

# A Survey on Event Prediction Methods from a Systems Perspective: Bringing Together Disparate Research Areas

JANIK-VASILY BENZIN, Technical University of Munich, Germany; TUM School of Computation, Information and Technology, Department of Informatics, Germany

STEFANIE RINDERLE-MA, Technical University of Munich, Germany; TUM School of Computation, Information and Technology, Department of Informatics, Germany

Event prediction is the ability of anticipating future events, i.e., future real-world occurrences, and aims to support the user in deciding on actions that change future events towards a desired state. An event prediction method learns the relation between features of past events and future events. It is applied to newly observed events to predict corresponding future events that are evaluated with respect to the user's desired future state. If the predicted future events do not comply with this state, actions are taken towards achieving desirable future states. Evidently, event prediction is valuable in many application domains such as business and natural disasters. The diversity of application domains results in a diverse range of methods that are scattered across various research areas which, in turn, use different terminology for event prediction methods. Consequently, sharing methods and knowledge for developing future event prediction methods is restricted. To facilitate knowledge sharing on account of a comprehensive classification, integration, and assessment of event prediction methods, we combine taxonomies and take a systems perspective to integrate event prediction methods into a single system, elicit requirements and assess existing work with respect to the requirements. Based on the assessment, we identify open challenges and discuss future research directions.

CCS Concepts: • **Information systems** → **Data mining**; **Decision support systems**; • **Computing methodologies** → **Machine learning**.

Additional Key Words and Phrases: Event prediction, artificial intelligence, predictive process monitoring, anomaly prediction, systems engineering

## ACM Reference Format:

Janik-Vasily Benzin and Stefanie Rinderle-Ma. 2023. A Survey on Event Prediction Methods from a Systems Perspective: Bringing Together Disparate Research Areas. *ACM Comput. Surv.* 1, 1 (February 2023), 43 pages. <https://doi.org/XXXXXXX.XXXXXXX>

## 1 INTRODUCTION

Given the rapid growth of data [140] and major advances in enabling technologies such as cloud computing [249], Internet of Things [265], data processing and machine learning [94], it is no surprise that event prediction is researched to a great extent. Event prediction is the ability of anticipating events, i.e., future real-world occurrences. To anticipate events, event prediction processes past events together with further relevant data for learning an event prediction method to map the processed input to predicted future events that are of interest to the respective application domain. Being

---

Authors' addresses: Janik-Vasily Benzin, Technical University of Munich, Germany; TUM School of Computation, Information and Technology, Department of Informatics, Boltzmannstr. 3, Garching, Germany, 85748, [janik.benzin@tum.de](mailto:janik.benzin@tum.de); Stefanie Rinderle-Ma, Technical University of Munich, Germany; TUM School of Computation, Information and Technology, Department of Informatics, Boltzmannstr. 3, Garching, Germany, 85748, [stefanie.rinderle-ma@tum.de](mailto:stefanie.rinderle-ma@tum.de).

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2023 Association for Computing Machinery.

Manuscript submitted to ACM

Manuscript submitted to ACM

1

able to anticipate events is highly beneficial to many application domains such as business, healthcare, transportation, crime and natural disasters [337]. The value of knowing and ideally understanding predicted future events today lies in the ability of a user in the application domain to act on the predicted future events, e.g., by focusing police resources in areas with high crime risk. To achieve the value, challenges regarding event prediction such as knowledge on the true relationships between causes and effects of events [337], heterogeneous multi-output predictions, e.g., predicting the time and location of future events [248], complex dependencies among the predicted future events as they may interact with each other [198], real-time stream of past events that requires continuous monitoring [256, 267], and challenges in the past event data regarding data properties and quality [256] have to be solved [337].

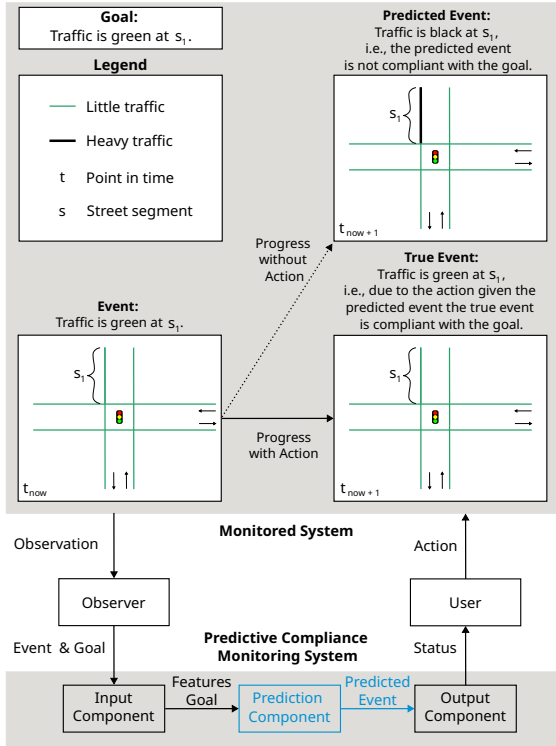


Fig. 1. Event Prediction as a Component of a Predictive Compliance Monitoring System.

Table 1. List of Abbreviations

Abbreviation	Meaning
PPM	Predictive process monitoring
PCM	Predictive compliance monitoring
CM	Compliance monitoring

Table 2. Application domains of covered articles

Domain	Short	Share	Count
Business	B	0.42	109
Engineering Systems	E	0.12	31
Healthcare	H	0.09	23
Political Events	P	0.09	23
Cyber Systems	I	0.08	22
Media	M	0.07	19
Crime	C	0.06	15
Natural Disasters	N	0.04	10
Transportation	T	0.03	8
$\Sigma$		1.00	260

Existing work on event prediction has proposed a multitude of different solutions to those challenges. However, the solutions are often only developed within the respective application domain and corresponding research area. Because the type of event and the laws and rules governing the occurrence of events greatly differ among the application domains, e.g., large-scale weather events governed by the laws and rules of meteorology in the domain of natural disasters or small-scale business events of a business process governed by the respective business regulations and internal procedures in the domain of business [9], event prediction research typically approaches the problem in an event-type-and domain-dependent manner [256, 337]. Despite apparent advantages of this approach to event prediction, the result complicates exchanging event prediction methods and solutions for the event prediction challenges with

respect to different application domains and corresponding research areas. Different terminology for event prediction such as predictive process monitoring (PPM) (cf. Table 1) in the field of business process management [72, 211, 290] or anomaly prediction in the field of engineering systems [180] exacerbates the problem of cumbersome method and knowledge exchange among various event prediction research areas. By taking a systems perspective on event prediction leading to a system for event prediction and placing the diverse range of event prediction methods in that system allows to integrate and, thus, foster knowledge and method exchange. Moreover, the system perspective enables the elicitation of general requirements for event prediction methods that serve as a scheme for assessing existing work. A comprehensive assessment of existing work sheds light on potential blind spots missed so far and establishes a well-founded understanding of the current research status for open challenges and future research directions.

To better understand event prediction through a systems perspective, Figure 1 depicts an intersection example from the transportation domain. The intersection has a road segment  $s_1$  that is depicted at two points in time, the current point in time  $t_{\text{now}}$  and a future point in time  $t_{\text{now}+1}$ . Our goal for the intersection with road segment  $s_1$ , the *monitored system* in this example, is to maintain green traffic for  $s_1$ , i.e., on road segment  $s_1$  little traffic should be maintained. Currently, event “Traffic is green at  $s_1$ ” at  $t_{\text{now}}$  *complies with* our goal. This event together with the goal is observed by the *observer*, e.g., a video camera directed at road segment  $s_1$  and a user interface for setting goals. In general, the observer is an abstract interface between the monitored system in the real-world and the system for event prediction, i.e., it typically consists of a set of sensors, user interfaces and/or application programming interfaces.

Then, the observer sends an event and goal to the *input component* of the generalized *predictive compliance monitoring* (PCM) system (cf. Figure 1) to process the event and goal for subsequent prediction by the *prediction component*. *Predictive compliance monitoring* is concerned with continuously anticipating whether the monitored system complies to a goal or set of goals in the future and supporting the user in understanding and acting on the anticipated compliance status [256]. Thus, a generalized PCM system conceptualizes event prediction through its input component and prediction component, i.e., the event prediction method is conceptualized as the prediction component. After pre-processing and encoding the event in the input component, the prediction component predicts “Traffic is black at  $s_1$ ” at point in time  $t_{\text{now}+1}$ . Hence, the future progress of the monitored system is predicted to be non-compliant with our goal.

The *output component* of the generalized PCM system prepares a status for the user, e.g., an urban traffic control center, that decides on the necessary action to change the future towards a compliant status, e.g., rerouting incoming traffic for road segment  $s_1$ . Due to the action, the ground-true real event (true event) at  $t_{\text{now}+1}$  is “Traffic is green at  $s_1$ ” and, thus, compliant with the goal instead of the predicted “Traffic is black at  $s_1$ ” that is non-compliant. To sum up, the user of the application domain is able to anticipate the future of the monitored system with respect to the goal and act accordingly on the account of the generalized PCM system.

The example indicates how the system perspective benefits the integration of event prediction methods by giving the method a general, yet practical context that covers important parts such as goals and actions. To comprehensively classify, integrate, and assess existing work on event prediction methods with the aim of facilitating knowledge sharing and deriving open challenges and research directions, this survey is guided by the following research questions:

**RQ1** How can existing event prediction methods be classified?

**RQ2** How can existing event prediction methods be assessed in a general, consistent, and systematic manner?

**RQ3** To what extent do existing event prediction methods fulfill assessment criteria?

**RQ4** Which open challenges and research directions remain in the field of event prediction?

## 1.1 Related Surveys

Due to its integration aim, this survey is related to different existing surveys. We structure the related surveys along the three pillars of our research methodology as depicted in Figure 2.

Starting with the event prediction survey pillar, [337] surveys the literature on event prediction with the aim of a systematic taxonomy and summary of existing approaches and application domains as well as standardizing event prediction evaluation for facilitating future development of event prediction. Considering the scope of [337], our survey focuses on the event prediction method instead of event prediction (resulting in 107 of the 216 references as depicted in Figure 2) and extends the scope in the direction of uncovered event prediction method research (67 added articles) and uncovered PPM research (86 added articles). Aside from the scope, [337] focuses on the techniques to solve the identified event prediction research problems, e.g., regression, point process and survival analysis technique for solving the problem of predicting a continuous point in time of future events. Hence, the existing approaches are classified and summarized along inherent properties of the proposed solutions. With respect to classifying existing approaches, our survey extends the classification in [337] by abstracting from the respective technique and integrating taxonomies proposed in other surveys [72, 109, 211, 256]. With respect to summarizing existing approaches, our survey extends the summary in [337] by summarizing existing work along the requirements for event prediction methods that we elicit through a systems perspective, i.e., our summary takes an external, systems perspective on event prediction compared to the internal, methodological perspective of [337]. Furthermore, the requirements are used to assess existing event prediction methods, resulting in a comprehensive assessment that further extends [337].

The event prediction search pillar includes eleven surveys [20, 109, 120, 130, 135, 205, 209, 230, 250, 281, 283]. Each of these surveys has a certain scope, e.g., [205, 209] focus on time-series data, [281, 283] focus on event prediction of business processes, and [20] on unstructured text events. When compared to the survey at hand, on top of a narrow scope, the eleven surveys neither take a system perspective to generally place event prediction method in a context, nor elicit requirements for assessing the methods, nor comprehensively assess existing work. [109] proposes a taxonomy for event prediction methods with a concept of domain-specificity that is integrated in our taxonomy.

The PPM research pillar includes 22 surveys [71, 72, 116, 141, 154, 157, 160, 167, 207, 211, 216, 222, 247, 256, 273, 275, 281, 286, 290, 297, 316, 318]. Similar to the surveys in the remaining event prediction pillar, the scope of the existing 22 surveys is limited. While the scope in the remaining event prediction pillar is heterogeneous, in the PPM pillar it is homogeneous and exclusively focuses on next events (e.g., [281]) or key performance indicators of business processes (e.g., [297]). [256] is the only PPM survey that takes a systems perspective for eliciting requirements and assessing existing PPM methods. The scope in [256] includes compliance monitoring (CM) (cf. Table 1), i.e., the ability to continuously monitor the compliance status of a business process with respect to a set of compliance constraints, to investigate the potential benefits of combining CM and PPM with respect to the established CM functionalities framework [193]. The investigation result is PCM as a combination of PPM and CM from a systems perspective. Existing work on CM and PPM is assessed with respect to an extended CM functionalities framework with a particular focus on prediction. Our survey generalizes the PCM concept to event prediction, develops a conceptual PCM system and extends and integrates the requirements in [193, 256] with the challenges in [337] and requirements contained in the 260 articles of our existing work selection. [72, 211] propose a similar concept of domain-specificity for classifying PPM methods that is generalized and integrated in our taxonomy, while the distinction in [256] into methods that directly predict the goal of PPM and methods that indirectly predict the goal of PPM via an intermediate representation is generalized and integrated in our taxonomy. For instance, the PPM goal is the late shipment status of online shop

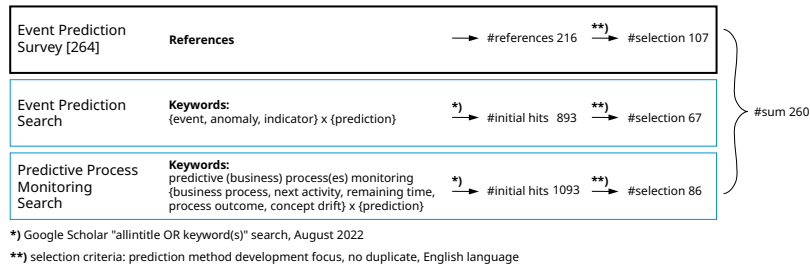


Fig. 2. Methodology for Literature Compilation

delivery packages. Predicting the late shipment status of an online shop delivery package is directly predicting the PPM goal and conceptually similar to the traffic status in the intersection example (cf. Figure 1). Predicting all future events pertaining to the delivery package and, then, evaluating the shipment status on the future events is indirectly predicting the PPM goal and conceptually similar to predicting the number of vehicles on road segment  $s_1$  to subsequently derive the traffic status in the intersection example.

All in all, our survey integrates various scopes of existing surveys and takes a system perspective for placing event prediction in a general context, eliciting requirements and assessing existing event prediction methods.

## 1.2 Research methodology

We follow a twofold methodology for selecting relevant articles in the event prediction literature. First, we take the recent survey in [337] as a baseline on existing work. Second, we extend the baseline by searches in two directions: event prediction to capture uncovered event prediction method research of the baseline and predictive process monitoring as event prediction literature that uses different terminology. The twofold methodology is illustrated by black and blue color respectively in Figure 2 together with keywords for searches and the results.

Then, we apply the same filter denoted by \*\*) in Figure 2 on each stream of literature resulting in selections that represent significant advancements in prediction method development, i.e., work that focuses on novel event prediction methods instead of only applying a classical machine learning algorithm to a new application domain or focusing on domain-specific feature engineering. Lastly, we take the union of the three selections yielding 260 articles covered in this survey. These articles span a wide array of different application domains each having a large enough absolute article count for proper representation (cf. Table 2).

## 1.3 Contribution and Structure

Taking a systems perspective on event prediction methods to integrate disparate areas of research results in the following four contributions of this survey:

- **Comprehensive classification of event prediction methods from a diverse range of application domains and disparate research areas.** Event prediction method research spans areas like social sciences, medicine, engineering, and computer science. Researchers at the intersection of these areas and data mining / machine learning propose a large collection of existing event prediction methods. This collection is classified using a taxonomy to support practitioners in finding suitable methods and researchers in developing new methods.

- **Integration of disparate research areas through the generalized PCM system and requirements elicitation for event prediction methods.** By taking a systems perspective, we identify an event prediction method as a component of a generalized PCM system. The generalized PCM system further integrates contributions from the various research areas on event prediction methods. The aim of this integration is to elicit similar requirements for event prediction methods and to facilitate the mutual exchange of method features. Method features are properties of the event prediction method that meet the requirements.
- **Qualitative assessment of existing methods with respect to the requirements and thorough analysis of current research status.** A qualitative assessment of all existing methods gives an in-depth view on the current research status. The current research status is illustrated with examples from the literature, analysed with respect to potential dependencies between requirements and discussed in detail.
- **Open challenges and discussion of future research directions.** The assessment result sheds light on blind spots that embody open challenges to be solved by future research. We identify five challenges and discuss the possible future research direction for each challenge.

The remaining survey is structured into five sections of which the first four correspond to the four research questions respectively. To begin with, Section 2 introduces a notion of an event prediction method, a taxonomy and according classification of existing event prediction methods ( $\mapsto$  RQ1). Section 3 gives the reasoning for placing an event prediction method as an integral component into the generalized PCM system, presents the generalized PCM system in detail and explains the requirements for assessing event prediction methods ( $\mapsto$  RQ2). Then, Section 4 shows to what extent existing prediction methods fulfill requirements and pinpoints exemplary method features that meet the respective requirement ( $\mapsto$  RQ3). Next, Section 5 highlights major open challenges and distills research directions ( $\mapsto$  RQ4). Lastly, Section 6 concludes and identifies limitations of the survey.

## 2 A NOTION AND TAXONOMY FOR EVENT PREDICTION METHODS

This section first introduces a notion of the event prediction method (Section 2.1). Then, it presents the rationale for our taxonomy of event prediction methods and the classification of event prediction methods based on our taxonomy (Section 2.2).

### 2.1 Event Prediction Method Formulation

In the beginning, event prediction requires a goal or a set of goals that is either formulated by researchers during their motivation for advancement of event prediction methods or by practitioners during implementation and application of event prediction. Without the prediction goal, we do not know what we predict and lack a purpose for predicting. Prediction goals are as diverse as event prediction domains are. For example, "The traffic is green at  $s_1$ " for the intersection example depicted in Figure 1 is an event prediction goal or "Reduce re-hospitalization and increase treatment outcome of heart failure patients" [80] is a set of event prediction goals. Event prediction goals define the monitored system implicitly or explicitly and how we perceive various current and future states of the monitored system by differentiating desired or compliant from undesired or non-compliant states. In the case of the intersection example, the goal implicitly defines the monitored system to be the depicted intersection and that green traffic is regarded as compliant, whereas black traffic is regarded as non-compliant. In the case of hospital patients with previously diagnosed heart failures,

the goal defines the monitored system to be hospital heart failure patients and that future states with reduced re-hospitalization counts or rates and improved treatment outcome statistics are desired in contrast to, e.g., states with constant re-hospitalization.

**Definition 2.1** (Event Prediction Goal). An event prediction goal  $g$  is a well-formed formula in some logic  $G$  with atomic propositions over the domain of the monitored system that evaluates to either true or false given some predicted instantiation of its variables. If the event prediction goal  $g$  evaluates to true, we say the monitored system future state is compliant, otherwise it is non-compliant.

For the sake of simplicity, we did not state the presented examples as actual formulas of a logic such as propositional logic or linear-time logic and point towards existing methods on translating natural language to formulas, e.g., of linear-time logic [31]. The two important points here are the existence of at least one event prediction goal and that it is stated in terms of atomic propositions over the domain of the monitored system (implicitly or explicitly).

Based on the event prediction goal and the monitored system, we identify two kinds of data used for event prediction<sup>1</sup>: Historical event data  $Y_0 \subseteq \mathcal{T}^- \times \mathcal{L} \times \mathcal{S}$  and historical indicator data  $X \subseteq \mathcal{T}^- \times \mathcal{L} \times \mathcal{F}_I$ , where  $\mathcal{T}$  is the time domain,  $\mathcal{T}^- \equiv \{t | t \leq t_{\text{now}}, t \in \mathcal{T}\}$  are all times up to the current time  $t_{\text{now}}$ ,  $\mathcal{T}^+ \equiv \{t | t > t_{\text{now}}, t \in \mathcal{T}\}$  are all times after the current time,  $\mathcal{L}$  is the location domain,  $\mathcal{S}$  is the event semantic domain and  $\mathcal{F}_I$  is the indicator feature domain that does not include the time and location domain [337]. Data types  $Y_0$  and  $X$  are generated by the observer (see Figure 1) through observing the real-world, e.g., the observer generates an event  $y = ("05.01.2023\ 13:00", "s_1", "Traffic = green")$  by observing that two vehicles have stopped in front of the red light and an indicator  $x = ("05.01.2023\ 13:00", "s_1", "Workday = Yes")$ . This data is sent alongside the event prediction goal to the input component of the PCM system (cf. Figure 1 and Section 1) that maps the data to suitable features  $F \subseteq \mathcal{T}^- \times \mathcal{F}_Y \times \mathcal{F}_X$ , where  $\mathcal{F}_Y$  is the domain of features based on historical event data and  $\mathcal{F}_X$  is the domain of features based on historical indicator data, by means of preprocessing and encoding. The input component outputs the features and prediction goal to the prediction component that is realized by an event prediction method, e.g.,  $f = ("05.01.2023\ 13:00", (1, 0), 1)$  for a one-hot encoded traffic value and identifying the workday by 1.

**Definition 2.2** (Event Prediction Method [337]). Given features  $F \subseteq \mathcal{T}^- \times \mathcal{F}_Y \times \mathcal{F}_X$  and an event prediction goal  $g \in G$  or set of goals, an event prediction method outputs predicted future events  $\hat{Y} \subseteq \mathcal{T}^+ \times \mathcal{L} \times \mathcal{S}$  such that for each future predicted event  $\hat{y} = (t, l, s) \in \hat{Y}$  for  $t > t_{\text{now}}$  it holds that (C) either the goal  $g$  or the set of goals is itself a future predicted event or it can be evaluated on the future predicted events.

In the traffic example, let us assume that we learn a neural network model that takes the most recent feature  $f$  and outputs  $\hat{y} = ("05.01.2023\ 14:00", "s_1", "Traffic = black")$  as the predicted future event (cf. Figure 1), i.e., the goal is itself the predicted event. Consequently, the future state of the road segment  $s_1$ 's traffic is predicted to be non-compliant, so an action by the user is required.

Following our scope and aim (cf. Section 1.1 and Section 1.2), we separate feature engineering from event prediction and focus on the remaining event prediction method. We further add condition (C) on prediction goals such that the event prediction method has an explicit aim expressing our reasons why we are interested in predicting the future events and how we perceive the possible predicted future events, namely as compliant with our desired future state or as non-compliant. Defining event prediction methods with an explicit goal has three benefits. First, all existing work in

<sup>1</sup>In the following, we adopt the existing definition of the event prediction problem and method in [337] as close as our different perspective and scope allows.

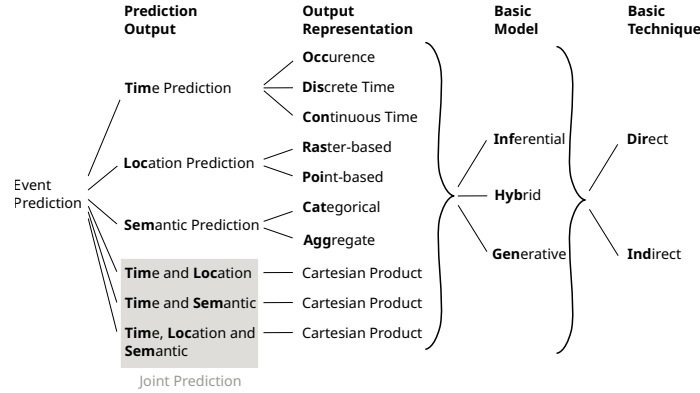


Fig. 3. Taxonomy of event prediction methods [109, 256, 337]. Abbreviations for each class highlighted in bold.

our selection state a goal or set of goals in their motivation that expresses why predicted future events of the method are of interest to the user and how a specific predicted future event is perceived, e.g., [80] predict future patient treatment outcomes in clinics for a proactive treatment effect analysis. Hence, including condition (C) in the definition combines the *how* (i.e., the event prediction method) and the *why* (i.e., the prediction goal) for event prediction. Second, it links predicted future events to actions. Since an action aims to change a monitored system towards a desired state, linking actions to future predicted events requires a notion of a desired state. With prediction goals the notion of desired states is defined by compliant or desired future states and by non-compliant or undesired future states. Based on this notion, an action can aim for a desired future state of the monitored system. Third, condition (C) emphasizes how we can decide whether a specific event prediction method is successful in predicting future events.

