PREDICTIVE MACHINE LEARNING FOR RECOMMENDATION SYSTEM IN BIG DATA UNSTRUCTURED BUSINESS PROCESSES

[1]SAMIA CHEHBI-GAMOURA, [2]RIDHA DERROUICHE, [3]HALIL-IBRAHIM KORUCA, [4]HIZIA KERROUCHE

[1]EM Strasbourg Business School Strasbourg University France,
[2]Isprata University, Turkey, [3]Hassiba Ben Bouali University, Chlef, Algeria
E-mail: [1]samia.gamoura@em-strasbourg.eu, [2]ridha.derrouiche@em-strasbourg.eu, [3]halilkoruca@sdu.edu.tr, [4]ramygassi@gmail.com

Abstract - In the current era of Big Data, mutations in Business Process Management (BPM) remain poorly understood where organizations are confronted to a growing complexity of Business Processes (BP). Due to continuous and incessant of unexpected changes, Unstructured Business Processes (UBP) become the most crucial issues in the area of Big Data business management. The proposed approach in this paper introduces a new Machine Learning (ML) approach, able to optimize unexpected exceptions, and reduce time-consuming in UBP. A new Reinforcement Learning algorithm is proposed to predict the best action to undertake and avoid unexpected paths in a recommending aid-to-system architecture. For empirical proof, a simulation case study is applied with key findings and validation outcomes. The results reveal how the method can preserve robustness despite unpredicted alterations. Moreover, this paper provides large spectra of academic bibliography as an interesting background for UBP, which still a rarely discussed topic in research works.

Index Terms - Prediction, Machine Learning, Big Data, Business Process Management.

I. INTRODUCTION

Irregularity, changing workflows, and multifaceted paths are the most hurdles in developing and executing Business Processes (BPs) [1]. Furthermore, with the advent of velocity in Big Data environment there is more complexity in executing automated workflows of BPs[2]. Continuous fluctuations in business behavior[3], integration of man-madetasks[4], transformations in internal services [5], changes in laws [6], adaptation of security rules [7], structural reforms in organizations [8], adaptation to unexpected situations are the main causes of this complexity, and for forth[9]. These issues have given way to a research problem called “Unstructured Business Process (UBP)” which has been the concern of some academic works in recent years ([1], [10], [6]). However, up-to-now, research efforts still insufficient and scare [11]. The purpose of this paper is to bond this break in academic context.

Picking the optimum action in a set of options in a BP has to be articulated via a strategy that aims to improve one or many indicators, such as the achievement ratio of the expected results [12], the Quality of Service (QoS) [13], and the time-consuming tasks [14]. To deal with these issues, academic efforts are releasing event-log orientedapproaches and process-oriented mining methodsthat have been usedincreasingly since years[15]. However, although their prevailing in the escaping biased execution paths and reducing changes-integration costs[3], these algorithms struggle in situations of high velocityand high-scaled of BPs inBig Data environments[16].

To face both problematics of UBP and Big Data velocity, this paper proposesa predictive approach based on Reinforcement Learning (RL), one of a wide-used Machine Learning (ML) techniques. The heuristic has the ability to predict the best action to perform to improveBP under changing. In divergence with the reactive approaches that are based on event-logs systems mostly used in Business Process Reengineering (BPR)[17].Our approach is predictive and is connected with an engine of recommending aid-to-system architecture to assist BP stakeholders.

The sections of the paper are as following: Section 1presents an overview of relevant literature and research lacks in the topic. Section 2 describes the proposed approach. Section 4 illustrates experimentation with a case study and key findings. The final section concludes with outcomes and open views.

II. BACKGROUND AND RESEARCH GAP

A. Unstructured Business Process

Authors in [18] cited “Real-world processes are sometimes executed with little structure, imperfect informationand unforeseen exceptions, leading to the emergence of UBP”. Because of the human participation in decision practices, and the adjustment of automation by irregularitiesand workarounds, frequent alterationshappen in BP workflows [1]. Real-world examples are amply: employmentin human resources services, Information Technology (IT) incidents, insurance claims management, safety and security rules management, medical emergency service, and so forth. This motivated the advent of research worksto deal with the issue of unstructured BPs as we synthetize in the following:Fluctuations in execution processes[1],diminutionof processes’ visibility[19], diminishingof efficiency [21],volatileprioritizing of activities[22], inability to sustain rapid and real-time changes [23],and so forth.

