International Journal of Informatics and Applied Mathematics
e-ISSN:2667-6990 Vol. 5, No. 1, 27-51

A Road-map for Mining Business Process Models via Artificial Intelligence Techniques

Afifi Chaima and Khebizi Ali

LabSTIC, University of 8 Mai 1945,Algeria, Guelma
affifichaima@gmail.com
ali.khebizi@gmail.com

Abstract. Nowadays, the size of data recorded and stored in enterprises information systems (IS) is increasing every second. To face to this phenomenon, contemporary technologies play a major role for gathering, analyzing, storing, and distributing data that enables organizations to make smart decisions and to take full control of their activities. The traditional Business Process (BP) mining techniques were intensively used to discover, monitor, and optimize processes from event-logs without needing any priory model. However, the majority of the suggested algorithms have exhibited their limits (such as discovering nested loops, managing duplicate and hidden tasks as well as dealing with concurrent processes). In parallel, recent advances in the Artificial Intelligence (AI) discipline have generated a great deal of enthusiasm in a large spectrum of research area. In this perspective, AI methods emerge as one of the pillars to overcome the drawbacks of the conventional approaches allowing anomalies detection, prediction and recommendation tasks on ongoing process instances at run-time. The aim of this work is to explore towards the use of AI techniques in the field of business process mining by presenting a state-of-the-art review ranging from traditional PM approaches to AI ones, as well as outlining a prospective road-map for mining business process models basing on AI techniques.

Keywords: Business process · Business Process Management · Business Process Modeling · Process Mining · Machine Learning · Deep Learning ·

1 Introduction

Currently, the massive amount of data generated every second have forced the companies to operate in a more competitive environment by expressing a pressing need to evaluate their activities and inventing new ways for performing their business responsibilities. Thus, regardless of the nature of their operations, enterprises must be able to govern and maintain their Business Processes (BP) by swiftly designing new BPs or adjusting, updating, and improving those that are currently in use. This challenge has provided the emergence of Business Process Management (BPM) technology as a holistic framework that has fundamentally impacted the way Information Systems (IS) are built and deployed. In reality, business process management (BPM) technologies constitute the cornerstone of today's information systems.

A business process (BP) is considered as a set of interconnected operations that culminate in the delivery of a service or product to a customer. It must have a single output and well specified inputs [1]. The BP model is a significant notion in the BPM ecosystem for specifying the organization's operations. It is an abstract representation that explains how the needed resources are allocated at each level of the BP progression, as well as the business logic that drives the current activity. This indicates that each stage of the operation requires a distinct participant with a certain function, time, and other material resources in order to be completed successfully [2]. Indeed, the amount of data created during BP execution and kept in firms' IS rises dramatically during the BP invocation, resulting in a strong and useful new resource for the company.

A pragmatic approach to benefit from this BP data explosion consists to explore the various data sources containing log files, which must be integrated adequately in permanent storage supports. Then, applying techniques that allow data exploration, by deploying suitable mechanisms for extracting data of interest [3], all while ensuring BP evolution and maintenance. Moreover, the BP dynamicity aspect caused by the changes occurring in enterprises environments, involves organizations to become open systems tightly connected with an unpredictable and increasingly turbulent and inconstancy environment. In fact, many reasons may lead to BPs evolution, such as: changes in laws and regulations, changes in organizational structures, business processes optimization and re-engineering as well as rapid growth and changes in the Web ecosystem. Consequently, organizations must constantly update and enhance their BPs for responding to the growing rate of change in the environment. To overcome these issues, BP mining appears as a new discipline that seeks to be a promising and quickly increasing technology for assisting businesses and discovering strategies to uncover real processes that occurred in the firms. Currently, this young discipline has reached an acceptable maturity level and it performs the majority of tasks associated with the BPM life-cycle, starting from the design (by models' discovery) step to the monitoring one. The fundamental concept behind these approaches is to start with information system-generated event logs and gather knowledge from them. In practice, the discovery, the conformance checking and enhancement steps represent the pillars of BP mining approaches. First, the BP model is building automatically from the event log data on the discovery step without priory model. Second, techniques for assessing conformance Compare an existing process model to the actual process produced from the event log and determine whether or not the real process conforms to the current model. Lastly, based on actual events, the improvement approaches may be used to expand or improve the designed model. However, the majority of these business process mining algorithms, on the other hand, have limits. Loops, duplicate tasks, hidden tasks, and concurrent processes are examples of structures that cannot be mined using all approaches. As a result, existing process mining algorithms have issues with inaccurate BP model structures (e.g., The effective and valid model representation). To address the aforementioned gap, the artificial intelligence emerges as one of the pillars to power the Process Mining Algorithms and it leads to a new class of algorithms called Intelligent Process Mining Algorithms (IPMA) employing machine learning and deep learning techniques. The capabilities of IPMA are divided into five categories: Descriptive Process Mining [4] is used to gain a complete understanding of what happened in the past. Diagnostic Process Mining [5] is employed to filter out why something happened in the past. Predictive Process Mining [6] involves looking at projections of what is most likely to happen in the future. Prescriptive Process Mining [7] entails obtaining recommended steps to avoid future problems. The next important paradigm transition is toward cognitive process mining analysis [8], which may get wiser and more effective over time via its interactions with data and humans. This paper presents a systematic overview of BP mining area and the way in which it was impacted by AI techniques, as well as the relevant related works of AI-powered approaches. In top of that, this work is primarily based on a reengineering context which consists to a bottom-up approach leading to rebuild the abstract BP models, by examining only execution data. In this perspective, this article exposes several approaches ranging from conventional approaches (addressing process discovery, conformance checking and enhancement of BP models) to the AI powers PM methods (allowing anomalies detection, prediction and recommendation tasks on ongoing process instances at run-time). Finally, this investigation offers new study directions, providing a road map and recommendations for future research works.

The remainder of the paper is organized as follows: the second section is dedicated to the presentation of basic concepts and definitions useful to make the paper self-contained. In section 3, we expose the junction between AI techniques and BPs area, by focusing on AI-based process mining related approaches. Section 4 is dedicated to a detailed analysis of recent AI-based works and a discussion is initiated in order to identify potential research tracks. Section 5 concludes the paper and it traces a roadmap for future researches in the field.

2 Preliminary and background

This section introduces the definition of the business process model and its basic notions, followed by the fundamental concepts and definitions related to business process mining.

2.1 Business processes

The phenomenal growth of connectivity and the widespread use of the internet, as well as the need of companies to develop their activities, have forced them to perform their BPs in a more competitive environment in which they must exhibit a high level of agility. Thus, regardless of the nature of their operations, companies' have to strive rapid reconfiguration by building up new BPs and updating the existing ones. In these perspectives, the BP paradigm, which is the lifeblood of any business, becomes inevitable in order to simplify individual activities and assure efficient resource use. This target contributed to the appearance of the BPM technology as a comprehensive and a systematic approach, methods and techniques for completely altering the way of this firms by discovering, modeling, analyzing, improving, and optimizing their BPs.
In what follows, we give a concise definition of BP and its model, and then we formalize the notions of events and traces.

Definition 1. ***(Business Process)*** *A BP is simply defined as: A set of activities undertaken by one or more organizations in pursuit of some particular business goals [9].*

As sample examples of BP, booking a flight, ordering goods or managing retirement applications are real-life BPs, composed of a set of steps each of which integrates different business rules reflecting the business logic of the company. To facilitate the management and the maintenance of these BPs, they must respect a life-cycle principal, as dictated by BPM technology. The life-cycle [10] is built around four consecutive stages, as shown in Figure 1. To begin, processes are frequently represented using a formal or a graphical tool that is both understandable and executable. Second, when they have been established, documented, and simulated, the processes will be incorporated into the information system. Third, because they are placed in an execution environment, they must be managed and monitored. Finally, after a certain period of deployment, the collected data may be used to analyze how the process operates. For managing these crucial steps of the BPM, Business Process Management Systems (BPMS) should be used as a software application and a collection of technologies suitable for defining, automating and analyzing BPs. The BP models produced in the first step of the BPM life-cycle exemplify the crucial concept in the BPM ecosystem, since they are abstract specifications that describe the business logic supported by the present procedure of the firm. The next steps of the life-cycle present the crucial complementary stages useful for deploying, evaluating, monitoring, as well as optimizing the models and its associated data in order to ensure the agility of the real BP.

