

RESEARCH

Open Access



# Predictive process monitoring: concepts, challenges, and future research directions

Paolo Ceravolo<sup>1†</sup>, Marco Comuzzi<sup>2\*†</sup>, Jochen De Weerd<sup>3†</sup>, Chiara Di Francescomarino<sup>4†</sup> and Fabrizio Maria Maggi<sup>5†</sup>

<sup>†</sup>Paolo Ceravolo, Marco Comuzzi, Jochen De Weerd, Chiara Di Francescomarino, and Fabrizio Maria Maggi contributed equally to this work.

\*Correspondence: mcomuzzi@unist.ac.kr

<sup>1</sup> Università degli Studi di Milano, Milan 20100, Italy

<sup>2</sup> Ulsan National Institute of Science and Technology, Ulsan 44919, Korea

<sup>3</sup> KU Leuven, Leuven 3000, Belgium

<sup>4</sup> Università di Trento, Povo 38123, Italy

<sup>5</sup> Free University of Bozen-Bolzano, Bolzano 39100, Italy

## Abstract

Predictive Process Monitoring (PPM) extends classical process mining techniques by providing predictive models that can be applied at runtime during the execution of a business process, for example, to predict the sequence of the next event(s) in a case, its outcomes, or performance-related aspects such as the remaining processing time. These predictive models go beyond process mining's inherent descriptive nature that is offered by typical process discovery, conformance checking, and model enhancement techniques. The growing interest in PPM is driven by its additional value proposition, i.e., delivering real-time information regarding the future execution of business process instances, thus allowing for better informed operational decision making and performance analysis. The rapid growth of PPM during the last ten years has left the field lacking a cohesive taxonomy and an explicit recognition of the prevailing challenges. This paper aims at closing this gap, by comprehensively defining PPM using a unified approach to the key terms and by discussing challenges and opportunities in the field. Specifically, we propose three overarching research challenges and nine research directions, which have been validated through a survey with PPM researchers.

**Keywords:** Predictive process monitoring, Process mining, Predictive analytics, Data science pipelines

## Introduction

Predictive Process Monitoring (PPM) is an emerging subfield of Process Mining (PM) (van der Aalst 2016). It widens traditional descriptive methods, such as process discovery, conformance checking, and model enhancement with monitoring capabilities based on the application of predictive models to ongoing cases of a process (Di Francescomarino and Ghidini 2022). This “forward-looking” approach leverages historical event data collected by information systems, not just to observe organizational behavior but to project future scenarios (Pourbafrani and van der Aalst 2022). The capability to foresee future events, remaining time, or outcome of a case can be of crucial importance during decision-making (Maggi et al. 2014). For example, the allocation of resources or the intervention into a procedure can be effectively optimized by applying predictive models (Kim et al. 2022). PPM encompasses a variety of statistical techniques from data mining, predictive modeling, and machine learning. These methods can be integrated in

various ways based on their objectives, data preparation procedures, and the algorithms exploited throughout the learning or monitoring phases.

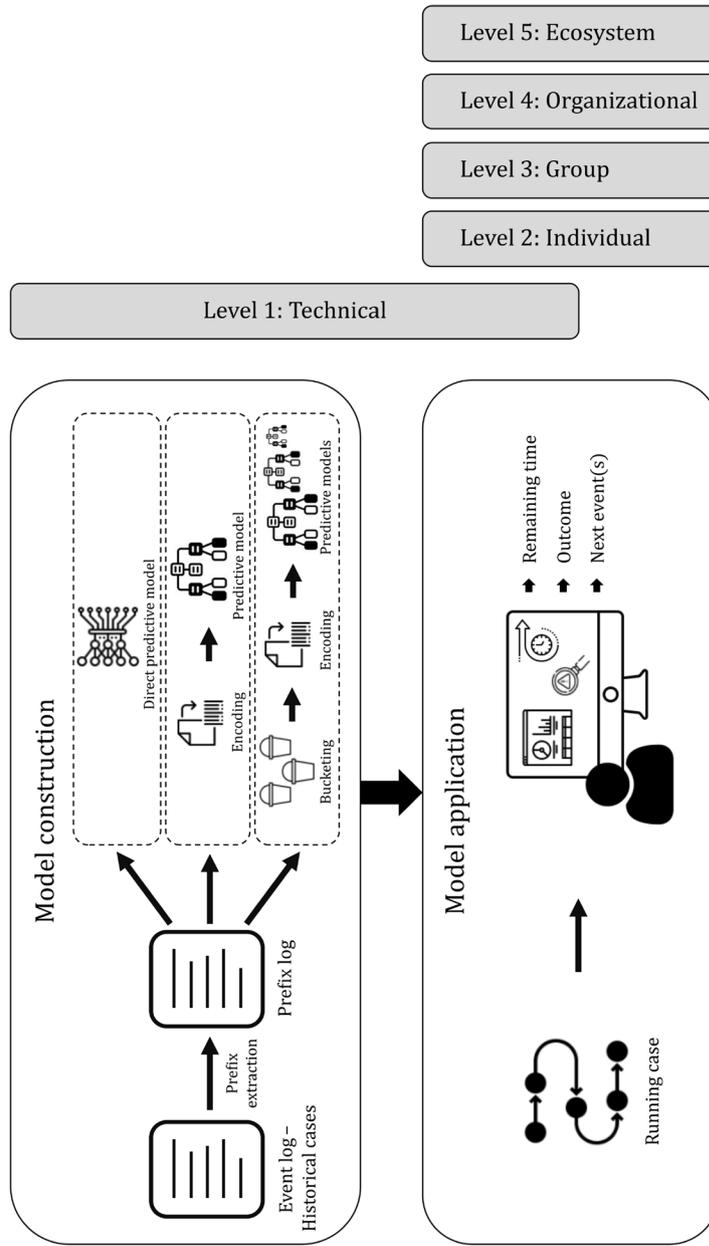
Despite the proliferation of PPM methods in the BPM literature, there is a lack of cohesive taxonomy and explicit recognition of the prevailing challenges. The first objective of this paper is to provide standardized terminology to ensure consistency, clarity, and effective communication within the field. A unified approach to the terms used facilitates better understanding and collaboration among researchers, reducing ambiguity. For instance, the term “feature encoding” has been used in different works for denoting either the feature engineering or the actual numerical encoding of the features. Additionally, we explicitly recognize and discuss the challenges in the PPM field, providing a comprehensive understanding of the current research landscape and suggesting potential solutions to overcome these challenges. This contribution aims to advance and improve research practices in PPM. We identify three major challenges for PPM, leading to nine specific future research directions. Accordingly, this paper serves as a foundational guide for researchers, both new and experienced. For those new to PPM, the paper provides important insights into the field’s objectives, achievements, and challenges. For experienced researchers, it acts as a comprehensive resource to consolidate foundational concepts, gain new perspectives, and stay updated on the evolving landscape of the field.

The remainder of the paper is structured as follows. The next section introduces the field, its main concepts, and a detailed description of the PPM workflow. Then, in [The future of the predictive process monitoring field](#) section, three main challenges are presented together with the identified future research directions. [Conclusions](#) are drawn in the last section.

### **The field of predictive process monitoring**

To guide our discussion, we present a comprehensive workflow that interlinks the activities inherent to PPM. The workflow unfolds in two phases, as illustrated in Fig. 1. Initially, there is a phase of *model construction*, in which a predictive model is learned based on a business process event log that contains historical cases. Subsequently, in the second phase, the learned model is deployed to make predictions during the real-time execution of new, previously unobserved running cases (*model application*). These predictions can then be extracted and presented to the end users, possibly in accordance with decision making protocols. Diverse objectives steer PPM, encompassing: (i) predicting a performance measure, such as the overall duration or cost of a case, or its remaining execution time; (ii) predicting an outcome, a category or state to be assigned to the executing case, like a decision, a class of risk, or the final effect produced; and (iii) predicting the sequence of the next event(s) to be observed during the execution.

In Fig. 1, we map the PPM workflow against the five levels of analysis as put forward in the framework for research on process mining by vom Brocke et al. (2021). The *technical* level refers to the design of process mining technology. The *individual* and *group* levels refer to the impact of process mining on the perceptions and behaviors of users and teams, respectively. The remaining two levels refer to how process mining affects operations and business value (*organizational*) and inter-organizational relations (*ecosystem*). A significant portion of PPM research concentrates on the technical level,



**Fig. 1** A high-level overview of the PPM workflow, including its two main phases: model construction and model application

primarily driven by design contributions in algorithmic engineering. However, there has been comparatively less emphasis on the model application phase, particularly regarding its integration into data-driven decision-making processes within organizations. Nonetheless, it is apparent that the adoption of PPM will generate impact at various levels, including the individual, group, organizational, and ecosystem levels. In [The future of the predictive process monitoring field](#) section, we will elaborate on our identified research directions mapped across these levels.