Furthermore, the condition (C) captures the fact that not all prediction methods predict occurrence times, locations and semantics of future events simultaneously, as many focus on a single domain that is sufficient for the prediction goal, e.g., [78] focuses on the time domain by predicting the occurrence of abnormal deformations for a single machine tool of a production facility, [237] focuses on the location domain by predicting the area that a particular crime will occur and [200] focus on the semantic domain by predicting the future activity of a business process instance [337].

## 2.2 Classification of Event Prediction Methods

Event prediction is a challenging and complex endeavour that spans many, diverse research areas and application domains. Plenty of methods can predict future events, yet we can navigate through the large body of existing work in a quick and structured manner by means of a simple taxonomy as depicted in Figure 3. First, we describe and explain the taxonomy in Section 2.2.1 and then present the classification of the 260 selected articles aggregated by the first level of the taxonomy tree in Section 2.2.2.

**2.2.1 Taxonomy of Event Prediction Methods.** In the following, we present a taxonomy of event prediction methods in Figure 3 that aims at integrating and generalizing existing taxonomies in a consistent and concise manner. The first two levels of the taxonomy, the prediction output and output representation, are based on [337], while the third level, *basic model*, generalizes ideas from [72, 109, 211] and the fourth level, *basic technique*, generalizes [256] and captures the two ways a prediction goal is related to an event prediction method (cf. condition (C) in Definition 2.2).



**Prediction output.** Event prediction method outputs are heterogeneous due to many different prediction goals that refer to heterogeneous domains such as healthcare, cyber systems or transportation. Nevertheless, four broad classes of prediction outputs are sufficient to represent this heterogeneity: *time* (tim), *location* (loc), *semantic* (sem) and *joint* prediction as depicted in Figure 3. Event prediction methods that focus on the time domain of future events, i.e., when the future event is predicted to occur, are classified into the *time* prediction output, e.g., forecasting the future progress of a machine sensor time-series [13] or predicting the time-to-event for patient treatment outcomes [103]. If the focus of the method is on the location domain of future events, i.e., where the future event is predicted to occur, the method is classified into the *location* prediction output, e.g., predicting the next place of social media users [308]. Event prediction methods classified into the *semantic* prediction output focus on the semantic domain of future events, i.e., what the future event means or how the event, that typically involves symbols or natural language, can be described, e.g., predicting the next activity of a business process [200]. The joint prediction class is the only class that focuses on simultaneously predicting multiple prediction domains of future events such as the time and location domain.

**Output representation.** Depending on the prediction output type, research has represented the respective output differently, i.e., each prediction output class can be further distinguished by its output representation [337]. The time prediction output has been represented as either a binary classification task that signifies whether the predicted event occurs in the future or not (*occurrence* (occ) in Figure 3), or a discretization of the future time domain into time bins with the output being the specific time bin that the future event is predicted to occur (*discrete time* (dis)) or a straightforward, continuous value in the time domain (*continuous time* (con)). The location prediction output has been represented as a grid cell that is the result of rasterizing a continuous space into spatial regions (e.g., integer-valued longitude latitude coordinates in the geographic coordinate system [48]) yielding a *raster-based* (ras) representation or as a point in the continuous space (*point-based* (ras)). The semantic prediction output has been represented as a categorical variable, e.g., the activity label of a business process event, or as an aggregate variable, e.g., a set of key-value pairs for sentences consisting of a subject, object and verb.

The next two levels of the taxonomy tree focus on the similarities of existing work in their developed event prediction method, i.e., despite the heterogeneity of domains, methods, outputs and output representations, they all share only three types of basic models and two types of basic techniques.

**Basic model.** The basic model underlying every event prediction method is distinguished into *inferential* (inf), *hybrid* (hyb) and *generative* (gen) [109]. This distinction may be called differently, e.g., *implicit* vs. *explicit* [72] or *process-unaware* vs. *process-aware* [211], yet in every instance identifies the type of model employed for developing the method with varying degree. The differentiating property of the model is the degree of knowledge the modeller has or can theorize about the (prediction) goal-dependent salient factors of the system under research. For example, predicting the patient treatment outcome in clinics results in the goal-dependent salient factors to be known risk factors for a negative treatment outcome.

If the degree is high, in the *generative* model class, the modeller applies specific knowledge of the goal-dependent salient factors, e.g., of the nature and the rules, guiding the monitored system that determine or significantly influence the evolution of events to design more specific, customized models and explicitly leaves out factors that can be shown to be irrelevant for the evolution of events. Thus, the modeller develops a counterfactual explanation of the event evolution in the monitored system expressed through the model [254]. The result is often a computational model that can be executed or simulated [314]. For example, business processes can be modelled by the Business Process Modelling Notation (BPMN) [51] or a similar business process modelling notation that carries the knowledge about business processes such as the observation of decision points for process instances. Comparing this knowledge and type of

model to a weather system, we cannot, for example, ascertain similar decision points in the weather system, i.e., the business process model is specific to properties of the application domain and, thus, only explains next activities of business processes and not what the weather on the next day will be.

If the degree is low, the monitored system is either not researched in detail yet or its nature is too complex in order to pinpoint a fitting class of knowledgeable models for its behavior. Consequently, in the *inferential* model the modeller applies generic, mathematical models [314], e.g., a regression model for predicting the time bins of future anomalies in long term evolution (LTE) mobile networks [115], that learn patterns in the data observed from the monitored system to predict its behavior. Here, the knowledge about goal-dependent salient factors may still enter the overall PCM system, but through domain-specific feature engineering and not through the event prediction method design and its underlying basic model type.

Lastly, in the *hybrid* case the modeller develops an event prediction method whose model can be divided into a part for which the generative approach is applied and a part for which the inferential approach is applied, e.g., learning a Markov chain model (*inferential*) for each state of a BPMN process model of a business process (*generative*) for predicting probabilities of next activities executed for a specific process instance [159].

**Basic technique.** The basic technique underlying every event prediction method is distinguished into *direct* (dir) and *indirect* (ind). Any event that can be generated through observing the real-world is generated by means of a similar formula as the prediction goal (cf. Definition 2.1). That is, at some point before the prediction, we introduce a formal notion of what constitutes an event given the real-world or the recorded data. This formal notion is expressed through the formula, e.g., green traffic is constituted by up to two vehicles on the monitored road segment before the intersection or a crime is constituted by the respective law that is used to convict the guilty. In the latter case, the evaluation of the law to either true or false may be up to the respective court, but in the data this ambiguity does not usually show up. Overall, an event is defined by a formal notion and the prediction goal is defined by another formal notion.

If the output of the prediction method given the features is the prediction goal itself, the prediction goal is *directly* predicted as a future event (cf. first part of condition (C) in Definition 2.2). This is the case for the example in Section 2.1 in which the neural network predicts orange traffic. If the output of the prediction method is a predicted future event or multiple future events on which we evaluate the prediction goal, then the prediction goal is *indirectly* predicted through future events (cf. second part of condition (C) in Definition 2.2). For example, [13] first predict the evolution of machine properties observed by sensors and then evaluate prediction goals defining certain types of anomalies on the predicted time-series. To put it in other words, the direct technique defines the output of the prediction method as the value of the prediction goal, whereas the indirect technique defines the output of the prediction method as an intermediate notion of an event on which the prediction goal can be evaluated. The choice of basic technique has major consequences on the architecture of the PCM system and its properties, as we will see in Section 3.1.

**2.2.2 Classification of Existing Work.** Overall, 154 articles belong to the class of time event prediction, while 10 articles, 46 articles and 49 articles belong to the class of location, semantic and joint prediction respectively. Hence, most of the existing work on event prediction focuses solely on the time domain. If we consider the importance of the time domain for prediction, this is no surprise. By focusing on a single or small set of prediction goals and a tractable set of events such as a certain set of machine anomalies for engineering systems, treatment outcomes for hospital patients or crimes in a certain geographical area, research on time prediction breaks the complexity of event prediction down to a bearable level while maintaining its attractiveness.

As a consequence of this active research on time prediction, most of the classes are occupied as depicted in Table 3, i.e., most of the various combinations of output representation, basic model and basic technique have been at least researched once. The largest single class is directly predicting the occurrence of the prediction goal with an inferential model. Here, the majority of work chose the business domain as its application. This finding is a direct consequence of active research on predicate prediction for outcome-oriented PPM with the use of standard machine learning methods, i.e., predicting the process outcome or remaining time of process instances expressed through the prediction goal without predicting further domains such as the next activity in the semantic domain is the most active research field in the BPM community [256].

Interestingly, the indirect technique for generative models is consistently missing in time (cf. Table 3), location (cf. Table 4), semantic (cf. Table 5) and joint prediction (cf. Table 6). Taking these missing classes together with a decreasing number of methods from inferential to generative model type, the research direction on integration of mechanistic knowledge and data-driven models in [337] is illustrated.

Considering the existence of many classes with very few work, 19 classes with only a single work and nine applications domains, a large area of possible research remains open despite active and already considerable research on event prediction. In Section 5, we will deepen our insights on possible areas for further research and discuss how this relates to commonly interesting prediction goals in the various application domains. For now, taking the classification results as a starting point, researchers can quickly search for conceptually related work in other application domains that often also implies a different research area and community, while practitioners can use this classification to look for suitable work on their respective problem.

### 3 COMPREHENSION OF EVENT PREDICTION METHODS THROUGH THE GENERALIZED PCM SYSTEM

If we want to fully understand event prediction in all of its heterogeneous aspects, multi-scale dimensions and overall complexity, we must achieve to grasp the set of systematic relationships of the system in which event prediction is conceptualized or implemented [84, 85, 110, 155, 254]. This feature of understanding emphasizes the need for a systems perspective on this topic. In general, considering the diverse nature of prediction goals throughout the various application domains, the type of system for event prediction must be a decision support system rather than an autonomous system, i.e., the user that decides what action is carried out is a human person (cf. Figure 1 and Figure 4). For instance, deciding on the treatment plan of patients based on predicted future treatment outcomes must be done by physicians [65]. Furthermore, the utility of event prediction critically depends on the lead time, the time from the prediction of an event to the true event time. This property necessitates a prediction at run-time and, thus, an online, stream-based conceptualization of the event prediction system [267]. The source streams have to be continuously monitored such that the lead time is as large as possible [337]. Altogether, we need a decision support system that continuously monitors and predicts with the aim of distinguishing non-compliant, action-demanding from compliant future states.

To that end, we present a conceptual PCM architecture that entails these properties in Section 3.1 and elicit assessment requirements in the form of requirements from the system architecture in Section 3.2.

#### 3.1 Widening the Scope: Event Prediction Method in the PCM system

Before we can present the PCM system in detail, we elaborate on how it applies to event prediction. Through our elaboration, we understand how multiple research areas work towards the same goal with more or less similar methods,

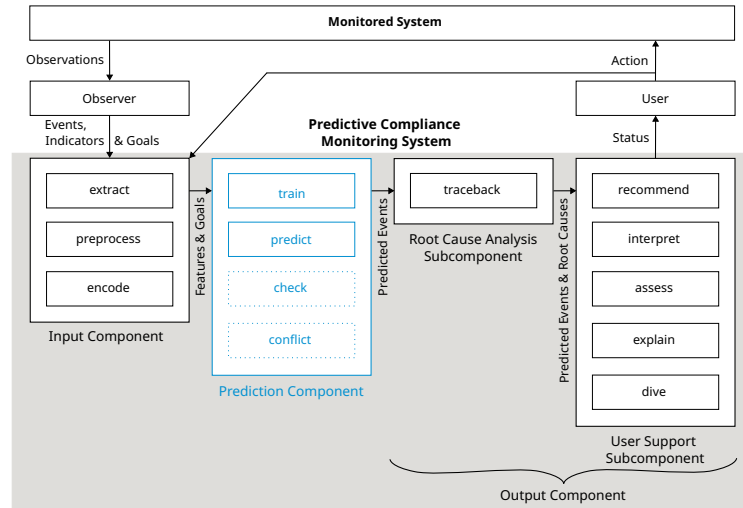


Fig. 4. Generalized Predictive Compliance Monitoring System - Conceptual and Structural View

yet different terminology. Moreover, we illustrate the community- and application-domain-spanning characteristic of event prediction.

In the field of business process management, a PCM system aims to help companies monitor and understand the future compliance status of their business processes to act proactively towards a desired future compliance status [193, 256]. To that end, a PCM system integrates research on CM and PPM. CM aims to evaluate compliance constraints on the current and future events or key performance indicators of a business process [193], while PPM aims to predict the future events or key performance indicators of a business process [72, 211]. By comparing CM, PPM, and the event prediction method definition (cf. Definition 2.2), we identify the following similarities: compliance constraints are prediction goals, future key performance indicators of a business process correspond to future predicted events (cf. condition (C) in Definition 2.2), PPM is an event prediction method, and the business process is a monitored system.

All in all, if we generalize

- (1) the monitored business process and its events as a monitored system and its events,
- (2) the prediction of the future compliance status as event prediction methods,
- (3) the anticipated compliance status as the evaluation result of an event prediction goal,
- (4) and the proactive actions as any action an user can take with the support of a decision support system,

then the generalized PCM system depicted in Figure 4 is a conceptual model of decision support systems featuring event prediction. Not only does this generalization place event prediction inside a generic system necessary for understanding its merits more deeply, it also represents an integration of research on the various aspects of event prediction and monitoring into a common theme. Furthermore, the generalization of a PCM system from business process management to a decision support system for event prediction in general enables us to mutually transfer requirements and solutions between the respective research areas.

A PCM system consists the input, prediction, and output component (cf. Figure 1) which are refined in Figure 4. Moreover, a feedback cycle is added from the user to the input component that contains the action as a more direct and efficient way of accounting for the effects of decided upon actions onto the monitored system in the PCM system [337].

**Input component.** The input component *extracts* data from the observer, *preprocesses* the data and *encodes* it into the respective feature space required by the event prediction method (cf. Definition 2.2). In practice, the input component has to implement the various functionalities of *data collection* proposed by the data management community, in particular data acquisition and data labeling [260]. Within data acquisition, the tasks of augmentation through data integration / data fusion are particularly important, while for the case of input event data lacking an event notion, the task of data programming or event extraction [75] become additionally relevant. Considering that the observer is only an abstraction for multiple, heterogeneous data sources, the input component's *extract* functionality has to deal with challenges of raw data ingestions from various data sources [199].

The input component's *preprocess* functionality covers the transformation of ingested data to a common representation based on a series of rules [139], data cleaning [132], schema matching for relational data [246], deduplication and record linkage [86, 107, 214], and, finally, canonical column selection for relational data [276]. In the case of non-relational data such as continuous, unlabelled sensor readings, schema matching is replaced by data programming or event extraction [75] to define an event notion for event sources. In case of low-level events coming from the event source, the task of event abstraction [294] may be necessary. Since events from multiple sources may actually be the same event, these events have to be matched to finalize the data integration. As a last step in the input component, the *encode* functionality maps the integrated data to the event and indicator feature space. Here, it is important to follow the principles of feature engineering [345], while the resulting features are typically application domain- and/or prediction goal-specific.

**Prediction component.** The prediction component is realized by an event prediction method that *trains* a prediction model and *predicts* the future events. In Definition 2.2, the future event can either be (i) the prediction goal or (ii) an intermediate event on which the prediction goal can be evaluated. Case (i) does not require the additional functionalities *check* and *conflict*. For case (ii), the evaluation of the prediction goal is realized by *checking* the compliance of the intermediate event with respect to the prediction goal. If a monitored system is subject to a set of prediction goals, these goals may be in conflict [193]. The PCM takes this possibility into account when checking the set of prediction goals on the intermediate event and includes this information in its output.

**Output component.** The output component is refined into two subcomponents; The **root cause analysis component** and the **user support component** to underline the need for model transparency and interpretability (in short explainability), accountability, and prescriptiveness [193, 256, 337]. The ability of the PCM to *trace* the non-compliance of the future events *back* to the root cause(s) contributes to an explainable and actionable prediction. The prediction with root cause analysis is explainable, because a causal relationship between a root cause and its anticipated result exists [254]. Furthermore, root cause analysis makes the prediction actionable, since the user can derive countermeasures for non-compliant future events through acting on the root causes. Researchers in the application domains of engineering and cyber systems, i.e., IT systems, typically develop root cause analysis for explaining the root causes of machine faults or system anomalies to the operation manager [83, 195].

Finally, the PCM system supports the user in understanding the results and acting on them. To that end, the user support subcomponent *recommends* actions and countermeasures for non-compliant predicted future events [229], *interprets* the impact of the proposed actions on the predicted future events [111], and *assesses* the overall uncertainty that remains throughout all prior functionalities [131]. Separately showing all the results to users may quickly overwhelm, distract from major points, and hinder understanding, which is why the gist of the results are *explained* to them. In

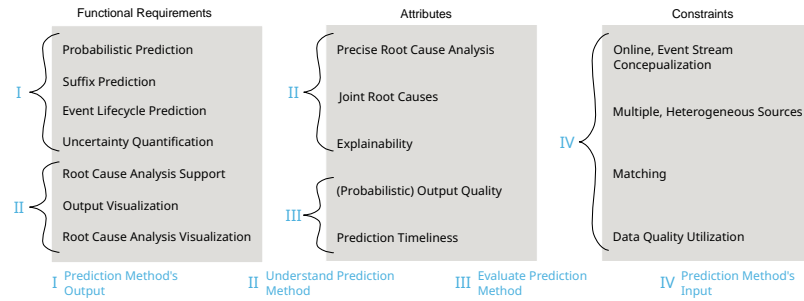


Fig. 5. Requirements engineering result for event prediction methods in the context of predictive compliance monitoring systems categorized by functional requirements, attributes and constraints [108] and annotated by the requirements respective target.

Section 5, we discuss a potential approach to this challenging functionality. If the gist is first explained to the user, but more detailed information is necessary for understanding or analysing, the user can *dive* into the respective results.

All in all, principles of systems engineering applied to the design of a conceptual decision support system featuring event prediction result in a new perspective on event prediction [24]. This new systems perspective encompasses a holistic view on the many facets of event prediction, yet is simple enough to guide us in our understanding of functionalities surrounding the actual prediction that are necessary to hone its value.

### 3.2 Assessment Scheme for Event Prediction Methods

After our refined understanding of event prediction through its conceptualization as a component of PCM systems, we take the systems perspective one step further by eliciting requirements for event prediction methods from existing surveys [193, 256, 337] and iteratively refining them during the assessment of the 260 articles in our selection. Aside from advancing our account of event prediction, the requirements serve as an assessment scheme for assessing the selection of existing work.

Following a concern-based taxonomy of requirements [108], the 16 event prediction method requirements in Figure 5 are classified into seven functional and nine non-functional requirements in the form of five attributes and four constraints. Attributes specify qualities of the method, whereas constraints limit the solution space from which methods can be instantiated. Furthermore, each requirement is assigned to one of four specification targets: (I) the prediction method's output, i.e., the properties of predicted future events, (II) understanding the prediction method, i.e., properties or qualities of the method that help to grasp the systematic relationships between the prediction and its input through, for example, visualization [8, 208], (III) evaluating the prediction method, i.e., what qualities are presented for evaluation, and (IV) the prediction method's input, i.e., what properties of the input were considered in developing the method.

**Functional requirements.** Considering that the real-world is complex, chaotic, and imperfect, a deterministic prediction of future events idealizes and keeps important nuances in the future events from the user [254, 256, 337]. Therefore, a *probabilistic prediction* that includes information on the likelihood of the predicted future events actually occurring is required. If we want to enable the user to optimize decisions with respect to a comprehensive anticipation of the future, predicting the next event is not enough. A *suffix prediction* specifies the need for anticipating more events following the next event, e.g., knowing of the next three heavy rainfalls instead of the next one, officials can more efficiently allocate resources. Instead of increasing the anticipated time horizon as suffix prediction does, *event lifecycle prediction* refines the granularity of the anticipated event time horizon. Although events are typically conceptualized as

real-world occurrences happening at a certain point of time, each real-world event has a duration and, thus, at least a *start* and an *end* point of time [165, 193]. If we do not predict the lifecycle of events, we are, for example, not aware of the time period affected by the event. Our travel plans remain unchanged, if the traffic jam is likely to vanish by the time we arrive at its location.

As the last functional requirement specifying the method’s output (I), *uncertainty quantification* aims at reflecting the epistemic and aleatoric uncertainty of the method’s output [131]. Epistemic uncertainty consists of model and approximation uncertainty and is the result of our lack of knowledge on the true relationships of the monitored system necessary for the design of a perfect prediction method. As such, it is reducible by further research. Aleatoric uncertainty is an irreducible uncertainty that captures the non-deterministic nature of real-world occurrences. Recently, method development acknowledged the existence of both types of uncertainty and incorporated their quantification in its output [317]. Uncertainty quantification can build trust for the prediction method by measuring confidence metrics [337] that accompany probabilistic outputs and improves the subsequent actions, as the user can decide to wait for more certain prediction results that are more likely to occur.

Despite the fact that root cause analysis and visualization are separated from prediction in the PCM system (cf. Figure 4), it is important that the method itself enables these functionalities or already presents a proof of concept on how these functionalities can be done. Otherwise, it will be more difficult for researchers and practitioners to develop these functionalities given only the method itself. Thus, *root cause analysis support* (RCA support), *output visualization* and *root cause analysis visualization* (RCA visualization) indicate whether the proposed method clearly defines how to achieve root cause analysis or visualization on top of the method respectively or already presents a proof of concept.

**Attributes.** The first set of attributes are the specific qualities *precise root cause analysis* (precise RCA), *joint root causes* (joint RCs) and *explainability* that are essential for an event prediction method, since this first set of attributes helps the user to quickly grasp the situation and act accordingly. The method should enable the precise identification of root causes and detect that multiple true root causes for different predicted future events may, in fact, be the same true root cause. As it is out of the scope of this survey to assess the full variety of how explainability can be achieved or conceptualized [16, 34] and there is no widely accepted definition [34, 104, 254], we limit the attribute to a distinction between *black box* models that are intransparent, *white box* models that are transparent, and a combination of both [34]. Taken together, these three qualities determine the accountability of the method to a large extent. If the user can explain the predicted future events as prerequisites to an optimal strategy for acting on the true root causes, the method and the user both should not be held accountable in case something goes wrong.

The second set of attributes are performance attributes that evaluate the event prediction method itself. [337] presents how the output quality of an event prediction method is evaluated. Considering the requirement on a *probabilistic output quality* (probabilistic quality), we further require a probabilistic evaluation, e.g., with the Brier score [18] or some other appropriate scoring rule [41]. *Prediction timeliness* pertains to the speed a method is able to predict. It determines the lead time, but as a result of concept drift, the training time also has an effect on the lead time. Concept drift due to changing environments such as the COVID-19 pandemic lead to outdated historic data [228, 258, 337]. Consequently, the prediction model must be kept up-to-date. Having to update prediction models results in the training time of the method playing an additional role for a timely prediction, yet we do not know the significance of this role. Moreover, it is not clear what the optimal time for updating is. Aside from these unknowns, training and prediction time comparisons are only reasonable in a fair benchmark that goes beyond the scope of this survey. Hence, prediction timeliness is not assessed in Section 4.

