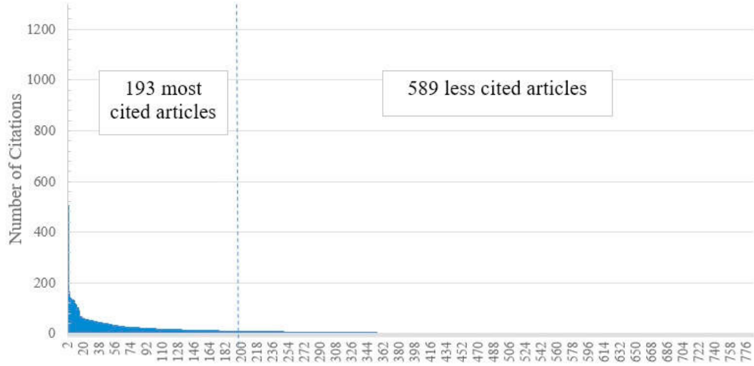

Fig. 1. Cut-off Value according to Citations.

After the condition's application, the found articles will only be approved after the abstract filter application.

Thus, 305 articles from the P repository were analyzed according to the A condition, since the publishing year was between 2017 and 2018. As with the most cited articles that went through the abstract filter, the same occurred with these articles. Hence, the application of the abstract filter resulted in 20 approved remaining articles for the A condition in P repository.

Before the analysis of the 284 articles belonging to the 2014 to 2016 publishing years, it was necessary to define which authors from the K repository would be considered at this stage. So, from the 12 articles selected in K repository, an author base was created with 57 authors. Using these 57 authors as a parameter to select new articles based on the B condition, 7 articles with the same authorship were found. However, none of these have been approved after the abstract filter application.

In conclusion for the Filters Application, 12 articles remained in the K repository and 20 in the P repository. Thus, the bibliographic portfolio ended with 32 articles for the bibliometric analysis.

3.2 Bibliometric Analysis

The bibliometric analysis allows the interpretation and the evaluation of the bibliographic portfolio. It establishes the journals' relevance regarding its impact factor, the scientific recognition of the articles, the publishing year, the highlighted authors and keywords.

For the 32 selected papers in the bibliographic portfolio, the most cited article [12] has 36 citations. These articles encompass 57 authors, and four of these authors have more than one publication. These authors are Juan-Ignacio Latorre-Biel, Emilio Jimenez-Macias, Mengshu Zhou and Hernan Ponce de Leon.

The analysis results of the Journals Relevance show that the journals: Expert Systems with Applications, and Procedia CIRP, are the most relevant journals for the research topic, with each one having three articles included in the bibliographic portfolio. Also, it is concluded that the two most relevant journals are not considered the best according to the scientific journal rankings (SJR), which are the science evaluation resources to assess worldwide universities and research-focused institutions.

An increase in the publishing rates in the last five years relating to Machine Learning and Process Mining applied to Process Optimization was detect through analysis of the publishing trends per year relating to the research topic. This is evidence that this topic has been well highlighted in academia.

Since the bibliographic portfolio was completed and the bibliometric analysis was concluded, the systemic analysis could be started as the final step of the methodology.

3.3 Systemic Analysis

The systemic analysis focused on analysing the articles' content from the bibliographic portfolio. This step works as a representativity test through a detailed evaluation to analyze their adherence to the research topic alignment and to look for knowledge opportunities for new and relevant developments.

The systemic analysis process was mainly structured according to [13]. All the articles were searched in their full version and this step was completely successful because all 32 articles from the bibliographic portfolio were available. During the full readings, unfortunately, three articles were eliminated since they were considered misaligned with the research topic. Thus, each remained article was fully read in order to address: *(i)* problem area, *(ii)* goal, *(iii)* methodology, *(iv)* main results, *(v)* future recommendations, *(vi)* research opportunities, and *(vii)* existing gaps.

3.3.1. Identified Research Problems

After completely reading the articles, the problems raised by the authors could be identified. The majority are described as follows.

(i)Process Management: Problems of process management and resource allocation [1] have been growing and creating several hardships for the companies that do not invest in the development of intelligent production and manufacturing process [14]. Nowadays most organizations store their data in one or more database management systems, but they do not have the capability to analyze and gain useful insights from this data [15]. The absence of adequate process planning, or the development of insufficient data mapping, can cause waste of time and costly resources [16]–[19].