A common thread among tasks within the PPM field is the emphasis on the generation of predictions at the individual case level. Consequently, it is appropriate to categorize PPM as a subset of the broader domain of predictive analytics research, as expounded by Shmueli and Koppius (2011). Although predictive analytics techniques find application in various contexts, PPM stands out for its primary objective of providing predictions for active cases, with the ultimate aim of facilitating case-level decision making throughout the process execution. It is worth noting that predictive analytics in process management is not confined solely to individual case analysis, as researchers have also developed methods to forecast entire process models in the future (De Smedt et al. 2023).

The input of any PPM application is an event log. An **event log** is a structured collection of data capturing specific events that occur within a business process. Each event in the log typically includes (i) a unique identifier for a specific instance of the process (case ID); (ii) the name or type of the event, representing a specific step in the process (activity); (iii) the exact time when the event occurred (timestamp); (iv) information about who or what performed the event (resource); (v) additional data related to the event, such as cost, location, or status. While model application is intrinsically easy to understand, model construction involves multiple steps, as discussed below. In the initial step, the event log undergoes a transformation called **prefix extraction**, resulting in a collection of prefixes. Given the sequence of events belonging to the same instance of a process, a.k.a. trace, the prefixes are all the sub-sequences of consecutive events that includes the first one, i.e., the one characterized by the earliest timestamp. For example, if we have a complete case with 5 events, we can consider up to 4 prefixes: after the first event, after the first and second events, and so on.

Using all possible prefixes could lead to some problems. Firstly, a large number of prefixes compared to the number of traces may slow down the training of the predictive models. Second, if the length of the original cases is very heterogeneous, the longer traces produce many more prefixes than the shorter ones, and therefore the predictive model is biased towards the longer cases. Consequently, one mitigation strategy is to only consider prefixes up to a certain number of events or to filter the prefix log using so-called *gaps* (Di Francescomarino et al. 2017). Instead of retaining all prefixes up to a certain length, only prefixes with a length equal to a base number (e.g., 1) plus multiples of a gap (e.g., 1, 6, 11, 16, 21 for a gap of 5) are retained.

The prefixes extracted serve as the basis for **learning predictive models**. As depicted in Fig. 1, three primary strategies exist for learning predictive models from a prefix log. These strategies are derived from the prevailing approaches to develop machine learning pipelines in data science. The first strategy involves direct modeling, where prefixes are directly fed into the predictive algorithm as sequences, without any transformations.

This typically necessitates the use of deep learning models, such as recurrent neural networks (RNNs), to handle the complex structure of input data. In contrast, the other two strategies adopt a more indirect approach, employing **feature selection and encoding** techniques to transform prefixes into vectors. The two indirect strategies differ in terms of the presence of a **bucketing** step, in which prefixes are first grouped according to some criteria before a predictive model is trained for each bucket separately. The main advantage of indirect strategies is that, once such a representation is obtained, one can rely on a plethora of powerful predictive analytics algorithms to obtain the predictive model. The next three sub-sections discuss more in detail the steps required for learning a predictive model, i.e., bucketing, encoding, and building the predictive model.

### Bucketing

The bucketing step allows to manage diversity within process executions and it consists in grouping similar historical prefix traces into buckets in order to allow separate predictive models to be developed for each bucket. At runtime, the most appropriate bucket for the ongoing case is determined and the corresponding predictive model is used to make predictions. This can help in creating more focused and accurate predictive models, ultimately leading to better process insights and decision making. Various bucketing approaches exist:

- **State-based:** In state-based approaches, a process model is derived from the event log. Then, relevant *states*, a.k.a. decision points, are determined from the process model and a predictive model is trained for each of them. At runtime, the current state of the running case is determined, and the respective predictive model is used to make a prediction for the running case.
- **Clustering-based:** In the clustering-based approach, the buckets are determined by applying a clustering algorithm to the encoded prefix traces. This results in a number of clusters that, differently from the state-based approach, do not exhibit any transitional structure. Then, for each cluster, a predictive model is trained leveraging only the historical prefix traces that fall into that particular cluster. At runtime, the cluster of the running case is determined based on its similarity to each of the existing clusters, and the respective predictive model is applied.
- **Prefix length-based:** In this approach, each bucket contains only partial traces of a specific length. For example, one bucket contains traces where only the first event has been executed, another bucket contains those where the first and the second event have been executed, and so on. A predictive model is built for each possible prefix length.
- **Domain knowledge-based:** While the bucketing methods described so far can detect buckets through an automatic procedure, it is possible to define a bucketing function that is based on manually constructed rules. In such an approach, input from a domain expert is needed. The resulting buckets may, for instance, refer to *context categories* (Ghattas et al. 2014) or *execution stages* (Castellanos et al. 2005; Schwegmann et al. 2013).
- **Runtime similarity:** In this bucketing approach, the preparatory training phase is skipped and the buckets are determined at runtime. For example, for each running

prefix trace, its  $k$  nearest neighbors are selected from the historical prefix traces and a predictive model is trained (at runtime) based on these  $k$  neighbors (Maggi et al. 2014). This means that the number of buckets (and predictive models) is not fixed, but grows with each event executed at runtime.

### Encoding

The encoding step deals with transforming the prefixes into fixed-length numerical feature vectors. It involves both feature engineering, i.e., deciding which features to generate from the prefix data, and the actual numerical encoding of these features. The strategies devised in the literature for feature engineering are various and can be combined as required (Back and Simonsen 2023; Tavares et al. 2023):

- **Last state:** In this case, a prefix is abstracted into the last  $n$  states that it reached. A simple concrete implementation of this strategy is to generate features using the last  $n$  events seen in a prefix. The approach is similar to the  $n$ -gram technique, commonly used in text mining applications (Bose and Van der Aalst 2009).
- **Aggregation:** In this case, the prefixes are abstracted aggregating the data of all events from the beginning of the case. The frequency of occurrence can be used to aggregate activities performed or control flow sequences (Appice and Malerba 2015) that occur in a prefix, possibly considering also the position in the prefix in which they appear (Ceravolo et al. 2017). For numerical attributes, general statistics have been used, such as average, maximum, minimum, and sum.
- **Index:** By neglecting the order of the events within a prefix, the last state and aggregation strategies incur information loss. The index-based strategy (Leontjeva et al. 2015) is lossless, since it uses all possible information in a prefix, including the order of events, generating one feature per each event attribute per each executed event (i.e., each *index*).
- **Case-level:** The strategies described above apply to the event-level attributes. Other features may be generated from the case-level attributes. Typically, a feature is generated for each case-level attribute (Leontjeva et al. 2015). The values of these attributes are most often constant across the prefixes and, therefore, yield constant-valued features that are appended at the beginning of a feature vector.
- **Inter-case:** While the strategies above generate features based only on the data of the prefix that is encoded, additional features may be generated by considering data from other prefixes as well. Examples are features that capture system load (Sendrovich et al. 2019), for example, how many other cases are active while a prefix that is encoded is executed, or resource-aware features (Kim et al. 2022), for example, features that capture the past experience of a resource in executing the same event types in a prefix that is encoded, or time related features (Grinvald et al. 2021), e.g., features related to the average elapsed time or delay of an activity.
- **Embeddings:** An embedding transforms discrete or categorical data into distributed, lower-dimensional, continuous representations, that capture the semantic relationships and contextual information between different categories or entities (Azzini et al. 2021). In PPM, they have outperformed other feature engineering strate-

gies (De Koninck et al. 2018; Tavares and Barbon 2020) in classification problems. The essence of this superiority lies in the ability of embeddings to reveal nuanced patterns and dependencies within the data, thereby enhancing the model's predictive capabilities.

The choice of an encoding strategy is not without trade-offs. Lossy strategies, such as the last-state and aggregation, do not maintain information about the order of events in a prefix, which could be a limitation particularly when a prediction on the order of future activities must be made. The inter-case strategy may lead to generating a high number of features for a prefix, resulting in lower-performing models due to the curse of dimensionality. Calculating embeddings can be computationally costly. Moreover, embeddings are associated with a loss of transparency. Unlike some other feature engineering methods, the resulting feature vectors from embeddings are organized within a discovered latent space. While this space is capable of capturing complex relationships, it lacks a direct reference to the original event log attributes or their statistical properties. Therefore, the interpretability of the model is compromised to some extent, making it difficult to intuitively understand the reasoning behind certain predictions.