**Constraints.** The event prediction method is constrained by four properties of the input. First, *online, event stream conceptualization* (online conceptualization) is a consequence of the need for monitoring and timely prediction [256, 337]. A method may fail to show how it can be applied to an event stream at run-time. Since there is a difference between online prediction based on an event stream and offline prediction based on an ex-post dataset, e.g., in terms of efficiently storing and updating already received events from the event stream [54], offline event prediction methods may be unsuitable for monitoring.

Second, *multiple, heterogeneous sources* (multiple sources) limit the method solution space to methods that are aware of or actively deal with the challenges of data integration (cf. input component in Section 3.1). Although overcoming these challenges is the responsibility of the input component in a PCM system, the solution may affect the event prediction method, e.g., by necessitating multiple prediction models [248].

Third, the challenge of *matching* events from multiple data sources is emphasized as a separate constraint (cf. input component in Section 3.1). The matching strategy for the input should be consistent with the strategy for matching events in prediction goals to predicted future events and the strategy for matching events for evaluation purposes. Multiple matching strategies exist [337], but may have to employ different equivalence notions that underlie the matching of events [256], e.g., matching based on equivalent machine identifiers for multiple sensors attached to a machine in a production facility.

Fourth, *data quality utilization* limits the solution space to methods that not only acknowledge the existence of data quality issues such as missing data in their design, but also improve the prediction by actively exploiting the existence of data quality issues. [190], for example, develops a self-supervised technique to fully utilize electronic health records that inherently lack a certain type of label.

Conjointly, 15 of the 16 discussed requirements (excluding the *prediction timeliness* requirement) serve as the assessment scheme for existing work on event prediction methods in Section 4.

## 4 ASSESSMENT OF EXISTING EVENT PREDICTION METHODS

This section presents the analysis of 260 papers from the literature compilation in Figure 2. Each paper is analysed with respect to the 15 requirements contained in the assessment scheme introduced in Section 3.2, resulting in an overview for each requirement depicted in Figure 6. The goal of the analysis is to understand the current status of event prediction methods from a systems perspective (cf. Section 1 and Section 3) and to pinpoint open research challenges with their corresponding research directions (cf. Section 5). The results are presented overall in Section 4.1 with details per type of requirement in Section 4.2 - Section 4.4.

The assessment of the requirements uses an ordinal scale: + means that the requirement is met (for the explainability requirement, this assessment corresponds to "white box" cf. Section 3.2), ~ means that the requirement is partly met (for the explainability requirement, this assessment corresponds to both "white box" and "black box") and – means that the requirement is not met (for the explainability requirement, this assessment corresponds to "black box").

### 4.1 Overall Results

The results for each assessment requirements in Figure 6 give a nuanced view on research progress in the field of event prediction methods, as the overall article shares vary across the assessment requirements. By applying a threshold  $\lambda = 20\%$  (corresponds to 52 articles) on the ~ assessment result, we get two sets of requirements: One set that contains requirements with article shares above the threshold (i) and another set that contains requirements below the threshold (ii). In set (i), existing work has proposed methods that meet the requirements *probabilistic prediction*, *RCA support*,



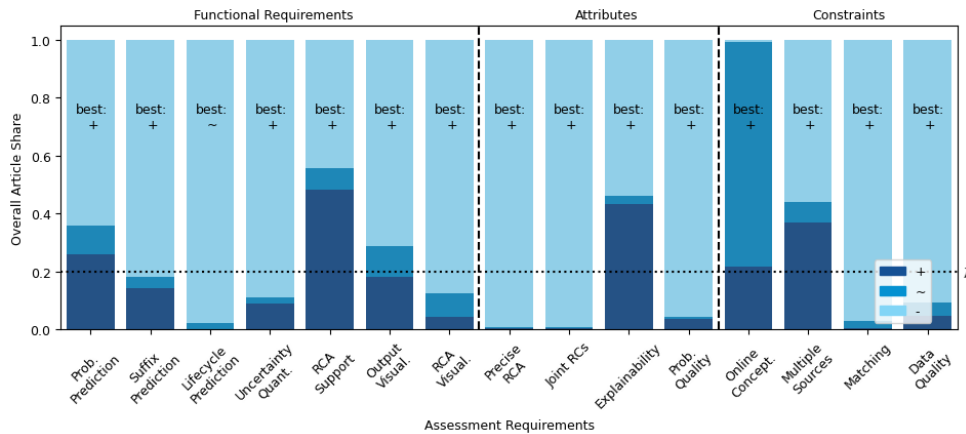


Fig. 6. Assessment results for all articles per requirement.

output visualization, explainability, online conceptualization and multiple sources. The majority of methods meeting those requirements emphasize the need for monitoring, causality, explanation, and visualization and acknowledge that relevant data is typically distributed over multiple data sources. In set (ii), existing work misses the requirements *suffix prediction*, *lifecycle prediction*, *uncertainty quantification*, *RCA visualization*, *precise RCA*, *joint RCs*, *probabilistic quality*, *matching* and *data quality* to a large extent. Despite the very low shares of articles meeting the requirements *precise RCA*, *joint RCs*, and *matching*, still some approaches fully meet them (+). *Lifecycle prediction* is the only requirement with a ~ as the best assessment.

The distribution of the + assessment count for each paper is presented as a boxplot in Figure 7. It reveals that the majority of papers focuses on one to three of the assessment requirements, i.e., the design of event prediction methods is typically aimed towards a few, particular characteristics. Although a method design with a clear focus is beneficial during method design, the result leaves the problem of system design and combining various characteristics of methods for meeting more requirements to the practitioner that wants to use the method. Moreover, the distribution shows that the mostly positive best assessment in the requirements-view is due to various papers, i.e., there is no single paper that is close to the best possible assessment.

In the following, we aim to support a more holistic development of event prediction methods meeting more requirements by illustrating the method’s feature(s)<sup>2</sup> for each requirement with an example from existing work on each requirement. The examples are selected based on how straightforward, concise and representative their features for meeting the requirements are and how well they reflect the diversity of prediction goals and application domains. The illustration shows how key design aspects are responsible for features of the method that are in turn responsible for meeting the requirement. Please note that we overload the meaning of features as the encoded input data (cf. Definition 2.2) and

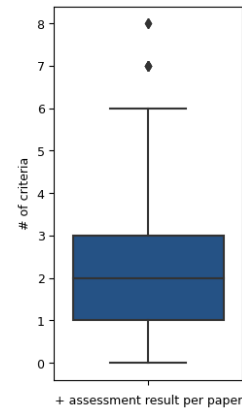


Fig. 7. Distribution of + assessment result per paper.

<sup>2</sup>Some features are illustrated with definitions or equations from the original work that is cited at the beginning of the example, i.e., if the illustration states a definition or equation, then it is taken from the original work and its key part for illustration is extracted and adapted for an as consistent as possible presentation using our survey’s notation.

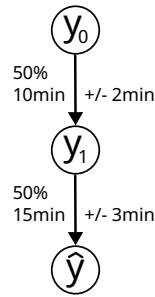


Fig. 8. An event graph for events  $y_0$ ,  $y_1$  and  $\hat{y}$  that follow each other in system logs of a supercomputing cluster node [98]. The probability of following, the average time delay and the standard deviation of time delay is added for each relation between events.

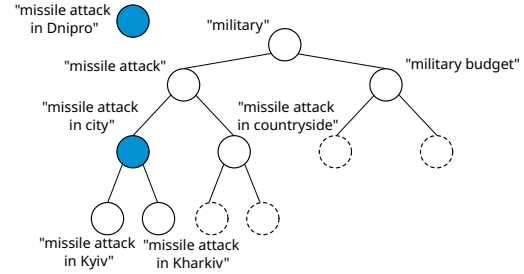


Fig. 9. A hierarchy of most general to most specific observed events [245]. On receiving a current event of a missile attack in Dnipro, the most specific event in the hierarchy that semantically matches the current event is used for prediction.

the properties of a method that are responsible for meeting a requirement. The presentation of examples is structured along the three groups of requirements: functional requirements, attributes and constraints. It is accompanied by an analysis of how the requirement may be related to other requirements and a discussion of the respective approaches.

#### 4.2 Results for Functional Requirements

There are seven requirements that are functional for the event prediction method (cf. Figure 6).

**Probabilistic prediction.** [98] predicts the time and event failure type of high-performance computing cluster nodes by extracting event graphs from system logs and computing the confidence of the correlation between events, the average time delay between events and its standard deviation. Figure 8 shows a small, exemplary event graph with three events  $y_0$ ,  $y_1$  and  $\hat{y}$ . For a *probabilistic prediction*, the event prediction method in [98] predicts the probability of event  $\hat{y}$  that follows event  $y_0$  and  $y_1$  with Bayes law to be  $50\% * 50\% = 25\%$  and the time of event  $\hat{y}$  to be 25 minutes +/-  $(2 * 4 \text{ minute} + 2 * 3 \text{ minute}) = 14 \text{ minutes}$  with 95% probability after  $y_0$  has occurred.

Leveraging the event graph, Bayes law and the assumption of normally distributed time delays leads to a straightforward, probabilistic conceptualization of event prediction. For the same reasons, the prediction can be traced back to one or a set of root causes in the event graph such that the method meets the requirements of *RCA support*. Furthermore, the simplicity of the method helps domain experts to understand its output, yet may fail to perform well when facing complex dependencies between events in multiple dimensions, e.g., a patient's clinical history and future treatment outcome. This indicates a trade-off between simplicity of the probabilistic conceptualization and its applicability to complex monitored systems, their application domains and corresponding prediction goals.

Methods with a probabilistic prediction all conceptualize the method with conditional probabilities, Bayesian or Markovian settings, but differ in their degree of statistical rigor. Statistical rigor captures the number of assumptions on the distribution of the input data, how restraining the distribution assumptions are and how formal the derivation of the event prediction method is given those assumptions. For example, [12] exhibits a high degree of statistical rigor, as it features a high number of assumptions on the distribution of the data that restrain it to a large degree (e.g., the input data is the result of a Gaussian autoregressive process of the second order) and the whole event prediction method is formally derived from those assumptions. [98] exhibit a medium degree of statistical rigor, as they assume normally distributed time delays and apply Bayes law, but do not formally derive the overall event prediction method. [236]

exhibits a low degree of statistical rigor, as they explicitly aim at designing the event prediction method to have no assumptions and do not formally derive any part of the method.

**Suffix prediction.** [142] predict the full sequence of future events for an ongoing manufacturing cycle, i.e., all the events that are likely to occur until the production is complete, with the goal of predicting key performance indicators such as product quality (the prediction goal). To the end of *suffix prediction*, [142] assumes that the future progress of the ongoing cycle  $s$  with length  $n$  can be similar to the progress of  $k$  historical cycles  $l_1, \dots, l_k$  after  $n$  events of each historical cycle  $l_i$  has been observed. Based on this assumption, [142] apply k-Nearest-Neighbor clustering with Euclidean distance on the  $n$ -length prefix of historical cycles and the ongoing cycle resulting in  $k$  suffixes for the ongoing trace. Then, [142] apply a support vector machine classifier on the  $k$  suffixes to predict the key performance indicator. At last, the  $k$  predicted key performance indicators are aggregated into a single indicator.

Although this indirect basic technique can show the operations manager in the production facility additional information on how the current cycle will progress, this method likely fails to accurately predict for a production line with hundreds of thousands of different cycle variations. In contrast to [142], many other methods such as [179] or [73] predict the suffix by repeatedly applying a next event prediction method. Although this can be done for any event prediction method, the major challenge of looping through events that are likely has to be tackled by methods repeatedly applying the event prediction method for *suffix prediction*.

**Lifecycle prediction.** [179] predict the next events of business processes by learning a recurrent neural network that takes multiple event attributes into account. Accordingly, the event prediction problem is stated for an ongoing process instance  $y = \langle (e_1, a_1^1, \dots, a_1^m), \dots, (e_n, a_n^1, \dots, a_n^m) \rangle$  with event activities  $e_1, \dots, e_n$  and event attributes  $a_1^1, \dots, a_n^m$  as predicting the next event  $(e_{n+1}, a_{n+1}^1, \dots, a_{n+1}^m)$  of the ongoing process instance  $s$ . Hence, [179] can predict lifecycles of events, since the lifecycle of an event is an attribute of the event. Nevertheless, lifecycle carries special semantics for event prediction, as, for example, an *abort* lifecycle event for the delivery activity of a logistics company signifies that the *end* will not occur at all and, therefore, the delivery will fail [256]. Also, the lifecycle determines the duration of the activity recorded through multiple events, e.g., an activity paused once for 5 minutes has a duration of 5 minutes lower than the time between its *start* and *end*. For prediction goals that are evaluated on the duration, e.g., efficiency of the resource executing the activity, the special semantics of the lifecycle has to be taken into account. Consequently, the prediction method should not stop at predicting the lifecycle as an attribute, but incorporate its semantics.

Similarly, the other proposals partly meeting lifecycle prediction conceptualize event prediction with multiple, unspecified event attributes and do not incorporate lifecycle semantics. If the data source for the evaluation does not provide lifecycle information, this should not be taken as the reason for excluding it in a method design, as this would not allow the method to be generally applied to data sources with event lifecycle information. Due to the importance of lifecycle information for prediction, if the data source does not provide lifecycle information, researchers should recommend the additional recording of this information to those data sources.

**Uncertainty quantification.** [320] predict the time-to-event for treatment outcomes of patients in clinics by extracting features with recurrent neural networks and learning a Parametric Predictive Gaussian Process (PPGP) Regressor in a survival analysis setting. The combination of the PPGP Regressor and survival analysis results in the survival function:

$$S(y_i | \mathbf{h}_i) = 1 - F \left( \frac{y_i - \mu_{\mathbf{f}}(\mathbf{h}_i)}{\sigma_{\mathbf{f}}(\mathbf{h}_i) + \sigma_{\text{obs}}} \right)$$

where  $S$  is a survival function,  $y_i$  the logarithm of the time-to-event for patient  $i$ ,  $\mathbf{h}_i$  the latent representation extracted by the recurrent neural network,  $F$  the normal cumulative distribution function,  $\mathbf{f} \in \mathbb{R}^n$  a vector of Gaussian process

function values,  $\sigma_f^2(\mathbf{h}_i)$  the input-dependent predictive variance and  $\sigma_{\text{obs}}^2$  the input-independent observational noise for the logarithmic time-to-event  $y_i$ . The corresponding objective that [320] propose puts more weight on the input-dependent predictive variance  $\sigma_f^2(\mathbf{h}_i)$  than earlier proposals that combine Gaussian Processes and survival analysis. Hence, the formulation of survival analysis with PPGP Regressors and the corresponding uncertainty-aware objective not only quantifies predictive variance, i.e., epistemic uncertainty<sup>3</sup>, and observational noise, i.e., aleatoric uncertainty, but also strengthens the positive relationship between predictive variance and prediction error. Consequently, the less confident the uncertainty-aware event prediction method is, the less accurate its predictions are.

In addition to quantifying the uncertainty of the prediction, appropriately integrating it to the objective of the event prediction method yields a desirable relationship between uncertainty and prediction quality. The information that more confident predictions correspond to more accurate predictions is valuable to practitioners [131]. This information likely results in more trust towards the method. By further differentiating between epistemic and aleatoric uncertainty, the quantified uncertainty also shows whether a currently low confidence could be improved with more data.

Existing proposals mostly quantify uncertainty by means of a probabilistic conceptualization that allows to inherently estimate uncertainty as some form of variance in the method and, thus, focuses on the method, e.g. [320]. Depending on the degree of statistical rigor (cf. *probabilistic prediction*), these estimates are more or less justified by statistical theory. Another option is to extrinsically quantify uncertainty by approximating it using, e.g. characteristics of the input data. [57] propose an uncertainty metric that also considers how many activities of a business process instance have been observed and what time has elapsed since the last activity has been observed. The rationale for their extrinsic uncertainty quantification is that more and more recent data on the monitored system relate to a more certain prediction.

**RCA support, output visualization and RCA visualization.** While [98] apply Bayes law on descriptive statistics of the data without a statistical analysis formulation, [169] combine decision rules with Bayesian analysis into Bayesian Rule Lists (BRL):

```

if antecedent1 then  $y \sim \text{Multinomial}(\Theta_1)$ ,  $\Theta_1 \sim \text{Dirichlet}(\Phi + \mathbf{N}_1)$ 
:
else if antecedentm then  $y \sim \text{Multinomial}(\Theta_m)$ ,  $\Theta_m \sim \text{Dirichlet}(\Phi + \mathbf{N}_m)$ 
else  $y \sim \text{Multinomial}(\Theta_0)$ ,  $\Theta_0 \sim \text{Dirichlet}(\Phi + \mathbf{N}_0)$ 

```

where antecedents are well-formed formulas with  $c_i$ ,  $i \in \{1, \dots, m\}$  atomic propositions over the feature vector that do not overlap,  $y$  is the prediction goal value,  $\Theta_i$  are parameters of the respective distribution,  $\Theta_0$  a default rule parameter for observations that do not satisfy any of the antecedents,  $\Phi$  a prior of the prediction goal values and  $\mathbf{N}_i$  the respective observation counts satisfying antecedent  $i$ . By limiting the parameters  $m$  and  $c_i$  to small integers as well as deriving a mean point estimate and credible interval for the prediction goal value  $y$ , BRLs represent a simple, interpretable prediction model. Potential root causes are represented as antecedents of the BRL (*RCA support*). The prediction output and the RCA are visualized by showing the BRL and traversing its rules from top to bottom given an observation. [169] compares this BRL to a medical score that is frequently used to predict the stroke risk by physicians, identifies antecedents with known root causes or risk fusers and concludes that the BRL is as interpretable as the medical score, while improving the accuracy.

Methods meeting all three requirements *RCA support*, *output visualization*, and *RCA visualization* are designed in their functionality to support the user in understanding the method and its prediction. A simple prediction model as

<sup>3</sup>Model uncertainty does not exist due to an implicit assumption that the true predictor is part of the chosen model hypothesis space [131].

in [98, 169] underlying the event prediction method support the user in tracing the prediction back to its root causes. This meaning of simplicity is captured by the "white box" assessment of *explainability*, but it is not the only sufficient condition for meeting *RCA support*, e.g., a hybrid basic approach can still support RCA in spite of a "black box" model due to the generative part of the model [181]. Due to a lack of a widely accepted definition on *explainability* (cf. Section 3.2), it is not clear whether further relationships with root causes exist. Considering visualization, some methods present a straightforward way of visualizing the output or RCA as in [169], while others need specifically designed visualization techniques, e.g. neural networks.

### 4.3 Results for Attributes

There are four requirements that are attributes of the event prediction method (cf. Figure 6).

**Precise RCA, joint RCs and Explainability.** [245] predict future events that can be caused by a current news event. To that end, they choose an aggregate semantic event representation for an event  $e = (P, O_1, \dots, O_4, t)$ , where  $P$  is a state or temporal action exhibited by the event's objects,  $O_1$  a set of users,  $O_2$  a set of objects on which the action  $P$  was performed,  $O_3$  a set of instruments utilized by the action,  $O_4$  a set of locations and  $T$  a timestamp. Given a set of historic events  $\{\langle e_1, g(e_1) \rangle, \dots, \langle e_n, g(e_n) \rangle\}$  for an unknown causality function  $g$  mapping historic cause event to its effects, [245] aim to approximate  $g$  by learning a causality graph of events with the help of existing ontologies on actions such as VerbNet [150] and on objects such as the LinkedData ontology [22].

Using existing ontologies, the method can evaluate the precision of identified root causes (*precise RCA*). For identifying *joint RCs*, the method proposes a generalization event by generalizing over its actions and objects. The result is a hierarchy that detects that similar root causes are in fact joint root causes. Figure 9, for example, depicts a generalization hierarchy of observed events that identifies similar root causes as joint, more general root causes. Since the prediction is computed solely on the causality graph through matching the current event with an existing event in the causality graph (denoted in green in Figure 9) and applying a prediction rule associated with the event, a user can traverse the predicted event back to its matched, historic event. Thus, the method employs a "white box" model (*explainability*).

Despite their existence, ontologies are seldomly used in event prediction. With the help of ontologies, the method can establish a concept of causality and meaning. Moreover, event prediction methods rarely exploit that events are related by a semantic generalization relationship. However, other semantic relationships between events are exploited by recent methods, e.g. the disease co-occurrence [188] or compatible related treatment outcomes relationships [103].

**Probabilistic Quality.** [309] predict crime frequencies in the spatio-temporal domain by combining variational autoencoders and sequence generative neural networks. Hence, a predicted crime frequency is  $\hat{y}_{i,j,k}^{(t)}$  for the geographical grid indexes  $i, j$ , crime type  $k$  in the time slot  $t$ . To evaluate these probabilistic predictions, [309] first normalize the true and predicted crime frequencies:

$$p_{i,j,k}^{(t)} = \frac{y_{i,j,k}^{(t)}}{\sum_{i=1}^M \sum_{j=1}^N y_{i,j,k}^{(t)}}, \quad \hat{p}_{i,j,k}^{(t)} = \frac{\hat{y}_{i,j,k}^{(t)}}{\sum_{i=1}^M \sum_{j=1}^N \hat{y}_{i,j,k}^{(t)}}$$

and then compute the Jensen-Shannon divergence with laplacian smoothing:

$$D_{JS}(P, \hat{P}) = \frac{1}{2} D_{KL}(P \parallel Q) + \frac{1}{2} D_{KL}(\hat{P} \parallel Q)$$

where  $Q = \frac{1}{2}(P + \hat{P})$  is the average of true and predicted probability. Next to the Jensen-Shannon divergence, there exist at least 200 further scores for evaluating probability estimates [41]. Thus, analysing its properties with respect to

event prediction and proposing a standardized set of scores goes beyond the scope of this survey. For this survey, the important point is that the evaluation is done using a metric that scores probability estimates, as is the case for [309].

For evaluating probabilistic predictions, the predictions must come with probability estimates. Thus, meeting the requirement of *probabilistic quality* implies meeting *probabilistic prediction*. Yet, these two requirements are not equivalent, because many articles do not report results of scoring probability estimates. While roughly 25% of the 260 articles propose methods meeting *probabilistic prediction*, less than 5% evaluate the predicted probabilities with an appropriate score (cf. Figure 6).

#### 4.4 Results for Constraints

There are four requirements that are constraints for the event prediction method (cf. Figure 1).

**Online conceptualization.** [248] predict civil unrest events across ten Latin American countries through a continuous, automated system that runs 24/7 and constantly processes events coming from various open sources such as social media. Consequently, the event prediction method consisting of an ensemble of methods is proposed to continuously learn from daily data, e.g., a logistic regression models is learned using daily tweet counts while a rule-based method detects keywords contained in social media posts. The system has a throughput of 200-2000 messages/sec, predicts roughly 40 events/day and can ingest up to 15 GB of messages on a given day. Therefore, it features an *online conceptualization* and can monitor the ten Latin American countries for social unrest events.