Patterns of optimization in UBP are manifold, such as the facility of tasks to circumvent automated streams of processes to generateunusual branches [17], and the possibilityto parallelize actionsin processes[6]. Through the examination of the state-of-the-art, we noticethere is a lack in a unitedconsent in UBPresearch works, comparing with the plenty publications in Business Process Improvement (BPI) ([23], [24]), Business Process Reengineering (BPR) ([17], [25]), and Business Process Model and Notation (BPMN) ([26], [27]). Indeed, publications in UBP are scarce and the most relevant we provideare reviewed in Table 1.

Table 1. The most relevant research works in UBP

Reference	Description	Contribution
[24]	Critical analysis of integrating process	Survey
[5]	Methodology to optimize UBP	Methodology/approach
[25]	Framework to integrate data models and BPs	Framework/model
[26]	Business mining method based on Particle Swarm Optimization (PSO)	Methodology/approach
[4]	Business operations based on artificial intelligence (AI)	Theoretical approach
[27]	Knowledge management in UBP	Theoretical approach
[20]	Process mining with Artificial Neural Networks (ANN) and Support Vector Machines (SVM)	Literature review
[7]	Approach in formal language (Object Z) and Petri Net	Framework/model
[1]	Approach ofBehavioral Process Mining (BPM) to optimise sub-processes	Framework/model
[19]	Predictive approach improve process changes.	Methodology/approach

As shown in table 1, different approaches exist to face the challenge of unstructured workflows in UBP, however, there is none clear-cut and comprehensive contribution. The most considerable among these them may be the analytical methods such as predictive approaches [19], [26], [32], [4], and[20]. As proposed in these works, we propose in this paper a prediction-oriented approach based on Machine Learning (ML).

B. Prediction in Unstructured Business Process

Predictive approaches aim to foresee actions before they occur so that problemscan be avoided and/or resolved proactively [33]. Proactivity is useful in facing the challenge of unstructured workflows in UBP where paths maybe predictively analysed and recommended by the algorithm. However, when investigating the academic literature in this topic, we distinguish two eras:

1. The era before the advent of Big Data Analytics (BDA) where research in applying predictive approacheswereclearly weakin BPM ([19]).Research in thiseraaredeal with techniques that acquire knowledge from experience (Schwartz., 2014). Such techniques have the ability to make predictions on data generated by BPs execution threads and events,and accordinglyadjust the behaviour of BPs in future actions[34].
2. The era Big Data Analytics where machine learning approaches play an important role in analysing BPM issues in connection with data science[35]. In Big Data context, predictive approaches offerautomation and proactivity with more facility and lower costs compared to the traditional era before the Big Data where the main purpose is industrialisingand optimizing BPs. In Big Data context, predictive machine learning approaches allow agilityin integrating in a continuous way. The management mode is driven not just by history but also by the analytical predictive know-how[35].

III. PROPOSED APPROACH

A. Proposed Version-Oriented Q-learning Algorithm

The policy selection in alternative options in UBPcan beprocessed as a multiple-stages decision-making problem [33], [14]. Indeed, RL belongs to the ML approaches that are suitable to address multiple-stages decision-making concerns to recommend the best next action to perform through experience[34]. In the mapping of RL algorithm to perform in modelling UBP, we usually define self-behaviour agents in decision points of tasks. The agenthas the ability to select the best path among the set of optional paths based on its experience. In the complete solution, we propose a recommendation system for UBP users with the using of an engine based on autonomous agents with RL algorithm.

The engine is embedded in the Information System (IS) where the BPM thread is executed. It follows two stages to be performed in parallel as illustrated in figure 1:

3. Training stage: The agents learn from system through experience (paths in the decision points

of the UBP),

4. Service stage: the IS uses the recommendations coming from the agents in the engine (outcome of the training stage).

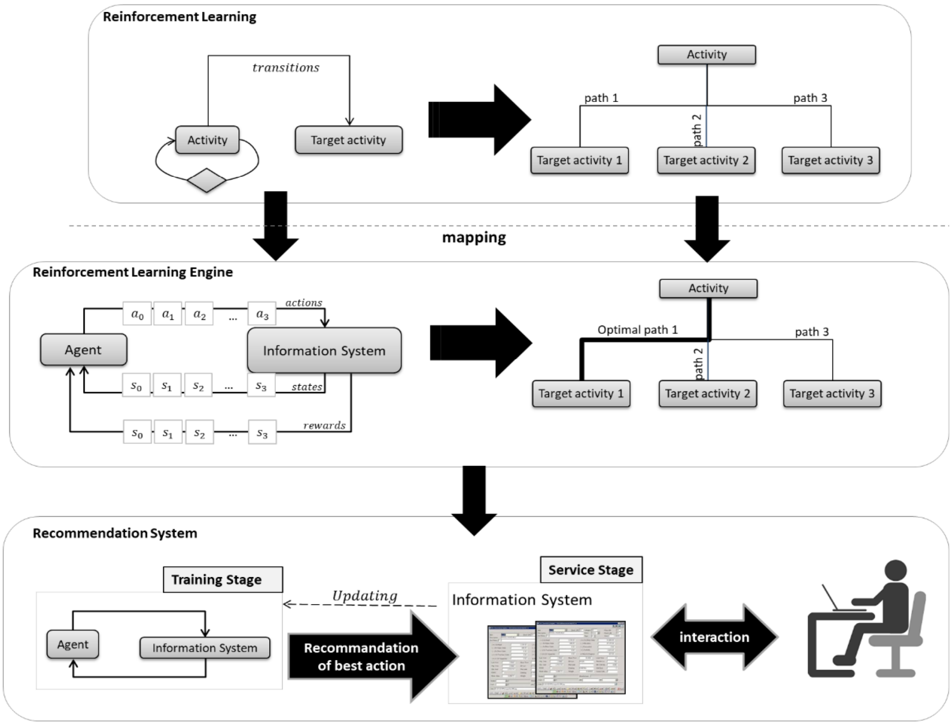

Figure 1. Proposed solution based on RL approach (mapping and recommendation engine)

Q-learning is the most common RL algorithm [35] despite the advent ofan assortment of variants in literature such as deep Q-Learning in [36], double Q-learning in [37], distributed Q-learning in [38]. As a reminder, in the following the formal equations of the q-learning algorithm:

$$Q \leftarrow Q + \alpha \,.\, (r + \gamma \,.\, \max_{history}(Q) - Q_{old}) \quad \ldots\ldots\ldots\ldots\ldots(2)$$

Where:

$\alpha \in [-1,1]$ is the learning rate,

$\gamma \in [-1,1]$ is the discount of rewards.

However, in the application onUBP, the situation is challenging as the learnt information may be overturned because of alterationsof new versions in the workflow of the BP. Therefore, in UBP, the space of agent's states may be vast and rewards may be infrequent. In this condition, the learningstage may be disparate. For this reason, we propose a new variant to adapt the q-learning algorithm to the changing versions of the UBP workflows and thus we propose a version-oriented q-learning.

The purpose in our version-oriented algorithm is that the learning rate α is reinitialized at the points where the workflow changes.

B. Proposed Recommending Engine

As aforementioned, we propose a framework of recommending as an engine based on a set of self-behavior agents that are able to learn from the system (invoving human users in the UBP) and thus optimize the decision policy (figure 1). Agents in the engine have the facility to handle inactive objects (e.g, emails, files, etc.) and activatespecific events (e.g, 'sending an email', 'pushing a notification', and 'recording a voice', etc.), after whichthey obtain rewards.

About the assessment of rewards, it is common to associate rewards with Key Performance Indicators (KPI) as proposed in [39], and [40]. Likewise, we propose to integrate the KPIs in the rewarding system as formularized in the following:

$$r = \sum_i^n f_i(KPI_i) \quad \ldots\ldots\ldots\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

Where :

$\{KPI_1, KPI_2, \ldots, KPI_n\}$isthe set of KPIs in the UBP.

$\{f_1, f_2, \ldots, f_n\}$are linear positive functions.

By introducing formula (3) in the main formula of Q-learning (2), we obtain the formula 4 to compute the Q values:

$$Q \leftarrow Q + \alpha_v \,.\, (\sum_i^n f_i(KPI_i) + \gamma \,.\, \max_{history}(Q) - Q) \ldots\ldots\ldots\ldots(4)$$

Where

$\alpha_v \in [-1,1]$is the variable learning rate with$\alpha_{v'} = 1$ if$v' > v$.

An overview of the engine's structure is illustrated through the UML class diagram in figure 2. The chart sketches the main classes: agents, events, roles, learning, and the system.