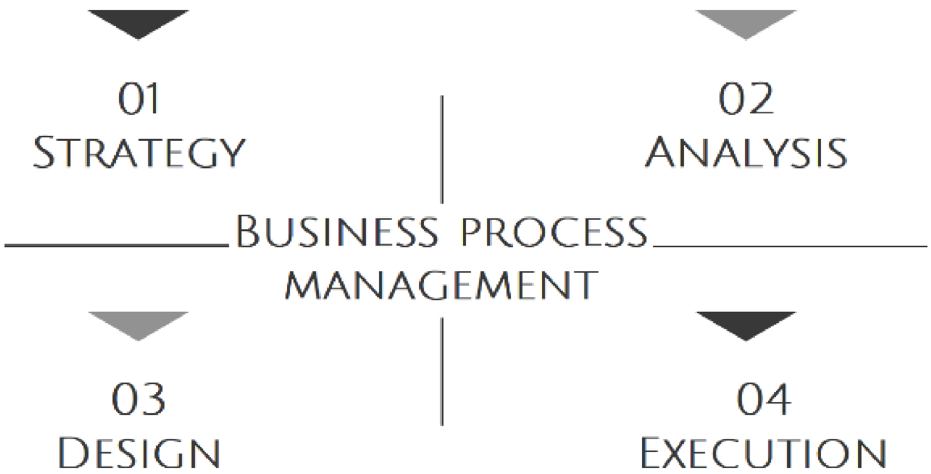

Fig. 1. BPM life-cycle [10]

To represent BPs, formal and abstract tools, such as Labeled Transitions Systems (LTS), graphs, Petri nets, finite state machines as well as graphic representations (UML diagrams), are intensively used in the literature to captures a set of constraints such as order and time ones [11]. Given this variety of BP representation models, picking the most expressive process model formalism is crucial. As Business Process Model and Notation (BPMN) formalism is the most well formal and popular tool, its used in our work to represent BP models and their related constraints.
Example 1: In figure 2, a BPNM formalism expressing an online order management BP is shown [12]. This e-commerce application has been widely adopted around the world, particularly in 2021 with the Corona pandemic and it will be used throughout the paper to illustrate various aspects of BPs management and mining. In the figure, each node represents a step of the delivery product

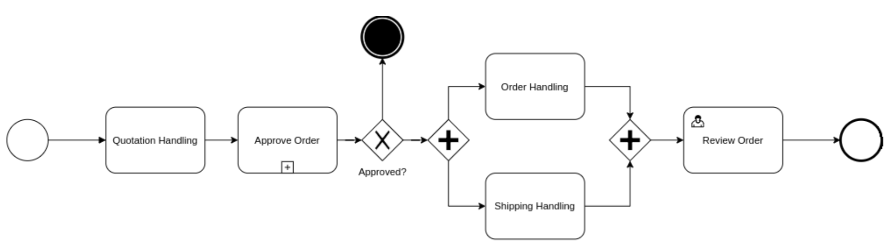

Fig. 2. BPMN of online order management application's [12]

procedure, while relation between nodes (Quotation handling, approve order) corresponds to activities (actions) to be accomplished to pass through nodes.
This model can be deployed by any company for managing the online order fulfillment process. After its deployment, such a BP is executed by an important number of customers (human, machines or invoked by others BPs). Each invocation of the BP generates a particular case or trace. Every trace is made of a series of events, each of which symbolizes the completion of a certain process activity. Three characteristics are required for each event: 1- The trace identification: indicates which trace created the event,
2- The activity name: indicates which activity the event relates to,
3- The timestamp: indicates the activity's completion time.
4- The additional than the ones listed above, we refer to all other attributes as performance attributes.
Example 2: Table 1, bellow represents two traces (or cases) containing the event logs of the BP model of the e-commerce order management applications of figure 2. A trace in this table is characterized with four attributes: Id (identifier case), customer (the name of the customer). In addition, each trace is composed of various events which are described with other properties, such as Activity name, Start time, and End time. Optionally, the trace can be enriched by other attributes, such as its textual description, the employee (who execute the operation) and other needed resources attributes. These attributes set is designated as performance parameters. Furthermore, actions inside the same trace are ordered and they constitute a control flow. Let's now define formally the aforementioned

Table 1. The event log of the online order handling application [12]

Case Attributes		Event Attributes			Performance Attributes
Trace ID	Customer	Activity Name	Start time	End Time	The employee
1	A	Quotation handling	1/1/2020 9:12	1/1/2020 9:15	Chaima
		Approve order	2/1/2020 10:12	2/1/2020 10:15	chaima
		Order handling	2/1/2020 16:12	2/1/2020 17:11	Chaima
		Review order	3/1/2020 16:12	3/1/2020 16:12	chaima
2	B	Quotation handling	1/3/2021 9:12	2/3/2021 9:20	Hadjer
		Approve order	4/3/2021 8:44	8/3/2021 11:15	Hadjer
		Shipping handling	8/3/2021 12:12	8/3/2021 16:00	Hadjer
		Review order	8/3/2021 18:20	9/3/2021 9:10	Hadjer

concepts of events and traces.

Definition 2. ***(Event)*** *An event is a tuple (a, c, t, (d1, v1), , (dm, vm)) where a is the activity name, c is the trace ID, t is the timestamp and (d1, v1), . . ., (dm, vm) (where m 0) are the event or trace attributes and their values [13].*

Example 3: Table 1, bellow shows a set of events and traces. The set of traces forms the event log data (in other word: the process variant or the execution trace). According to definition 3, a is the activity, c is the trace ID, t is the start time and (A1, V1) , (Am, Vm) (where m 0) are the event or trace attributes and their values. According to table 1, an example of event is: (Quotation handling, 1, 01/01/2020 9:12:57, (Completion time, 01/01/2020 9:15:00), (Employee, Chaima)). A trace or a case is the sequence of events produced by the execution of a particular procedure. Over the time, the various executions of the activities of the deployed BP, capture generated events reflecting historical execution traces. These executions lead to the formation of database containing the historical data or execution traces (The event log data). The formal definition of a trace is the following.

Definition 3. ***(Trace)*** *A trace E is a non-empty sequence = [e1,, en] of events such that i [1...n], ei E, and j, j [1...n] ei.c = ej.c. In other words, all events in the trace refer to the same trace [13].*

Example 4: In table 1, the first trace (Trace ID=1) stores an event log composed with a set of the following data: [(Quotation handling, 1, 01/01/2020 9:12:57, (Completion time, 01/01/2020 9:15:00), (Employee, Chaima)), (Approve order, 1, 02/01/2020 10:12:57, (Completion time, 02/01/2020 10:15:32), (Employee, Chaima)), (Order handling, 1, 02/01/2020 16:12:57, (Completion time, 02/01/2020 17:11:00), (Employee, Chaima)), (Review order, 1, 03/01/2020 16:12:57, (Completion time, 03/01/2020 16:12:57), (Employee, Chaima))] After having introduced the basic concepts of BP, in what follows we present the process mining domain.