Given its daily cycle and the ensemble of prediction methods, a daily, full retraining of the prediction models is feasible. However, a full retraining may become infeasible for event prediction methods in light of updating cycles not longer than a few minutes. For this reason, [258] investigate strategies for updating the prediction model that go beyond full retraining; Do nothing, i.e., no update; retraining with no hyperparameter optimization, i.e., a lightweight retraining; full retraining, i.e., train on all available events; and incremental update, i.e., applying incremental learning algorithms. The results show that full retrain and incremental update are the best strategies for predicting business process events.

Still, not all prediction models allow for incremental updates through incremental learning algorithms. Thus, [210] investigates six strategies to handle updates to the prediction model by means of data selection strategies: Baseline strategy, i.e., no update to the training data; cumulative, i.e., update training data on every new event; non-cumulative, i.e., update training data on every new event by keeping the  $k$  most recent events; ensemble, i.e., non-cumulative and keeping all models for ensemble prediction; sampling, i.e., update training data by sampling all available events; and drift, i.e., non-cumulative with drift detection such that only training data after a drift is selected. The results show that the ensemble strategy performs best for predicting business process events.

Despite these pointers to suitable strategies for updating prediction models, a best strategy for all use cases cannot be determined due to a case-dependent trade-off between the resources required to improve the quality of prediction through updating and its benefits.

**Multiple sources.** [4] predict the time-to-event for future failures of engineering system devices. Sensor data  $X_p \in \mathbb{R}^{d \times c_p}$  from  $d$  sensors for each device  $p$  observed until time  $c_p$  and event data  $y_p \in \{0, 1\}$  is used to predict the time-to-event  $T$  for the future failure  $y_p = 1$  of device  $p$ . For prediction, [4] propose a multi-task learning framework based on neural networks with a task  $p$  for each device. Hence,  $p$  data sources are used in the event prediction method (*multiple sources*) and the integration of these data sources is implicitly learned by the neural network within the multi-task learning framework.

We can distinguish two approaches to data integration. The *implicit* approach to data integration combines the data integration with the learning of a prediction model in the event prediction method such that the data integration is tightly coupled with the event prediction method, e.g., the data integration of the  $p$  devices is implicitly learned by the neural network in [4]. The *explicit* approach to data integration separates data integration from the event prediction method such that both of them become interchangeable. To that end, the explicit approach implements necessary functionality for data integration (cf. input component in Section 3.1) before the event prediction method is applied to the data. [248] not only illustrates the *online conceptualization*, but also the explicit approach to data integration. Many different sources such as NASA satellite meteorological data, Bloomberg financial news and Twitter’s public API are used. Each source is ingested by a specialized routine to convert the data input to JSON and add identifiers. Before the event prediction method is applied, the JSON input is enriched through, e.g. geocoding, data normalization and entity extraction. Consequently, the event prediction method does not need to learn how it can integrate the data. In comparison, the implicit approach intertwines the input component of a PCM system (cf. Section 3.1) with the prediction component, whereas the explicit approach keeps them separated. The former has the advantage, that it does not need to design data integration functionalities, but may experience worse performance for heterogeneous data sources and consuming more resources for training. The latter has the advantage, that it has direct control and knowledge on the specific data integration routines, but comes at the cost of more design and maintenance effort.

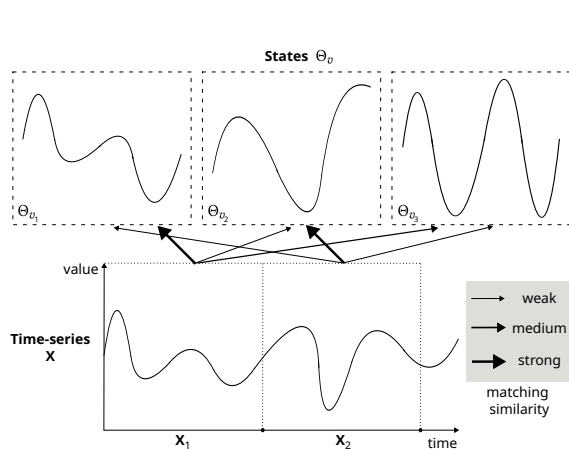


Fig. 10. Time-series  $X$  is segmented into  $X_1$  and  $X_2$ . Each time-series segment is recognized with states  $\Theta_{v_1}$ ,  $\Theta_{v_2}$  and  $\Theta_{v_3}$  [123]. Recognized states and the similarities are used in the event prediction method proposed in [123] instead of the original time-series.

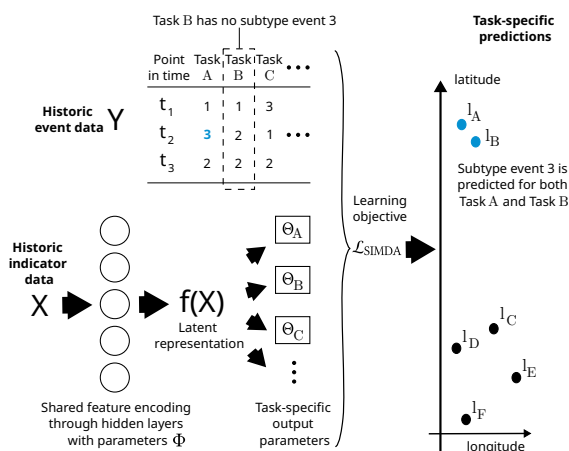


Fig. 11. Spatial Incomplete Multi-Task Deep Learning (SIMDA) Framework [106]. Despite no recorded subtype event 3 in the historic event data  $Y$  for task  $B$ , SIMDA allows to predict subtype event 3 in task  $B$  that corresponds to location  $l_B$ , since location  $l_A$  is geographically close to location  $l_A$  and location  $l_A$  has recorded subtype event 3 in the historic event data of task  $A$ .

**Matching.** [123] predict a future event in a time-series sequence, e.g. a future anomalous watt-hour meter clock event of the State Grid of China in weekly sensor readings. To that end, a future event  $y_{t+1} \in \mathbb{Z}$  is predicted based on historic events  $y_t$  and  $T$  segments  $X_t \in \mathbb{R}^{\tau \times d}$  of the time-series sequence  $X \in \mathbb{R}^{N \times d}$  with length  $\tau$ :

$$\langle X_{1:T}, y_{1:T} \rangle = \{(X_1, y_1), \dots, (X_T, y_T)\}$$

Before feeding the historic data  $\langle X_{1:T}, y_{1:T} \rangle$  into a neural network, the proposed event prediction method recognizes representative time-series patterns called *states* and denoted as  $\Theta_v \in \mathbb{R}^{r \times d}$  for segments of the data (cf. Figure 10). Hence, redundant or very similar time-series patterns from multiples sources are recognized by the same state (*matching*).

[123] is the only method proposed that features an explicit matching mechanism for data coming from multiple sources. Other proposals for matching [2, 121, 143, 166, 177, 248, 307, 334] are not explicitly including the mechanism in their method.

**Data quality.** [106] predict the location of event subtypes, e.g. instead of predicting the air pollution of cities as future events, the respective air pollutant subtype (Carbon monoxide, nitrogen dioxide, fine dust pollution etc.) is predicted. While the predicted event subtype allows for more fine-grained actions, e.g. allocating appropriate resources per air pollutant subtype, incomplete, historical data for event subtypes in many locations becomes a major data quality challenge for predicting event subtypes for all locations.

Figure 11 depicts the data quality challenge for predicting event subtypes and the proposed method, Spatial Multi-task Deep leArning (SIMDA), to overcome the challenge. Each depicted prediction task  $A, B$  and  $C$  corresponds to a geographical location  $L = \{l_A, l_B, l_C, \dots\}$ . Historic event data  $Y_{l,t} \in S = \{1, 2, 3, \dots\}$  and  $d$ -dimensional, historic indicator data  $X_{l,t} \in \mathbb{R}^{1 \times d}$  for date  $t$  is available for each location  $l \in L$ . However, for the location  $l_B$  corresponding to task  $B$ , no event subtype  $s = 3$  was recorded, e.g. a city in the past has not recorded fine dust pollution (cf. Figure 11). Hence, the specific model for task  $B$  will not predict event subtype  $\hat{y}_{l_B, \tau} = 3$  at location  $l_B$  for  $\tau = t + p, p > 0$ , although the geographically close location  $l_A$  in task  $A$  shows that event subtype  $s = 3$  has occurred. Since geographical heterogeneity across all locations does not allow for a single task approach, learning from other locations such as  $l_A$  for location  $l_B$  has to be achieved differently. On account of the first law of geography stating that geographically closer locations will be more similar to each other than farther locations [61], [106] state that two close locations  $i, j \in L$  have similar conditional probabilities for an event subtype  $s$ :

$$P(Y_{i,t} = s | X_{i,t}) \approx P(Y_{j,t} = s | X_{j,t})$$

This relationship is exploited to formulate a SIMDA objective that enforces event subtype patterns to be similar for geographically close tasks:

$$\mathcal{L}_{\text{SIMDA}} = \mathcal{L}_d(\Phi, \Theta) + \frac{\beta}{2} \sum_l^L \sum_{s,r}^{S^2} \|f(X_l)(\Theta_{l,s} - \Theta_{l,r})^T - \frac{1}{N_l} \sum_c^L \text{adj}(l, c) f(X_c)(\Theta_{c,s} - \Theta_{c,r})^T\|_2^2$$

where  $\mathcal{L}_d(\Phi, \Theta)$  is the general multi-task deep learning objective with  $d$  dimensions,  $\Phi$  the weight parameters of the shared hidden layer,  $\Theta$  the weights of the task-specific output layer,  $\beta$  a hyperparameter,  $f(\cdot)$  the computation of the shared hidden layers,  $\Theta_{l,s}$  the weights of task  $l$  for predicting event subtype  $s$ ,  $N_l = \sum_c^L \text{adj}(l, c)$  the normalization term for location  $l$ , and  $\text{adj}(\cdot, \cdot)$  some physical distance function. Hence, [106] propose to add a regularization term to the general multi-task deep learning objective  $\mathcal{L}_d(\Phi, \Theta)$  that enforces geographically close locations  $i, j$  to have similar event subtype patterns  $X_i(\Theta_{i,s} - \Theta_{i,r})^T \approx X_j(\Theta_{j,s} - \Theta_{j,r})^T$ . For the example as depicted in Figure 11, this allows SIMDA to predict fine dust pollution  $s = 3$  for location  $l_B$  through task  $B$ , as the geographically close location  $l_A$  exhibits a lot of fine dust pollution.

Analogous to [106], methods that actively utilize data quality issues so far have exploited the following semantic similarity relationships between components of the monitored system (e.g., close locations, comparable patients etc.) to improve the prediction while keeping the component-level granularity in the prediction model (e.g. a task for each



location): semantic similarity based on proximity [105, 106, 339, 343], hierarchy of geographical regions [339, 343], logging granularity of cyber system host sources [319], patient disease histories [190, 236, 271], hierarchy of medical codes [190], hierarchy of mobile health study participants [68], individual units of engineering systems [67], feature missingness patterns [304] and structural tweet content [338]. In contrast to preprocessing techniques such as imputation for missingness, these proposals make the method aware of the data quality and base the awareness on the semantic similarity relationships.

Since data quality issues should lead to more uncertain predictions, this requirement is related to *uncertainty quantification*. For example, [271] propose a method that abstains from predicting in case of high uncertainty of the prediction due to missingness of data. Hence, the missingness of data has a direct impact on the prediction in their proposed method. All in all, the link between data quality and uncertainty can be used to guide the prediction (e.g., no prediction in case of high uncertainty due to data quality issues) or the data quality issue can be alleviated by exploiting semantic relationships leading to more certain predictions.

#### 4.5 Summary

Overall, event prediction approaches broadly support the requirements for building a PCM system concerning the input and output of the prediction method as well as its evaluation and understanding (cf. Figure 5). Looking more closely at the requirements support, the following conclusions can be drawn:

- (1) So far, existing approaches have been fragmented with respect to research area (e.g., event prediction and PPM) as well as with respect to the application area (cf. Figure 1).
- (2) The concept of the PCM system enables the integration of disparate event prediction approaches and applications.
- (3) Event prediction approaches support requirements through a combination of the prediction goal and event prediction method.
- (4) None of the approaches supports event life cycle prediction.
- (5) None of the approaches supports all requirements; most of the approaches support 1 – 3 requirements (cf. Figure 7) and at most 8 out of 16 requirements.

Conclusions 1 – 5 underpin the importance of the PCM system concept for holistic event prediction and point to open challenges and research directions, which will be discussed in Section 5.

## 5 OPEN CHALLENGES AND RESEARCH DIRECTIONS

Section 4 shows the considerable progress of event prediction methods regarding requirements of a PCM system. In spite of these advancements, the landscape of existing work is fragmented and faces the following open challenges with potential future research directions.

### 5.1 Holistic Event Prediction Method

*Challenge:* No event prediction method simultaneously meets all requirements. Hence, selecting a single event prediction method leaves us with the necessity of trading off various sets of features with one another. The challenge is to develop event prediction methods that support more or all requirements by the method's respective features.

*Research direction:* For a holistic approach to method development, future research can either combine existing features from various methods into a single method, develop new solutions to the respective existing problems for a single method or combine existing methods and, thus, their features in a generic framework. Independent of the chosen option,

a conceptual model of event prediction that identifies the systematic relationships between application domains, domain knowledge, prediction goals, input data, feature engineering, prediction methods and the action space is desirable. This conceptual model particularly represents our understanding on how an application domain, the existing domain knowledge and given prediction goals require a certain set of data sources, a certain set of features engineered on the data, appropriate prediction methods (cf. Figure 3) and the possible action space. Such a model extends the structural system perspective on event prediction and its surrounding components embodied by the PCM system (cf. Figure 4) and the methodological perspective on event prediction in [337] with a behavioral and operational system perspective.

## 5.2 Semantics, Knowledge and Explainability

*Challenge:* Event prediction aims at anticipating the future of the monitored system by means of algorithms. Monitored systems come from a diverse range of application domains and, thus, have a diverse range of semantics. At some point in the PCM system, the algorithms used to develop event prediction should take the semantics, e.g. of the event lifecycle, into account. The same applies to existing domain knowledge. Aside from potentially positive effects on accuracy, the way semantics and knowledge are taken into account determines the explainability of the method and its generality: If the method is very abstract and generic, it is typically not explainable to the domain expert, whereas a very concrete and domain-specific method is typically explainable, but does not generalize. The challenge is to shift the trade-off between explainability and generality and develop new ways to integrate semantics and knowledge in the method for an improved explainability while maintaining generality.

*Research direction:* Starting with a clear definition of explainability for event prediction, e.g. is it rather a counterfactual or the more strict causal definition [254], and the necessary content type, communication means and target group of the explanation [34], future research can evaluate the correctness of explanations. Explanation-guided learning [104] demonstrates promising approaches to explainability and correctness evaluation, yet it pre-determines the way of how the method integrates semantics and domain knowledge by stating them as explanation supervision and regularization in the learning objective. Other or additional ways of integration could be parametrization, feature engineering or a certain hypothesis space for explainable models. It is highly beneficial to explore various combinations of integration and test for generality of the method such that we can improve our understanding on how semantics and domain knowledge can be used for explainability while maintaining generality.

## 5.3 Flexibility and Robustness

*Challenge:* Considering the heterogeneity of application domains, data sources and prediction goals, event prediction methods must find a balance between flexibility and robustness. On the one hand, they have to be flexible with respect to (i) domain-specific properties and requirements, e.g. a very dynamic, clinical environment in the healthcare domain with a particularly strong requirement on uncertainty vs. a relatively stable environment of a standardized goods manufacturing facility with a relatively weak requirement on uncertainty, (ii) the data properties and quality issues of multiple data sources, e.g. different granularities of the data such as city- vs. country-level and missingness of data, and (iii) prediction goal properties, e.g. a single, stable prediction goal such as "treatment outcome of a patient is positive" vs. multiple, dynamic prediction goals such as the set of regulatory constraints a bank's business processes have to adhere to. On the other hand, event prediction methods have to be robust with respect to adversarial attacks and/or noise of the input data [57, 153, 337], since event prediction in a PCM system has a direct real-world consequence through the action that the user chooses.

*Research direction:* One way to deal with the flexibility challenge is to limit the scope of the method to a certain problem

space for which flexibility is not required anymore. So far, this approach dominates research on event prediction. Although this approach can solve the challenge by removing it from the equation to a large extent, it results in a very fragmented set of event prediction methods and transfers the challenge to the user. Future methods should abstain from definitions that simplify the problem in this way. For (i) domain-specific properties and requirements and (ii) data properties and quality issues existing work supports the required flexibility (cf. assessment for *uncertainty quantification*, *online conceptualization* and *multiple sources* in Section 4). Thus, the combination of these is a promising starting point for future work on meeting flexibility for (i) and (ii) together. While [193] called for flexibility regarding various sets of prediction goals, flexibility for (iii) prediction goal properties is usually still abstracted from or left unmentioned. Considering the multi-objective nature of event prediction [337] and the abundance of regulatory requirements on monitored systems such as banks [99], clinics [29, 47] or manufacturing facilities [119], event prediction methods that are flexible with respect to prediction goal heterogeneity are highly beneficial. Since a more flexible event prediction leaves more room for adversarial attacks, research on flexible event prediction should study how these mechanisms interact with adversarial attack models. With respect to noise, methods are desirable that can reliably distinguish variances in the input data that require flexibility from noise that should be captured as aleatoric uncertainty.

#### 5.4 User Interface

*Challenge:* As event prediction research focuses on the method and its rather technical evaluation, it often does not provide a simple and clear user interface of relevant output. Moreover, probabilistic predictions with uncertainty estimates, multiple prediction goals, root causes and potential actions aggravate the challenge to present the output in a simple and clear way to the user. Although the user interface corresponds to the output component of a PCM system that is separated from the prediction component, the method design should already have the subsequent user interface in mind to ease its design and implementation, e.g., by suggesting certain visualization techniques. Lastly, some methods have started proposing the integration of prediction with decision theory to include actions and their impact [28, 271, 305], but these are goal-dependent and domain-specific and, thus, cannot be easily transferred to other goals and application domains. Furthermore, [28, 271, 305] lack to provide how the additional information should be integrated to present the output to the user.

*Research direction:* By adopting principles of user-centred design [113], future research can explore means of presenting and visualizing the complex output of event prediction and the PCM system to the user. This challenge should not be taken lightly, as it significantly determines the adoption of the method by users and may affect the user's trust in the system. Research that incorporates actions through a decision theoretic framework further supports the user in deciding what to do based on the anticipated future. By allowing for parametrizing the action space, future work on event prediction can increase the flexibility of the method (cf. Section 5.3) and enables the user in setting the action space. The development of a comprehensive user interface for a PCM system that streamlines the relevant information to the user and takes subsequent actions into account is key for building trust and increasing the accessibility for users.

#### 5.5 Evaluation

*Challenge:* While [337] presents the evaluation design with respect to matching predicted with real events and metrics of effectiveness for event prediction, it neither considers probabilistic metrics of effectiveness nor the performance in terms of training and prediction time. However, both are crucial for a standardized and valuable evaluation of event prediction.

*Research direction:* [41] reviewed 201 probabilistic metrics of effectiveness that are candidates for event prediction.

Both an analysis as well as standard proposal for event prediction is needed. This proposal can be directly applied in a comprehensive, empirical comparison of existing methods within comparable classes of our taxonomy for event prediction methods (cf. Figure 3). With respect to the time performance, the comparison can utilize appropriate benchmark case studies to additionally compare the training and prediction times.

## 6 CONCLUSION

This work presents a comprehensive survey on event prediction methods that span multiple, disparate research areas corresponding to the nine application domains for event prediction. To bring the work on event prediction methods from these disparate research areas together, we take a systems perspective on event prediction and identify four research questions. The first question aims at a taxonomy of event prediction methods for integrating the diverse selection of event prediction methods. To facilitate the design of the taxonomy and to reflect the systems perspective, we formulate the event prediction method with an event prediction goal. For the design of the taxonomy, we additionally followed the aim of integrating existing taxonomies for event prediction. After the comprehensive classification of event prediction methods with the taxonomy, the second question aims at designing an assessment scheme for event prediction that takes the multi-faceted nature of event prediction into account. This question leads us to leverage the systems perspective further to propose the PCM system as a decision support system for event prediction. In proposing the PCM system, we generalize ideas from the research areas PPM and CM such that the respective contributions can be mutually exchanged with other research areas for event prediction and structurally identify event prediction as a component of a decision support system. Based on the PCM system and existing work on event prediction methods, we elicit requirements for event prediction methods. The third question aims at the qualitative assessment result of existing event prediction methods with respect to the requirements. The assessment shows that not only the research areas for event prediction are fragmented, but also the event prediction methods in terms of meeting the requirements. Furthermore, the assessment illustrates how a specific combination of a prediction goal and an event prediction method are used by existing work to achieve specific requirements. The fourth question aims to derive open challenges and future research directions for event prediction given the assessment of existing work. To that end, we find five challenges and corresponding research directions for event prediction that, inter alia, cover a holistic approach to event prediction methods as a direct consequence of the PCM system interpretation (cf. Section 5.1).

## ACKNOWLEDGMENTS

This work has been partly funded by the Austrian Research Promotion Agency (FFG) via the “Austrian Competence Center for Digital Production” (CDP) under the contract number 854187 and by the Deutsche Forschungsgemeinschaft (DFG) under the contract number GRK 2201. This work has been supported by the Pilot Factory Industry 4.0, Seestadtstrasse 27, Vienna, Austria.

## REFERENCES

- [1] Ahmad Aburomman, Manuel Lama, and Alberto Bugarin. 2019. A Vector-Based Classification Approach for Remaining Time Prediction in Business Processes. *IEEE Access* 7 (2019), 128198–128212.
- [2] Saurav Acharya, Byung Suk Lee, and Paul Hines. 2017. Causal Prediction of Top-k Event Types Over Real-Time Event Streams. *Comput. J.* 60, 11 (2017), 1561–1581.
- [3] Bijaya Adhikari, Xinfeng Xu, Naren Ramakrishnan, and B. Aditya Prakash. 2019. EpiDeep: Exploiting Embeddings for Epidemic Forecasting. In *Proceedings of SIGKDD*. 577–586.
- [4] Karan Aggarwal, Onur Atan, Ahmed K. Farahat, Chi Zhang, Kosta Ristovski, and Chetan Gupta. 2018. Two Birds with One Network: Unifying Failure Event Prediction and Time-to-failure Modeling. In *Proceedings of IEEE Big Data*. 1308–1317.