(ii)Complex models: A few systems, methods and methodologies have already been used for process optimization, but due to the technology evolution, these strategies have become less effective for factories [12]. [20] described that up-to-date models and systems are essential for production control but challenging to maintain. Continuous improvement is very important [21], since companies may reach new levels of competitiveness, reliability and accuracy [22], [23]. The use of a new system or even the combination of more systems may overcome many weaknesses from past studies [24], [25].

(iii)Decision-making support: Decision strategies in dynamic environments do not always get either desired outcomes or optimizations, and this has been a huge problem outlined by several authors [2], [3], [26], [27]. Deviations from the standard operating procedures, uncertainties, and delays [28] are examples of unforeseen events on current and future states without proper supervision. Logistics [29], [30], supply chain [31] and distributed business processes [2], [32]–[36] were identified as applications of decision support by associated authors, since the execution of real processes can be highly unpredictable and vulnerable to disruptive events [15], mainly due to either lack of maintenance [36] or human error [37].

The next section presents the proposed goals defined by the authors, encompassing these problems, as well as resource used for reaching them.

3.3.2 Proposed Goals and Resources

Resources are either actions or strategies which may be adopted by several applications for getting better outcomes. All the resources, such as tools, methods, approaches, systems and frameworks will be presented along with the proposed goals.

(i)Process Management and Modelling: AI methodologies have been constantly applied to improve strategies and eliminate failures. [1], [14], [16], [20] used the modelling language Petri Net (PN) for modelling production systems in order to support analysis and availability optimization as well as supporting resources and performance evaluation. [14] built three PN models and simulated them for modelling interactions and self-organization.

[17] used PM techniques for modelling the production planning process of a manufacturing company. They discovered the process model with both heuristic miner and analysis.

[18] extracted anomalous frequent patterns from historical logging data. Consequently, it was possible to get more accurate analysis to exhibit parallel behaviors and correlate recurrent deviations which occurred in different portions of the process.

[19] used an PN model for both path planning and collision avoidance of a machine system. [30] took advantage of a methodological approach which can be used by managers for easily mapping, visualizing unforeseen and undesired events, supply chain and understanding potential weaknesses. Similarly [28] used PM and data mining techniques for investigating and identifying possible factors that cause delay in the production system.

[29] and [30] developed comprehensible methodologies for applying both PM and ML to add context awareness for unstructured event data in logistics (i.e., the execution of exploratory, performance and conformance analysis to get the potential to improve the results).

[33] used PN to solve complex problems in an efficient way (i.e., integrating different decisions into the same optimization problem).

(ii)New approaches, frameworks and combinations: [15] presented an approach for discovering probabilistic belief networks from practical real-life application event logs.

[12] used a hybrid PN by embedding a neural network algorithm which can model a runtime environment and collaborate in making adaptation decisions while the system is running.

Similar to these proposals, [24] integrated a five-stage framework of data mining, process improvement and process ontology for managing and improving processes with high volume of data, and [25] designed and implemented a PN based Generic Genetic Algorithm (GGA) framework which can be used for optimizing any given business processes modeled in Color Petri Nets (CPN) and exploiting simulation outputs.

[22] incorporated negative information in process discovery of complex systems for deriving less complex, fitting and precise process models, as well as being very good of generalizing the right behavior for an event log. [23] modelled a new architecture using production flow schemas (PFS) and their dynamic behaviors which are validated by PN models. This system was used to improve the communication between machines and products in a modular production system.

[3] filled a gap for the lack of tools which enable an intuitive and direct comparison of multiple management cohorts (i.e., a coherent group of process instances with one or more shared characteristics). The framework supports the cohort's selection and a tridimensional visualization for comparing their performance metrics. Among their resources, there is a set of plug-ins which includes a PM open source framework called PROM with two real-life datasets.

Due to the limited time required to obtain a solution, either the manufacturing control or decision-making problems may require unaffordable computations resources. Thus, [27] proposed an efficiency improvement throughout simulation-based optimization, PN models and genetic algorithm metaheuristics.