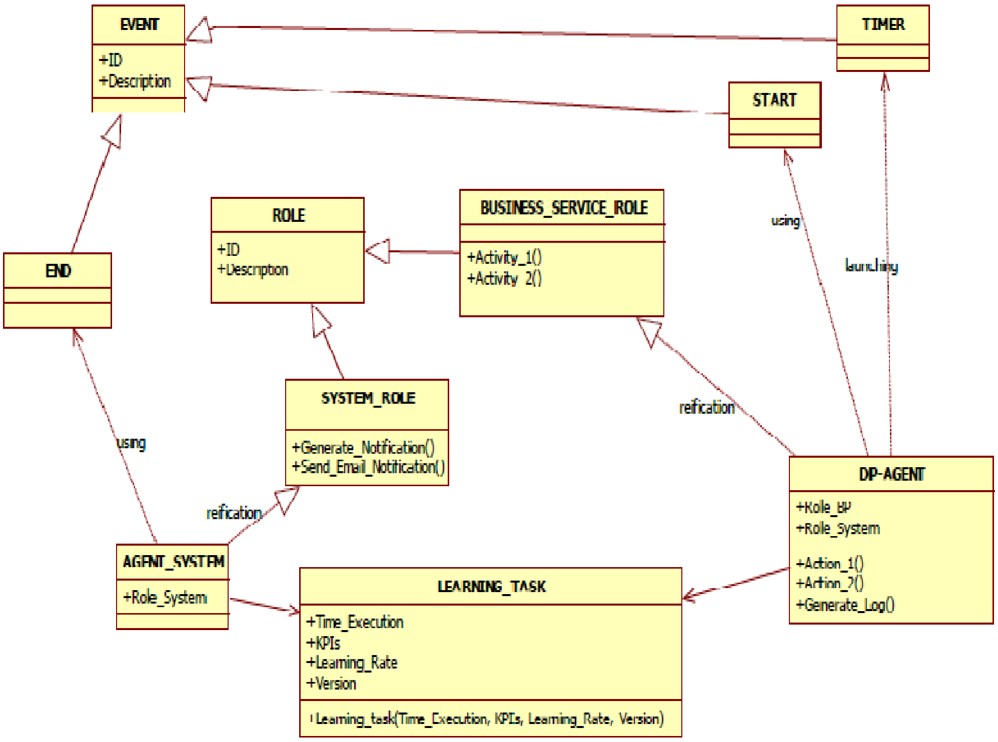

Figure 2. UML class diagram of the proposed engine

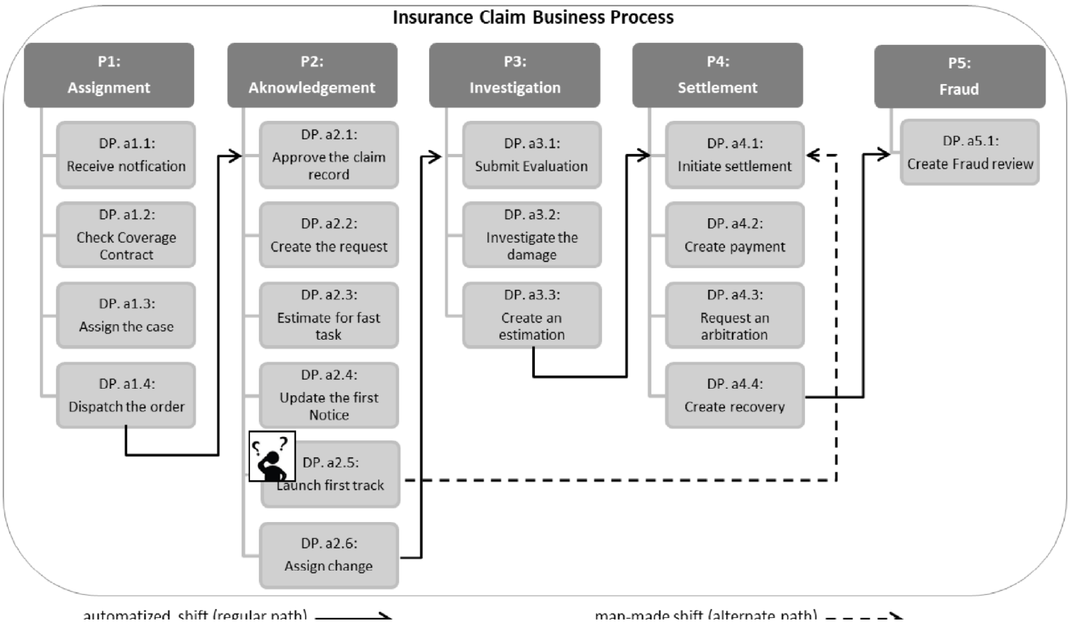

Figure 3. Illustration of the case study (insurance claim BP)