- [5] Karan Aggarwal, Georgios Theocharous, and Anup B. Rao. 2020. Dynamic Clustering with Discrete Time Event Prediction. In *Proceedings of SIGIR*. ACM, Virtual Event China, 1501–1504.
- [6] Fabrizio Albertetti and Hatem Ghorbel. 2020. *Workload Prediction of Business Processes – An Approach Based on Process Mining and Recurrent Neural Networks*. Technical Report. arXiv: 2002.11675. Retrieved from <http://arxiv.org/abs/2002.11675>.
- [7] Elias Alevizos, Alexander Artikis, and Georgios Paliouras. 2022. Complex event forecasting with prediction suffix trees. *The VLDB Journal* 31, 1 (2022), 157–180.
- [8] Syed Mohd Ali, Noopur Gupta, Gopal Krishna Nayak, and Rakesh Kumar Lenka. 2016. Big data visualization: Tools and challenges. In *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*. 656–660.
- [9] James Allan, Ron Papka, and Victor Lavrenko. 1998. On-line new event detection and tracking. In *Proceedings of SIGIR*. 37–45.
- [10] Renato Marinho Alves, Luciana Barbieri, Kleber Stroeh, Sarajane Marques Peres, and Edmundo Roberto Mauro Madeira. 2022. Context-Aware Completion Time Prediction for Business Process Monitoring. In *Information Systems and Technologies*. 355–365.
- [11] Giuseppe Amodeo, Roi Blanco, and Ulf Brefeld. 2011. Hybrid models for future event prediction. In *Proceedings of CIKM*. 1981.
- [12] M. Antunes, M. A. Amaral Turkman, and K. F. Turkman. 2003. A Bayesian Approach to Event Prediction. *J. Time Ser. Anal.* 24, 6 (2003), 631–646.
- [13] Mulugeta Weldezgina Asres, Grace Cummings, Aleko Khukhunaishvili, and Pavel Parygin et al. 2022. Long Horizon Anomaly Prediction in Multivariate Time Series with Causal Autoencoders. In *PHM Society European Conference*, Vol. 7. PHM Society, 21–31.
- [14] Long Bai, Saiping Guan, Jiafeng Guo, Zixuan Li, Xiaolong Jin, and Xueqi Cheng. 2021. Integrating Deep Event-Level and Script-Level Information for Script Event Prediction. In *EMNLP*. 9869–9878.
- [15] Jie Bao, Pan Liu, and Satish V. Ukkusuri. 2019. A spatiotemporal deep learning approach for citywide short-term crash risk prediction with multi-source data. *Accid. Anal. Prev.* 122 (January 2019), 239–254.
- [16] Alejandro Barredo Arrieta, Natalia Diaz-Rodríguez, Javier Del Ser, Adrien Bénéttot, Siham Tabik, Alberto Barbado, Salvador Garcia, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera. 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* 58 (June 2020), 82–115.
- [17] Yana A. Bekeneva, Ivan I. Kholod, Evgenia S. Novikova, Denis Shevelev, and Andrey V. Shorov. 2019. Event Pattern Constructing Technique for Prediction of Violations in Technological Processes. In *Russian Young Researchers in Electrical and Electronic Engineering*. 182–186.
- [18] Riccardo Benedetti. 2010. Scoring Rules for Forecast Verification. *Mon. Weather Rev.* 138, 1 (2010), 203–211.
- [19] Antonio Bevaqua, Marco Carnuccio, Francesco Folino, Massimo Guarascio, and Luigi Pontieri. 2014. A Data-Driven Prediction Framework for Analyzing and Monitoring Business Process Performances. In *Enterp. Inf. Syst.* 100–117.
- [20] Krishnanjan Bhattacharjee, S. ShivaKarthik, Swati Mehta, Ajai Kumar, and Rushikesh Kothawade et al. 2021. Survey and Gap Analysis on Event Prediction of English Unstructured Texts. In *Machine Learning for Predictive Analysis*. 491–501.
- [21] Marin Biloš, Bertrand Charpentier, and Stephan Günemann. 2019. Uncertainty on asynchronous time event prediction. In *NeurIPS*.
- [22] Christian Bizer, Tom Heath, and Tim Berners-Lee. 2011. Linked data: The story so far. In *Semantic services, interoperability and web applications: emerging concepts*. IGI global, 205–227.
- [23] Robert A Blair and Nicholas Sambanis. 2020. Forecasting civil wars: Theory and structure in an age of “Big Data” and machine learning. *J. Confl. Resolut.* 64, 10 (2020), 1885–1915.
- [24] Benjamin S Blanchard, Wolter J Fabrycky, and Walter J Fabrycky. 1990. *Systems engineering and analysis*. Vol. 4. Prentice hall Englewood Cliffs, NJ.
- [25] Alfredo Bolt and Marcos Sepúlveda. 2014. Process Remaining Time Prediction Using Query Catalogs. In *BPM Workshops*. 54–65.
- [26] Mohammad Al Boni and Matthew S. Gerber. 2016. Area-Specific Crime Prediction Models. In *Proceedings of ICMLA*. 671–676.
- [27] Lydia Bouzar-Benlabiod, Lila Méziani, Stuart H. Rubin, Kahina Belaidi, and Nour Elhouda Haddar. 2019. Variational Encoder-Decoder Recurrent Neural Network (VED-RNN) for Anomaly Prediction in a Host Environment. In *2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI)*. 75–82.
- [28] Zahra Dasht Bozorgi, Irene Teinema, Marlon Dumas, Marcello La Rosa, and Artem Polyvyanyy. 2021. Prescriptive Process Monitoring for Cost-Aware Cycle Time Reduction. In *2021 3rd International Conference on Process Mining (ICPM)*. 96–103.
- [29] Jeffrey Braithwaite, Kate Churruca, and Louise A Ellis. 2017. Can we fix the uber-complexities of healthcare? *J. R. Soc. Med.* 110, 10 (2017), 392–394.
- [30] Patrick T Brandt, John R Freeman, and Philip A Schrodt. 2011. Real time, time series forecasting of inter-and intra-state political conflict. *Confl. Manag. Peace Sci.* 28, 1 (2011), 41–64.
- [31] Andrea Brunello, Angelo Montanari, and Mark Reynolds. 2019. Synthesis of LTL Formulas from Natural Language Texts: State of the Art and Research Directions. In *26th International Symposium on Temporal Representation and Reasoning (TIME 2019)*. 17:1–17:19.
- [32] Jens Brunk, Matthias Stierle, Leon Papke, Kate Revoredo, Martin Matzner, and Jörg Becker. 2021. Cause vs. effect in context-sensitive prediction of business process instances. *Inf. Syst.* 95 (January 2021), 101635.
- [33] Zaharah A. Bukhsh, Aaqib Saeed, and Remco M. Dijkman. 2021. *ProcessTransformer: Predictive Business Process Monitoring with Transformer Network*. Technical Report. arXiv: 2104.00721. Retrieved from <http://arxiv.org/abs/2104.00721>.
- [34] Nadia Burkart and Marco F. Huber. 2021. A Survey on the Explainability of Supervised Machine Learning. *J. Artif. Intell. Res.* 70 (2021), 245–317.
- [35] Tim Bäßler and Marc Eulerich. 2022. Three Lines 4.0 - Predictive Process Monitoring for Internal Audit. *SSRN Electronic Journal* (2022).
- [36] Kristof Böhmer and Stefanie Rinderle-Ma. 2018. Probability Based Heuristic for Predictive Business Process Monitoring. In *Proceedings of OTM*. 78–96.

- [37] Kristof Böhmer and Stefanie Rinderle-Ma. 2020. LoGo: Combining Local and Global Techniques for Predictive Business Process Monitoring. In *Advanced Information Systems Engineering*. 283–298.
- [38] Cristina Cabanillas, Claudio Di Ciccio, Jan Mendling, and Anne Baumgrass. 2014. Predictive Task Monitoring for Business Processes. In *Business Process Management*. 424–432.
- [39] Raffaella Calabrese and Johan A. Elkind. 2016. Estimating Binary Spatial Autoregressive Models for Rare Events. In *Spatial Econometrics: Qualitative and Limited Dependent Variables*. Advances in Econometrics, Vol. 37. 145–166.
- [40] Rui Cao, Weijian Ni, Qingtian Zeng, Faming Lu, Cong Liu, and Hua Duan. 2021. Remaining time prediction for business processes with concurrency based on log representation. *China Communications* 18, 11 (2021), 76–91.
- [41] Arthur Carvalho. 2016. An Overview of Applications of Proper Scoring Rules. *Decis. Anal.* 13, 4 (2016), 223–242.
- [42] Flávia Dias Casagrande, Jim Tørresen, and Evi Zouganeli. 2018. Sensor Event Prediction using Recurrent Neural Network in Smart Homes for Older Adults. In *2018 International Conference on Intelligent Systems (IS)*. 662–668.
- [43] Flávia Dias Casagrande, Jim Tørresen, and Evi Zouganeli. 2019. Prediction of the next sensor event and its time of occurrence in smart homes. In *International Conference on Artificial Neural Networks*. 462–473.
- [44] Michelangelo Ceci, Pasqua Fabiana Lanotte, Fabio Fumarola, Dario Pietro Cavallo, and Donato Malerba. 2014. Completion Time and Next Activity Prediction of Processes Using Sequential Pattern Mining. In *Discovery Science*. 49–61.
- [45] Recep Firat Cekinel and Pinar Karagoz. 2022. Event prediction from news text using subgraph embedding and graph sequence mining. *World Wide Web* 25, 6 (2022), 2403–2428.
- [46] Mehmet Celenk, Thomas Conley, James Graham, and John Willis. 2008. Anomaly prediction in network traffic using adaptive Wiener filtering and ARMA modeling. In *2008 IEEE International Conference on Systems, Man and Cybernetics*. 3548–3553.
- [47] Center for Regulatory Strategy Americas. 2020. *2020 health care regulatory outlook*. Technical Report.
- [48] Kang-Tsung Chang. 2008. *Introduction to geographic Inf. Syst.* Vol. 4. McGraw-Hill Boston.
- [49] Meng Chen, Xiaohui Yu, and Yang Liu. 2018. PCNN: Deep Convolutional Networks for Short-Term Traffic Congestion Prediction. *IEEE Trans. Intell. Transp. Syst.* 19, 11 (2018), 3550–3559.
- [50] Seong-Pyo Cheon, Sungshin Kim, So-Young Lee, and Chong-Bum Lee. 2009. Bayesian networks based rare event prediction with sensor data. *Knowl. Based. Syst.* 22, 5 (2009), 336–343.
- [51] Michele Chinosi and Alberto Trombetta. 2012. BPMN: An introduction to the standard. *Comput. Stand. Interfaces* 34, 1 (2012), 124–134.
- [52] Andrea Chiorrini, Claudia Diamantini, Alex Mircoli, and Domenico Potena. 2021. A Preliminary Study on the Application of Reinforcement Learning for Predictive Process Monitoring. In *Process Mining Workshops*. 124–135.
- [53] Andrea Chiorrini, Claudia Diamantini, Alex Mircoli, and Domenico Potena. 2022. Exploiting Instance Graphs and Graph Neural Networks for Next Activity Prediction. In *Process Mining Workshops*. 115–126.
- [54] Chung-Wen Cho, Yi-Hung Wu, Show-Jane Yen, Ying Zheng, and Arbee L. P. Chen. 2011. On-line rule matching for event prediction. *The VLDB Journal* 20, 3 (2011), 303–334.
- [55] Alexandre Checoli Choueiri, Denise Maria Vecino Sato, Edson Emilio Scalabrin, and Eduardo Alves Portela Santos. 2020. An extended model for remaining time prediction in manufacturing systems using process mining. *J. Manuf. Syst.* 56 (July 2020), 188–201.
- [56] Ryan Compton, Craig Lee, Jiejun Xu, Luis Artieda-Moncada, Tsai-Ching Lu, Lalindra De Silva, and Michael Macy. 2014. Using publicly visible social media to build detailed forecasts of civil unrest. *Security Informatics* 3, 1 (2014), 4.
- [57] Marco Comuzzi, Alfonso E. Marquez-Chamorro, and Manuel Resinas. 2019. Does Your Accurate Process Predictive Monitoring Model Give Reliable Predictions?. In *Service-Oriented Computing – IC3OC 2018 Workshops*. 367–373.
- [58] Raffaele Conforti, Sven Fink, Jonas Manderscheid, and Maximilian Röglinger. 2016. PRISM – A Predictive Risk Monitoring Approach for Business Processes. In *Business Process Management*. 383–400.
- [59] Stefano Coniglio, Anthony J. Dunn, and Alain B. Zemkoho. 2021. *Infrequent adverse event prediction in low carbon energy production using machine learning*. Technical Report. arXiv: 2001.06916. Retrieved from <http://arxiv.org/abs/2001.06916>.
- [60] Bitzel Cortez, Berny Carrera, Young-Jin Kim, and Jae-Yoon Jung. 2018. An architecture for emergency event prediction using LSTM recurrent neural networks. *Expert Syst. Appl.* 97 (May 2018), 315–324.
- [61] Noel Cressie. 2015. *Statistics for spatial data*. John Wiley & Sons.
- [62] Alfredo Cuzzocrea, Francesco Folino, Massimo Guarascio, and Luigi Pontieri. 2019. Predictive monitoring of temporally-aggregated performance indicators of business processes against low-level streaming events. *Inf. Syst.* 81 (March 2019), 236–266.
- [63] Sina Dami, Ahmad Abdollahzadeh Barforoush, and Hossein Shirazi. 2018. News events prediction using Markov logic networks. *J. Inf. Sci.* 44, 1 (2018), 91–109.
- [64] Arno De Caigny, Kristof Coussement, and Koen W. De Bock. 2018. A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *Eur. J. Oper. Res.* 269, 2 (2018), 760–772.
- [65] Matthew DeCamp and Jon C Tilburt. 2019. Why we cannot trust artificial intelligence in medicine. *The Lancet Digital Health* 1, 8 (2019), e390.
- [66] Tom Decroos, Vladimir Dzyuba, Jan Van Haaren, and Jesse Davis. 2017. Predicting Soccer Highlights from Spatio-Temporal Match Event Streams. In *Proceedings of AAAI*, Vol. 31.
- [67] Akash Deep, Dharmaraj Veeramani, and Shiyu Zhou. 2020. Event Prediction for Individual Unit Based on Recurrent Event Data Collected in Teleservice Systems. *IEEE Trans. Rel.* 69, 1 (March 2020), 216–227.

- [68] Walter H. Dempsey, Alexander Moreno, Christy K. Scott, Michael L. Dennis, David H. Gustafson, Susan A. Murphy, and James M. Rehg. 2017. iSurvive: An Interpretable, Event-time Prediction Model for mHealth. In *PMLR*. 970–979.
- [69] Xiaolei Di, Yu Xiao, Chao Zhu, Yang Deng, Qinpei Zhao, and Weixiong Rao. 2019. Traffic Congestion Prediction by Spatiotemporal Propagation Patterns. In *2019 20th IEEE International Conference on Mobile Data Management (MDM)*. 298–303.
- [70] Chiara Di Francescomarino, Marlon Dumas, Marco Federici, Chiara Ghidini, Fabrizio Maria Maggi, Williams Rizzi, and Luca Simonetto. 2018. Genetic algorithms for hyperparameter optimization in predictive business process monitoring. *Inf. Syst.* 74 (2018), 67–83.
- [71] Chiara Di Francescomarino and Chiara Ghidini. 2022. Predictive process monitoring. In *Process Mining Handbook*. Vol. 448. Springer, 320–346.
- [72] Chiara Di Francescomarino, Chiara Ghidini, Fabrizio Maria Maggi, and Fredrik Milani. 2018. Predictive Process Monitoring Methods: Which One Suits Me Best?. In *Business Process Management*. 462–479.
- [73] Chiara Di Francescomarino, Chiara Ghidini, Fabrizio Maria Maggi, Giulio Petrucci, and Anton Yeshchenko. 2017. An Eye into the Future: Leveraging A-priori Knowledge in Predictive Business Process Monitoring. In *Business Process Management*. 252–268.
- [74] Nicola Di Mauro, Annalisa Appice, and Teresa M. A. Basile. 2019. Activity Prediction of Business Process Instances with Inception CNN Models. In *AI\*IA 2019 – Advances in Artificial Intelligence*. 348–361.
- [75] Kiarash Diba, Kimon Batoulis, Matthias Weidlich, and Mathias Weske. 2020. Extraction, correlation, and abstraction of event data for process mining. *Wiley Interdiscip. Rev.: Data Min. Knowl. Discov.* 10, 3 (2020), e1346.
- [76] Steven X. Ding, Shen Yin, Kaixiang Peng, Haiyang Hao, and Bo Shen. 2013. A Novel Scheme for Key Performance Indicator Prediction and Diagnosis With Application to an Industrial Hot Strip Mill. *IEEE Trans. Ind. Inform.* 9, 4 (2013), 2239–2247.
- [77] Zuohua Ding, Yuan Zhou, Geguang Pu, and MengChu Zhou. 2018. Online Failure Prediction for Railway Transportation Systems Based on Fuzzy Rules and Data Analysis. *IEEE Trans. Rel.* 67, 3 (2018), 1143–1158.
- [78] Giovanni Diraco, Pietro Siciliano, and Alessandro Leone. 2023. Machine Learning-Based Anomaly Prediction for Smart Manufacturing. In *Sensors and Microsystems*. 239–244.
- [79] Paulo do Carmo and Ricardo Marcacini. 2021. Embedding propagation over heterogeneous event networks for link prediction. In *Proceedings of IEEE Big Data*. 4812–4821.
- [80] Huilong Duan, Zhoujian Sun, Wei Dong, Kunlun He, and Zhengxing Huang. 2020. On Clinical Event Prediction in Patient Treatment Trajectory Using Longitudinal Electronic Health Records. *IEEE J. Biomed. Health Inform.* 24, 7 (2020), 2053–2063.
- [81] Manisha Dubey, Ragja Palakkadavath, and P. K. Srijith. 2022. *Bayesian Neural Hawkes Process for Event Uncertainty Prediction*. Technical Report. arXiv: 2112.14474. Retrieved from <http://arxiv.org/abs/2112.14474>.
- [82] Juan C. Duenas, Jose M. Navarro, Hugo A. Parada G., Javier Andion, and Felix Cuadrado. 2018. Applying Event Stream Processing to Network Online Failure Prediction. *IEEE Commun. Mag.* 56, 1 (2018), 166–170.
- [83] Eduardo e Oliveira, Vera L. Miguéis, and José L. Borges. 2022. Automatic root cause analysis in manufacturing: an overview & conceptualization. *J. Intell. Manuf.* (February 2022).
- [84] Catherine Elgin. 2007. Understanding and the facts. *Philos. Stud.* 132, 1 (2007), 33–42.
- [85] Catherine Z Elgin. 2017. *True enough*. MIT Press.
- [86] Ahmed K. Elmagarmid, Panagiotis G. Ipeirotis, and Vassilios S. Verykios. 2007. Duplicate Record Detection: A Survey. *IEEE Trans. Knowl. Data Eng.* 19, 1 (2007), 1–16.
- [87] Tae Hyun Eom and Hyukjoon Lee. 2022. A WTTE-RNN Based Hypoglycemic Event Prediction. In *2022 International Conference on Information Networking (ICOIN)*. 378–379.
- [88] Şeyda Ertekin, Cynthia Rudin, and Tyler H. McCormick. 2015. Reactive point processes: A new approach to predicting power failures in underground electrical systems. *Ann Appl Stat* 9, 1 (2015), 122–144.
- [89] Lina Fahed and Philippe Lenca. 2020. Distant Event Prediction Based on Sequential Rules. *Data Science and Pattern Recognition* 4, 1 (2020), 1–23.
- [90] Stephan A. Fahrenkrog-Petersen, Niek Tax, Irene Teinmaa, Marlon Dumas, Massimiliano de Leoni, Fabrizio Maria Maggi, and Matthias Weidlich. 2022. Fire now, fire later: alarm-based systems for prescriptive process monitoring. *Knowledge and Inf. Syst.* 64, 2 (2022), 559–587.
- [91] Mahtab J. Fard, Sanjay Chawla, and Chandan K. Reddy. 2016. Early-Stage Event Prediction for Longitudinal Data. In *Advances in Knowledge Discovery and Data Mining (Lecture Notes in Computer Science)*. Springer International Publishing, 139–151.
- [92] Alexandre G L Fernandes, Thais R Neubauer, Marcelo Fantinato, and Sarajane M Peres. 2022. Impact of non-fitting cases for remaining time prediction in a multi-attribute process-aware method. In *CEUR Workshop Proceedings*. 13.
- [93] Iman Firouzian, Morteza Zahedi, and Hamid Hassanpour. 2019. Real-time Prediction and Synchronization of Business Process Instances using Data and Control Perspective. *Int. J. Nonlinear Anal. Appl.* 10, 1 (2019), 217–228.
- [94] Ioannis Flouris, Nikos Giatrakos, Antonios Deligiannakis, Minos Garofalakis, Michael Kamp, and Michael Mock. 2017. Issues in complex event processing: Status and prospects in the big data era. *J. Syst. Softw.* 127 (2017), 217–236.
- [95] Jon Fox and Donald E Brown. 2012. Using temporal indicator functions with generalized linear models for spatial-temporal event prediction. *Procedia Comput. Sci.* 8 (2012), 106–111.
- [96] Chiara Di Francescomarino, Marlon Dumas, Fabrizio Maria Maggi, and Irene Teinmaa. 2019. Clustering-Based Predictive Process Monitoring. *IEEE Trans. Serv. Comput.* 12, 6 (November 2019), 896–909.
- [97] Ilenia Fronza, Alberto Sillitti, Giancarlo Succi, Mikko Terho, and Jelena Vlasenko. 2013. Failure prediction based on log files using Random Indexing and Support Vector Machines. *J. Syst. Softw.* 86, 1 (2013), 2–11.