(iii)Predictive systems: [34] used a combination of ML techniques and PM features for predicting performance, and [35] applied both deep learning and recurrent neural networks for predicting future events in a business process.

[36] used a new system for tracking maintenance production along with a domain maintenance ontology. [2] used recommendation systems for incorporating factorization machines in event prediction tasks. [38] created a new predictive modelling technique based on both previous weaker biases, PM and grammatical inference which accurately predicts and provides comprehensible results. [26], [32] used a methodology based on ML techniques which may detect errors and predict failures with high accuracy. [37] used PM techniques to estimate human error probabilities when required tasks are conducted. For the presented solutions and results, a few opportunities were identified, which will be discussed in the next section.

3.3.3. Unfilled Gaps Identification

For the raised problems and their solutions, a few unfilled gaps could be identified.

(i)Additional Features and Programming: There are many improvement opportunities from proposed solutions which the researchers could not completely implement, and suggested should be the focus of additional research. [26] declared big potential for their ML approach improvement, PROCEDO, along with data mining. Similarly, [2], [3], [20], [23] and [31] suggested additional features for reaching better performance, response time, service availability and richer understanding of their process vulnerabilities. PM applications for logistician intelligence

support are considered to have several opportunities for enhancing logistics process transparency, strengthening the internal control of logistics firms and improving performance [29], [31].

(ii)Smart method adaptation: Several combinations of ML workflows, PM algorithms and PN language improvement may have potential for full execution of the activities [17], [21], [25], [27]. One of the most interesting developments of smart systems has been applied for logistics and supply chain [29]–[31], and this has opened possibilities for new improved manufacturing methods (e.g., Lean Manufacturing Method and Total Quality Management).

(iii)Environmental Adaptability: Smart systems applications which have had good outcomes in one specific environment may be advantageous in other environments [1]. Methodologies such as [28] and [34], through suitable adaptations may be tried out in different environments and processes for getting new accurate outcomes.

(iv)Smart Validation Dataset: Difficulties and limitations could be identified regarding the method validation process (i.e., in real applications there was not enough collected data for accurately simulating and validating them). Also, the study could be subject to insufficient conclusions since relevant knowledge and information may not have been considered. [14], [34], [35]. There are other developments which have obtained great results, however they still demonstrated dependency on manual experts' intervention in either validating or making final decisions [15], [36].

Therefore, the study conducted highlights the scope for continued research and further improvement based on the unfilled gaps.

4. Conclusions

The study has aimed a structured methodology called ProKnow-C in a bibliographic and systemic analysis of Machine learning and process mining applied to process optimization.

Firstly, the bibliographic portfolio was created, whereby from 3562 published articles found in five databases between 2014 and 2018, only 32 articles relating to the topic were approved.

Secondly, the bibliometric analysis allowed the interpretation of the bibliographic portfolio and the evaluation, through comparative charts, of the journals' relevance regarding its impact factor, scientific recognition of the articles, publishing year, highlighted authors and keywords.

Thirdly, the systemic analysis carried out thorough reading of all selected articles in order to identify the main research problems, proposed goals and resources, and the unfilled gaps. The research opportunities identified were: *(i)* smart method adaptation; *(ii)* environmental adaptability; *(iii)* additional features and programming; and *(iv)* smart validation dataset.

ProKnow-C methodology has been very effective for the understanding of current problems, goals, gaps and opportunities in several areas. However, it demands a lot of time and dedication from the researcher, due to the huge number of articles found in the databases and the application of filters, which caused many challenges during the analyses. In future studies deeper research is recommended into references from each article in this bibliographic portfolio. This will ensure more relevant articles and identify not only the recent authors but the ones who coined the concepts relating to this research topic.