### Predictive model building

The final step in the phase of model construction of the PPM workflow (see Fig. 1) deals with creating the predictive model. The choice of the modeling algorithm typically depends on the prediction target. The literature has mainly considered three aspects to be predicted:

- **Process outcomes** (Márquez-Chamorro et al. 2017; Teinemaa et al. 2019; Junior et al. 2020): In this setting, the outcome of a case may refer to whether it led to a customer complaint or not, or to a product return or other claims, or whether a case was completed on time or not. Having a reliable model that can predict the outcomes of a business process can be exploited to take proactive actions when the outcome of a process is likely to be negative, e.g., allocating more resources to double-check a case that is likely to lead to a customer complaint.
- **Next event(s) in a running case** (Tax et al. 2017; Camargo et al. 2019; Tama and Comuzzi 2019; Bukhsh et al. 2021; Gunnarsson et al. 2023): In this setting, the objective is to predict the next event(s) that will be executed in the currently running process cases. Having a reliable model that can predict the next tasks (and data payloads) can be exploited to take proactive actions depending on the type of task that is predicted to happen. For instance, if the next task is an unwanted exception, then the current case can be double-checked to avoid this exception.
- **Time-related aspects** (Tax et al. 2017; Verenich et al. 2019; Camargo et al. 2019; Bukhsh et al. 2021; Ni et al. 2022): In this setting, the objective is to predict the timestamp of the next tasks that are predicted to happen in a case and, in particular, the one of the last task, which is used to predict the remaining case duration. Time-related information can be exploited for a variety of proactive actions, for instance, sending warnings to customers when a case that is serving them is predicted to run later than expected.

Consider a simple example involving a loan application process to illustrate the three aspects. We could be interested in predicting whether a loan application will be approved or rejected (process outcomes). We could be interested in predicting the next activity in a specific execution of the loan application process (next event in a running case). We could also be interested in predicting the remaining time until the loan application is fully processed (time prediction).

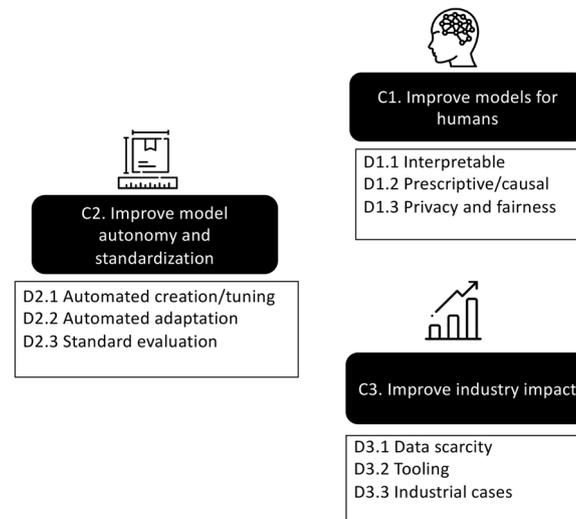
Typically, classification algorithms are used for outcome-oriented predictions, regression algorithms for time-related predictions, and deep learning techniques for next event(s) and time-related predictions. Furthermore, for each type of prediction, different options are available. For example, for outcome-oriented PPM, the most popular choice has been the decision tree (DT), which has obvious benefits in terms of the interpretability of the results (Maggi et al. 2014). Tree-based ensemble methods, such as random forest (RF) and gradient-boosted trees, have also become popular, usually achieving better prediction accuracy than a single decision tree. However, such models are much harder to interpret (Di Francescomarino et al. 2017). Similarly, time-related predictions can be tackled with different regression models, among which regression trees (van Dongen et al. 2008).

In the last few years, deep learning algorithms have also been widely used, in particular RNNs. The Long Short-Term Memory (LSTM) neural network is the most adopted due to its ability to work with sequential data. This architecture has been used to predict the next event(s) (Tax et al. 2017; Camargo et al. 2019; Tama and Comuzzi 2019; Taymouri et al. 2020; Weytjens and De Weerd 2022; Gunnarsson et al. 2023) and case duration (Tax et al. 2017; Camargo et al. 2019; Ni et al. 2022; Taymouri et al. 2021; Gunnarsson et al. 2023) with increasing refinements that improve the performance obtained in the experimental results. More recently, the Transformer architecture has been applied to PPM tasks, obtaining further performance improvements (Bukhsh et al. 2021).

In general, the experimental analysis in the literature clearly shows that, while known to be limited in terms of the interpretability of the results, generalization (Peeperkorn et al. 2023), and adaptation of the learned models to previously unseen data sets (Kim et al. 2020; Peeperkorn et al. 2023), deep learning classifiers substantially outperform classical machine learning classifiers with respect to accuracy (Kratsch et al. 2021).

### **The future of the predictive process monitoring field**

Following our overview of the PPM field in the previous section, our attention now turns to envisioning its future trajectory. We highlight the three main *challenges* that currently characterize the field. Within the framework of these three challenges, we identify a collection of 9 specific *directions* that should steer the efforts of PPM researchers in the coming years. The challenges and directions have been devised by the authors based on their research expertise. Their mapping to the 5-level process mining research framework (vom Brocke et al. 2021) is discussed in [Mapping the directions to the process mining research framework](#) section, whereas [Assessing challenges and future directions in PPM: a panel evaluation](#) section presents their empirical evaluation by a panel of research experts in PPM. To anchor the discussion, Fig. 2 summarizes the challenges and future directions that we have identified. These are discussed in detail next.



**Fig. 2** Challenges and directions for the future of the PPM field

### C1. Improve models for humans

The PPM research field has historically focused on the definition of new PPM use cases and new techniques to approach them. These new techniques are evaluated mainly by comparing their performance against the state-of-the-art. In respect of the 5-level process mining research framework (vom Brocke et al. 2021), such contributions can be classified as *design exaptations*, when a new PPM use case is devised, or *design improvements* (Mendling et al. 2021), when new techniques achieving higher performance are devised for existing PPM use cases. Only recently PPM has been perceived as a research field aiming at supporting process-related decision making (Dumas et al. 2023). Since these decisions are normally taken by humans, it is becoming crucial in the PPM field to understand how humans make sense of and use the PPM models and how we can develop better PPM models for human decision makers.

#### D1.1. Interpretable models

One of the main current challenges of machine learning predictive models is the one of their interpretability, that is (Du et al. 2019) (p.1) “the lack of transparency behind their behaviors, which leaves users with little understanding of how particular decisions are made by these models”. Users should not trust a model without understanding why specific predictions have been made. This is critical in many decision-making situations where resolutions require legal responsibility or organizational commitment.

In PPM, the issue of model interpretability has emerged recently, particularly with the increasing adoption of deep learning approaches. The literature on PPM has mainly adopted post-hoc explainers to add an interpretability (or explainability) layer to the predictive models, as, for example, in Mehdiyev and Fettke (2021a); Galanti et al. (2020); Stevens et al. (2022). Other works have focused on the interpretability of

PPM deep learning models, which by definition require post-hoc explainers. Shapley values and ICE (Individual Conditional Expectation) are used by Mehdiyev and Fettke (2021b), while counterfactual explanations are used by Huang et al. (2021), Hsieh et al. (2021), Buliga et al. (2023) and Hundogan et al. (2023). Complementary to interpretability, providing end users with more insight into the uncertainty of a prediction could also impact the adoption of PPM in practice (Weytjens and De Weerd 2022).

We call on PPM research to refine the techniques available to provide the interpretability of PPM models, while at the same time focusing on the role of the *human-in-the-loop* in decision-making supported by PPM models. Thus, we see two main directions for future research in this context. The first concerns the development of approaches for *model interpretability specifically tailored to the predictive monitoring* of business processes. This may involve effective visualization techniques for the post-hoc explainers in the context of PPM, but also deeper and ad-hoc extensions of existing explainability techniques able to fit specific PPM characteristics. A second direction concerns *assessing the impact of interpretable PPM models* on the decisions supported by such models, evaluating, for instance, the actual usefulness and ease of use of these techniques in practical decision-making contexts.

#### ***D1.2. Prescriptive and causal process monitoring models***

PPM is motivated by the need to enable process owners to make decisions in a proactive way. In this context, prescriptive process monitoring (Kubrak et al. 2022) represents the natural evolution of PPM. It deals with incorporating the effects of what occurs during the execution of a business process, for example, as the result of human actions, while predicting the value of aspects of interest in the long term, such as outcomes (Weinzierl et al. 2020; Teinmaa et al. 2018). While PPM models are obtained by adapting the typical machine learning pipeline to event log data, prescriptive monitoring of business processes requires ad-hoc approaches that are aware of the type of actions (often domain-specific) that can be executed in a process and their effect on the process execution. The point in a running trace where a prediction is made is also crucial in PPM. A prediction may be more or less accurate depending on how early it is made. Also, the execution of a predictive service is not free and should be launched when the number of events or their duration may significantly change the likelihood of the predictions provided by a model. Unlike other fields, such as the selection of control points in project management (Raz and Erel 2000), the PPM literature has only partially investigated the problem of the identification of optimal prediction points (Metzger et al. 2019). Recently, causal reasoning (Shoush and Dumas 2022; Leemans and Tax 2022) has also been applied in this direction. By highlighting the causal relationships between features and predicted aspects, causal reasoning allows decision makers to act on those process features that are more likely to determine the value of monitored process aspects in the future.