2.2 Process Mining

Process mining is a study field that combines BPM with data analytics techniques to derive meaningful insights from process execution data [14]. Its approaches aim to support different stages of the BPM life cycle, such as process discovery, analysis, and monitoring [15]. In fact, it targets to identify, monitor, and improve real-world processes by extracting knowledge from event logs [16], which are widely available in today's information systems. More precisely;

Definition 4. ***(Process Mining)*** *Process mining is a set of approaches that connects the areas of data science and BP management to help analyze operational business processes using event logs [17].*

As depicted in figure 3, the process mining ecosystem is perceived into three complementary dimensions, based on its input and output. The first aspect consists to discover BP models from an event log without using any previous knowledge. The second one allows the conformance checking of BPs, by operating comparison between an event log and an abstract BP model (extracted from an event log or developed manually from scratch). Thus, the conformance checks if the reality is model-compliant as recorded in the log and vice versa. The third kind is related to the BP model enhancement and it aims to improve current process models by using execution information recorded in the event log, or to detect

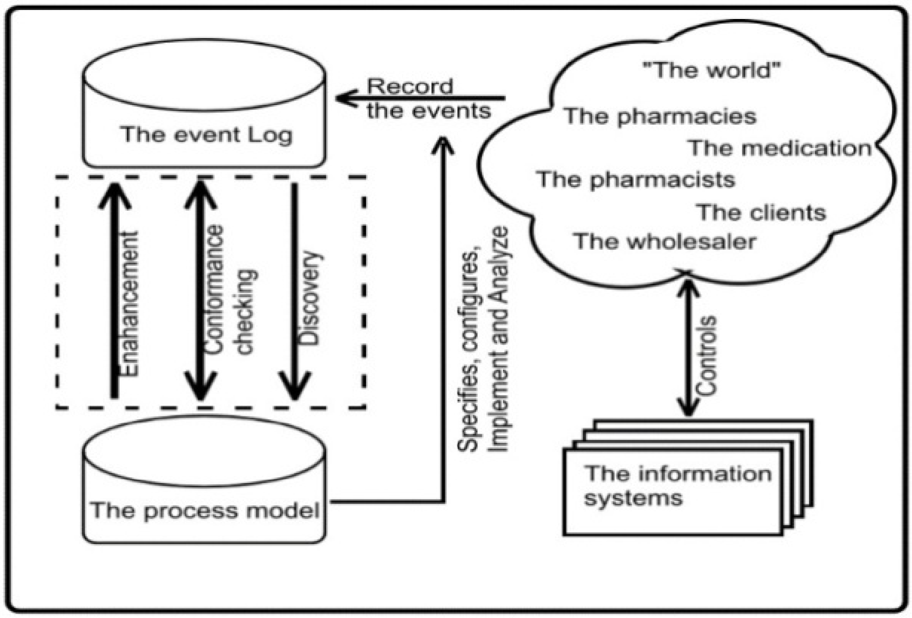

Fig. 3. Business process mining lifecycle [18]

the malfunctions and anomalies induced through conformity checking.
In the following, we focus on each aspect and provide the basic definitions that explain the innovative perspectives of process mining analytics systems.

Process discovery approaches Process discovery [19] is the initial step of process mining and is one of the most difficult activities. The general question of BP discovery has been extensively addressed and the recent research literature is very rich in approaches and techniques that tackle different facets of BP design basing on log file analysis and exploration [20, 21, 22, 23]. For example, the figure 2 shows a Petri-net graph expressing an online pharmaceutical BP model which can be extracted from event logs using process discovery approaches. In fact, these approaches exploit the event log as input to construct a model without requiring any priori data. However, facing the difficulties related to extracting abstract behavioral specification reveals a variety of problems of different natures. Van der Aalst [24] pinpoint these key challenging obstacles that remain in the field of business process mining ranging from noise, hidden tasks, duplicate tasks, mining loops, concurrent processes and the local vs global search. In general, these shortcomings led to the fact that all the conventional process discovery algorithms operate in a similar manner, scanning through the events recorded in the event log and progressively developing an abstract model that best explains the observed behavior. It either matches the present process model or necessitates certain changes to the model that have been identified thus far. When the algorithm reaches the end of the event log, it has observed all conceivable variants of the system's behavior and, as a result, produced a process model that represents the system's behavior. In the top of that, as will be revealed in the next sections, the generative models [25], the caps Net [26] and the graph nets [27] are the challenging AI-based BP models that aim to enrich and perform the conventional PM approaches by giving a novel ingenious viewpoint of the BP models for developing a features map elucidating the characteristics of the BP without designing the abstract representation of the BP models.

Conformance checking approaches Conformance checking [28] is a method for comparing event logs or the resulting process to the current reference model of the process (target model). This method is used to check if the target process is the same as the real one and it verifies its compliance. In other word, it solves a basic issue: matching a real-world process model to a theoretical one. For example, checking whether the petri net model of the online pharmacy application presented in the figure 2 conforms to the reality or the event recorded in the log of the table 1. The delta analysis [24, 29] is the vital challenge problem presented by (Van der Aalst (2001-2020)) for comparing the process model and the referenced one, by checking the similarity and disparity between them. In this case, several applications occur as a result of this challenge, such as calculating the degree of observed behavior that is contained in the process model (fitness) or the amount of modeled behavior that is really seen in the event log (precision). Regardless, the need of AI-based approaches for fulfilling a smart conformance checking mechanism is extremely prominent especially with the generative models that attempt to learn a small space in which the event log data instances can be well rebuilt. As will be underlined in the next parts, given that these types of generative models are divided into two parts: the generative part to produce BP models as well as a discriminator part, that employ the back propagation and the training logs to verify that the target output regenerated matches the real event log data for identifying fakes in the BP models [30].

Process enhancement The extension or improvement of an existing process model using pertinent information about the actual process recorded in some event logs [31] is the most cited definition of process enhancement within the process mining field. The findings of the improvement model represent the analysis quality and provide a benchmark for future analysis.
For example, the online pharmacy application presented in the figure 2 may be improved by modifying (adding, moving or deleting) specific constructs of the initial BP model. Such enhancements are induced by other occurring knowledge and could be extracted from the performance attributes or by integrating new constraints. As an illustration, reviewing timestamps in the event logs, and finding out about bottlenecks, frequencies of purchases as well sells statistics are potential factors which can impact the specification of the initial BP by involving its enhancement. These environmental variables are handled in an en-

hanced model using techniques of process discovery and analysis of conformity, where many process shortcomings include flash experiences, process loops, and undesirable process differences was the target of various AI-based BP mining solutions.

3 When Artificial intelligence meets business process mining

To achieve one of the fundamental objectives of the BP mining discipline, for discovering the BP model from execution logs, various conventional algorithms, ranging from alpha algorithm to fuzzy [32], genetic [33] and heuristic ones [34] have been proposed to cope with the related issues (e.g., the noise, the concurrent processes, the loops structures etc.). Whereas, while all these conventional techniques still suffer from innumerable limits by producing an overloaded and confused workflow model (spaghetti effects), the spectacular advances of the artificial intelligence techniques occur to overcome these drawbacks and to promote the BP mining area with additional benefits. In fact, AI can provide considerable services and achievements for the BP discipline. Considering that Machine Learning (ML) and Deep Learning (DL) resides at the root of these AI approaches, the next section gives a brief description of these techniques and discusses the impact of AI on BP mining, by exposing the induced approaches and methods.

3.1 The fundamentals of artificial intelligence

Data science is at its basis a field of study that tries to extract knowledge and insights from data using a scientific method. It is a big data concept that entails data representation, preparation, processing, and analysis through the use of the AI [35] techniques. ML and DL resides at the base of these AI approaches. The simplest way to think about their connection is to see them as overlapping rings, with AI as the earliest and greatest notion, followed by machine learning, which bloomed later, and lastly deep learning, which is driving today's AI explosion and fits inside both. In what follows, we present a brief definition of AI, ML as well as DL.