References

[1] Y. Zeng and Y. Yin, Virtual and Physical Systems Intra-referenced Modelling for Smart Factory, Procedia CIRP, vol. 63, pp. 378–383, 2017.
[2] W. L. J. Lee, D. Parra, J. Munoz-Gama, and M. Sepúlveda, Predicting Process Behavior Meets Factorization Machines, Expert Syst. Appl., 2018.
[3] M. T. Wynn *et al.*, ProcessProfiler3D: A visualisation framework for log-based process performance comparison, Decis. Support Syst., vol. 100, pp. 93–108, 2017.
[4] D. Paschek, C. T. Luminosu, and A. Draghici, Automated business process management – in times of digital transformation using machine learning or artificial intelligence, MATEC Web Conf., vol. 121, p. 04007, 2017.
[5] A. R. C. Maita, L. C. Martins, C. R. López Paz, M. Fantinato, and S. M. Peres, Process mining through artificial neural networks and support vector machines: A systematic literature review, Bus. Process Manag. J., vol. 21, no. 6, pp. 1391–1415, Oct. 2015.
[6] A. C. Gil, Como Elaborar Projetos de Pesquisa, 3rd ed. Belo Horizonte: Atlas, 1996.
[7] L. Ensslin, O Design Na Pesquisa Quali-Quantitativa Em Engenharia De Produção – Questões Epistemológicas the Design in the Quali-Quantitative Research in the Production Engineering – Epistemological Issues, Rev. Produçao line, vol. 8, no. 48, p. 16, 2008.
[8] M. B. M. A. J. Eduardo Tasca, L. Ensslin, S. Rolim Ensslin, An approach for selecting a theoretical framework for the evaluation of training