- [98] Joshi Fullop, Ana Gainaru, and Joel Plutchak. 2012. Real time analysis and event prediction engine. *Cray User Group (CUG) (2012)*.
- [99] Prasanna Gai, Malcolm Kemp, Antonio Sánchez Serrano, and Isabel Schnabel. 2019. Regulatory Complexity and the Quest for Robust Regulation. *SSRN Electronic Journal* (2019).
- [100] Ana Gainaru, Franck Cappello, Joshi Fullop, Stefan Trausan-Matu, and William Kramer. 2011. Adaptive event prediction strategy with dynamic time window for large-scale HPC systems. In *SLAML '11*. 1–8.
- [101] Riccardo Galanti, Bernat Coma-Puig, Massimiliano de Leoni, Josep Carmona, and Nicolò Navarin. 2020. Explainable Predictive Process Monitoring. In *2020 2nd International Conference on Process Mining (ICPM)*. 1–8.
- [102] Luis Galdo Seara and Renata Medeiros de Carvalho. 2019. An Approach for Workflow Improvement based on Outcome and Time Remaining Prediction. In *Proceedings of the 7th International Conference on Model-Driven Engineering and Software Development*. 473–480.
- [103] Yan Gao and Yan Cui. 2022. Clinical time-to-event prediction enhanced by incorporating compatible related outcomes. *PLOS Digital Health* 1, 5 (2022).
- [104] Yuyang Gao, Siyi Gu, Junji Jiang, Sungsoo Ray Hong, Dazhou Yu, and Liang Zhao. 2022. *Going Beyond XAI: A Systematic Survey for Explanation-Guided Learning*. Technical Report. arXiv: 2212.03954. Retrieved from <http://arxiv.org/abs/2212.03954>.
- [105] Yuyang Gao and Liang Zhao. 2018. Incomplete Label Multi-Task Ordinal Regression for Spatial Event Scale Forecasting. In *Proceedings of AAAI*, Vol. 32.
- [106] Yuyang Gao, Liang Zhao, Lingfei Wu, Yanfang Ye, Hui Xiong, and Chaowei Yang. 2019. Incomplete Label Multi-Task Deep Learning for Spatio-Temporal Event Subtype Forecasting. In *Proceedings of AAAI*, Vol. 33. 3638–3646.
- [107] Lise Getoor and Ashwin Machanavajjhala. 2012. Entity resolution: theory, practice & open challenges. *Proceedings of the VLDB Endowment* 5, 12 (2012), 2018–2019.
- [108] Martin Glinz. 2007. On Non-Functional Requirements. In *15th IEEE International Requirements Engineering Conference (RE 2007)*. 21–26.
- [109] Fatma Ezzahra Gmati, Salem Chakhar, Wided Lejouad Chaari, and Mark Xu. 2019. A Taxonomy of Event Prediction Methods. In *Advances and Trends in Artificial Intelligence. From Theory to Practice (Lecture Notes in Computer Science)*. Springer International Publishing, Cham, 12–26.
- [110] Stephen R. Grimm. 2006. Is Understanding A Species Of Knowledge? *Br. J. Philos. Sci.* 57, 3 (2006), 515–535.
- [111] Thomas Grisold, Jan Mendling, Markus Otto, and Jan vom Brocke. 2020. Adoption, use and management of process mining in practice. *Bus. Process Manag. J.* 27, 2 (January 2020), 369–387.
- [112] Xiaohui Gu and Haixun Wang. 2009. Online Anomaly Prediction for Robust Cluster Systems. In *Data Engineering*. 1000–1011.
- [113] Jan Gulliksen, Bengt Göransson, Inger Boivie, Stefan Blomkvist, Jenny Persson, and Åsa Cajander. 2003. Key principles for user-centred systems design. *Behav. Inf. Technol.* 22, 6 (2003), 397–409.
- [114] Björn Rafn Gunnarsson, Sepp K. L. M. vanden Broucke, and Jochen De Weerd. 2019. Predictive Process Monitoring in Operational Logistics: A Case Study in Aviation. In *Business Process Management Workshops*. 250–262.
- [115] Imed Hadj-Kacem, Sana Ben Jemaa, Sylvain Allio, and Yosra Ben Slimen. 2020. Anomaly prediction in mobile networks : A data driven approach for machine learning algorithm selection. In *NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium*. 1–7.
- [116] Nitin Harane and Sheetal Rathi. 2020. Comprehensive Survey on Deep Learning Approaches in Predictive Business Process Monitoring. In *Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough: Latest Trends in AI*. Springer International Publishing, Cham, 115–128.
- [117] Maximilian Harl, Sven Weinzierl, Mathias Stierle, and Martin Matzner. 2020. Explainable predictive business process monitoring using gated graph neural networks. *J. Decis. Syst.* 29, sup1 (August 2020), 312–327.
- [118] Kai Heinrich, Patrick Zschech, Christian Janiesch, and Markus Bonin. 2021. Process data properties matter: Introducing gated convolutional neural networks (GCNN) and key-value-predict attention networks (KVP) for next event prediction with deep learning. *Decis. Support Syst.* 143 (2021).
- [119] Germán Herrera Vidal and Jairo Rafael Coronado Hernández. 2021. Complexity in manufacturing systems: a literature review. *Prod. Eng.* 15, 3-4 (June 2021), 321–333.
- [120] Rachel Heyard, Jean-François Timsit, Leonhard Held, and Combacte-Magnet Consortium. 2020. Validation of discrete time-to-event prediction models in the presence of competing risks. *Biom. J.* 62, 3 (2020), 643–657.
- [121] Andrew Hoegh, Scotland Leman, Parang Saraf, and Naren Ramakrishnan. 2015. Bayesian Model Fusion for Forecasting Civil Unrest. *Technometrics* 57, 3 (July 2015), 332–340.
- [122] Linmei Hu, Juanzi Li, Liqiang Nie, Xiao-Li Li, and Chao Shao. 2017. What happens next? future subevent prediction using contextual hierarchical LSTM. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. 3450–3456.
- [123] Wenjie Hu, Yang Yang, Ziqiang Cheng, Carl Yang, and Xiang Ren. 2021. Time-Series Event Prediction with Evolutionary State Graph. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. ACM, Virtual Event Israel, 580–588.
- [124] Chao Huang, Chuxu Zhang, Jiashu Zhao, Xian Wu, Dawei Yin, and Nitesh Chawla. 2019. MiST: A Multiview and Multimodal Spatial-Temporal Learning Framework for Citywide Abnormal Event Forecasting. In *The World Wide Web Conference*. 717–728.
- [125] Fan Fang Huang, Li Qun Zhang, and Zhao Li. 2014. Time Prediction Based on Process Mining Taking Concept Drift into Consideration. In *Materials Science, Computer and Information Technology*. 2101–2105.
- [126] Huiqun Huang, Xi Yang, and Suining He. 2021. Multi-Head Spatio-Temporal Attention Mechanism for Urban Anomaly Event Prediction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 104:1–104:21.
- [127] Shaohan Huang, Carol Fung, Kui Wang, Polo Pei, Zhongzhi Luan, and Depei Qian. 2016. Using recurrent neural networks toward black-box system anomaly prediction. In *2016 IEEE/ACM 24th International Symposium on Quality of Service (IWQoS)*. 1–10.



- [128] Zhenyu Huang, Yongjun Wang, Hongzuo Xu, Songlei Jian, and Zhongyang Wang. 2021. Script event prediction based on pre-trained model with tail event enhancement. In *2021 5th International Conference on Computer Science and Artificial Intelligence*. 242–248.
- [129] G.F. Hughes, J.F. Murray, K. Kreutz-Delgado, and C. Elkan. 2002. Improved disk-drive failure warnings. *IEEE Trans. Rel.* 51, 3 (2002), 350–357.
- [130] Kristian Wang Høiem, Vemund Santi, Bendik Nybakk Torsæter, Helge Langseth, Christian André Andresen, and Gjert H. Rosenlund. 2020. Comparative Study of Event Prediction in Power Grids using Supervised Machine Learning Methods. In *2020 International Conference on Smart Energy Systems and Technologies (SEST)*. 1–6.
- [131] Eyke Hüllermeier and Willem Waegeman. 2021. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Mach. Learn.* 110, 3 (2021), 457–506.
- [132] Ihab F. Ilyas and Xu Chu. 2015. Trends in Cleaning Relational Data: Consistency and Deduplication. *Found. Trends Databases* 5, 4 (October 2015), 281–393.
- [133] Salman Jahani, Shiyu Zhou, Dharmaraj Veeramani, and Jeff Schmidt. 2021. Multioutput Gaussian Process Modulated Poisson Processes for Event Prediction. *IEEE Trans. Rel.* 70, 4 (December 2021), 1569–1580.
- [134] Abdulrahman Jalayer, Mohsen Kahani, Asef Pourmasoumi, and Amin Beheshti. 2022. HAM-Net: Predictive Business Process Monitoring with a hierarchical attention mechanism. *Knowl. Based. Syst.* 236 (January 2022), 107722.
- [135] Fakhruzy Shakirin Jamaludin and R. Kanesaraj Ramasamy. 2020. Systematic review on event prediction models. *Indones. J. Electr. Eng. Comput. Sci.* 19, 3 (2020), 1490.
- [136] Bram Jans, Steven Bethard, Ivan Vulic, and Marie-Francine Moens. 2012. Skip n-grams and ranking functions for predicting script events. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2012)*. 336–344.
- [137] Renhe Jiang, Xuan Song, Dou Huang, Xiaoya Song, Tianqi Xia, Zekun Cai, Zhaonan Wang, Kyoung-Sook Kim, and Ryosuke Shibasaki. 2019. DeepUrbanEvent: A System for Predicting Citywide Crowd Dynamics at Big Events. In *Proceedings of SIGKDD*. 2114–2122.
- [138] Song Jiang, Minjie Lian, Caiwu Lu, Qinghua Gu, Shunling Ruan, and Xuecai Xie. 2018. Ensemble Prediction Algorithm of Anomaly Monitoring Based on Big Data Analysis Platform of Open-Pit Mine Slope. *Complexity* 2018 (August 2018), 1–13.
- [139] Zhongjun Jin, Michael R. Anderson, Michael Cafarella, and H. V. Jagadish. 2017. Foofah: Transforming Data By Example. In *Proceedings of the 2017 ACM International Conference on Management of Data (SIGMOD '17)*. Association for Computing Machinery, New York, NY, USA, 683–698.
- [140] John Rydning. 2022. *Worldwide IDC Global DataSphere Forecast, 2022–2026: Enterprise Organizations Driving Most of the Data Growth*. Technical Report.
- [141] B. Kamala and B. Latha. 2022. Process Mining and Deep Neural Network approach for the Prediction of Business Process Outcome. In *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)*. 1–4.
- [142] Bokyoung Kang, Dongsoo Kim, and Suk-Ho Kang. 2012. Periodic performance prediction for real-time business process monitoring. *Ind. Manag. Data Syst.* 112, 1 (January 2012), 4–23.
- [143] Wei Kang, Jie Chen, Jiuyong Li, Jixue Liu, Lin Liu, Grant Osborne, Nick Lothian, Brenton Cooper, Terry Moschou, and Grant Neale. 2017. Carbon: Forecasting Civil Unrest Events by Monitoring News and Social Media. In *Advanced Data Mining and Applications*. 859–865.
- [144] Bill Karakostas. 2016. Event Prediction in an IoT Environment Using Naïve Bayesian Models. *Procedia Comput. Sci.* 83 (January 2016), 11–17.
- [145] Ahmed Kattan, Shaheen Fatima, and Muhammad Arif. 2015. Time-series event-based prediction: An unsupervised learning framework based on genetic programming. *Inf. Sci.* 301 (April 2015), 99–123.
- [146] Hojjat Kaveh and Hassan Salarieh. 2020. A new approach to extreme event prediction and mitigation via Markov-model-based chaos control. *Chaos Solitons Fractals* 136 (July 2020), 109827.
- [147] Naveed Khan, Sally McClean, Zulfqar Ali, Aftab Ali, Darryl Charles, Paul Taylor, and Detlef Nauck. 2019. Predictive Process Monitoring using a Markov Model Technique. In *2019 International Conference on Computing, Electronics & Communications Engineering (iCCECE)*. 193–196.
- [148] Amin Vahedian Khezerlou, Xun Zhou, Ling Tong, Yanhua Li, and Jun Luo. 2021. Forecasting Gathering Events through Trajectory Destination Prediction: A Dynamic Hybrid Model. *IEEE Trans. Knowl. Data Eng.* 33, 3 (March 2021), 991–1004.
- [149] Jongchan Kim, Marco Comuzzi, Marlon Dumas, Fabrizio Maria Maggi, and Irene Teinemaa. 2022. Encoding resource experience for predictive process monitoring. *Decis. Support Syst.* 153 (February 2022), 113669.
- [150] Karin Kipper, Anna Korhonen, Neville Ryant, and Martha Palmer. 2006. Extending VerbNet with novel verb classes. In *Proceedings of LREC*.
- [151] Hirokazu Kiyomaru, Kazumasa Omura, Yugo Murawaki, Daisuke Kawahara, and Sadao Kurohashi. 2019. Diversity-aware Event Prediction based on a Conditional Variational Autoencoder with Reconstruction. In *Commonsense Inference in Natural Language Processing*. 113–122.
- [152] Eva L. Klijn and Dirk Fahland. 2020. Identifying and Reducing Errors in Remaining Time Prediction due to Inter-Case Dynamics. In *2020 2nd International Conference on Process Mining (ICPM)*. 25–32.
- [153] Christopher Klinkmüller, Nick R. T. P. van Beest, and Ingo Weber. 2018. Towards Reliable Predictive Process Monitoring. In *Information Systems in the Big Data Era*. 163–181.
- [154] Wolfgang Kratsch, Jonas Manderscheid, Maximilian Röglinger, and Johannes Seyfried. 2021. Machine Learning in Business Process Monitoring: A Comparison of Deep Learning and Classical Approaches Used for Outcome Prediction. *Business & Inf. Syst. Engineering* 63, 3 (2021), 261–276.
- [155] Jonathan L Kvanvig. 2003. *The value of knowledge and the pursuit of understanding*. Cambridge University Press.
- [156] Martin Käppel, Stefan Jablonski, and Stefan Schöning. 2021. Cost-Sensitive Predictive Business Process Monitoring. In *New Trends in Database and Information Systems*. 14–26.

- [157] Martin Käppel, Stefan Jablonski, and Stefan Schöning. 2021. *Evaluating Predictive Business Process Monitoring Approaches on Small Event Logs*. Technical Report. arXiv: 2104.00362. Retrieved from <http://arxiv.org/abs/2104.00362>.
- [158] Seyed Behzad Lajevardi and Behrouz Minaei-Bidgoli. 2008. Combination of Time Series, Decision Tree and Clustering: A Case Study in Aerology Event Prediction. In *2008 International Conference on Computer and Electrical Engineering*. 111–115.
- [159] Geetika T. Lakshmanan, Davood Shamsi, Yurdaer N. Doganata, Merve Unuvar, and Rania Khalaf. 2015. A markov prediction model for data-driven semi-structured business processes. *Knowledge and Inf. Syst.* 42, 1 (2015), 97–126.
- [160] Zineb Lamghari, Maryam Radgui, Rajaa Saidi, and Moulay Driss Rahmani. 2019. Predictive Process Monitoring related to the remaining time dimension: a value-driven framework. In *Smart Systems and Data Science (ICSSD)*. 1–6.
- [161] Rocco Langone, Alfredo Cuzzocrea, and Nikolaos Skantzos. 2020. Interpretable Anomaly Prediction: Predicting anomalous behavior in industry 4.0 settings via regularized logistic regression tools. *Data Knowl. Eng.* 130 (November 2020), 101850.
- [162] Srivatsan Laxman, Vikram Tankasali, and Ryen W. White. 2008. Stream prediction using a generative model based on frequent episodes in event sequences. In *Proceedings of SIGKDD*. 453.
- [163] Mai Le, Bogdan Gabrys, and Detlef Nauck. 2017. A hybrid model for business process event and outcome prediction. *Expert Systems* 34, 5 (2017).
- [164] Suhwan Lee, Marco Comuzzi, and Nahyun Kwon. 2022. Exploring the Suitability of Rule-Based Classification to Provide Interpretability in Outcome-Based Process Predictive Monitoring. *Algorithms* 15, 6 (2022), 187.
- [165] Sander J. J. Leemans, Sebastiaan J. van Zelst, and Xixi Lu. 2023. Partial-order-based process mining: a survey and outlook. *Knowledge and Inf. Syst.* 65, 1 (2023), 1–29.
- [166] Lei Lei, Xuguang Ren, Nigel Franciscus, Junhu Wang, and Bela Stantic. 2019. Event Prediction Based on Causality Reasoning. In *Intelligent Information and Database Systems*. 165–176.
- [167] Anna Leontjeva, Raffaele Conforti, Chiara Di Francescomarino, Marlon Dumas, and Fabrizio Maria Maggi. 2015. Complex Symbolic Sequence Encodings for Predictive Monitoring of Business Processes. In *Business Process Management*. 297–313.
- [168] Benjamin Letham, Cynthia Rudin, and David Madigan. 2013. Sequential event prediction. *Mach. Learn.* 93, 2-3 (2013), 357–380.
- [169] Benjamin Letham, Cynthia Rudin, Tyler H. McCormick, and David Madigan. 2015. Interpretable classifiers using rules and Bayesian analysis: Building a better stroke prediction model. *Ann Appl Stat* 9, 3 (September 2015), 1350–1371.
- [170] Eldon Y. Li, Chen-Yuan Tung, and Shu-Hsun Chang. 2016. The wisdom of crowds in action: Forecasting epidemic diseases with a web-based prediction market system. *Int. J. Med. Inform.* 92 (August 2016), 35–43.
- [171] Hongcheng Li, Yuan Gao, Bing Wang, Yuewei Ming, and Zhongwei Zhao. 2022. Network Anomaly Sequence Prediction Method Based on LSTM and Two-layer Window Features. *J. Phys. Conf. Ser.* 2216, 1 (2022), 012063.
- [172] Hua Li, Héctor Muñoz-Avila, Diane Bramsen, Chad Hogg, and Rafael Alonso. 2009. Spatial Event Prediction by Combining Value Function Approximation and Case-Based Reasoning. In *Case-Based Reasoning Research and Development*. Vol. 5650. Springer Berlin Heidelberg, 465–478.
- [173] Hui Li, Jun-Ling Yu, Le-An Yu, and Jie Sun. 2014. The clustering-based case-based reasoning for imbalanced business failure prediction: a hybrid approach through integrating unsupervised process with supervised process. *Int. J. Syst. Sci.* 45, 5 (2014), 1225–1241.
- [174] Manling Li, Sha Li, Zhenhailong Wang, Lifu Huang, Kyunghyun Cho, Heng Ji, Jiawei Han, and Clare Voss. 2021. The future is not one-dimensional: Complex event schema induction by graph modeling for event prediction. In *EMNLP*. 5203–5215.
- [175] Yang Li. 2014. Reflection: enabling event prediction as an on-device service for mobile interaction. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM, Honolulu Hawaii USA, 689–698.
- [176] Yan Li, Tingjian Ge, and Cindy Chen. 2020. Data stream event prediction based on timing knowledge and state transitions. *Proceedings of the VLDB Endowment* 13, 10 (2020), 1779–1792.
- [177] Zhongyang Li, Xiao Ding, and Ting Liu. 2018. Constructing narrative event evolutionary graph for script event prediction. In *Proceedings of AAAI*. 4201–4207.
- [178] Renjie Liao, Xueyao Wang, Lun Li, and Zengchang Qin. 2010. A novel serial crime prediction model based on Bayesian learning theory. In *2010 International Conference on Machine Learning and Cybernetics*, Vol. 4. 1757–1762.
- [179] Li Lin, Lijie Wen, and Jianmin Wang. 2019. MM-Pred: A Deep Predictive Model for Multi-attribute Event Sequence. In *ICDM*. 118–126.
- [180] Xin-Xue Lin, Phone Lin, and En-Hau Yeh. 2021. Anomaly Detection/Prediction for the Internet of Things: State of the Art and the Future. *IEEE Network* 35, 1 (2021), 212–218.
- [181] Benjamin Lindemann, Nasser Jazdi, and Michael Weyrich. 2020. Anomaly detection and prediction in discrete manufacturing based on cooperative LSTM networks. In *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*. 1003–1010.
- [182] Chen Liu and Xiaoqi Li. 2021. Mining Temporal Dependency among Proactive Data Services and Its Delivery to System-level Anomaly Prediction. In *2021 IEEE International Conference on Web Services (ICWS)*. 657–662.
- [183] Daoyuan Liu, Yu Guo, Shaohua Huang, Weiguang Fang, and Xu Tian. 2022. A stacking denoising auto-encoder with sample weight approach for order remaining completion time prediction in complex discrete manufacturing workshop. *Int. J. Prod. Res.* (June 2022), 1–14.
- [184] Hua Liu and D.E. Brown. 2004. A new point process transition density model for space-time event prediction. *IEEE Trans. Syst., Man, Cybern. C* 34, 3 (August 2004), 310–324.
- [185] Qianlan Liu and Budan Wu. 2018. Prediction of Business Process Outcome based on Historical Log. In *Proceedings of the 10th International Conference on Computer Modeling and Simulation*. 118–122.

- [186] Sicen Liu, Xiaolong Wang, Yang Xiang, Hui Xu, Hui Wang, and Buzhou Tang. 2022. Multi-channel fusion LSTM for medical event prediction using EHRs. *J. Biomed. Inform.* 127 (March 2022), 104011.
- [187] Zhiyuan Liu, Hongming Cai, Han Yu, Bingqing Shen, and Lihong Jiang. 2021. Constructing the Sequential Event Graph for Event Prediction towards Cyber-Physical Systems. In *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. 1292–1297.
- [188] Chang Lu, Tian Han, and Yue Ning. 2022. Context-Aware Health Event Prediction via Transition Functions on Dynamic Disease Graphs. *Proceedings of AAAI* 36, 4 (June 2022), 4567–4574.
- [189] Chang Lu, Chandan K. Reddy, Prithwish Chakraborty, Samantha Kleinberg, and Yue Ning. 2021. *Collaborative Graph Learning with Auxiliary Text for Temporal Event Prediction in Healthcare*. Technical Report. arXiv: 2105.07542. Retrieved from <http://arxiv.org/abs/2105.07542>..
- [190] Chang Lu, Chandan K. Reddy, and Yue Ning. 2021. Self-Supervised Graph Learning With Hyperbolic Embedding for Temporal Health Event Prediction. *IEEE Trans. Cybern.* (2021), 1–13.
- [191] Ke Lu, Xinjian Fang, and Xianwen Fang. 2022. Business Process Outcome Prediction Based on Deep Latent Factor Model. *Electronics* 11, 9 (2022).
- [192] Wenjuan Luo, Han Zhang, Xiaodi Yang, Lin Bo, Xiaoqing Yang, Zang Li, Xiaohu Qie, and Jieping Ye. 2020. Dynamic Heterogeneous Graph Neural Network for Real-time Event Prediction. In *Proceedings of SIGKDD*. 3213–3223.
- [193] Linh Thao Ly, Fabrizio Maria Maggi, Marco Montali, Stefanie Rinderle-Ma, and Wil M. P. van der Aalst. 2015. Compliance monitoring in business processes: Functionalities, application, and tool-support. *Inf. Syst.* 54 (2015), 209–234.
- [194] Laura María López Castro, Sonia Geraldine Martínez, Nestor Eduardo Rodriguez, and Luna Violeta Lovera et al. 2021. Development of a Predictive Process Monitoring Methodology in a Self-organized Manufacturing System. In *Service Oriented, Holonic and Multi-Agent Manufacturing Systems for Industry of the Future*. 3–16.
- [195] Qiuping Ma, Hongyan Li, and Anders Thorstenson. 2021. A big data-driven root cause analysis system: Application of Machine Learning in quality problem solving. *Comput. Ind. Eng.* 160 (October 2021), 107580.
- [196] Fabrizio Maria Maggi, Chiara Di Francescomarino, Marlon Dumas, and Chiara Ghidini. 2014. Predictive Monitoring of Business Processes. In *Advanced Information Systems Engineering*. 457–472.
- [197] Marco Maisenbacher and Matthias Weidlich. 2017. Handling Concept Drift in Predictive Process Monitoring. In *2017 IEEE International Conference on Services Computing (SCC)*. 1–8.
- [198] Yasuko Matsubara, Yasushi Sakurai, Christos Faloutsos, Tomoharu Iwata, and Masatoshi Yoshikawa. 2012. Fast mining and forecasting of complex time-stamped events. In *Proceedings of SIGKDD*. 271–279.
- [199] John Meehan, Cansu Aslantas, Stan Zdonik, Nesime Tatbul, and Jiang Du. 2017. Data Ingestion for the Connected World.. In *CIDR*.
- [200] Nijat Mehdiyev, Joerg Evermann, and Peter Fettke. 2017. A Multi-stage Deep Learning Approach for Business Process Event Prediction. In *2017 IEEE 19th Conference on Business Informatics (CBI)*, Vol. 01. 119–128.
- [201] Nijat Mehdiyev, Joerg Evermann, and Peter Fettke. 2020. A Novel Business Process Prediction Model Using a Deep Learning Method. *Business & Inf. Syst. Engineering* 62, 2 (2020), 143–157.
- [202] Nijat Mehdiyev and Peter Fettke. 2021. Explainable Artificial Intelligence for Process Mining: A General Overview and Application of a Novel Local Explanation Approach for Predictive Process Monitoring. *Interpretable Artificial Intelligence: A Perspective of Granul. Comput.* 937 (2021), 1.
- [203] Nijat Mehdiyev and Peter Fettke. 2021. *Local Post-Hoc Explanations for Predictive Process Monitoring in Manufacturing*. Technical Report. arXiv: 2009.10513. Retrieved from <http://arxiv.org/abs/2009.10513>..
- [204] Sheyda Kiani Mehr, Wenting Sun, Xuancheng Fan, Nikita Butakov, and Nicolas Ferlans. 2020. Graph Machine Learning for Anomaly Prediction in Distributed Systems.. In *SEKE*. 204–209.
- [205] Soheila Mehrmolaie and Mohammad Reza Keyvanpour. 2015. A Brief Survey on Event Prediction Methods in Time Series. In *Artificial Intelligence Perspectives and Applications*. 235–246.
- [206] Andreas Metzger and Felix Föcker. 2017. Predictive Business Process Monitoring Considering Reliability Estimates. In *Advanced Information Systems Engineering*. 445–460.
- [207] Andreas Metzger, Philipp Leitner, Dragan Ivanović, and Eric Schmieders et al. 2015. Comparing and Combining Predictive Business Process Monitoring Techniques. *IEEE Trans. Syst., Man, Cybern., Syst.* 45, 2 (February 2015), 276–290.
- [208] Luay Thamer Mohammed, AbdAllah A. AlHabshy, and Kamal A. ElDahshan. 2022. Big Data Visualization: A Survey. In *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*. 1–12.
- [209] Soheila Mehr Molaie and Mohammad Reza Keyvanpour. 2015. An analytical review for event prediction system on time series. In *2015 2nd International Conference on Pattern Recognition and Image Analysis (IPRIA)*. 1–6.
- [210] Alfonso E. Márquez-Chamorro, Isabel A. Nepomuceno-Chamorro, Manuel Resinas, and Antonio Ruiz-Cortés. 2022. Updating Prediction Models for Predictive Process Monitoring. In *Advanced Information Systems Engineering*. 304–318.
- [211] Alfonso Eduardo Márquez-Chamorro, Manuel Resinas, and Antonio Ruiz-Cortés. 2018. Predictive Monitoring of Business Processes: A Survey. *IEEE Trans. Serv. Comput.* 11, 6 (November 2018), 962–977.
- [212] Alfonso E. Márquez-Chamorro, Manuel Resinas, Antonio Ruiz-Cortés, and Miguel Toro. 2017. Run-time prediction of business process indicators using evolutionary decision rules. *Expert Syst. Appl.* 87 (November 2017), 1–14.
- [213] Alfonso E. Márquez-Chamorro, Kate Revoredo, Manuel Resinas, Adela Del-Río-Ortega, Flávia M. Santoro, and Antonio Ruiz-Cortés. 2020. Context-Aware Process Performance Indicator Prediction. *IEEE Access* 8 (2020), 222050–222063.
- [214] Felix Naumann and Melanie Herschel. 2010. *An Introduction to Duplicate Detection*. Springer International Publishing, Cham.