Artificial intelligence AI is a wide field of study in computer science covering a set of tools, techniques, methods and approaches for allowing machines to reason by imitating the human intelligence and reasoning principles [36].
The ultimate goal of AI is to reproduce and facilitate humans daily tasks by automating activities and allowing machines to reflect human intelligence. There are three types of AI systems.
1.Artificial Narrow Intelligence (ANI): is oriented and trained to do a particular job.
2.Artificial General Intelligence (AGI): is the ability of machines to learn, comprehend, and behave in situations where they are indistinguishable from humans.

3.Artificial Super Intelligence (ASI) is a hypothetical AI in which robots can outperform the smartest humans in intelligence.

Machine learning Machine Learning is a branch of AI that uses statistical learning techniques to create intelligent systems and it aims to answer the question: how can we build computer systems that automatically evolve with experience, and what are the fundamental laws that govern all learning processes? [37].

In general, ML tries to grasp data structures and integrate them into models that humans can use and comprehend. It enables systems to learn and evolve naturally from experience without having to be explicitly programmed [38]. In fact, ML mechanism consists of three major steps: pre-processing, feature extraction, and classification or regression. One of the key factors in the success of ML algorithms is extracting data characteristics by transforming raw data into a vector of features to train our learning algorithm how to recognize the object's features.

ML can be classified based on their intended purpose into three categories: the supervised learning (the machine learns from a well labeled training data, i.e.; some data is already tagged with the correct answer [39]), the unsupervised learning (the model works on its own to discover information, i.e.; it mainly deals with the unlabeled data [40].) and the reinforcement learning (the system learns an optimal, or nearly-optimal, policy that maximizes the "reward function" or other user-provided reinforcement signal that accumulates from the immediate rewards [41]).

The crucial success factor of the traditional ML systems relies on feature extraction to turn raw data into a characteristic vector initially designated by the ML model creator. However, identifying and formalizing characteristics of systems is not always an obvious task.

Deep learning Nowadays, Due to the data explosion, the aspect of extracting the relevant features has weakened as a result of the difficulty of the ML creator's for extracting such characteristics. This contributes to the appearance of the Deep Learning (DL) as a subset of ML techniques for automating the feature extraction step. In a nutshell, DL is defined by Yann LeCun et al as a technique that allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction [42]. The DL network architectures are ranging from the supervised algorithms as Convolutional Neural Networks (CNN) [43], Capsule Networks (Caps Net) [44], Recurrent Neural Networks (RNN) [45], transformers [46], Graph Networks (GNN) [47], to the unsupervised ones: generative models [48], as well as the deep reinforcement learning [49].

From an operational point of view that is related to the context of BP mining, AI techniques aims to consolidate BP mining field by ensuring a solid foundation and allows it benefiting from advances and advantages of AI techniques. In what follows, we expose the junction between these two fields.

3.2 AI powers business processes mining

Discovering BP models aims to construct the abstract specification of a BP from real execution data. According to conventional BP discovery techniques, as mentioned in the beginning of section 3, the target model may contain noise, hidden and redundant tasks, mining loops structures and concurrent processes. As a result, the produced model is not always an optimal one, in the sense that it may ignore strategic structures or integrate fictive ones. Consequently, the obtained model leads to tactical and strategic execution errors. AI-powered BP mining approaches use ML and DL techniques to overcome the aforementioned gap and strengthening the capabilities of the traditional process mining techniques, while making them benefit from the latest achievements of AI. Such an incursion of the AI in the field of BP has led to the appearance of the following five BP mining categories.

Descriptive process mining For activities that might benefit from the knowledge acquired from summarizing data in novel and interesting ways, a descriptive model [50] is employed. As previously stated, the first form of process mining is discovery, which is a descriptive approach for identifying and characterizing patterns while gaining a thorough understanding of real-world company activities. A statistical study of the event data was used to determine the ability to build a visual process flowchart which emphasizes the most common process flows [51]. It also enables the addition of a detail level by enhancing the model and capturing any uncommon process deviations or anomalies. In the top of list, Deep anomaly detection learning techniques as well as generative models [52, 53] aims at learning representations of the functionalities or anomaly scores using neural networks in order to identify anomalies. A vast number of techniques have been created that show much better results in dealing with difficult detections in a wide range of real-world applications and business processes than traditional anomaly detection.

Diagnostic process mining Diagnostic analytics [54] are focused on the past, such as descriptive analytics, by explaining why something happened. In other words, the next step after detecting a problem of the process is to investigate why it has happened in the business process management systems. The aim is to compare past occurrences in order to discover reasons to guide: Identify anomalies [55], patterns, and relationships. Diagnostic process mining assists to find the solution by analyzing all the process mining data accessible in the model. Analysts may apply more advanced analytics to extract the relationships between causes and effects, using the theory of probability, regression analysis and time series. The key functions and useful machine learning methods are used in the Root Cause Analysis area [56] by determining the underlying reasons of any identified process issues.

Predictive process mining The two preceding levels of smart process mining showed how problems can be recognized (descriptive process mining) and why they occurred (diagnostic process mining). The goal is to predict future issues using ML models in this third stage. Because the process mining model already has information on every completed case [57], these predictions are possible. Using this knowledge, the ML algorithm can predict the outcomes of each case in a manner that is very similar to what a human expert would predict. As the semantics represent one of the key factors for success of the Recurrent Neuronal Network (RNNs), Long Short-Term Memory (LSTMs), transformers and predictive DL methods in general which is not readily available in the majority of event logs. The challenging conception for BP prediction is employing text mining methods and methodologies of natural language processing where incoming text is transformed into word vectors using techniques such as word embedding.

Perspective process mining As soon as we have a prediction, prescriptive process mining introduces the concepts of recommender systems [58], which have been playing a vital and indispensable role in various information access systems to boost business [59] and facilitate decision-making processes [60] and are pervasive across numerous web domains, such as e-commerce [61] and/or media websites, where continuously improving business processes are critical to a company's success [62]. To accomplish this primary goal, BP modeling, which is a graphical presentation of processes in an organization, is employed [63]. Traditionally, recommender systems have relied on approaches such as clustering [64], closest neighbor [65], and matrix factorization [66] to overcome the aforementioned constraint. However, in recent years, we have seen remarkable success with what we term ML and DL for recommendation systems that are categorized into three types based on how the suggestions are generated [67]: Recommender Systems Based on Content, Recommender Systems Based on Collaborative Filtering and Recommender Systems in Hybrid Form.

Cognitive process mining The next major radical change that has impacted BP area is cognitive analysis, which takes advantage of huge advances in high-performance computers by combining a variety of clever technology generally used for specialized activities by human intelligence, such semantics, AI algorithms, and DL. By extending process management beyond process logic to business logic, a cognitive approach to BPM enables flexibility, agility, and adaptation in complex and dynamic business ecosystems. The way that BPs handle complex tasks while simultaneously maintaining continual awareness of circumstances and making real-time decisions demands a concurrently providing between humans and the technologies with which they operate [68]. Cognitive computing systems acquire experience and refine procedures over time, and use

AI and ML algorithms to manage both organized and unstructured BP data [69, 70]. They may also adapt in real time to new and ever-present changes that occur in the environment. A cognitive BP enables the BP management system to perceive, analyze and interpret business events and evolution to make BPM dynamic and agile.
Before closing this section, it's important to summarize the situation on the incursion of AI techniques in the field of PB, by reformulating the initial question induced by the use of AI techniques in order to improve BP mining approaches, while overcoming their limits. Perceiving the main challenging research question induced by smart BP mining can be expressed as follows:
How can a robust method extract behavioral models basing only on execution data and how prediction, generalization and recommender system can be built by designing an ideal feature map with much higher characteristics that carefully consider the maximum amount of information embedded in log data?
To address such concerns, intensive research works basing on AI techniques have been carried out over the past decade. An overview of these methods is presented in the next section.