For this direction, we call on future PPM research to develop novel *causal reasoning techniques* for PPM and to investigate techniques to identify the optimal point in time in which predictions should be made. By relying on causal relationships, which are more understandable by practitioners than, for instance, weights capturing the feature importance, causal analysis may help closing the loop between the theoretical model

development and practical applications of predictive monitoring. Similarly, understanding the *best point in time* in which predictions should be made answers a crucial concern when developing decision making pipelines in practical settings.

### **D1.3. Privacy and fairness in PPM**

In addition to interpretability, data privacy and fairness represent open issues in the so-called responsible process mining (Mannhardt 2022) research landscape, as event logs can contain information protected by privacy regulations. A recent stream of works has started developing privacy-preserving algorithms for process mining (Elkoumy et al. 2022; Mannhardt et al. 2019), using state-of-the-art techniques such as group-based approaches (Li et al. 2006) or differential privacy (Dwork 2008).

We call on future PPM research to address the theoretical concerns that have limited the application of process mining data privacy and fairness techniques in PPM. First, these techniques produce information loss hampering the prediction interpretability and, as a consequence, the users' trust, thus making the privacy issue even more challenging in PPM. A balance between interpretability and privacy must be found. Additionally, these techniques produce event log distributions that may include trace variants patently inconsistent with a given business process. This affects the model accuracy and provides a means of identifying traces generated for privacy-preserving reasons (Fahrenkrog-Petersen et al. 2021). A call for solutions to deal with *event-log-specific data distribution* is then open. Finally, guaranteeing *fair conclusions* and hence decisions represents a poorly explored issue for the whole process mining field (Qafari and van der Aalst 2019). This is even more challenging in PPM, where predictions and recommendations are usually derived only from data. Indeed, learning predictive models from event log data could lead to predictions based on unfair factors (e.g., work experience or age when predicting the outcome of a production process) and may then result in discriminating conclusions and decisions.

## **C2. Improve model autonomy and standardization**

In the real world, PPM models are not stand-alone entities but they live in complex data extraction and engineering pipelines, ideally continuously feeding new information to the decision makers. At the same time, the business processes to which PPM models refer may continuously change to address new business requirements, such as changing market conditions or global economic downturns.

This creates a dual challenge for BPM researchers. On the one hand, repeatable and standardized methods to develop and tune the PPM models should be investigated to improve the models' fitness with real-world data engineering pipelines. On the other hand, PPM models should have the ability to adapt to changing process execution conditions and, most importantly, concept drift in business processes. PPM researchers should also strive to have widely agreed-upon ways to evaluate their models. While the issue of standard model evaluation is recognised in the machine learning literature (Hernández-Orallo 2017), in practice, each individual PPM approach seems to adopt its own evaluation protocol.

### ***D2.1. Automated model creation and tuning***

Every ML technique used to develop a PPM model requires to be configured to find the appropriate values of the hyperparameters for a specific dataset. Hyperparameter values, in fact, can greatly impact the performance of predictive models (Luo 2015). Selecting an appropriate ML technique and a configuration of hyperparameter values that maximizes its performance for a given dataset is, however, a nontrivial task. For non-experts, these choices often result in arbitrary (or default-case) choices (Thorn-ton et al. 2013). Practically, training and validating multiple configurations of a model to find the best one may require a high computational effort. Moreover, comparing the outcomes of different models may not be straightforward for non-expert users, such as business analysts and process owners.

In the PPM field, there are only a few works that look at the problem of automated model creation and tuning. For instance, Teinemaa et al. (2019) and Verenich et al. (2019) provide benchmarks of different ML techniques in outcome-based and time-related PPM, respectively. However, both benchmarks propose only generic guidelines regarding the effectiveness of different ML techniques in specific PPM scenarios. Other approaches (Di Francescomarino et al. 2018; Kwon and Comuzzi 2022) equip a PPM environment with a hyperparameter optimization method based on genetic algorithms.

To this end, we call for future PPM research to address this issue more organically. Specifically, the suitability of existing automated machine learning (AutoML) (Karmaker et al. 2021) frameworks for the PPM field should be empirically evaluated. Additionally, *AutoML approaches specifically tailored* to the PPM field should be developed. These should account for design choices during model development specific to the PPM field, which are not considered by general-purpose AutoML frameworks, such as how to encode an event log, whether or not to choose bucketing, as well as the optimal prediction point.

### ***D2.2. Automated model adaptation***

Process changes, driven, for instance, by seasonality, changing market conditions, or simply by the need to quickly address inefficiencies, are often the rule in modern organizations. The Process Mining Manifesto (van der Aalst et al. 2011a) highlights that “concept drift is of primary importance in the management of processes” (p.187), and every process mining technique should cope with the situation in which a process changes while being analyzed. In PPM, this translates into the need to develop predictive models that can adapt to concept drift. One straightforward solution is to adopt incremental ML techniques (Chefrour 2019), which produce models that can be quickly re-trained, even with every new observation.

Also in this direction, the number of existing research works in the PPM field is limited. The work by Maisenbacher and Weidlich (2017) is one of the first to exploit the traditional incremental ML in the context of PPM. The paper provides, in particular, an evaluation on synthetic logs with different types of concept drifts. Rizzi et al. (2022) investigate four different model update strategies (including the case of the *incremental* update) both in terms of accuracy of the results and the time required

to update the models. In the case of processes with concept drift, incremental ML techniques emerge as the most effective ones.

We call for future PPM research to go beyond the simple application of incremental machine learning techniques into existing PPM solutions, focusing specifically on the interplay between general purpose concept drift detection in process mining and PPM. A crucial aspect is how existing concept drift detection techniques in process mining can be applied along the PPM pipeline. *Accurate concept drift* detection allows to identify the specific points where a model should be re-trained, avoiding the need to revert to re-training a model with every new observation using incremental techniques. At the same time, insights from PPM models may inform the design of more accurate concept drift techniques. For example, a sudden drop in model accuracy may signal a change in the underlying data distribution, i.e., a change in the business process on which the model is making predictions.

### **D2.3. Standard evaluation of PPM models**

PPM research currently lacks standard evaluation methods, datasets, and metrics. PPM approaches are usually evaluated using different dataset preprocessing and splitting procedures (Weytjens and De Weerd [2021](#)), resulting in different input data for the PPM approach, some of which are far from a realistic scenario. In terms of datasets, the dominant evaluation approach in PPM relies on the event logs that have been published by the Business Process Intelligence Challenge over the past decade. These event logs were originally curated for analysis using traditional process mining techniques, often lacking relevance in a PPM context. Notably, the trace outcome labels associated with these logs typically signify the satisfaction or violation of declarative constraints, raising questions about their meaningfulness as process outcomes in a PPM context. This concern becomes particularly pronounced when striving for interpretable or prescriptive models if the attributes of an event log and the labels to be predicted are not meaningful from a decision-making perspective. Finally, standard performance metrics for classification and regression are usually adopted to collect the results. However, these may be limited in specific PPM contexts. For example, in outcome prediction, the labels can distinguish between positive and negative case outcomes. Naturally, negative outcomes are likely to be considerably less frequent than positive ones, leading to highly imbalanced classification problems. Traditional metrics such as accuracy or AUC, which are often used in PPM, may not be the most appropriate for highly imbalanced classification problems. The combination of PPM-specific performance metrics and standard evaluation datasets facilitates the development of reference analyses to continue to evaluate solutions developed for different PPM problems. In turn, standard benchmarks facilitate the identification of the best-performing techniques to be adopted in real-world PPM scenarios.

In response to these considerations, we call for future research to investigate an appropriate *evaluation framework* and/or specific guidelines for the evaluation of PPM approaches, possibly including new performance metrics for PPM. Additionally, the establishment of *benchmark* datasets tailored for PPM is imperative. These datasets should be made readily available to the community, enriched with relevant annotations for various PPM tasks. Existing initiatives in this direction, such as those outlined

in Weytjens and De Weerd (2021); Pauwels and Calders (2020); Marques Tavares et al. (2019), mark initial efforts, but a more comprehensive collective effort is essential.

### **C3. Improve the industry impact of PPM**

Following the footsteps of other process mining use cases, like process discovery and conformance checking, one of the aims of the research in PPM should be to demonstrate the impact of PPM in practical industry scenarios. This is likely to foster the adoption of PPM techniques by commercial process mining solutions.

We argue that the key to increasing industry adoption lies in (i) demonstrating the benefits of applying PPM techniques in practical scenarios and (ii) improving the tooling support for PPM techniques. Practically, the challenge of increasing industry adoption can be addressed by developing standard datasets to compare PPM models, embedding PPM techniques into open source process mining tools, and making PPM case studies and success stories widely available to academic and industry audiences.