programs, J. Eur. Ind. Train., vol. 34, pp. 631–655, 2010.
[9] Google Scholar. [Online]. Available: http://scholar.google.com.
[10] M. E. Newman, Power laws, Pareto distributions and Zipf's law, Contemp. Phys., vol. 46, pp. 323–351, 2005.
[11] C. L. Philip Chen and C.-Y. Zhang, Data-intensive applications, challenges, techniques and technologies: A survey on Big Data, Inf. Sci. (Ny)., vol. 275, pp. 314–347, 2014.
[12] Z. Ding, Y. Zhou, and M. Zhou, Modeling Self-Adaptive Software Systems With Learning Petri Nets, IEEE Trans. Syst. Man, Cybern. Syst., vol. 46, no. 4, pp. 483–498, 2016.
[13] T. R. Stewart, Improving Reliability of Judgmental Forecasts, in *Principles of Forecasting: A Handbook for Researchers and Practitioners*, J. S. Armstrong, Ed. Boston, MA: Springer, 2001, pp. 81–106.
[14] F. Long, P. Zeiler, and B. Bertsche, Modelling the production systems in industry 4.0 and their availability with high-level Petri nets, IFAC-PapersOnLine, vol. 49, no. 12, pp. 145–150, 2016.
[15] T. Savickas and O. Vasilecas, Belief network discovery from event logs for business process analysis, Comput. Ind., vol. 100, pp. 258–266, 2018.
[16] J.-I. Latorre-Biel, J. Faulin, A. A. Juan, and E. Jimenez-Macias, Petri Net Model of a Smart Factory in the Frame of Industry 4.0 BT - 9th Vienna International Conference on Mathematical Modelling, IFAC-PapersOnLine, vol. 51, no. 2, pp. 266–271, 2018.
[17] M. ER, N. Arsad, H. M. Astuti, R. P. Kusumawardani, and R. A. Utami, Analysis of production planning in a global manufacturing company with process mining, J. Enterp. Inf. Manag., vol. 31, no. 2, pp. 317–337, Jan. 2018.
[18] L. Genga, M. Alizadeh, D. Potena, C. Diamantini, and N. Zannone, Discovering anomalous frequent patterns from partially ordered event logs, pp. 1–44, 2018.
[19] J. Li, X. Meng, M. Zhou, and X. Dai, A Two-Stage Approach to Path Planning and Collision Avoidance of Multibridge Machining Systems, IEEE Trans. Syst. Man, Cybern. Syst., vol. 47, no. 7, pp. 1039–1049, 2017.
[20] P. Denno, C. Dickerson, and J. A. Harding, Dynamic production system identification for smart manufacturing systems, J. Manuf. Syst., 2018.
[21] R. Vrabič, D. Kozjek, and P. Butala, Knowledge elicitation for fault diagnostics in plastic injection moulding: A case for machine-to-machine communication, CIRP Ann., vol. 66, no. 1, pp. 433–436, 2017.
[22] H. Ponce-de-Leon, J. Carmona, and S. K. L. M. vanden Broucke, Incorporating negative information in process discovery BT - 13th International Conference on Business Process Management, BPM 2015, August 31, 2015 - September 3, 2015, 2015, vol. 9253, pp. 126–143.
[23] M. A. Pisching, M. A. O. Pessoa, F. Junqueira, D. J. dos Santos Filho, and P. E. Miyagi, An architecture based on RAMI 4.0 to discover equipment to process operations required by products, Comput. Ind. Eng., 2018.
[24] M. Khanbabaei, F. M. Sobhani, M. Alborzi, and R. Radfar, Developing an integrated framework for using data mining techniques and ontology concepts for process improvement, J. Syst. Softw., vol. 137, pp. 78–95, 2018.
[25] Y.-W. Si, V.-I. Chan, M. Dumas, and D. Zhang, A Petri Nets based Generic Genetic Algorithm framework for resource optimization in business processes, Simul. Model. Pract. Theory, vol. 86, pp. 72–101, 2018.
[26] G. Meyer *et al.*, A machine learning approach to improving dynamic decision making, Inf. Syst. Res., vol. 25, no. 2, pp. 239–263, 2014.
[27] J.-I. Latorre-Biel, E. Jiménez-Macías, M. des Pérez de la Parte, J. Blanco-Fernández, and E. Martínez-Cámara, Control of Discrete Event Systems by Means of Discrete Optimization and Disjunctive Colored PNs: Application to Manufacturing Facilities., Abstr. Appl. Anal., pp. 1–16, Jan. 2014.
[28] R. Gerhardt, J. F. Valiati, and J. V. Canto dos Santos, An Investigation to Identify Factors that Lead to Delay in Healthcare Reimbursement Process: A Brazilian case, Big Data Res., 2018.
[29] Y. Wang, F. Caron, J. Vanthienen, L. Huang, and Y. Guo, Acquiring logistics process intelligence: Methodology and an application for a Chinese bulk port, Expert Syst. Appl., vol. 41, no. 1, pp. 195–209, 2014.
[30] T. Becker and W. Intoyoad, Context Aware Process Mining in Logistics BT - 50th CIRP Conference on Manufacturing Systems, CIRP CMS 2017, May 3, 2017 - May 5, 2017, 2017, vol. 63, pp. 557–562.
[31] J. Blackhurst, M. J. Rungtusanatham, K. Scheibe, and S. Ambulkar, Supply chain vulnerability assessment: A network based visualization and clustering analysis approach, J. Purch. Supply Manag., vol. 24, no. 1, pp. 21–30, 2018.
[32] M. Borkowski, W. Fdhila, M. Nardelli, S. Rinderle-Ma, and S. Schulte, Event-based failure prediction in distributed business processes, Inf. Syst., 2017.
[33] J.-I. Latorre-Biel, E. Jiménez-Macías, and M. Pérez-Parte, Sequence of decisions on discrete event systems modeled by Petri nets with structural alternative configurations, J. Comput. Sci., vol. 5, no. 3, pp. 387–394, 2014.
[34] R. Umer, T. Susnjak, A. Mathrani, and S. Suriadi, On predicting academic performance with process mining in learning analytics, J. Res. Innov. Teach. Learn., vol. 10, no. 2, pp. 160–176, Jul. 2017.
[35] J. Evermann, J.-R. Rehse, and P. Fettke, Predicting process behaviour using deep learning, Decis. Support Syst., vol. 100, pp. 129–140, 2017.
[36] M.-H. Karray, B. Chebel-Morello, and N. Zerhouni, PETRA: Process Evolution using a TRAce-based system on a maintenance platform, Knowledge-Based Syst., vol. 68, pp. 21–39, 2014.
[37] J. Park, J.-Y. Jung, G. Heo, Y. Kim, J. Kim, and J. Cho, Application of a process mining technique to identifying information navigation characteristics of human operators working in a digital main control room – feasibility study., Reliab. Eng. Syst. Saf., vol. 175, pp. 38–50, Jul. 2018.
[38] D. Breuker, M. Matzner, P. Delfmann, and J. Becker, Comprehensible Predictive Models for Business Processes, MIS Q., vol. 40, no. 4, pp. 1009-A9, Dec. 2016.