4 Review and trend analysis of AI Powers BP mining approaches

This section is dedicated to the presentation of a panorama, of process mining methods and approaches having made use of AI techniques, as well a general discussion which highlights the new trends of research in this field.

4.1 Related works

A variety of approaches and methods based on AI techniques are employed at different levels of BP mining. Each of which focuses on particular aspect of the BP mining and it use its own ML or DL mechanism. In what follows, a systematic review of recent works having deal with such concerns is exposed and summarized in Table 2 bellow. Each work number in the review refers to its entry in the table.
1.Although one of the pivotal concerns of PM approaches consists to discover the BP models from the recorded event log data, evaluating and filtering these logs is a vital step towards developing an optimal and realistic model. In this regard, the cornerstone is analyzing the execution traces in real-time for reacting properly and for acquiring the right BP model as well as reducing losses and risks in organizations' activities. In spite these benefits, the real-time management can cause several troubles in the event log data as the noise and the ambiguous information...etc. In this case, managing such invalid data can lead to erroneous findings during subsequent real-time processing. For this reason, one of the difficult and crucial AI and ML tasks is to identify such confused dataset and to

discard them from the analysis process in order to construct a more accurate real-time BP. The fundamental of these AI anomaly detection approaches are the probability, distance-based, reconstruction-based, and domain-based novelty detection. In which reconstruction-based techniques are built on the concept of training neural networks that can reconstruct normal behavior but fail with abnormal ones too. To achieve such goals, the first study [71] suggest a BINET, which is a Recurrent Neural Network (RNN) trained to anticipate the following event and has attractivities that focus to identify anomalies in the multivariate detection of information in discrete events.

2.Furthermore, as previously stated, the PM conformance checking aims to provide methodologies to detect anomalies from the event log, by comparing the discovered models to the reference ones, or by reconstructing the event log from the generated model in the absence of reference ones. To fix the second case, a novel category of anomaly detection methods arises for rebuilding the event log itself through generative models such Generative Adversial Network (GAN) [72] and Variational Auto-Encoder (VAE) [73] architectures, where irregularities are determined by following the fundamental assumptions: no preceding process information, training in anomalies, no required model references, no required label and the algorithm has to eliminate the anomalous activity. In this scenario, authors in the 2nd study [74] propose a method for identifying and analyzing abnormalities in the execution of BP using an auto-encoder by reconstructing the event logs. The first step of the method encodes each activity and user using one-hot encoding. Each action in the event log is represented by an n-dimensional vector, where n is the number of different activities in the log. To encode one activity, just set the appropriate dimension of the one-hot vector to a constant value of one while leaving the other dimensions at zero. In the same way, encoding the user event attribute. Finally, a single hot-encoded vector is obtained. The auto-encoder may be trained using a one-hot event encoding log utilizing, both the input and the label with the backward propagation technique. The unique noise layer adds Gaussian noise before the auto-encoder is added. Only during training this layer is active. The auto-encoder is now trained to replicate its input, for minimizing the mean square error from input to output.

3.In the top of that, the authors of the third work [75] use deep generative models for unsupervised anomaly detection in process event data for an online real-time operational system. The major contribution of this work allows training Variational Auto-Encoder (VAE) with normal and abnormal data. Although the output model in the 2nd and 3rd previous papers is, more or less, consistent with the reality contained in the training data, this work doesnt include log dependencies during process discovery. Consequently, the target model may be incomplete and which means that some erroneous behavior cannot be identified, whereas Recurrent Neural Networks (RNNs), long short-term memory (LSTMs), and transformers with VAE and GAN can address this problem for identifying complex behaviors.

4.In addition to this descriptive BP, forecasting and advising the future activity of a running BP execution trace is the second kernel AI-powered BP approaches.

Several innovative ML methods, like as Deep Learning (DL), Neural Network (NN), RNN, LSTM, Transformers, and CNN were developed to deal with various prediction problems as: outcome prediction, remaining time prediction and suffix prediction. Although RNNs are well suited for event log data due to their sequential nature, the fourth study [76] proved that using a 1-dim CNNs can compete with RNNs for future activity prediction utilizing stacked inception CNNs modules.

5.In spite of these benefits, such machine learning techniques can include hundreds to millions of parameters to estimate, requiring a large amount of labeled training data in order to generalize well and uncover hidden patterns. However, the small size of the real-life event log available for training motivates the GAN proposed by [77] for a novel adversarial training framework to address the problem of next event prediction, whereas the proposed approach by the researches of article 5 [78] gives more relevant results by having two main parts: In the first stage, data are prepared as prefixes for the prediction task and the encoding necessary to categorize variables is adopted. The encoding of the proposed adversarial net employs a one-hot encoding. The second section provides a minimal game of fake and real prefixes between the generator and the discriminator. The real prefixes are those in the training set, and the output of the generator produces false prefixes. Since the adversarial network's generator and discriminator compete with each other, the objective of this technique is to reduce its error in discriminating between actual and false prefixes.

6.Ongoing with prediction problems, the previous log dependencies issue is still strongly present which can be solved using the text mining and natural language processing methodologies. In this context, the aim of the work 6 [79] is to forecast the process and predict the events based on the previously completed event log, as well as to evaluate the predictive performance of the next process event. One-hot encoding and min-max scaling of the data were done in the preprocessing stage to enhance prediction accuracy and allow categorical data to be used. On the basis of GPT-2 that can create a context processed representation by according attention to different parts of the input sentence, the POP-ON model adds a linear network which exploits the event attribute for the residual structure within the Transformer Decoder Layer to reflect elements in the prediction. In this case, the article 7 [80] offers another approach known as the transformer process to solve the problems of changing and complex time-specific data sequences because of multiple control flows in real-life event logs. The main contribution is a process transformer technology for learning high-level representations, which can be utilized to reason through a long sequence of networks like as LSTM from sequence event log data with minimal preprocessing step.

Finally, the enhancement step represents another key concept of AI-powers BP mining approach summarized in the optimization and stochastic techniques, among others [81]. In this scenario, the reinforcement learning is a suitable option, if there is a certain amount of uncertainty in the environment of the agent who exposes some degree of uncertainty by exploiting learned behavior vs. exploring new behaviors which might improve the result. As a consequence, the complexity of BPs expands and more states may be required to describe and optimize the environment. The research 8 [82] demonstrates that finding a rule of thumb for parameter settings in the field of business process optimization is a critical component of an efficient solution. While the article 9 [83] emphasizes that resource allocation is crucial for process development, it did not receive much attention at the time. However, as indicated [84], the topic has gotten significantly greater attention in the recent decade, as evidenced by the number of published scientific publications. This paper's research used double Deep Reinforcement Learning (DRL) to allocate resources in business operations. In order to maximize resource allocation for many processes and resources at the same time, as is required in real-world scenarios. Table 2. summarizes the most

Table 2. Related works of AI-based process mining approaches

Order	Type	Title	Approaches	Publication year
1	Descriptive BP Machine learning and Anomaly detection	BINET: Multivariate Business Process Anomaly Detection Using Deep Learning[71]	BIN et RNN	Aug. 2018
2		Analyzing business process anomalies using auto-encoders[74]	Auto-encoders	2018
3		Variational Auto-encoder for Anomaly Detection in Event Data in Online Process Mining [75]	VAE	2021
4	Predictive and recommender BP machine learning	Activity Prediction of Business Process Instances with Inception CNN Models [76]	CNN	Nov. 2019
5		Predictive Business Process Monitoring via Generative Adversarial Nets: The Case of Next Event Prediction [78]	GAN	April 2020
6		POP-ON: Prediction of Process Using One-Way Language Model Based on NLP Approach [79]	Fully attention-based transformer	2021
7		Process Transformer: Predictive Business Process Monitoring with Transformer Network [80]	Transformer Network	April 2021
8	Cognitive BP Optimization	Business Process Optimization with Reinforcement Learning [82]	DRL	July 2019
9		Deep Reinforcement Learning for Resource Allocation in Business Processes [83]	DRL	Mar. 2021

interesting works that have addressed the intelligent mining of BP, i.e.; which use one or more AI techniques to deal with the aspects of the domain. These related works are classified according to the concern of BP mining, as discussed in subsection 3.2 (descriptive BP, predictive and recommender BP, cognitive BP and optimization).