#### ***D3.1. Data scarcity***

In practical settings, the application of machine learning-based techniques, such as PPM, often has to overcome a lack of labeled data to train and test the models. In PPM, such a scarcity of labeled data can be attributed to various factors, including recent process deployments, infrequent executions, or outcomes that are difficult to access. One way to address this shortcoming is generating of synthetic data, representing scenarios missing in the training dataset, but yet theoretically plausible. This could involve simulating processes with varying activity durations, employing either discrete event simulation models or deep learning models. Discrete event models involve discovering a model from an event log and adjusting parameters to maximize similarity to existing traces. Deep learning models use PPM to predict future events and generate new traces. While empirical results show that discrete event simulation models perform better on small datasets and deep learning models on larger datasets (Camargo et al. 2021; Kumar et al. 2022), recent approaches have demonstrated the efficacy of combining simulation and prediction for more accurate event log generation (Camargo et al. 2023; Meneghello et al. 2023). Additionally, *zero-shot learning*, an approach leveraging semantic information or textual descriptions to learn new concepts without specific examples (Kecht et al. 2021), represents a valuable strategy for overcoming the challenge of data scarcity in PPM.

We call for future research to address the issue of labeled *data scarcity* in PPM. As mentioned above, future research should investigate how increasingly accurate process simulation models can be used to produce data that enhance the accuracy of PPM models. At the same time, however, future research should also consider how novel learning approaches that do not rely on synthetic data generation can be customized to PPM.

#### ***D3.2. Tooling***

The adoption of process mining in industry crucially has relied on tools that initially were developed by the academic research community. For example, many commercially available process mining tools have evolved from plug-ins developed by researchers for the ProM process mining toolkit.

Within the extensive array of ProM plugins (van Dongen et al. 2005), several incorporate techniques for the prediction of outcomes, e.g. (Maggi et al. 2014), the prediction of numerical values, e.g. (van der Aalst et al. 2011b; de Leoni et al. 2016), as well as the prediction of next activity sequences, e.g. (Polato et al. 2018). As far as commercial process mining tools are concerned, Apromore (La Rosa et al. 2011) provides a PPM plugin performing outcome-based, numeric-based prediction, as well as next event predictions (Verenich et al. 2018). Celonis provides tools to automatically create classification problems for decision mining. Nirdizati (Rizzi et al. 2019) stands out as a tool explicitly designed for PPM, enabling users to construct, compare, and analyze predictive models for future case developments. In the open-source domain, processpredictR (Esin et al. 2023), part of the BupaR process mining toolkit, facilitates prediction of case outcomes, next activities, and remaining time.

Despite these efforts, the current landscape of PPM-specific tools remains somewhat constrained. There is a pressing need for comprehensive and extensible tools and applications tailored to PPM within the broader process mining community. Advocating for the wider integration of PPM techniques into open-source and other process mining tools, our specific call is for the development of extensible packages within established frameworks such as PM4Py (Berti et al. 2023) and BupaR. These packages should be designed for potential integration into libraries with general machine learning capabilities, fostering a more unified and expansive ecosystem.

An encouraging step in this direction is showcased in Oyamada et al. (2023), where an extension of the Scikit-learn library, seamlessly integrated with PM4Py, is presented. This integration aims to standardize pre-processing procedures and learning workflows, promoting accessibility for researchers and practitioners. The incorporation of process mining capabilities into widely used libraries, such as Scikit-learn, not only enhances tool accessibility but also establishes consistent methodologies, addressing the challenge of reproducibility and the current absence of benchmarking resources within the process mining community.

### ***D3.3. Industrial cases***

Demonstrating the real-world impact of new process mining techniques is important, and this impact is best illustrated through authentic case studies. Unfortunately, such examples are largely absent in the context of PPM. While real-life case studies in process mining exist (Reinkemeyer 2020), only a single instance (Lillig 2020) mentions the application of predictive analytics techniques in conjunction with process mining. Only few instances in the literature showcase the assessment of PPM techniques through proprietary event logs case studies that actively engage process stakeholders. For example, while not explicitly framed as a case study, Galanti et al. (2020) explores a proprietary event log from a banking process, delving into discussions with process stakeholders about the interpretable PPM models derived. Furthermore, Gunnarsson et al. (2019) has demonstrated the application of PPM techniques in the context of airport logistics processes.

Recognising the lack of case studies incorporating PPM techniques, we call for future PPM research to forge closer collaborations with industry partners. Whenever possible, researchers should publish their findings, elucidating the practical implications of adopted or proposed PPM techniques. Bridging the gap between academia and industry

is crucial for advancing the field and facilitating the integration of PPM into actual business practices.

**Mapping the directions to the process mining research framework**

In this section we map our PPM research directions to the levels of the process mining research framework (vom Brocke et al. 2021) introduced in [The field of predictive process monitoring](#) section. Having presented the directions in depth in the previous section, our objective is now to contextualize them methodologically within the broader process mining research landscape. A summary of this mapping is shown in Table 1.

Direction D1.1 (Interpretability) significantly impacts both the *individual* and *organizational* levels. Interpretable PPM models improve decision-making capabilities for process owners and users, thereby fostering organizational adoption through enhanced comprehension and traceability of process-related decisions. Similarly, Direction D1.2 (Prescriptive/causal) also resonates with *individual* and *organizational* levels. Insights derived from prescriptive/causal analysis must be user-friendly and beneficial to stakeholders, potentially providing a competitive advantage at the organizational level if effectively implemented. Finally, Direction D1.3 (Privacy and fairness) spans across *individual*, *group* and *ecosystem* levels. Privacy considerations within PPM affect individuals, by providing a means to ensure appropriate treatment of their sensitive data. The fairness aspect affects the treatment of data of protected groups of process actors. Moreover, privacy-preserving techniques facilitate inter-organizational adoption at the (*ecosystem*) level, where process data may need to be shared among competing organizations. For all the directions within challenge C1, the mapping with the *technical* level concerns *design exaptations*, i.e., adapting existing explainable AI, prescriptive/causal, or privacy/fairness enabling techniques to existing PPM solutions.

Regarding Challenge C2, Directions D2.2 (Automated adaptation) and D2.3 (Standard evaluation) are primarily aligned with the *ecosystem* level. Automated PPM model adaptation aids organizations within an ecosystem in addressing evolving business

**Table 1** Mapping between the PPM directions and the framework levels from vom Brocke et al. (2021)

Challenge	Direction	Framework levels				
		Technical	Individual	Group	Organization	Ecosystem
C1. Improve models for humans	D1.1 Interpretability	✓	✓		✓	
	D1.2 Prescriptive/causal	✓	✓		✓	
	D1.3 Privacy and fairness	✓	✓	✓		✓
C2. Improve model autonomy and standardization	D2.1 Automated creation/tuning	✓				
	D2.2 Automated adaptation	✓				✓
	D2.3 Standard evaluation	✓				✓
C3. Improve industry impact	D3.1 Data scarcity	✓			✓	
	D3.2 Tooling	✓	✓		✓	
	D3.3 Industrial cases				✓	

requirements, while standardized evaluation methods promote comparability among PPM solutions, facilitating the robust adaptation of diverse solutions across different domains. Direction D2.1 (Automated creation/tuning) primarily corresponds to the *technical* level. Unlike other challenges, this mapping pertains to the facilitation of *knowledge contribution*, as automated and fine-tuned model creation enables fair comparisons between competing techniques. A similar rationale applies to the mapping of D2.3 to the *technical* level.

For Challenge C3, Direction D3.3 (Industrial cases) primarily maps to the *organizational* level. The availability of successful PPM case studies can help make the case for PPM adoption and garner top-management support. The availability of PPM tools (D3.2) is linked to the *individual* level — for evaluating their acceptance by process owners and users — and the *organizational* level — to evaluate their fit with organizational policies and intended users. Direction D3.1 (data scarcity) maps to the *organizational* level, as access to benchmark data enables more rigorous testing of PPM solutions. The mapping of C3.1 and C3.2 to the *technical* level, similarly, pertains to the level of *knowledge contributions*, because widely available tools and benchmark datasets can foster the comparison of existing solutions and the development of industry benchmarks.

To sum up, the identified directions highlight the necessity for PPM research to go beyond the *technical* level in process mining research. This suggests a need to emphasize the model application across the levels of individuals, groups, organizations, and ecosystems (see Fig. 1). As shown in Table 1, except for D3.3, all directions primarily align with the *technical* level of the framework. For most directions, this alignment relates to the *design* aspect, rather than to the *knowledge contribution* (Mendling et al. 2021) aspect, since most directions involve the development of different types of technical solutions to known or emerging problems in PPM.