IV. APPLICATIVE CASE STUDY: INSURANCE CLAIM BP

Figure 3outlinesour case study for experimentation trials. The case illustrates the insurance claim BP. Itshows the challenge of UBP with options and changing bifurcating paths produced by users (humans) as they participate in the decision-making points. During experimentation, we selected a data set of real-world customer files in an insurance claim service belonging to the company AXA (France). The assessed values of rewards following the formula (X) are illustrated in figure 4. Unfortunately, the detailed experimentation scenarios and the data set tables exceed the limit of this paper. Therefore, for illustrative reason and to avoid extensive details we picked up only 5 iterations for one customer file.

As illustrated in figure 3, 5 main processes (activities) exist $(P1, P2, \dots, P5)$with 18 Decision Points (DP) $(DP.a1.1, DP.a1.2, \dots, DP.a5.1)$. All the decision points are automated with the option for the user (human) to sidestep the transition a2.5t -to- a4.1. as modelled in the paths-tree in figure 4.

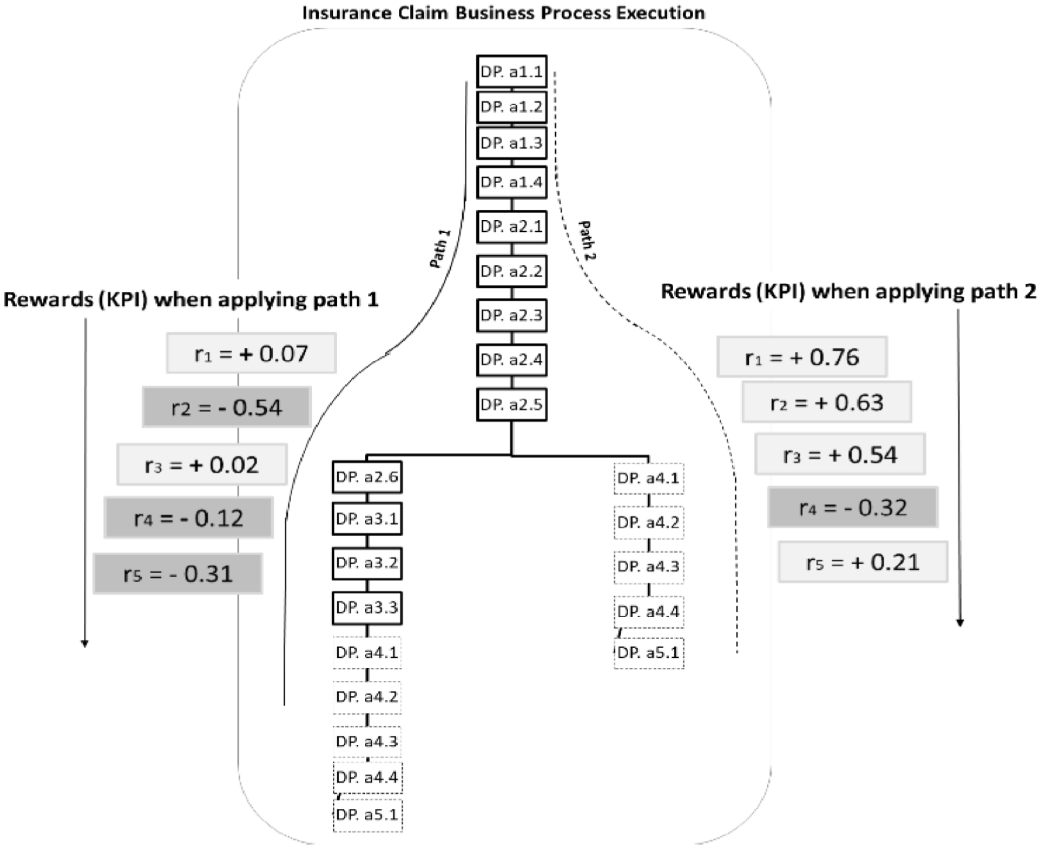

Figure 4. Tree-paths of the case study (5 first iterations)

As we observe, the rewards are different in following path 1 or path 2 for the illustrative 5 first iterations. This makes difference in the behavior of the agent at this decision point when it recommends to the user. In this illustrative situation, the agent recommends path 2 as it collects more positive rewards.