4.2 Discussion and trend analysis

Based on the related works of AI-powers PM approaches, we can argue that the effectiveness of the organization is heavily linked to their BP which must be continuously renewed, updated and, sometimes, completely redesigned. In this perspective, BP modeling is a crucial and unavoidable step, where once the abstract specification is designed (from scratch or from event logs), their conformance and optimization must be operated. Nevertheless, its not an easy task to achieve such goals, especially with the permanent explosion of the mass of data generated by the intensive deployment of BPs via the Web. Indeed, the wide range of the conventional approaches offered for extracting, updating and optimizing BPs, the generated models from the logs data still present different limits.

Nowadays, the spectacular growth of the AI techniques in the different domains might be of tremendous help to the field of BP. In the top of list, the ML and DL have been largely exploited by the BP mining leading to AI-powered BP mining area that aims to discover, enhance and optimize the BP. Such reinforced approaches aim firstly to preprocess the log data by removing the anomalies and extracting the relevant features necessary for predicting and suggesting the events, or any other type of information. Even though this discovered feature does not result in a graphical representation of the PM model, it indicates one of its powerful aspects, in which they absolutely have further knowledge to enhance the discovered PM model. Furthermore, despite the fact that optimization is an important component of any data analytics system, and getting an optimized BP model is another BP objective, few PM researchers have focused on it.

We outline that the AI-powered BP mining approaches are still far from producing perfect and optimized models that satisfy the performance criteria (growth of the model extraction time induced by the mass of data). In this case, finding an optimal model with sufficiently descriptive characteristics that take into account anomalies, semantic and hidden rules extracted from log data for robust prediction and recommendation system which, is extremely crucial, particularly with the data explosion enforced in recent years.

Finally, different perspectives can draw the road-map for future works:

1.Starting from the preprocessing step, how the relevant features can be extracted from event logs for classifying different BPs and identifying the underlying anomalies in order to filter-out a high-quality event data useful for the next steps?

2.How the basic generated model can be enhanced and enriched with further descriptive attributes such as: time constraints, further event/case attributes, the transactional constraints, as well as the performance attributes or quality of service (QoS) parameters? And what are the most expressive and pertinent attributes to be handled for enriching the BP model?

3.Whats the best form of the output discovery step? The graphical representation (a flowchart model suitable for collaboration, communication and composition) or a formal description expressing business rules (which is most adequate for operating reasoning actions), or both the two forms?
4.Whats the benefit of the semantic, hidden rules and the performance parameters included in the log data for designing the general model without over or under fitting the real BP?
5.Since the dynamicity of the PBs, what is the part that should be updated in the PB model in the case of presence and/or absence of the reference model?
6.Since the heterogeneity of BP data, how the complex behaviors can be extracted from the event log?
7.How can the prediction and recommendation beneficial to design and enhance the BPs models?
8.Since the real-time management of the BPs is pivotal for the success of the organization, how benefiting from the concepts of DRL mechanisms?
9.How can parallel and distributed computing be used to design an optimum system that works with a large amount of data?

5 Conclusion

For a rigorous understanding of firms BPs and an efficient analysis of execution traces stored in their information systems, a multi-perspective descriptive model (flowchart) is often constructed by using traditional process mining approaches. Innovative ML including the semantic, the organizational aspects, historical traces and firms business rules are recently deployed to overcome drawbacks of traditional methods. These algorithms are used in a variety of complimentary aspects of the BP, including BP extraction, prediction, recommendation, case error detection, feature detectors, and so on.
As a first step toward developing a unified and novel perception of the subject which takes into account limitations of the related techniques, this article offers a lucid review and an insightful and innovative insight survey, followed by future perspectives of BP smart management systems. In this regard, a new deep learning model, as well as novel formal and progressive imitation of human reasoning mechanism, will be used to find an optimal model with sufficient descriptive characteristics that takes into account anomalies, semantics, and hidden rules in log data. Such concerns target to elaborate a robust prediction and recommender system which is extremely crucial, particularly with the data explosion enforced in recent years. As another direction of our future works, we plan to integrate semantics in the target BP model by using ontological concepts. Further, we envisage using parallel and distributed computing (MapReduce, Hadoop) to experiment the performance of the conceived systems while deploying our approach on real-data originating from social networks.

6 References

1. Van Der Aalst, Wil MP. Business process management: a comprehensive survey. International Scholarly Research Notices, (2013). vol. 2013.
2. Xu, Jiajie, Liu, Chengfei, et Zhoa, Xiaohui. Resource allocation vs. business process improvement: How they impact on each other. International Conference on Business Process Management, Springer, Berlin, Heidelberg, (2008). p 228-243. 10.1007/978-3-540-85758-7-18
3. LI, Jingran, Tao, Fei, Cheng, Ying, et al. Big data in product lifecycle management. The International Journal of Advanced Manufacturing Technology, (2015), vol. 81, no 1, p. 667-684. https://doi.org/10.1007/s00170-015-7151-x
4. Lismont, Jasmien, Vanthienen, Jan, Baesens, Bart, et al. Defining analytics maturity indicators: A survey approach. International Journal of Information Management, (2017), vol. 37, no 3, p. 114-124.
5. Delen, Dursun et Ram, Sudha. Research challenges and opportunities in business analytics. Journal of Business Analytics, (2018), vol. 1, no 1, p. 2-12. https://doi.org/10.1080/2573234X.2018.1507324.
6. Mishra, Nishchol et Silakari, Sanjay. Predictive analytics: A survey, trends, applications, oppurtunities qnd challenges. International Journal of Computer Science and Information Technologies, (2012), vol. 3, no 3, p. 4434-4438.
7. Poornima, S. et Pushpalatha, M. A survey on various applications of prescriptive analytics. International Journal of Intelligent Networks, (2020), vol. 1, p. 76-84. https://doi.org/10.1016/j.ijin.2020.07.001
8. Handfield, Robert, Jeong, Seongkyoon, et Choi, Thomas. Emerging procurement technology: data analytics and cognitive analytics. International journal of physical distribution qnd logistics management, (2019). Vol. 49 No. 10, pp. 972-1002. https://doi.org/10.1108/IJPDLM-11-2017-0348
9. Tama, Bayu Adhi, Comuzzi, Marco, et Ko, Jonghyeon. An empirical investigation of different classifiers, encoding, and ensemble schemes for next event prediction using business process event logs. ACM Transactions on Intelligent Systems and Technology (TIST), (2020), vol. 11, no 6, p. 1-34.
10. Tenschert, Johannes, Tang, Willi, et MATZNER, Martin. Enterprise System Capabilities for Organizational Change in the BPM Life Cycle. In: International Conference on Business Process Management. Springer, Cham, (2020). p. 70-82
11. Szpyrka, Marcin, Matyasik, Piotr, Biernacki, Jerzy, et al. Hierarchical communication diagrams. Computing and Informatics, (2016), vol. 35, no 1, p. 55-83.
12. COSTA, Inaldo Capistrano et DE OLIVEIRA, Jos M. Parente. GO4SOA: Goal-Oriented Modeling for SOA. In : WEBIST (1). (2016). p. 247-254.
13. Taymouri, Farbod, La rosa, Marcello, Dumas, Marlon, et al. Business process variant analysis: Survey and classification. Knowledge-Based Systems, (2021), vol. 211, p. 106557.
14. Zhu, Li, Yu, Fei Richard, Wang, Yige, et al. Big data analytics in intelligent transportation systems: A survey. IEEE Transactions on Intelligent Transportation Systems, (2018), vol. 20, no 1, p. 383-398. 10.1109/TITS.2018.2815678.