### Assessing challenges and future directions in PPM: a panel evaluation

This section presents the findings of an empirical evaluation conducted with a panel of experts to assess the current challenges and future directions of PPM. Panelists were selected based on their expertise in PPM to ensure a diverse representation of perspectives. They were asked to respond to a survey that included questions regarding their expertise in PPM and their assessment of the importance of specific challenges and future directions identified in our previous research. The participants primarily hailed from Europe (68%), with additional representation from Asia Pacific (23%) and the Americas (9%). Most respondents reported over 15 years of experience in the field (32%), followed by those with 5 to 9 years (27%), 0 to 4 years (23%), and 10 to 14 years (14%) of experience. The panel consisted of 22 researchers. Each participant was provided with a concise description of the identified challenges and associated future research directions. They were then asked to rate the importance of these future directions using a Likert scale ranging from 1 to 7, where higher ratings denoted greater importance. Moreover, they were asked to assess the completeness of the identified research directions and challenges.

The results of the empirical evaluation provided valuable insights into the perceived importance of challenges and future directions in PPM. Table 2 presents the average ratings and standard deviations for each direction. Notably, all future directions received

**Table 2** Average Ratings and Standard Deviations for PPM Future Directions

Challenge	Direction	Average Rating	Standard Deviation
C1. Improve models for humans	D1.1 Interpretability	5.5	0.91
	D1.2 Prescriptive/causal	5.59	1.05
	D1.3 Privacy and fairness	5.0	1.15
C2. Improve model autonomy and standardization	D2.1 Automated creation/tuning	4.77	1.11
	D2.2 Automated adaptation	4.86	1.16
	D2.3 Standard evaluation	5.54	1.33
C3. Improve industry impact	D3.1 Data scarcity	5.95	0.99
	D3.2 Tooling	4.36	1.17
	D3.3 Industrial cases	5.45	1.22

high ratings indicating their perceived importance. This trend is particularly evident in challenge C1, where D1.1 and D1.2 received scores ranging from 4 to 7, and D1.3 ranged from 3 to 7. Challenge C2 received slightly lower scores, with D2.1 and D2.3 in the range of 3-7 and D2.2 in the range of 2-7, but with a single 2 and a single 7. Challenge C3 yielded mixed results, with D3.1 ranging from 3 to 7 with a single 3 and six 7s, D3.2 from 3 to 7 but with a single 7, and D3.3 from 3 to 7 with three 3s and three 7s. No specific concerns were raised on direction and challenge completeness.

The strong support for the identified research challenges in PPM is underlined by the results of our empirical evaluation. While the attention of the research experts is especially directed towards the interpretability and “actionability” of predictive models (D1.1 and D1.2), as well as towards issues related to the scarcity of standard evaluation approaches, data and industrial cases (D2.3, D3.1 and D3.3), with the average ratings of all challenges ranging between 4.36 and 5.95, it is evident that all of them are perceived as significant by experts in the field.

**Conclusions**

PPM has emerged as a forward-looking extension of process mining, in which event log data are used to create models to predict aspects of interests of process execution in the future. While it has seen extreme traction in the process mining research community in recent years, a comprehensive conceptualization of PPM and its scope, with a summary of existing proposals, as well as an analysis of challenges and directions, is lacking in the process mining literature.

This paper aimed to close this gap, especially by outlining the challenges and future directions to be addressed by the research community. We hope that this paper will help process mining researchers find a common conceptual ground when developing future PPM approaches, and that such a common ground will facilitate the development of new solutions able to cope with emerging challenges, the comparison of research results, as well as the adoption of advanced PPM techniques in commercial process mining tools.

**Authors’ contributions**

All authors contributed equally to writing and reviewing the manuscript.

**Funding**

M. Comuzzi was supported by the NRF Korea, Grant Number 2022R1F1A1072843.

**Availability of data and materials**

No datasets were generated or analysed during the current study.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Competing interests**

The authors declare no competing interests.

Received: 14 December 2023 Accepted: 31 July 2024

Published online: 27 September 2024

**References**

- Appice A, Malerba D (2015) A co-training strategy for multiple view clustering in process mining. *IEEE Trans Serv Comput* 9(6):832–845
- Azzini A, Barbon Jr S, Bellandi S et al (2021) *Advances in Data Management in the Big Data Era*. Springer International Publishing, Cham. [https://doi.org/10.1007/978-3-030-81701-5\\_4](https://doi.org/10.1007/978-3-030-81701-5_4)
- Back CO, Simonsen JG (2023) Comparing trace similarity metrics across logs and evaluation measures. In: Indulska M, Reinhartz-Berger I, Cetina C et al (eds) *Advanced Information Systems Engineering*. Springer Nature Switzerland, Cham, pp 226–242
- Berti A, van Zelst S, Schuster D (2023) Pm4py: A process mining library for python. *Softw Impacts* 17:100556. <https://doi.org/10.1016/j.simpa.2023.100556>
- Bose RJC, Van der Aalst WM (2009) Context aware trace clustering: Towards improving process mining results. In: *proceedings of the 2009 SIAM International Conference on Data Mining*. SIAM, pp 401–412
- Bukhsh ZA, Saeed A, Dijkman RM (2021) *Processtransformer: Predictive business process monitoring with transformer network*. arXiv preprint arXiv:210400721
- Buliga A, Di Francescomarino C, Ghidini C, et al (2023) Counterfactuals and ways to build them: Evaluating approaches in predictive process monitoring. In: Indulska M, Reinhartz-Berger I, Cetina C, et al (eds) *Advanced Information Systems Engineering - 35th International Conference, CAiSE 2023, Zaragoza, Spain, June 12-16, 2023, Proceedings, Lecture Notes in Computer Science*, vol 13901. Springer, pp 558–574. [https://doi.org/10.1007/978-3-031-34560-9\\_33](https://doi.org/10.1007/978-3-031-34560-9_33)
- Camargo M, Bårøn D, Dumas M et al (2023) Learning business process simulation models: A hybrid process mining and deep learning approach. *Inf Syst* 117:102248. <https://doi.org/10.1016/J.IS.2023.102248>
- Camargo M, Dumas M, González-Rojas O (2019) Learning accurate lstm models of business processes. In: *International Conference on Business Process Management*. Springer, Cham, pp 286–302
- Camargo M, Dumas M, González-Rojas O (2021) Discovering generative models from event logs: data-driven simulation vs deep learning. *PeerJ Comput Sci* 7:e577
- Castellanos M, Salazar N, Casati F, et al (2005) *Predictive business operations management*. In: *International Workshop on Databases in Networked Information Systems*. Springer, Cham, pp 1–14
- Ceravolo P, Damiani E, Torabi M, et al (2017) *Toward a new generation of log pre-processing methods for process mining*. In: *International Conference on Business Process Management*. Springer, Cham, pp 55–70
- Chefrour A (2019) Incremental supervised learning: algorithms and applications in pattern recognition. *Evol Intel* 12(2):97–112
- De Koninck P, vanden Broucke S, De Weerd J (2018) act2vec, trace2vec, log2vec, and model2vec: Representation learning for business processes. In: Weske M, Montali M, Weber I et al (eds) *Business Process Management*. Springer International Publishing, Cham, pp 305–321
- de Leoni M, van der Aalst WMP, Dees M (2016) A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs. *Inf Syst* 56:235–257. <https://doi.org/10.1016/j.is.2015.07.003>
- De Smedt J, Yeshchenko A, Polyvyanyy A et al (2023) Process model forecasting and change exploration using time series analysis of event sequence data. *Data Knowl Eng* 145:102145
- Di Francescomarino C, Ghidini C (2022) Predictive process monitoring. *Process Min Handb LNBIP* 448:320–346
- Di Francescomarino C, Dumas M, Maggi FM, Teinemaa I (2016) Clustering-based predictive process monitoring. *IEEE Trans Serv Comput* 12(6):896–909
- Di Francescomarino C, Dumas M, Federici M et al (2018) Genetic algorithms for hyperparameter optimization in predictive business process monitoring. *Inf Syst* 74(Part):67–83. <https://doi.org/10.1016/j.is.2018.01.003>
- Du M, Liu N, Hu X (2019) Techniques for interpretable machine learning. *Commun ACM* 63(1):68–77
- Dumas M, Fournier F, Limonad L et al (2023) *AI-augmented business process management systems: a research manifesto*. *ACM Trans Manag Inf Syst* 14(1):1–19
- Dwork C (2008) *Differential privacy: A survey of results*. In: *International conference on theory and applications of models of computation*. Springer, Cham, pp 1–19
- Elkoumy G, Fahrenkrog-Petersen SA, Sani MF et al (2022) Privacy and confidentiality in process mining: Threats and research challenges. *ACM Trans Manag Inf Syst* 13(1):11:1-11:17. <https://doi.org/10.1145/3468877>
- Esin I, Beloshitskiy D, Janssenswillen G (2023) processpredictr: predictive process monitoring using bupar. In: *Proceedings of the ICPM Doctoral Consortium and Demo Track 2023* (to appear). CEUR-WS.org
- Fahrenkrog-Petersen SA, Kabierski M, Rösel F, et al (2021) Sacofa: Semantics-aware control-flow anonymization for process mining. In: *2021 3rd International Conference on Process Mining (ICPM)*. IEEE, pp 72–79