- [215] Nicolo Navarin, Beatrice Vincenzi, Mirko Polato, and Alessandro Sperduti. 2017. LSTM networks for data-aware remaining time prediction of business process instances. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*. 1–7.
- [216] Dominic A. Neu, Johannes Lahann, and Peter Fettke. 2022. A systematic literature review on state-of-the-art deep learning methods for process prediction. *Artif. Intell. Rev.* 55, 2 (2022), 801–827.
- [217] Lukas Neumann, Andrew Zisserman, and Andrea Vedaldi. 2019. Future Event Prediction: If and When. In *Proceedings of CVPR*. 0–0.
- [218] Armelle Noulapeu Ngaffo, Walid El Ayeb, and Zied Choukair. 2019. An Event Prediction-based IP Multimedia Subsystem Service Exposure. In *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. 31–36.
- [219] An Nguyen, Srijeet Chatterjee, Sven Weinzierl, Leo Schwinn, Martin Matzner, and Bjoern Eskofier. 2020. Time matters: time-aware LSTMs for predictive business process monitoring. In *2020 2nd International Conference on Process Mining (ICPM)*. 112–123.
- [220] Yue Ning, Rongrong Tao, Chandan K. Reddy, Huzefa Rangwala, James C. Starz, and Naren Ramakrishnan. 2018. STAPLE: Spatio-Temporal Precursor Learning for Event Forecasting. In *Proceedings of SDM*. SIAM. 99–107.
- [221] Niyi Ogunbiyi, Artie Basukoski, and Thierry Chausalet. 2020. Investigating Social Contextual Factors in Remaining-Time Predictive Process Monitoring—A Survival Analysis Approach. *Algorithms* 13, 11 (2020), 267.
- [222] Niyi Ogunbiyi, Artie Basukoski, and Thierry Chausalet. 2021. Comparative analysis of clustering-based remaining-time predictive process monitoring approaches. *Int. J. Bus. Process Integr. Manag.* 10, 3-4 (2021), 230–241.
- [223] Niyi Ogunbiyi, Artie Basukoski, and Thierry Chausalet. 2021. Incorporating spatial context into remaining-time predictive process monitoring. In *Proc. ACM Symp. Appl. Comput.* ACM, New York, NY, USA, 535–542.
- [224] Maya Okawa, Tomoharu Iwata, Yusuke Tanaka, Takeshi Kurashima, Hiroyuki Toda, and Hisashi Kashima. 2022. Context-aware spatio-temporal event prediction via convolutional Hawkes processes. *Mach. Learn.* 111, 8 (2022), 2929–2950.
- [225] Motoyuki Oki, Koh Takeuchi, and Yukio Uematsu. 2018. Mobile Network Failure Event Detection and Forecasting With Multiple User Activity Data Sets. In *Proceedings of AAAI*, Vol. 32.
- [226] Chun Ouyang, Renuka Sindhgatta, and Catarina Moreira. 2021. *Explainable AI Enabled Inspection of Business Process Prediction Models*. Technical Report. arXiv: 2107.09767. Retrieved from <http://arxiv.org/abs/2107.09767>.
- [227] Suraj Pandey, Surya Nepal, and Shiping Chen. 2011. A test-bed for the evaluation of business process prediction techniques. In *7th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*. 382–391.
- [228] Kasey Panetta. 2021. Gartner Top 10 Data and Analytics Trends for 2021.
- [229] Gyunam Park and Wil M. P. van der Aalst. 2022. Action-oriented process mining: bridging the gap between insights and actions. *Prog. Artif. Intell* (2022).
- [230] Seo Young Park, Ji Eun Park, Hyungjin Kim, and Seong Ho Park. 2021. Review of Statistical Methods for Evaluating the Performance of Survival or Other Time-to-Event Prediction Models (from Conventional to Deep Learning Approaches). *Korean J. Radiol.* 22, 10 (2021), 1697–1707.
- [231] Vincenzo Pasquadibisceglie, Annalisa Appice, Giovanna Castellano, and Donato Malerba. 2022. A Multi-View Deep Learning Approach for Predictive Business Process Monitoring. *IEEE Trans. Serv. Comput.* 15, 4 (July 2022), 2382–2395.
- [232] Vincenzo Pasquadibisceglie, Annalisa Appice, Giovanna Castellano, Donato Malerba, and Giuseppe Modugno. 2020. ORANGE: Outcome-Oriented Predictive Process Monitoring Based on Image Encoding and CNNs. *IEEE Access* 8 (2020), 184073–184086.
- [233] Vincenzo Pasquadibisceglie, Giovanna Castellano, Annalisa Appice, and Donato Malerba. 2021. FOX: a neuro-Fuzzy model for process Outcome prediction and eXplanation. In *2021 3rd International Conference on Process Mining (ICPM)*. 112–119.
- [234] Stephen Pauwels and Toon Calders. 2021. Incremental Predictive Process Monitoring: The Next Activity Case. In *Business Process Management*. 123–140.
- [235] Marco Pegoraro, Merih Seran Uysal, David Benedikt Georgi, and Wil M. P. van der Aalst. 2021. *Text-Aware Predictive Monitoring of Business Processes*. Technical Report. arXiv: 2104.09962. Retrieved from <http://arxiv.org/abs/2104.09962>.
- [236] Kaci L. Pickett, Krithika Suresh, Kristen R. Campbell, Scott Davis, and Elizabeth Juarez-Colunga. 2021. Random survival forests for dynamic predictions of a time-to-event outcome using a longitudinal biomarker. *BMC Med. Res. Methodol.* 21, 1 (2021), 216.
- [237] Freddy Piraján, Andrey Fajardo, and Miguel Melgarejo. 2019. Towards a deep learning approach for urban crime forecasting. In *Workshop on Engineering Applications*. 179–189.
- [238] Mirko Polato, Alessandro Sperduti, Andrea Burattin, and Massimiliano de Leoni. 2014. Data-aware remaining time prediction of business process instances. In *2014 International Joint Conference on Neural Networks (IJCNN)*. 816–823.
- [239] Mirko Polato, Alessandro Sperduti, Andrea Burattin, and Massimiliano de Leoni. 2018. Time and activity sequence prediction of business process instances. *Computing* 100, 9 (2018), 1005–1031.
- [240] Pietro Portolani, Alessandro Brusaferrri, Andrea Ballarino, and Matteo Matteucci. 2022. Uncertainty in Predictive Process Monitoring. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*. 547–559.
- [241] Mahsa Pourbafrani, Shreya Kar, Sebastian Kaiser, and Wil M. P. van der Aalst. 2022. Remaining Time Prediction for Processes with Inter-case Dynamics. In *Process Mining Workshops*. 140–153.
- [242] Mohammad Pourhomayoun, Nabil Alshurafa, Bobak Mortazavi, Hassan Ghasemzadeh, Konstantinos Sideris, Bahman Sadeghi, Michael Ong, Lorraine Evangelista, Patrick Romano, Andrew Auerbach, Asher Kimchi, and Majid Sarrafzadeh. 2014. Multiple model analytics for adverse event prediction in remote health monitoring systems. In *2014 IEEE Healthcare Innovation Conference (HIC)*. 106–110.

- [243] Ioannis Prasadis, Nikolaos-Paraskevas Theodoropoulos, Alexandros Bousdekis, Georgia Theodoropoulou, and Georgios Miaoulis. 2021. Handling Uncertainty in Predictive Business Process Monitoring with Bayesian Networks. In *Information, Intelligence, Systems & Applications*. 1–8.
- [244] Zhi Qiao, Shiwang Zhao, Cao Xiao, Xiang Li, Yong Qin, and Fei Wang. 2018. Pairwise-Ranking based Collaborative Recurrent Neural Networks for Clinical Event Prediction. In *Proceedings of IJCAI*. 3520–3526.
- [245] Kira Radinsky, Sagie Davidovich, and Shaul Markovitch. 2012. Learning causality for news events prediction. In *Proceedings of the 21st international conference on World Wide Web*. 909–918.
- [246] Erhard Rahm and Philip A. Bernstein. 2001. A survey of approaches to automatic schema matching. *The VLDB Journal* 10, 4 (2001), 334–350.
- [247] Efred Rama-Maneiro, Juan Vidal, and Manuel Lama. 2021. Deep Learning for Predictive Business Process Monitoring: Review and Benchmark. *IEEE Trans. Serv. Comput.* (2021), 1–1.
- [248] Naren Ramakrishnan, Patrick Butler, Sathappan Muthiah, and Nathan Self et al. 2014. 'Beating the news' with EMBERS: forecasting civil unrest using open source indicators. In *Proceedings of SIGKDD*. 1799–1808.
- [249] Aaqib Rashid and Amit Chaturvedi. 2019. Cloud computing characteristics and services: a brief review. *Int. J. Comput. Sci. Eng.* 7, 2 (2019), 421–426.
- [250] Jonathan Rebane, Isak Karlsson, and Panagiotis Papapetrou. 2019. An Investigation of Interpretable Deep Learning for Adverse Drug Event Prediction. In *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*. 337–342.
- [251] David Reid, Abir Jaafar Hussain, Hissam Tawfik, Rozaida Ghazali, and Dhiya Al-Jumeily. 2018. Forecasting Natural Events Using Axonal Delay. In *2018 IEEE Congress on Evolutionary Computation (CEC)*. 1–6.
- [252] Honglei Ren, You Song, Jingwen Wang, Yucheng Hu, and Jinzhi Lei. 2018. A Deep Learning Approach to the Citywide Traffic Accident Risk Prediction. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. 3346–3351.
- [253] Lei Ren, Yuxin Liu, Di Huang, Keke Huang, and Chunhua Yang. 2022. MCTAN: A Novel Multichannel Temporal Attention-Based Network for Industrial Health Indicator Prediction. *IEEE Trans. Neural Netw. Learn. Syst.* (2022), 1–12.
- [254] Collin Rice. 2021. *Leveraging Distortions: Explanation, Idealization, and Universality in Science*. The MIT Press.
- [255] Pedro Rico, Félix Cuadrado, Juan C. Dueñas, Javier Andión, and Hugo A. Parada G. 2021. Business Process Event Prediction Through Scalable Online Learning. *IEEE Access* 9 (2021), 136313–136333.
- [256] Stefanie Rinderle-Ma, Karolin Winter, and Janik-Vasily Benzin. 2022. *Predictive Compliance Monitoring in Process-Aware Information Systems: State of the Art, Functionalities, Research Directions*. Technical Report. arXiv: 2205.05446. Retrieved from <http://arxiv.org/abs/2205.05446>.
- [257] Williams Rizzi, Chiara Di Francescomarino, and Fabrizio Maria Maggi. 2020. Explainability in Predictive Process Monitoring: When Understanding Helps Improving. In *Business Process Management Forum*. 141–158.
- [258] Williams Rizzi, Chiara Di Francescomarino, Chiara Ghidini, and Fabrizio Maria Maggi. 2022. How do I update my model? On the resilience of Predictive Process Monitoring models to change. *Knowledge and Inf. Syst.* 64, 5 (2022), 1385–1416.
- [259] Andreas Rogge-Solti and Mathias Weske. 2015. Prediction of business process durations using non-Markovian stochastic Petri nets. *Inf. Syst.* 54 (2015), 1–14.
- [260] Yuji Roh, Geon Heo, and Steven Euijong Whang. 2021. A Survey on Data Collection for Machine Learning: A Big Data - AI Integration Perspective. *IEEE Trans. Knowl. Data Eng.* 33, 4 (April 2021), 1328–1347.
- [261] Massimiliano Ronzani, Roger Ferrod, Chiara Di Francescomarino, Emilio Sulis, Roberto Aringhieri, Guido Boella, Enrico Brunetti, Luigi Di Caro, Mauro Dragoni, Chiara Ghidini, and Renata Marinello. 2022. Unstructured Data in Predictive Process Monitoring: Lexicographic and Semantic Mapping to ICD-9-CM Codes for the Home Hospitalization Service. In *AIXIA 2021 – Advances in Artificial Intelligence*. 700–715.
- [262] Cynthia Rudin, Benjamin Letham, and David Madigan. 2013. Learning theory analysis for association rules and sequential event prediction. *The Journal of Machine Learning Research* 14, 1 (2013), 3441–3492.
- [263] Cynthia Rudin, Rebecca J. Passonneau, Axinia Radeva, Haimonti Dutta, Steve Jerome, and Delfina Isaac. 2010. A process for predicting manhole events in Manhattan. *Mach. Learn.* 80, 1 (2010), 1–31.
- [264] Cao Rui, Zeng Qingtian, Ni Weijian, Lu Faming, Liu Cong, and Duan Hua. 2022. Explainable Business Process Remaining Time Prediction using Reachability Graph. *Chin. J. Electron.* (2022).
- [265] Mohammed Mohammed Sadeeq, Nasiba M Abdulkareem, Subhi RM Zeebaree, Dindar Mikael Ahmed, Ahmed Saifullah Sami, and Rizgar R Zebari. 2021. IoT and Cloud computing issues, challenges and opportunities: A review. *Qubahan Academic Journal* 1, 2 (2021), 1–7.
- [266] Ramendra K. Sahoo, Adam J. Oliner, Irina Rish, Manish Gupta, José E. Moreira, Sheng Ma, Ricardo Vilalta, and Anand Sivasubramaniam. 2003. Critical event prediction for proactive management in large-scale computer clusters. In *Proceedings of SIGKDD*. 426–435.
- [267] Felix Salfner, Maren Lenk, and Mirosław Malek. 2010. A survey of online failure prediction methods. *ACM Comput. Surveys* 42, 3 (2010), 1–42.
- [268] Arik Senderovich, Chiara Di Francescomarino, and Fabrizio Maria Maggi. 2019. From knowledge-driven to data-driven inter-case feature encoding in predictive process monitoring. *Inf. Syst.* 84 (September 2019), 255–264.
- [269] Seung Yeop Shin and Han-joon Kim. 2019. Autoencoder-based One-class Classification Technique for Event Prediction. In *Proceedings of the 2019 4th International Conference on Cloud Computing and Internet of Things*. ACM, Tokyo Japan, 54–58.
- [270] Siva S. Sivatha Sindhu, S. Geetha, Siva. S. Sivanath, and A. Kannan. 2006. A Neuro-genetic ensemble Short Term Forecasting Framework for Anomaly Intrusion Prediction. In *2006 International Conference on Advanced Computing and Communications*. 187–190.
- [271] Hossein Soleimani, James Hensman, and Suchi Saria. 2018. Scalable Joint Models for Reliable Uncertainty-Aware Event Prediction. *IEEE Trans. Pattern Anal. Mach. Intell.* 40, 8 (August 2018), 1948–1963.

- [272] Andrei Solomon and Marin Litoiu. 2011. Business process performance prediction on a tracked simulation model. In *Proceedings of the 3rd International Workshop on Principles of Engineering Service-Oriented Systems*. 50–56.
- [273] Florian Spree. 2020. Predictive Process Monitoring - A Use-Case-Driven Literature Review.. In *EMISA Forum: Vol. 40, No. 1*. 43–54.
- [274] Alexander Stevens, Johannes De Smedt, and Jari Peepkorn. 2022. Quantifying Explainability in Outcome-Oriented Predictive Process Monitoring. In *Process Mining Workshops*. 194–206.
- [275] Matthias Stierle, Jens Brunk, Sven Weinzierl, Sandra Zilker, Martin Matzner, and Jörg Becker. 2021. *Bringing Light Into the Darkness - A Systematic Literature Review on Explainable Predictive Business Process Monitoring Techniques*.
- [276] Michael Stonebraker and Ihab F. Ilyas. 2018. Data Integration: The Current Status and the Way Forward. *IEEE Data Eng. Bull.* 41 (2018), 3–9.
- [277] Zichun Su and Jialin Jiang. 2020. Hierarchical Gated Recurrent Unit with Semantic Attention for Event Prediction. *Future Internet* 12, 2 (2020), 39.
- [278] Xiaoxiao Sun, Wenjie Hou, Yuke Ying, and Dongjin Yu. 2020. Remaining Time Prediction of Business Processes based on Multilayer Machine Learning. In *2020 IEEE International Conference on Web Services (ICWS)*. 554–558.
- [279] Gian Antonio Susto, Andrea Schirru, Simone Pampuri, Seán McLoone, and Alessandro Beghi. 2015. Machine Learning for Predictive Maintenance: A Multiple Classifier Approach. *IEEE Trans. Ind. Informat.* 11, 3 (June 2015), 812–820.
- [280] Yashar Taghizadegan, Nader Jafarnia Dabanloo, Keivan Maghooli, and Ali Sheikhani. 2021. Obstructive sleep apnea event prediction using recurrence plots and convolutional neural networks (RP-CNNs) from polysomnographic signals. *Biomed. Signal Process. Control* 69 (2021), 102928.
- [281] Bayu Adhi Tama and Marco Comuzzi. 2019. An empirical comparison of classification techniques for next event prediction using business process event logs. *Expert Syst. Appl.* 129 (September 2019), 233–245.
- [282] Bayu Adhi Tama and Marco Comuzzi. 2022. Leveraging an Advanced Heterogeneous Ensemble Learning for Outcome-Based Predictive Monitoring Using Business Process Event Logs. *SSRN Electronic Journal* (2022).
- [283] Bayu Adhi Tama, Marco Comuzzi, and Jonghyeon Ko. 2020. An Empirical Investigation of Different Classifiers, Encoding, and Ensemble Schemes for Next Event Prediction Using Business Process Event Logs. *ACM Trans. Intell. Syst. Technol.* 11, 6 (2020), 1–34.
- [284] Yongmin Tan, Xiaohui Gu, and Haixun Wang. 2010. Adaptive system anomaly prediction for large-scale hosting infrastructures. In *Proceeding of the 29th ACM SIGACT-SIGOPS symposium on Principles of distributed computing - PODC '10*. 173.
- [285] Zeeshan Tariq, Darryl Charles, Sally McClean, Ian McChesney, and Paul Taylor. 2021. Proactive business process mining for end-state prediction using trace features. In *SmartWorld/SCALCOM/UIC/ATC/IOP/SCI*. 647–652.
- [286] Niek Tax, Irene Teinemaa, and Sebastiaan J. van Zelst. 2020. An interdisciplinary comparison of sequence modeling methods for next-element prediction. *Softw. Syst. Model.* 19, 6 (2020), 1345–1365.
- [287] Niek Tax, Ilya Verenich, Marcello La Rosa, and Marlon Dumas. 2017. Predictive Business Process Monitoring with LSTM Neural Networks. In *Advanced Information Systems Engineering*. 477–492.
- [288] Farbod Taymouri, Marcello La Rosa, Sarah Erfani, Zahra Dasht Bozorgi, and Ilya Verenich. 2020. Predictive Business Process Monitoring via Generative Adversarial Nets: The Case of Next Event Prediction. In *Business Process Management*. 237–256.
- [289] Irene Teinemaa, Marlon Dumas, Fabrizio Maria Maggi, and Chiara Di Francescomarino. 2016. Predictive Business Process Monitoring with Structured and Unstructured Data. In *Business Process Management*. 401–417.
- [290] Irene Teinemaa, Marlon Dumas, Marcello La Rosa, and Fabrizio Maria Maggi. 2019. Outcome-Oriented Predictive Process Monitoring: Review and Benchmark. *ACM Trans. Knowl. Discov. Data* 13, 2 (2019), 17:1–17:57.
- [291] Jonathan Tuke, Andrew Nguyen, Mehwish Nasim, Drew Mellor, Asanga Wickramasinghe, Nigel Bean, and Lewis Mitchell. 2020. Pachinko Prediction: A Bayesian method for event prediction from social media data. *Inf. Process Manag.* 57, 2 (2020), 102147.
- [292] Jay Ulfelder. 2013. A multimodel ensemble for forecasting onsets of state-sponsored mass killing. In *Proceedings of APSA*.
- [293] Amin Vahedian, Xun Zhou, Ling Tong, W. Nick Street, and Yanhua Li. 2019. Predicting Urban Dispersal Events: A Two-Stage Framework through Deep Survival Analysis on Mobility Data. In *Proceedings of AAAI*, Vol. 33. 5199–5206.
- [294] Sebastiaan J. van Zelst, Felix Mannhardt, Massimiliano de Leoni, and Agnes Koschmider. 2021. Event abstraction in process mining: literature review and taxonomy. *Granul. Comput.* 6, 3 (2021), 719–736.
- [295] Jarne Vandenebeele, Gilles Vermaut, Jari Peepkorn, and Jochen De Weerd. 2022. *Enhancing Stochastic Petri Net-based Remaining Time Prediction using k-Nearest Neighbors*. Technical Report. arXiv: 2206.13109. Retrieved from <http://arxiv.org/abs/2206.13109>..
- [296] Ilya Verenich, Marlon Dumas, Marcello La Rosa, Fabrizio Maria Maggi, and Chiara Di Francescomarino. 2016. Complex Symbolic Sequence Clustering and Multiple Classifiers for Predictive Process Monitoring. In *Business Process Management Workshops*. 218–229.
- [297] Ilya Verenich, Marlon Dumas, Marcello La Rosa, Fabrizio Maria Maggi, and Irene Teinemaa. 2019. Survey and Cross-benchmark Comparison of Remaining Time Prediction Methods in Business Process Monitoring. *ACM Trans. Intell. Syst. Technol.* 10, 4 (2019), 34:1–34:34.
- [298] R. Vilalta and Sheng Ma. 2002. Predicting rare events in temporal domains. In *2002 IEEE International Conference on Data Mining*. 474–481.
- [299] Pascal Vrignat, Manuel Avila, Florent Duculty, and Frédéric Kratz. 2015. Failure Event Prediction Using Hidden Markov Model Approaches. *IEEE Trans. Rel.* 64, 3 (September 2015), 1038–1048.
- [300] Nur Ahmad Wahid, Hyerim Bae, Taufik Nur Adi, Yulim Choi, and Yelita Anggiane Iskandar. 2021. Parallel-Structure Deep Learning for Prediction of Remaining Time of Process Instances. *Appl. Sci.* 11, 21 (January 2021), 9848.
- [301] Chi Wang and Jian Cao. 2021. Interval-Based Remaining Time Prediction for Business Processes. In *Service-Oriented Computing*. 34–48.
- [302] Dawei Wang and Wei Ding. 2015. A Hierarchical Pattern Learning Framework for Forecasting Extreme Weather Events. In *2015 IEEE International Conference on Data Mining*. 1021–1026.