15. Szelagowski, Marek. Evolution of the BPM Lifecycle. In: Federated Conference on Computer Science and Information Systems. (2018). p. 205-211.
16. Calvanese, Diego, Kalayci, Tahir Emre, Montali, Marco, et al. Ontology-based data access for extracting event logs from legacy data: The onprom tool and methodology. In: International Conference on Business Information Systems. Springer, Cham, (2017). p. 220-236.
17. Bozorgi, Zahra Dasht, Teinemaa, Irene, Dumas, Marlon, et al. Process mining meets causal machine learning: Discovering causal rules from event logs. In : 2020 2nd International Conference on Process Mining (ICPM). IEEE, (2020). p. 129-136. doi: 10.1109/ICPM49681.2020.00028.
18. Manoj Kumar, M. V., Thomas, Likewin, et Annappa, B. Capturing the sudden concept drift in process mining. Algorithms and Theories for the Analysis of Event Data (ATAED'15, Brussels, Belgium, (June 22-23, 2015), p. 132.
19. Cook, Jonathan E. et Wolf, Alexander L. Automating process discovery through event-data analysis. In: 1995 17th International Conference on Software Engineering. IEEE, (1995). p. 73-73. doi: 10.1145/225014.225021.
20. Verbeek, H. M. W., Buijs, J. C. A. M., Van Dongen, B. F., et al. Prom 6: The process mining toolkit. Proc. of BPM Demonstration Track, (2010), vol. 615, p. 34-39.
21. Gunther, C., Rozinat, Anne, Van Der Aalst, Wil MP, et al. Monitoring deployed application usage with process mining. BPM Center Report BPM-08-11, 2008, no 2008, p. 1-8.
22. Di Francescomarino, Chiara, Dumas, Marlon, Federici, Marco, et al. Genetic algorithms for hyperparameter optimization in predictive business process monitoring. Information Systems, (2018), vol. 74, p. 67-83.
23. Van Dongen, Boudewijn F., De Medeiros, Ak Alves, et Wen, Lijie. Process mining: Overview and outlook of petri net discovery algorithms. transactions on petri nets and other models of concurrency II, (2009), p. 225-242. https://doi.org/10.1007/978-3-642-00899-3-13
24. Van Der Aalst, Wil. Academic view: Development of the process mining discipline. In: Process Mining in Action. Springer, Cham, (2020). p. 181-196.
25. Kocaoglu, Murat, Snyder, Christopher, Dimakis, Alexandros G., et al. Causalgan: Learning causal implicit generative models with adversarial training, (2017), arXiv preprint arXiv:1709.02023.
26. Sabour, Sara, Frosst, Nicholas, et Hinton, Geoffrey E. Dynamic routing between capsules, (2017), arXiv preprint arXiv:1710.09829.
27. Leclair, Alexander, Haque, Sakib, Wu, Lingfei, et al. Improved code summarization via a graph neural network. In: Proceedings of the 28th International Conference on Program Comprehension, (2020), p. 184-195.
28. Pegoraro, Marco, Uysal, Merih Seran, et Van Der Aalst, Wil Mp. Conformance checking over uncertain event data. Information Systems, (2021), p. 101810. https://doi.org/10.1016/j.is.2021.101810.
29. Der Aalst, Van et Minig, W. Process. Discovery, Conformance and Enhancement of B.P. (2011). 1st edition.

30 Creswell, Antonia, White, Tom, Dumoulin, Vincent, et al. Generative adversarial networks: An overview. IEEE Signal Processing Magazine, (2018), vol. 35, no 1, p. 53-65.
31. Helm, Emmanuel et Paster, Ferdinand. First steps towards process mining in distributed health information systems. International Journal of Electronics and Telecommunications, (2015), vol. 61, p. 137-142.
32. Mannhard, Felix, De Leoni, Massimiliano, et Reijers, Hajo A. Heuristic Mining Revamped: An Interactive, Data-aware, and Conformance-aware Miner. In: BPM (Demos), (2017).
33. De Medeiros, Ak Alves, Van Dongen, Boudewijn F., Van Der Aalst, Wil Mp, et al. Process mining: Extending the alpha algorithm to mine short loops, (2004).
34. Weijters, A. J. M. M., Van Der Aalst, Wil Mp, et De Medeiros, AK Alves. Process mining with the heuristic's mineralgorithm. Technische Universiteit Eindhoven, Tech. Rep. WP, (2006), vol. 166, p. 1-34.
35. Sun, Shiliang, Cao, Zehui, Zhu, Han, et al. A survey of optimization methods from a machine learning perspective. IEEE transactions on cybernetics, (2019), vol. 50, no 8, p. 3668-3681. doi: 10.1109/TCYB.2019.2950779.
36. Anderson, Michael et Anderson, Susan Leigh. Machine ethics: Creating an ethical intelligent agent. AI magazine, (2007), vol. 28, no 4, p. 15-15.
37. Morency, Louis-Philippe et Baltrusaitis, Tadas. Multimodal machine learning: integrating language, vision and speech. In: Proceedings of the 55th Annual Meeting of the Assoc. for Computational Linguistics: Tutorial Abstracts, (2017). p. 3-5.
38. Campbell, Colin, Sands, Sean, Ferraro, Carla, et al. From data to action: How marketers can leverage AI. Business Horizons, (2020), vol. 63, no 2, p. 227-243.
39 Zhou, Zhi-Hua. A brief introduction to weakly supervised learning. National science review, (2018), vol. 5, no 1, p. 44-53.
40. Usama, Muhammad, Qadir, Junaid, Raza, Aunn, et al. Unsupervised machine learning for networking: Techniques, applications and research challenges. IEEE access, (2019), vol. 7, p. 65579-65615.
41. Hu, Junyan, Niu, Hanlin, Carrasco, Joaquin, et al. Voronoi-based multi-robot autonomous exploration in unknown environments via DRL. IEEE Transactions on Vehicular Technology, (2020), vol. 69, no 12, p. 14413-14423.
42. Lecun, Yann, Bengio, Yoshua, et Hinton, Geoffrey. Deep learning. nature, (2015), vol. 521, no 7553, p. 436-444.
43. Albawi, Saad, Mohammed, Tareq Abed, et Al-zawi, Saad. Understanding of a convolutional neural network. In: 2017 International Conference on Engineering and Technology (ICET). Ieee, (2017). p. 1-6.
44. Sabour, Sara, Frosst, Nicholas, et Hinton, Geoffrey E. Dynamic routing between capsules, (2017), arXiv preprint arXiv:1710.09829.
45. Yin, Wenpeng, Kann, Katharina, Yu, Mo, et al. Comparative study of CNN and RNN for natural language processing, (2017), arXiv preprint arXiv:1702.01923.
46. Vaswani, Ashish, Shazzer, Noam, Parmar, Niki, et al. Attention is all you need. In : Advances in neural information processing systems, (2017), p. 5998-6008.