- Galanti R, Coma-Puig B, de Leoni M, et al (2020) Explainable predictive process monitoring. In: 2020 2nd International Conference on Process Mining (ICPM). IEEE, pp 1–8
- Ghattas J, Soffer P, Peleg M (2014) Improving business process decision making based on past experience. *Decis Support Syst* 59:93–107
- Grinvald A, Soffer P, Mokryn O (2021) Inter-case properties and process variant considerations in time prediction: A conceptual framework. In: Augusto A, Gill A, Nurcan S, et al (eds) *Enterprise, Business-Process and Information Systems Modeling - 22nd International Conference, BPMDS 2021, and 26th International Conference, EMMSAD 2021*, Held at CAISE 2021, Melbourne, VIC, Australia, June 28–29, 2021, Proceedings, Lecture Notes in Business Information Processing, vol 421. Springer, pp 96–111. [https://doi.org/10.1007/978-3-030-79186-5\\_7](https://doi.org/10.1007/978-3-030-79186-5_7)
- Gunnarsson BR, vanden Broucke S, De Weerd J (2019) Predictive process monitoring in operational logistics: A case study in aviation. In: *Business Process Management Workshops: BPM 2019 International Workshops*, Vienna, Austria, September 1–6, 2019, Revised Selected Papers 17. Springer, Cham, pp 250–262
- Gunnarsson BR, vanden Broucke S, De Weerd J (2023) A direct data aware lstm neural network architecture for complete remaining trace and runtime prediction. *IEEE Trans Serv Comput* 16(4):2330–2342. <https://doi.org/10.1109/TSC.2023.3245726>
- Hernández-Orallo J (2017) Evaluation in artificial intelligence: from task-oriented to ability-oriented measurement. *Artif Intell Rev* 48:397–447
- Hsieh C, Moreira C, Ouyang C (2021) Dice4el: Interpreting process predictions using a milestone-aware counterfactual approach. In: 3rd International Conference on Process Mining (ICPM). pp 88–95. <https://doi.org/10.1109/ICPM53251.2021.9576881>
- Huang TH, Metzger A, Pohl K (2021) Counterfactual explanations for predictive business process monitoring. In: *European, Mediterranean, and Middle Eastern Conference on Information Systems*. Springer, pp 399–413
- Hundogan O, Lu X, Du Y, et al (2023) CREATED: generating viable counterfactual sequences for predictive process analytics. In: Indulska M, Reinhartz-Berger I, Cetina C, et al (eds) *Advanced Information Systems Engineering - 35th International Conference, CAISE 2023*, Zaragoza, Spain, June 12–16, 2023, Proceedings, Lecture Notes in Computer Science, vol 13901. Springer, pp 541–557. [https://doi.org/10.1007/978-3-031-34560-9\\_32](https://doi.org/10.1007/978-3-031-34560-9_32)
- Junior SB, Ceravolo P, Damiani E, et al (2020) Anomaly detection on event logs with a scarcity of labels. In: 2020 2nd International Conference on Process Mining (ICPM). IEEE, pp 161–168
- Karmaker SK, Hassan MM, Smith MJ et al (2021) Automl to date and beyond: Challenges and opportunities. *ACM Comput Surv (CSUR)* 54(8):1–36
- Kecht C, Egger A, Kratsch W, et al (2021) Event log construction from customer service conversations using natural language inference. In: 2021 3rd International Conference on Process Mining (ICPM). pp 144–151. <https://doi.org/10.1109/ICPM53251.2021.9576869>
- Kim B, Park J, Suh J (2020) Transparency and accountability in ai decision support: Explaining and visualizing convolutional neural networks for text information. *Decis Support Syst* 134:113302
- Kim J, Comuzzi M, Dumas M et al (2022) Encoding resource experience for predictive process monitoring. *Decis Support Syst* 153:113669
- Kratsch W, Manderscheid J, Röglinger M et al (2021) Machine learning in business process monitoring: a comparison of deep learning and classical approaches used for outcome prediction. *Bus Inf Syst Eng* 63(3):261–276
- Kubrak K, Milani F, Nolte A, Dumas M (2022) Prescriptive process monitoring: Quo vadis?. *Peer J Computer Science* 8:e1097
- Kumar S, Dixit AK, Akarte M (2022) Machine learning based decision support system for resilient supplier selection. In: XVIII International Conference on Data Science and Intelligent Analysis of Information. Springer, Cham, pp 33–43
- Kwon N, Comuzzi M (2022) Genetic algorithms for automl in process predictive monitoring. In: *International Conference on Process Mining*. Springer, Cham, pp 242–254
- La Rosa M, Reijers HA, van der Aalst WMP et al (2011) APROMORE: an advanced process model repository. *Expert Syst Appl* 38(6):7029–7040. <https://doi.org/10.1016/j.eswa.2010.12.012>
- Leemans SJ, Tax N (2022) Causal reasoning over control-flow decisions in process models. In: *International Conference on Advanced Information Systems Engineering*. Springer, Cham, pp 183–200
- Leontjeva A, Conforti R, Di Francescomarino C, et al (2015) Complex symbolic sequence encodings for predictive monitoring of business processes. In: *International Conference on Business Process Management*. Springer, Cham, pp 297–313
- Li N, Li T, Venkatasubramanian S (2006) t-closeness: Privacy beyond k-anonymity and l-diversity. In: 2007 IEEE 23rd international conference on data engineering. IEEE, pp 106–115
- Lillig G (2020) Telekom: Process mining in shared services. In: *Process Mining in Action*. Springer, Cham, pp 169–178
- Luo G (2015) Mlbc: a machine learning tool for big clinical data. *Health Inf Sci Syst* 3(1):1–19
- Maggi FM, Di Francescomarino C, Dumas M, et al (2014) Predictive monitoring of business processes. In: *International conference on advanced information systems engineering*. Springer, Cham, pp 457–472
- Maisenbacher M, Weidlich M (2017) Handling concept drift in predictive process monitoring. In: Liu XF, Bellur U (eds) 2017 IEEE International Conference on Services Computing, SCC 2017, Honolulu, HI, USA, June 25–30, 2017. IEEE Computer Society, pp 1–8. <https://doi.org/10.1109/SCC.2017.10>
- Mannhardt F (2022) Responsible process mining. In: van der Aalst WMP, Carmona J (eds) *Process Mining Handbook, Lecture Notes in Business Information Processing*, vol 448. Springer, pp 373–401. [https://doi.org/10.1007/978-3-031-08848-3\\_12](https://doi.org/10.1007/978-3-031-08848-3_12)
- Mannhardt F, Koschmider A, Baracaldo N et al (2019) Privacy-preserving process mining - differential privacy for event logs. *Bus Inf Syst Eng* 61(5):595–614. <https://doi.org/10.1007/s12599-019-00613-3>
- Marques Tavares G, Barbon Jr. S, Ceravolo P (2019) Synthetic event streams. <https://doi.org/10.21227/2kxd-m509>
- Márquez-Chamorro AE, Resinas M, Ruiz-Cortés A (2017) Predictive monitoring of business processes: a survey. *IEEE Trans Serv Comput* 11(6):962–977