CONCLUSION AND OUTLOOK

Althoughthe multiple efforts of automation and reengineering in BPM, research efforts still in earlyprogress where some challenging issuespersist such as unstructured business processes. To fill this gap, this paper proposes apredictive approach based on reinforcement learning. The technique is a version-oriented Q-learning algorithm able to provide flexibility and elasticity to face continuous changes in the UBP. The algorithm is reified by autonomous agents that are actively interacting in the core of an engine used as recommendation system for BPM users.

A case study of insurance claim BP in AXA company is considered. The case demonstrates the achievability of the proposed approach. The outcomes of this introductive research paper is very promising and contributes to follow in the investigations of applying predictive approaches in UBP. As well, in the continuation velocity of data in such environment would highlight the reliability of our approach in a high-scalable context.

REFERENCES

[1] C. Diamantini, L. Genga and D. Potena, “Behavioral process mining for unstructured processes,” Journal of Intelligent Information Systems, vol. 47, no. 1, pp. 5-32, 2016.

[2] S. Fosso Wamba and D. Mishra, "Big data integration with business processes: a literature review," Business Process Management Journal, vol. 3, no. 477-492, p. 23, 2017.

[3] D. Nadarajah, S. A. and S. L. Kadir, "Measuring Business Process Management using business process orientation and process improvement initiatives," Business Process Management Journal, vol. 6, no. 1069-1078, p. 22, 2016.

[4] A. Siurdyban and C. Møller, “Towards intelligent supply chains: A unified framework for business process design,” International Journal of Information Systems and Supply Chain Management (IJISSCM), vol. 5, no. 1, pp. 1-19, 2012.

[5] P. Soti and S. Pandey, “Business process optimization for RHIOs,” Journal of healthcare information management: JHIM, vol. 21, no. 1, pp. 40-47, 2007.

[6] A. Souza, L. G. Azevedo and F. M. Santoro, "Automating the identification of opportunities for business process improvement patterns application.," International Journal of Business Process Integration and Management, vol. 8, no. 4, pp. 252-272, 2017.

[7] A. Ahmadikatouli and H. Motameni, “Enrichment of Object Oriented Petri Net and Object Z Aiming at Business Process Optimization,” International Journal of Advanced Computer Science and Applications, vol. 6, no. 7, pp. 13-19, 2015.

[8] J. Vom Brocke, S. Zelt and T. Schmiedel, "On the role of context in business process management," International Journal of Information Management, vol. 3, no. 486-495, p. 36, 2016.

[9] B. Hitpass and H. Astudillo, "Industry 4.0 Challenges for Business Process Management and Electronic-Commerce," Journal of Theoretical and Applied Electronic Commerce Research, vol. 1, p. 14, 2019.

[10] Z. A. Bukhsh, M. Van Sinderen, K. Sikkel and D. A. Quartel, "Understanding Modeling Requirements of Unstructured Business Processes. In ICE-B," in Proceedings of 14th International Conference on e-Business, Madrid, Spain, Springer, 2017, pp. 17-27.

[11] J. Y. Yip, R. W. Lee and E. Tsui, "Examining knowledge audit for structured and unstructured business processes: a comparative study in two Hong Kong companies," Journal of knowledge management, vol. 19, no. 3, pp. 514-529, 2015.

[12] A. Yousfi, A. De Freitas, A. K. Dey and R. Saidi, "The use of ubiquitous computing for business process improvement," IEEE Transactions on Services Computing, vol. 4, no. 621-632, p. 9, 2016.

[13] W. Viriyasitavat, L. Da Xu, Z. Bi and A. Sapsomboon, " (2018). Blockchain-based business process management (BPM) framework for service composition in industry 4.0," Journal of Intelligent Manufacturing, no. 1-12.

[14] J. Ghattas, P. Soffer and M. Peleg, "Improving business process decision making based on past experience," Decision Support Systems, no. 93-107, p. 59, 2014.