47. Shi, Weijing et Rajkumar, Raj. Point-gnn: Graph neural network for 3d object detection in a point cloud. In : Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, (2020), p. 1711-1719.
48 Oussidi, Achraf et Elhassouny, Azeddine. Deep generative models: Survey. In : 2018 International Conference on Intelligent Systems and Computer Vision (ISCV). IEEE, (2018). p. 1-8.
49. Arulkumaran, Kai, Deisenroth, Marc Peter, Brundage, Miles, et al. Deep reinforcement learning: A brief survey. IEEE Signal Processing Magazine, (2017), vol. 34, no 6, p. 26-38.
50. Mendling, Jan, Baesens, Bart, Bernstein, Abraham, et al. Challenges of smart business process management: An introduction to the special issue. (2017).
51. Yeshchenko, Anton, Di Ciccio, Claudio, Mendling, Jan, et al. Visual Drift Detection for Sequence Data Analysis of Business Processes. IEEE Transactions on Visualization and Computer Graphics, (2021). doi: 10.1109/TVCG.2021.3050071.
52. Pang, Guansong, Shen, Chunhua, CAO, Longbing, et al. Deep learning for anomaly detection: A review. ACM Computing Surveys (CSUR), (2021), vol. 54, no 2, p. 1-38. https://doi.org/10.1145/3439950
53. Chalapathy, Raghavendra et Chawla, Sanjay. Deep learning for anomaly detection: A survey, (2019), arXiv preprint arXiv:1901.03407.
54. Liu, Yonghui, MAO, Guohua, LEIVA, Vctor, et al. Diagnostic analytics for an autoregressive model under the skewnormal distribution. Mathematics, (2020), vol. 8, no 5, p. 693.
55. Grahn, Kaj, Westerlund, Magnus, et Pulkkis, Gran. Analytics for network security: A survey and taxonomy. In : Information fusion for cyber-security analytics. Springer, Cham, (2017). p. 175-193.
56. Bhatt, Sanjeev Kumar, et al. Survey on Big Data Analytics: Domain Areas and Features. In : 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN). IEEE, (2020). p. 254-258.
57. Marquez-Chamorro, Alfonso Eduardo, Resinas, Manuel, et Ruiz-Cortes, Antonio. Predictive monitoring of business processes: a survey. IEEE Transactions on Services Computing, (2017), vol. 11, no 6, p. 962-977.
58. De Leoni, Massimiliano, Dees, Marcus, et Reulink, Laurens. Design and Evaluation of a Process-aware Recommender System based on Prescriptive Analytics. In : 2020 2nd International Conference on Process Mining (ICPM). IEEE, (2020). p. 9-16. 10.1109/ICPM49681.2020.00013.
59. Delias, Pavlos et Nguyen, Gia-Thi. Prototyping a business process improvement plan. An evidence-based approach. Information Systems, (2021), p. 101812. https://doi.org/10.1016/j.is.2021.101812.
60. Eili, Mansoureh Yari, Rezaeenour, Jalal, et Sani, Mohammadreza Fani. A Systematic Literature Review on ProcessAware Recommender Systems, (2021), arXiv preprint arXiv:2103.16654.
61. Akter, Shahriar et Wamba, Samuel Fosso. Big data analytics in E-commerce: a systematic review and agenda for future research. Electronic Markets, (2016), vol. 26, no 2, p. 173-194. 10.1007/s12525-016-0219-0.

62. Belhadi, Amine, Zkik, Karim, Cheherrafi, Anass, et al. Understanding big data analytics for manufacturing processes: insights from literature review and multiple case studies. Computers and Industrial Engineering, (2019), vol. 137, p. 106099. https://doi.org/10.1016/j.cie.2019.106099.
63. Galanti, Riccardo. Explainable Prescriptive Process Analytics. Computing, (2018), vol. 11, no 6, p. 962-977.
64. Seeliger, Alexander, Nolle, Timo, et Muhlhause, Max. ProcessExplorer: An Interactive Visual Recommendation System for Process Mining. In : KDD Workshop on Interactive Data Exploration and Analytics. (2018).
65. Weinzierl, Sven, Dunzer, Sebastian, Zilker, Sandra, et al. Prescriptive business process monitoring for recommending next best actions. In : International conference on business process management. Springer, Cham, (2020). p. 193-209.
66. Goldenberg, Dmitri, Kofman, Kostia, Albert, Javier, et al. Personalization in Practice: Methods and Applications. In : Proceedings of the 14th ACM International Conference on Web Search and Data Mining. (2021). p. 1123-1126. https://doi.org/10.1145/3437963.3441657.
67. Chicaiza, Janneth et Valdiviezo-Diaz, Priscila. A Comprehensive Survey of Knowledge Graph-Based Recommender Systems: Technologies, Development, and Contributions. Information, (2021), vol. 12, no 6, p. 232.
68. Wang, Hongbing Xiaojun, Hu, Xingguo, et al. A multi-agent reinforcement learning approach to dynamic service composition. Information Sciences, (2016), vol. 363, p. 96-119.
69. Dai, Hanjun, Tian, et al. Adversarial attack on graph structured data. In : International conference on machine learning. PMLR, (2018). p. 1115-1124.
70. Tovarnak, Daniel et Pitner, Tom. Normalization of unstructured log data into streams of structured event objects. In : 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM). IEEE, (2019). p. 671-676.
71. Nolle, Timo, Alexander, et Max. BINet: multivariate business process anomaly detection using deep learning. In : International Conference on Business Process Management. Springer, Cham, (2018). p. 271-287.
72. Pan, Zhaoqing, YU, Weijie, YI, Xiaokai, et al. Recent progress on generative adversarial networks (GANs): A survey. IEEE Access, (2019), vol. 7, p. 36322-36333.
73. Zhao, Dongfang, Liu, Shulin, Gu, Dan, et al. Enhanced data-driven fault diagnosis for machines with small and unbalanced data based on variational auto-encoder. Measurement Science and Technology, (2019), vol. 31, no 3, p. 035004.
74. Nolle, Timo, Stefan, Alexander, et al. Analyzing business process anomalies using autoencoders. Machine Learning, (2018), vol. 107, no 11, p. 1875-1893.
75. Krajsic, Philippe et Bogdan. Variational Autoencoder for Anomaly Detection in Event Data in Online Process Mining. (2021).
76. Di Mauro, Nicola, APPICE, Annalisa, et Basile, Teresa Ma. Activity prediction of business process instances with inception CNN models. In: Interna-

tional conference of the italian association for AI. Springer, Cham, (2019). p. 348-361.
77. Goodfellow, Ian, Pouget-Abadie, Jean, Mirza, Mehdi, et al. Generative adversarial nets. Advances in neural information processing systems, (2014), vol. 27.
78. Taymouri, Farbod, Marcello, Erfani, Sarah, et al. Predictive business process monitoring via generative adversarial nets: The case of next event prediction. In: International Conference on BPM. Springer, Cham, (2020). p. 237-256.
79. Moon, Junhyung, Park, Gyuyoung, et Jongpil. POP-ON: Prediction of Process Using One-Way Language Model Based on NLP Approach. Applied Sciences, (2021), vol. 11, no 2, p. 864.
80. Bukhsh, Zaharah A., Saeed, Aaqib, et Remco M. ProcessTransformer: Predictive Business Process Monitoring with Transformer Network, (2021), arXiv preprint arXiv:2104.00721.
81. Himanen, Lauri, Amber, FOSTER, Adam Stuart, et al. Datadriven materials science: status, challenges, and perspectives. Advanced Science, (2019), vol. 6, no 21, p. 1900808.
82. Silvander, Johan. Business process optimization with reinforcement learning. In: International Symposium on Business Modeling and Software Design. Springer, Cham, (2019). p. 203-212.
83. bikowski, Kamil, Micha, et Piotr. Deep Reinforcement Learning for Resource Allocation in Business Processes, (2021), arXiv preprint arXiv:2104.00541.
84. Arias, Michael, Saavedra, Rodrigo, Marques, Maira R., et al. Human resource allocation in business process management and process mining: A systematic mapping study. Management Decision, (2018).