- Mehdiyev N, Fettke P (2021a) Explainable artificial intelligence for process mining: A general overview and application of a novel local explanation approach for predictive process monitoring. *Interpretable Artif Intell Perspect Granul Comput* 1–28. [https://link.springer.com/chapter/10.1007/978-3-030-64949-4\\_1](https://link.springer.com/chapter/10.1007/978-3-030-64949-4_1)
- Mehdiyev N, Fettke P (2021b) Local post-hoc explanations for predictive process monitoring in manufacturing. *ECIS 2021 Research Papers* 35. [https://aisel.aisnet.org/ecis2021\\_rp/35](https://aisel.aisnet.org/ecis2021_rp/35)
- Mendingling J, Depaire B, Leopold H (2021) Theory and practice of algorithm engineering. arXiv preprint. arXiv:2107.10675
- Meneghello F, Di Francescomarino C, Ghidini C (2023) Runtime integration of machine learning and simulation for business processes. In: 5th International Conference on Process Mining, ICPM 2023, Rome, Italy, October 23–27, 2023. IEEE, pp 9–16. <https://doi.org/10.1109/ICPM60904.2023.10271993>
- Metzger A, Neubauer A, Bohn P, et al (2019) Proactive Process Adaptation Using Deep Learning Ensembles. In: Giorgini P, Weber B (eds) *Advanced Information Systems Engineering - 31st International Conference, CAISE 2019, Rome, Italy, June 3–7, 2019, Proceedings, Lecture Notes in Computer Science*, vol 11483. Springer, Cham, pp 547–562
- Ni W, Yan M, Liu T et al (2022) Predicting remaining execution time of business process instances via auto-encoded transition system. *Intell Data Anal* 26(2):543–562
- Oyamada RS, Marques Tavares G, Junior SB, Ceravolo P (2023). A Scikit-learn Extension Dedicated to Process Mining Purposes. In: *CEUR workshop proceedings*, vol 3552. CEURWS, pp 11–15
- Pauwels S, Calders T (2020) Bayesian network based predictions of business processes. In: *Business Process Management Forum: BPM Forum 2020, Seville, Spain, September 13–18, 2020, Proceedings* 18. Springer, Cham, pp 159–175
- Peeperkorn J, Broucke SV, De Weerd J (2023) Can recurrent neural networks learn process model structure?. *J Intell Inf Syst* 61(1):27–51
- Polato M, Sperduti A, Burattin A et al (2018) Time and activity sequence prediction of business process instances. *Computing* 100(9):1005–1031. <https://doi.org/10.1007/s00607-018-0593-x>
- Pourbafrani M, van der Aalst WMP (2022) Hybrid business process simulation: Updating detailed process simulation models using high-level simulations. In: Guizzardi R, Ralyté J, Franch X (eds) *Research Challenges in Information Science*. Springer International Publishing, Cham, pp 177–194
- Qafari MS, van der Aalst WMP (2019) Fairness-aware process mining. In: Panetto H, Debruyne C, Hepp M, et al (eds) *On the Move to Meaningful Internet Systems: OTM 2019 Conferences - Confederated International Conferences: CoopIS, ODBASE, C & TC 2019, Rhodes, Greece, October 21–25, 2019, Proceedings, Lecture Notes in Computer Science*, vol 11877. Springer, pp 182–192. [https://doi.org/10.1007/978-3-030-33246-4\\_11](https://doi.org/10.1007/978-3-030-33246-4_11)
- Raz T, Erel E (2000) Optimal timing of project control points. *Eur J Oper Res* 127(2):252–261
- Reikenmeyer L (2020) *Process Mining in Action. Principles, Use Cases and Outlook*. Springer, Cham
- Rizzi W, Di Francescomarino C, Ghidini C, Maggi FM (2022) How do I update my model? On the resilience of Predictive Process Monitoring models to change. *Knowl Inf Syst* 64(5):1385–1416
- Rizzi W, Simonetto L, Di Francescomarino C, et al (2019) Nirdizati 2.0: New features and redesigned backend. In: Depaire B, Smedt JD, Dumas M, et al (eds) *Proceedings of the Dissertation Award, Doctoral Consortium, and Demonstration Track at BPM 2019 co-located with 17th International Conference on Business Process Management, BPM 2019, Vienna, Austria, September 1–6, 2019, CEUR Workshop Proceedings*, vol 2420. CEUR-WS.org, pp 154–158
- Schwegmann B, Matzner M, Janiesch C (2013) A Method and Tool for Predictive Event-Driven Process Analytics. *Wirtschaftsinformatik Proc* 2013:46. <https://aisel.aisnet.org/wi2013/46>
- Senderovich A, Di Francescomarino C, Maggi FM (2019) From knowledge-driven to data-driven inter-case feature encoding in predictive process monitoring. *Inf Syst* 84:255–264
- Shmueli G, Koppius OR (2011) Predictive analytics in information systems research. *MIS Quart* 35(3):553–572
- Shoush M, Dumas M (2022) Prescriptive process monitoring under resource constraints: a causal inference approach. In: *International Conference on Process Mining*. Springer, pp 180–193
- Stevens A, De Smedt J, Peepkorn J (2022) Quantifying explainability in outcome-oriented predictive process monitoring. In: Munoz-Gama J, Lu X (eds) *Process Mining Workshops*. Springer International Publishing, Cham, pp 194–206
- Tama BA, Comuzzi M (2019) An empirical comparison of classification techniques for next event prediction using business process event logs. *Expert Syst Appl* 129:233–245
- Tavares GM, Barbon S (2020) Analysis of language inspired trace representation for anomaly detection. In: *ADBIS, TPD and EDA 2020 Common Workshops and Doctoral Consortium*. Springer, pp 296–308
- Tavares GM, Oyamada RS, Junior SB, Ceravolo P (2023) Trace encoding in process mining: A survey and benchmarking. *Eng Appl Artif Intell* 126:107028
- Tax N, Verenich I, La Rosa M, et al (2017) Predictive business process monitoring with lstm neural networks. In: *International Conference on Advanced Information Systems Engineering*. Springer, Cham, pp 477–492
- Taymouri F, La Rosa M, Erfani S, et al (2020) Predictive business process monitoring via generative adversarial nets: the case of next event prediction. In: *International Conference on Business Process Management*. Springer, pp 237–256
- Taymouri F, La Rosa M, Erfani SM (2021) A deep adversarial model for suffix and remaining time prediction of event sequences. In: *Proceedings of the 2021 SIAM International Conference on Data Mining (SDM)*. SIAM, pp 522–530
- Teinemaa I, Dumas M, La Rosa M et al (2019) Outcome-oriented predictive process monitoring: review and benchmark. *ACM Trans Knowl Discov Data (TKDD)* 13(2):1–57
- Teinemaa I, Tax N, de Leoni M, et al (2018) Alarm-based prescriptive process monitoring. In: *International Conference on Business Process Management*. Springer, Cham, pp 91–107
- Thornton C, Hutter F, Hoos HH, et al (2013) Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. In: *Proc. of KDD-2013*. ACM, pp 847–855
- van der Aalst WMP (2016) *Process Mining: Data Science in Action*, 2nd edn. Springer, Heidelberg. <https://doi.org/10.1007/978-3-662-49851-4>
- van der Aalst WMP, Adriansyah A, de Medeiros AKA, et al (2011a) Process mining manifesto. In: Daniel F, Barkaoui K, Dustdar S (eds) *Business Process Management Workshops - BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part I, Lecture Notes in Business Information Processing*, vol 99. Springer, pp 169–194
- van der Aalst WMP, Schonenberg MH, Song M (2011b) Time prediction based on process mining. *Inf Syst* 36(2):450–475

- van Dongen BF, de Medeiros AKA, Verbeek HMW, et al (2005) The prom framework: A new era in process mining tool support. In: Ciardo G, Darondeau P (eds) Applications and Theory of Petri Nets 2005, 26th International Conference, ICATPN 2005, Miami, USA, June 20-25, 2005, Proceedings, Lecture Notes in Computer Science, vol 3536. Springer, pp 444–454. [https://doi.org/10.1007/11494744\\_25](https://doi.org/10.1007/11494744_25)
- van Dongen BF, Crooy RA, van der Aalst WM (2008) Cycle time prediction: When will this case finally be finished? In: OTM Confederated International Conferences "On the Move to Meaningful Internet Systems". Springer, pp 319–336
- vom Brocke J, Jans M, Mendling J et al (2021) A five-level framework for research on process mining. *Bus Inf Syst Eng* 63(5):483–490. <https://doi.org/10.1007/s12599-021-00718-8>
- Verenich I, Dumas M, La Rosa M et al (2019) Survey and cross-benchmark comparison of remaining time prediction methods in business process monitoring. *ACM Trans Intell Syst Technol (TIST)* 10(4):1–34
- Verenich I, Möskovski S, Raboczi S, et al (2018) Predictive process monitoring in apromore. In: Mendling J, Mouratidis H (eds) Information Systems in the Big Data Era - CAiSE Forum 2018, Tallinn, Estonia, June 11-15, 2018, Proceedings, Lecture Notes in Business Information Processing, vol 317. Springer, pp 244–253. [https://doi.org/10.1007/978-3-319-92901-9\\_21](https://doi.org/10.1007/978-3-319-92901-9_21)
- Weinzierl S, Dunzer S, Zilker S, et al (2020) Prescriptive business process monitoring for recommending next best actions. In: International conference on business process management. Springer, pp 193–209
- Weytjens H, De Weerd J (2021) Creating unbiased public benchmark datasets with data leakage prevention for predictive process monitoring. In: Marrella A, Weber B (eds) Business Process Management Workshops - BPM 2021 International Workshops, Rome, Italy, September 6-10, 2021, Revised Selected Papers, Lecture Notes in Business Information Processing, vol 436. Springer, pp 18–29. [https://doi.org/10.1007/978-3-030-94343-1\\_2](https://doi.org/10.1007/978-3-030-94343-1_2)
- Weytjens H, De Weerd J (2022) Learning uncertainty with artificial neural networks for predictive process monitoring. *Appl Soft Comput* 125:109134. <https://doi.org/10.1016/j.asoc.2022.109134>

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.