[15] S. Pourmirza, R. Dijkman and P. Grefen, "Correlation miner: mining business process models and event correlations without case identifiers," International Journal of Cooperative Information Systems, vol. 02, no. 1742002, p. 26, 2017.

[16] L. Dezi, G. Santoro, H. Gabteni and A. C. Pellicelli, "The role of big data in shaping ambidextrous business process management: Case studies from the service industry," Business Process Management Journal, vol. 24, no. 5, pp. 1163-1175., 2018.

[17] H. L. Bhaskar, "Business process reengineering framework and methodology: a critical study," International Journal of Services and Operations Management, vol. 29, no. 4, pp. 527-556, 2018.

[18] J. Y. Yip, R. W. Lee and E. Tsui, "Examining knowledge audit for structured and unstructured business processes: a comparative study in two Hong Kong companies," Journal of knowledge management, vol. 19, no. 3, pp. 514-529, 2015.

[19] K. Shoenbill, Y. Song, N. L. Cobb, M. K. Drezner and E. A. Mendonca, “IRB Process Improvements: A Machine Learning Analysis,” Journal of Clinical and Translational Science, pp. 1-8, 2017.

[20] M. Lehnert, A. Linhart and M. Roeglinger, "Exploring the intersection of business process improvement and BPM capability development: A research agenda," Business Process Management Journal, vol. 23, no. 2, pp. 275-292, 2017.

[21] A. Al-Thuhli, M. Al-Badawi, Y. Baghdadi and A. Al-Hamdani, "A Framework for Interfacing Unstructured Data Into Business Process From Enterprise Social Networks," International Journal of Enterprise Information Systems, vol. 13, no. 4, pp. 15-30, 2017.

[22] R. Rialti, G. Marzi, M. Silic and C. Ciappei, "Ambidextrous organization and agility in big data era: the role of business process management systems," Business Process Management Journal, vol. 24, no. 5, pp. 1091-1109, 2018.

[23] A. Manfreda, A. Kovacic, M. I. Štemberger and P. Trkman, "Absorptive capacity as a precondition for business process improvement.," Journal of Computer Information Systems, vol. 54, no. 2, pp. 35-43, 2014.

[24] D. Nadarajah, S. A. and S. L. Kadir, "Measuring Business Process Management using business process orientation and process improvement initiatives," Business Process Management Journal, vol. 22, no. 6, pp. 1069-1078, 2016.

[25] A. Hakim, M. Gheitasi and F. Soltani, "Fuzzy model on selecting processes in business process reengineering," Business Process Management Journal, vol. 22, no. 6, pp. 1118-1138, 2016.

[26] A. Kalenkova, A. Burattin, M. de Leoni, W. van der Aalst and A. Sperduti, "Discovering high-level BPMN process models from event data," Business Process Management Journal, 2018.

[27] M. Ramos-Merino, L. M. Álvarez-Sabucedo, J. M. Santos-Gago and J. Sanz-Valero, "A BPMN Based Notation for the Representation of Workflows in Hospital Protocols," Journal of medical systems, vol. 42, no. 10, p. 181, 2018.

[28] T. H. Davenport, J. G. Harris and S. Cantrell, “Enterprise systems and ongoing process change,” Business Process Management Journal, vol. 10, no. 1, pp. 16-26, 2004.

[29] P. Samaranayake, “Business process integration, automation, and optimization in ERP: Integrated approach using enhanced process models. ,,” Business Process Management Journal, vol. 15, no. 4, pp. 504-526, 2009.

[30] X. Fang, X. Gao, Z. Yin and Q. Zhao, “An efficient process mining method based on discrete particle swarm optimization,” Information Technology Journal, vol. 10, no. 6, pp. 1240-1245, 2011.

[31] J. Recicar and J. Tomlain, “Knowledge of Enterprise As a Tool For Business Process Optimization Using Marketing Activities,” European Scientific Journal, ESJ, vol. 10, no. 7, 2014.

[32] A. R. C. Maita, L. C. Martins, C. R. López Paz, S. M. Peres and M. Fantinato, "Process Mining Through Artificial Neural Networks and Support Vector Machines: A Systematic Literature Review," Business Process Management Journal, vol. 21, no. 6, pp. 1391-1415, 2015.

[33] A. S. Aydiner, E. Tatoglu, E. Bayraktar, S. Zaim and D. Delen, "Business analytics and firm performance: The mediating role of business process performance," Journal of Business Research, vol. 96, pp. 228-237, 2019.