- [303] Jiaojiao Wang, Victor Chang, Dongjin Yu, Chang Liu, Xiaoyu Ma, and Dingguo Yu. 2022. Conformance-oriented Predictive Process Monitoring in BPaaS Based on Combination of Neural Networks. *J. Grid Comput.* 20, 3 (2022), 25.
- [304] Junxiang Wang, Yuyang Gao, Andreas Züfle, Jingyuan Yang, and Liang Zhao. 2018. Incomplete Label Uncertainty Estimation for Petition Victory Prediction with Dynamic Features. In *2018 IEEE International Conference on Data Mining (ICDM)*. 537–546.
- [305] Jingwen Wang, Jingxin Liu, Juntao Pu, Qinghong Yang, Zhongchen Miao, Jian Gao, and You Song. 2019. An anomaly prediction framework for financial IT systems using hybrid machine learning methods. *J. Ambient Intell. Humaniz. Comput.* (December 2019).
- [306] Jiaojiao Wang, Dongjin Yu, Chengfei Liu, and Xiaoxiao Sun. 2019. Outcome-Oriented Predictive Process Monitoring with Attention-Based Bidirectional LSTM Neural Networks. In *2019 IEEE International Conference on Web Services (ICWS)*. 360–367.
- [307] Lihong Wang, Juwei Yue, Shu Guo, Jiawei Sheng, Qianren Mao, Zhenyu Chen, Shenghai Zhong, and Chen Li. 2021. Multi-level Connection Enhanced Representation Learning for Script Event Prediction. In *Proceedings of the Web Conference 2021*. 3524–3533.
- [308] Mingjun Wang and Matthew S. Gerber. 2015. Using Twitter for Next-Place Prediction, with an Application to Crime Prediction. In *2015 IEEE Symposium Series on Computational Intelligence*. 941–948.
- [309] Qi Wang, Guangyin Jin, Xia Zhao, Yanghe Feng, and Jincai Huang. 2020. CSAN: A neural network benchmark model for crime forecasting in spatio-temporal scale. *Knowl. Based. Syst.* 189 (February 2020), 105120.
- [310] Zhongqing Wang, Yue Zhang, and Ching-Yun Chang. 2017. Integrating Order Information and Event Relation for Script Event Prediction. In *EMNLP. Association for Computational Linguistics, Copenhagen, Denmark*, 57–67.
- [311] Tingyi Wanyan, Hossein Honarvar, Suraj K. Jaladanki, Chengxi Zang, Nidhi Naik, Sulaiman Somani, Jessica K. De Freitas, Ishan Paranjpe, Akhil Vaid, Jing Zhang, Riccardo Miotto, Zhangyang Wang, Girish N. Nadkarni, Marinka Zitnik, Ariful Azad, Fei Wang, Ying Ding, and Benjamin S. Glicksberg. 2021. Contrastive learning improves critical event prediction in COVID-19 patients. *Patterns* 2, 12 (2021), 100389.
- [312] Sven Weinzierl, Sebastian Dunzer, Johannes Tenschert, Sandra Zilker, and Martin Matzner. 2021. Predictive Business Process Deviation Monitoring. In *Proceedings of ECIS*.
- [313] Sven Weinzierl, Sandra Zilker, Matthias Stierle, Martin Matzner, and Gyunam Park. 2020. From predictive to prescriptive process monitoring: Recommending the next best actions instead of calculating the next most likely events.. In *Wirtschaftsinformatik (Zentrale Tracks)*. 364–368.
- [314] Michael Weisberg. 2012. *Simulation and similarity: Using models to understand the world*. Oxford University Press.
- [315] Jeremy C. Weiss and David Page. 2013. Forest-Based Point Process for Event Prediction from Electronic Health Records. In *Machine Learning and Knowledge Discovery in Databases*. 547–562.
- [316] Hans Weytjens and Jochen De Weerd. 2020. Process Outcome Prediction: CNN vs. LSTM (with Attention). In *Business Process Management Workshops*. 321–333.
- [317] Hans Weytjens and Jochen De Weerd. 2022. Learning uncertainty with artificial neural networks for predictive process monitoring. *Appl. Soft Comput.* 125 (August 2022), 109134.
- [318] Frederik Wolf, Jens Brunk, and Jörg Becker. 2021. A Framework of Business Process Monitoring and Prediction Techniques. In *Innovation Through Information Systems*. 714–724.
- [319] Songyun Wu, Bo Wang, Zhiliang Wang, Shuhan Fan, Jiahai Yang, and Jia Li. 2022. Joint prediction on security event and time interval through deep learning. *Comput. Secur.* 117 (June 2022), 102696.
- [320] Zhiliang Wu, Yinchong Yang, Peter A. Fashing, and Volker Tresp. 2021. Uncertainty-Aware Time-to-Event Prediction using Deep Kernel Accelerated Failure Time Models. In *Proceedings of the 6th Machine Learning for Healthcare Conference*. PMLR, 54–79.
- [321] Yizhou Xu, Jiandong Wang, and Yan Yu. 2020. Alarm Event Prediction From Historical Alarm Flood Sequences Based on Bayesian Estimators. *IEEE Trans. Autom. Sci. Eng.* 17, 2 (April 2020), 1070–1075.
- [322] Cong Xue, Zehua Zeng, Yuanye He, Lei Wang, and Neng Gao. 2018. A MIML-LSTM neural network for integrated fine-grained event forecasting. In *Proceedings of 2018 International Conference on Big Data Technologies*. 44–51.
- [323] Yifei Xue and Donald E. Brown. 2006. Spatial analysis with preference specification of latent decision makers for criminal event prediction. *Decis. Support Syst.* 41, 3 (2006), 560–573.
- [324] Xiang-Bin Yan, Tao Lu, Yi-Jun Li, and Guang-Bin Cui. 2004. Research on event prediction in time-series data. In *Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.04EX826)*, Vol. 5. 2874–2878 vol.5.
- [325] Qiang Yang, Hui Wang, and Wei Zhang. 2002. Web-log mining for quantitative temporal-event prediction. *IEEE Comput. Intell. Mag.* 1, 1 (2002), 10–18.
- [326] Yiyang Yang, Zhongyu Wei, Qin Chen, and Libo Wu. 2019. Using External Knowledge for Financial Event Prediction Based on Graph Neural Networks. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 2161–2164.
- [327] Zezhang Yang, Jian Li, and Ping Yang. 2021. FedADMP: A Joint Anomaly Detection and Mobility Prediction Framework via Federated Learning. *ICST Transactions on Security and Safety* 8, 29 (October 2021).
- [328] Anton Yeshchenko, Fernando Durier, Kate Revoredo, Jan Mendling, and Flavia Santoro. 2018. Context-Aware Predictive Process Monitoring: The Impact of News Sentiment. In *On the Move to Meaningful Internet Systems. OTM 2018 Conferences*. 586–603.
- [329] Fei Yi, Zhiwen Yu, Fuzhen Zhuang, and Bin Guo. 2019. Neural Network based Continuous Conditional Random Field for Fine-grained Crime Prediction. In *Proceedings of IJCAI*. 4157–4163.
- [330] Fei Yi, Zhiwen Yu, Fuzhen Zhuang, Xiao Zhang, and Hui Xiong. 2018. An Integrated Model for Crime Prediction Using Temporal and Spatial Factors. In *2018 IEEE International Conference on Data Mining (ICDM)*. 1386–1391.

- [331] Shuqi Yu, Linmei Hu, and Bin Wu. 2019. DRAM: A Deep Reinforced Intra-attentive Model for Event Prediction. In *Knowledge Science, Engineering and Management*. 701–713.
- [332] Yuan Yuan, Shiyu Zhou, Crispian Sievenpiper, Kamal Mannar, and Yibin Zheng. 2011. Event log modeling and analysis for system failure prediction. *IIE Trans.* 43, 9 (2011), 647–660.
- [333] Zhuoning Yuan, Xun Zhou, and Tianbao Yang. 2018. Hetero-ConvLSTM: A Deep Learning Approach to Traffic Accident Prediction on Heterogeneous Spatio-Temporal Data. In *Proceedings of SIGKDD*. 984–992.
- [334] Hongming Zhang, Muhao Chen, Haoyu Wang, Yangqiu Song, and Dan Roth. 2020. Analogous Process Structure Induction for Sub-event Sequence Prediction. In *EMNLP*. 1541–1550.
- [335] Shengdong Zhang, Soheil Bahrapour, Naveen Ramakrishnan, Lukas Schott, and Mohak Shah. 2017. Deep learning on symbolic representations for large-scale heterogeneous time-series event prediction. In *ICASSP*. 5970–5974.
- [336] Yaru Zhang and Xijin Tang. 2022. Introducing Trigger Evolutionary Graph and Event Segment for Event Prediction. In *Knowledge and Systems Sciences*. 186–201.
- [337] Liang Zhao. 2022. Event Prediction in the Big Data Era: A Systematic Survey. *ACM Comput. Surveys* 54, 5 (June 2022), 1–37.
- [338] Liang Zhao, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. 2015. Spatiotemporal Event Forecasting in Social Media. In *Proceedings of SDM*. SIAM. 963–971.
- [339] Liang Zhao, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. 2016. Multi-resolution Spatial Event Forecasting in Social Media. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*. 689–698.
- [340] Liang Zhao, Jiangzhuo Chen, Feng Chen, Wei Wang, Chang-Tien Lu, and Naren Ramakrishnan. 2015. SimNest: Social Media Nested Epidemic Simulation via Online Semi-Supervised Deep Learning. In *2015 IEEE International Conference on Data Mining*. 639–648.
- [341] Liang Zhao, Olga Gkountouna, and Dieter Pfoser. 2019. Spatial Auto-regressive Dependency Interpretable Learning Based on Spatial Topological Constraints. *ACM Trans. Spat. Algorithms Syst.* 5, 3 (2019), 19:1–19:28.
- [342] Liang Zhao, Junxiang Wang, and Xiaojie Guo. 2018. Distant-Supervision of Heterogeneous Multitask Learning for Social Event Forecasting With Multilingual Indicators. In *Proceedings of AAAI*, Vol. 32.
- [343] Liang Zhao, Jieping Ye, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. 2016. Hierarchical incomplete multi-source feature learning for spatiotemporal event forecasting. In *Proceedings of SIGKDD*. 2085–2094.
- [344] Xinyu Zhao, Hao Yan, and Yongming Liu. 2021. Hierarchical Tree-based Sequential Event Prediction with Application in the Aviation Accident Report. In *2021 IEEE 37th International Conference on Data Engineering (ICDE)*. 1925–1930.
- [345] Alice Zheng and Amanda Casari. 2018. *Feature engineering for machine learning: principles and techniques for data scientists*. "O'Reilly Media, Inc."
- [346] Qiuzhen Zhong, Jingjing Wang, Xuejie Meng, Siqing Liu, and Jiancun Gong. 2019. Prediction Model for Solar Energetic Proton Events: Analysis and Verification. *Space Weather* 17, 5 (2019), 709–726.
- [347] Bo Zhou, Yubo Chen, Kang Liu, Jun Zhao, Jiexin Xu, Xiaojian Jiang, and Jinlong Li. 2021. Multi-Task Self-Supervised Learning for Script Event Prediction. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. ACM, 3662–3666.
- [348] Cheng Zhou, Boris Cule, and Bart Goethals. 2015. A pattern based predictor for event streams. *Expert Syst. Appl.* 42, 23 (2015), 9294–9306.
- [349] Lihua Zhou, Guowang Du, Ruxin Wang, Dapeng Tao, Lizhen Wang, Jun Cheng, and Jing Wang. 2019. A tensor framework for geosensor data forecasting of significant societal events. *Pattern Recognit.* 88 (April 2019), 27–37.
- [350] Xiaozhen Zhou, Shanping Li, and Zhen Ye. 2013. A Novel System Anomaly Prediction System Based on Belief Markov Model and Ensemble Classification. *Math. Probl. Eng.* 2013 (2013), 179390.
- [351] Jiarun Zhu, Xingyu Wu, Muhammad Usman, Xiangyu Wang, and Huanhuan Chen. 2022. Link Prediction in Continuous-Time Dynamic Heterogeneous Graphs with Causality of Event Types. *Int. J. Crowd Sci.* 6, 2 (June 2022), 80–91.
- [352] Sandra Zillker. 2022. Designing a Method for Resource-specific Next Activity Prediction. In *PACIS 2022 Proceedings*.



## A CLASSIFICATION OF EXISTING WORK

In the following, we present the classification of existing work into our taxonomy for event prediction methods (cf. Figure 3) as additional material for the electronic supplement. The classification is valuable for orientation by practitioners and searching for related work in other application domains by researchers, but is not necessary in its entire extent (cf. Table 3 - Table 6) for the main findings of this survey.

Table 3. Time event prediction article classification with abbreviations for classes as highlighted in bold in Figure 3 and the superscript for domains as introduced in Figure 2.

Output Repr.	Basic Model	Basic Techn.	Sources
Occ	Inf	Dir	[78] <sup>E</sup> [182] <sup>E</sup> [218] <sup>I</sup> [269] <sup>N</sup> [187] <sup>E</sup> [335] <sup>I</sup> [17] <sup>E</sup> [144] <sup>E</sup> [262] <sup>B</sup> [175] <sup>B</sup> [324] <sup>B</sup> [42] <sup>E</sup> [168] <sup>B</sup> [191] <sup>B</sup> [303] <sup>B</sup> [164] <sup>B</sup> [233] <sup>B</sup> [282] <sup>B</sup> [274] <sup>B</sup> [35] <sup>B</sup> [261] <sup>B</sup> [210] <sup>B</sup> [96] <sup>B</sup> [296] <sup>B</sup> [149] <sup>B</sup> [257] <sup>B</sup> [226] <sup>B</sup> [202] <sup>B</sup> [117] <sup>B</sup> [101] <sup>B</sup> [90] <sup>B</sup> [268] <sup>B</sup> [70] <sup>B</sup> [197] <sup>B</sup> [258] <sup>B</sup> [203] <sup>B</sup> [232] <sup>B</sup> [306] <sup>B</sup> [185] <sup>B</sup> [312] <sup>B</sup> [206] <sup>B</sup> [289] <sup>B</sup> [196] <sup>B</sup> [212] <sup>B</sup> [173] <sup>B</sup> [321] <sup>E</sup> [45] <sup>M</sup> [242] <sup>H</sup> [190] <sup>H</sup> [348] <sup>E</sup> [64] <sup>B</sup> [105] <sup>P</sup> [304] <sup>P</sup> [346] <sup>N</sup>
	Hyb	Dir	[133] <sup>E</sup> [285] <sup>B</sup> [189] <sup>H</sup> [291] <sup>P</sup> [188] <sup>H</sup> [326] <sup>B</sup>
		Ind	[62] <sup>B</sup>
	Gen	Dir	[58] <sup>B</sup> [129] <sup>I</sup>
Dis	Inf	Dir	[270] <sup>I</sup> [350] <sup>I</sup> [284] <sup>I</sup> [138] <sup>N</sup> [204] <sup>I</sup> [161] <sup>E</sup> [171] <sup>I</sup> [127] <sup>I</sup> [253] <sup>E</sup> [91] <sup>H</sup> [100] <sup>I</sup> [60] <sup>C</sup> [82] <sup>I</sup> [158] <sup>N</sup> [7] <sup>T</sup> [266] <sup>I</sup> [176] <sup>H</sup> [299] <sup>E</sup> [344] <sup>E</sup> [59] <sup>E</sup> [123] <sup>E</sup> [325] <sup>B</sup> [38] <sup>B</sup> [322] <sup>M</sup> [3] <sup>H</sup> [146] <sup>N</sup> [186] <sup>H</sup> [279] <sup>E</sup> [225] <sup>I</sup>
		Ind	[305] <sup>I</sup> [115] <sup>I</sup> [46] <sup>I</sup> [13] <sup>E</sup> [112] <sup>I</sup> [27] <sup>I</sup> [227] <sup>B</sup> [251] <sup>N</sup> [280] <sup>H</sup> [49] <sup>T</sup> [145] <sup>P</sup>
	Hyb	Dir	[50] <sup>N</sup> [4] <sup>E</sup> [102] <sup>B</sup> [80] <sup>H</sup> [343] <sup>P</sup>
		Ind	[181] <sup>E</sup>
Gen	Dir	[76] <sup>E</sup>	
Con	Inf	Dir	[213] <sup>B</sup> [183] <sup>E</sup> [264] <sup>B</sup> [317] <sup>B</sup> [241] <sup>B</sup> [19] <sup>B</sup> [272] <sup>B</sup> [328] <sup>B</sup> [152] <sup>B</sup> [223] <sup>B</sup> [215] <sup>B</sup> [300] <sup>B</sup> [114] <sup>B</sup> [25] <sup>B</sup> [278] <sup>B</sup> [81] <sup>C</sup> [311] <sup>H</sup> [217] <sup>M</sup> [236] <sup>H</sup>
		Ind	[313] <sup>B</sup> [301] <sup>B</sup> [33] <sup>B</sup>
	Hyb	Dir	[332] <sup>E</sup> [21] <sup>E</sup> [10] <sup>B</sup> [295] <sup>B</sup> [1] <sup>B</sup> [55] <sup>B</sup> [238] <sup>B</sup> [40] <sup>B</sup> [125] <sup>B</sup> [87] <sup>H</sup> [103] <sup>H</sup> [67] <sup>E</sup> [315] <sup>H</sup> [244] <sup>H</sup> [320] <sup>H</sup> [293] <sup>P</sup> [88] <sup>E</sup>
		Ind	[271] <sup>H</sup> [259] <sup>B</sup> [93] <sup>B</sup>
Gen	Dir	[12] <sup>C</sup> [92] <sup>B</sup> [221] <sup>B</sup>	

Received TODO; revised XXX; accepted XXX

Table 4. Location event prediction article classification with abbreviations for classes as highlighted in bold in Figure 3 and the superscript for domains as introduced in Figure 2.

Output Repr.	Basic Model	Basic Techn.	Sources
Ras	Inf	Dir	[178] <sup>C</sup> [26] <sup>C</sup> [308] <sup>P</sup> [237] <sup>C</sup>
	Hyb	Dir	[339] <sup>P</sup> [329] <sup>C</sup>
Poi	Inf	Dir	[39] <sup>P</sup> [106] <sup>P</sup>
	Hyb	Dir	[220] <sup>M</sup>
		Ind	[341] <sup>H</sup>

Table 5. Semantic event prediction article classification with abbreviations for classes as highlighted in bold in Figure 3 and the superscript for domains as introduced in Figure 2.

Output Repr.	Basic Model	Basic Techn.	Sources
Cat	Inf	Dir	[163] <sup>B</sup> [200] <sup>B</sup> [89] <sup>B</sup> [151] <sup>M</sup> [5] <sup>B</sup> [277] <sup>M</sup> [347] <sup>M</sup> [118] <sup>B</sup> [352] <sup>B</sup> [2] <sup>E</sup> [97] <sup>I</sup> [169] <sup>H</sup> [298] <sup>I</sup> [136] <sup>M</sup> [122] <sup>M</sup>
		Ind	[53] <sup>B</sup> [240] <sup>B</sup> [231] <sup>B</sup> [201] <sup>B</sup> [74] <sup>B</sup> [32] <sup>B</sup> [156] <sup>B</sup> [194] <sup>B</sup> [134] <sup>B</sup> [243] <sup>B</sup> [234] <sup>B</sup> [147] <sup>B</sup> [66] <sup>P</sup>
	Hyb	Dir	[192] <sup>E</sup> [174] <sup>C</sup>
		Ind	[159] <sup>B</sup>
Agg	Inf	Dir	[177] <sup>M</sup> [331] <sup>M</sup> [11] <sup>M</sup> [14] <sup>M</sup> [310] <sup>M</sup> [336] <sup>M</sup> [307] <sup>M</sup> [128] <sup>M</sup> [245] <sup>P</sup>
		Ind	[73] <sup>B</sup> [179] <sup>B</sup>
	Hyb	Dir	[63] <sup>M</sup> [338] <sup>P</sup>
		Ind	[162] <sup>B</sup>
Gen	Dir	[166] <sup>M</sup>	

Table 6. Joint event prediction article classification with abbreviations for classes as highlighted in bold in Figure 3 and the superscript for domains as introduced in Figure 2.

Pred. Output	Output Repr.	Basic Model	Basic Techn.	Sources
Tim and Loc	Occ and Ras	Inf	Dir	[172] <sup>C</sup> [263] <sup>E</sup>
	Dis and Ras	Inf	Dir	[327] <sup>B</sup> [126] <sup>T</sup> [252] <sup>T</sup> [302] <sup>N</sup> [292] <sup>P</sup> [330] <sup>C</sup> [137] <sup>P</sup>
			Ind	[342] <sup>P</sup> [15] <sup>T</sup> [309] <sup>C</sup> [333] <sup>T</sup> [124] <sup>C</sup> [69] <sup>T</sup>
		Hyb	Dir	[148] <sup>T</sup>
		Hyb	Dir	[95] <sup>C</sup> [23] <sup>P</sup>
Dis and Poi	Inf	Dir	[224] <sup>N</sup>	
Con and Poi	Gen	Dir	[184] <sup>C</sup> [323] <sup>C</sup>	
Tim and Sem	Occ and Agg	Inf	Ind	[142] <sup>B</sup>
	Dis and Cat	Inf	Dir	[43] <sup>E</sup> [30] <sup>P</sup>
			Ind	[37] <sup>B</sup> [36] <sup>B</sup>
		Hyb	Ind	[77] <sup>E</sup>
	Dis and Agg	Inf	Ind	[52] <sup>B</sup> [198] <sup>B</sup>
			Hyb	Dir
	Con and Cat	Inf	Dir	[255] <sup>B</sup> [319] <sup>I</sup> [351] <sup>E</sup> [98] <sup>I</sup>
Ind			[44] <sup>B</sup> [235] <sup>B</sup>	
	Hyb	Dir	[68] <sup>H</sup>	
Con and Agg	Inf	Ind	[288] <sup>B</sup> [287] <sup>B</sup> [219] <sup>B</sup>	
		Hyb	Ind	[239] <sup>B</sup> [6] <sup>B</sup>
Tim, Loc and Sem	Occ, Ras and Cat	Inf	Dir	[79] <sup>P</sup>
	Dis, Ras and Cat	Inf	Dir	[56] <sup>P</sup>
			Ind	[349] <sup>N</sup>
		Gen	Dir	[170] <sup>H</sup>
Dis, Ras and Agg	Inf	Dir	[248] <sup>P</sup> [143] <sup>P</sup> [121] <sup>P</sup>	