[34] C. Combi and M. Gambini, “Flaws in the flow: The weakness of unstructured business process modeling languages dealing with data. In OTM Confederated International Conferences "On the Move to Meaningful Internet Systems",” Berlin, Heidelberg, Germany, 2009.

[35] S. Schulte, C. Janiesch, S. Venugopal, I. Weber and P. Hoenisch, "Elastic Business Process Management: State of the art and open challenges for BPM in the cloud," Future Generation Computer Systems, vol. 46, pp. 36-50, 2015.

[36] D. P. Bertsekas and H. Yu, “Q-learning and enhanced policy iteration in discounted dynamic programming,” Journal of Mathematics of Operations Research, pp. 37(1), 66-94, 2012.

[37] T. Hester, M. Vecerik, O. Pietquin, M. Lanctot, T. Schaul, B. Piot, D. Horgan, J. Quan, A. Sendonaris, I. Osband, G. Dulac-Arnold, J. Agapiou, J. Z. Leibo and G. A., "Deep q-learning from demonstrations.," in Proceedings of Thirty-Second AAAI Conference on Artificial Intelligence, 2018.

[38] H. V. Hasselt, "Double Q-learning.," in In Advances in Neural Information Processing Systems, 2010, pp. 2613-2621.

[39] A. Galindo-Serrano and L. Giupponi, "Distributed Q-learning for aggregated interference control in cognitive radio networks," IEEE Transactions on Vehicular Technology, vol. 59, no. 4, pp. 1823-1834, 2010.

[40] M. Y. Samiri, M. Najib, A. El Fazziki and J. Boukachour, "Toward a Self-Adaptive Workflow Management System Through Learning and Prediction Model," Arabian Journal for Science and Engineering, vol. 42, no. 2, pp. 897-912, 2017.

[41] D. Scrimieri, N. Antzoulatos, E. Castro and S. M. Ratchev, "Automated experience-based learning for plug and produce assembly systems," International Journal of Production Research, vol. 55, no. 13, pp. 3674-3685, 2017.

[42] Z. Huang, W. M. van der Aalst, X. Lu and H. Duan, “Reinforcement learning based resource allocation in business process management,” Data & Knowledge Engineering, vol. 70, no. 1, pp. 127-145, 2011.

[43] F. Niedermann, S. Radeschütz and B. Mitschang, "Business process optimization using formalized optimization patterns.," Business Information Systems, pp. 123-135., 2011.

[44] M. Brinch, "Understanding the value of big data in supply chain management and its business processes: Towards a conceptual framework," International Journal of Operations & Production Management, vol. 38, no. 7, pp. 1589-1614, 2018.

[45] K. Fertalj and M. Matejaš, “Using agility in building Business Process Management solutions. World academy of science, engineering and technology,” International Journal of Social,

Behavioral, Educational, Economic, Business and Industrial Engineering, vol. 9, no. 11, pp. 3700-3706, 2015.

[46] X. Liu, J. Chen, Y. Ji and Y. Yu, “Q-learning algorithm for task allocation based on social relation,” in Communications in Computer and Information Science, J. Cao, L. Wen and X. Liu, Eds., Berlin Heidelberg, Germany, Springer, 2014, pp. 49-58.

[47] S. Tayade and V. Chavan, "Flexibility analysis in Business Process reengineering with theory of constraint using intelligent dynamic workflow system," International Journal of Computer Applications, vol. 80, no. 12, pp. 32-39, 2013.

[48] M. E. H. Souidi, P. Songhao, L. Guo and C. Lin, “Multi-agent cooperation pursuit based on an extension of AALAADIN organisational model,” Journal of Experimental & Theoretical Artificial Intelligence, vol. 28, no. 6, pp. 1-14, 2015.

[49] H. M. Schwartz., Multi-Agent Machine Learning : a reinfocement approach, Hoboken, New Jersey, U.S.A.: Wiley, 2014.

[50] E. Zámečníková and J. Kreslíková, "Time series prediction based on event driven Business Process Management," International Journal of Computational Engineering Research, vol. 4, no. 4, pp. 76-81, 2014.

[51] J. Krumeich, B. Weis, D. Werth and P. Loos, "Event-driven Business Process Management: Where are we now? A comprehensive synthesis and analysis of literature," Business Process Management Journal, vol. 20, no. 4, pp. 615-633, 2014.

★★★