Putting Decision Mining into Context: A Literature Study

Chapter · September 2020

DOI: 10.1007/978-3-030-47355-6_3

Sam Leewis
HU University of Applied Sciences Utrecht

Koen Smit
HU University of Applied Sciences Utrecht

Putting Decision Mining into Context: A Literature Study

Sam Leewis, Koen Smit, and Martijn Zoet

Abstract The value of a decision can be increased through analyzing the decision logic, and the outcomes. The more often a decision is taken, the more data becomes available about the results. More available data results into smarter decisions and increases the value the decision has for an organization. The research field addressing this problem is Decision mining. By conducting a literature study on the current state of Decision mining, we aim to discover the research gaps and where Decision mining can be improved upon. Our findings show that the concepts used in the Decision mining field and related fields are ambiguous and show overlap. Future research directions are discovered to increase the quality and maturity of Decision mining research. This could be achieved by focusing more on Decision mining research, a change is needed from a business process Decision mining approach to a decision focused approach.

Keywords Decision mining · Data mining · Process mining · Business intelligence

1 Introduction

Decisions in the modern world are often made in fast-changing, sometimes unexpected, situations [1]. Such situations require the selection of the right decision maker and supplying them with the necessary data. Decision mining solves this problem by estimating data quality and interpretation of their semantics and relevance, the actual meaning, and unit of measurement [1]. The second major advantage is the classification of decisions which allow the discovery of correspondence between decision

S. Leewis (✉) · K. Smit
HU University of Applied Sciences Utrecht, Utrecht, The Netherlands
e-mail: Sam.Leewis@hu.nl

K. Smit
e-mail: Koen.Smit@hu.nl

M. Zoet
Zuyd University of Applied Sciences, Sittard, The Netherlands
e-mail: Martijn.Zoet@zuyd.nl

R. Agrifoglio et al. (eds.), *Digital Business Transformation*, Lecture Notes in Information Systems and Organisation 38, https://doi.org/10.1007/978-3-030-47355-6_3

makers and their roles through the development of decision models and (semi) automatic decision analysis techniques [1]. Decision mining can be segmented into three types: Discovery, Conformance checking, and Improvement (similar as in Process mining research [2]. An often-used definition of Decision mining, and also referred to as decision point analysis, is *"aims at the detection of data dependencies that affect the routing of a case*" [3]. This focus leaves out other decision elements embedded in decision trees, database tables etc., such as business rules, business decision tables, or executable analytic models. Therefore, we define Decision mining as "*the method of extracting and analyzing decision logs with the aim to extract information from such decision logs for the creation of business rules, to check compliance to business rules and regulations, and to present performance information*". The system supporting and improving decision making is known as a Decision Support System (DSS). DSSs is *"the area of the information systems (IS) discipline that is focused on supporting and improving managerial decision making"* [4]. The current DSSs give insufficient insight into how decisions are executed, this is especially the case for executing multiple decisions at once (a group of decisions), thereby lacking the transparency [5, 6]. Methods of Data mining are used in Decision mining and DSSs for the following purposes: finding associative rules between decisions and the factors affecting them, user clustering using decision trees and neural networks, recognition of common users' features or interests [7–9].

To the knowledge of the authors, little research exists on the topic of Decision mining, especially that of a comprehensive literature study. Furthermore, to the knowledge of the authors, no research exists where the concepts of Decision mining, Data mining and Process mining are compared. The Decision mining field and its related fields lack the use of unambiguous concepts. Conducting a literature study on the Decision mining field and its related fields creates a clear overview of which concepts and definitions are used and if any overlap between these concepts exists. This research will focus on more than just the business process aspect when considering mining decisions compared to previous research [3, 10]. To do so, we aim to answer the following research question: "*What is the current state of the Decision mining research field?*"

The remainder of the paper is structured as follows: First, the research method that was utilized to collect and review the literature for the theoretical review. This is followed by the results of the literature review depicting Decision mining and its context, resulting in the conceptual framework. Finally, we discuss the conclusions of our research and provide a discussion about this research and the results, which is followed by possible future research directions.

2 Research Method

The goal of this study is to evaluate the current state of Decision mining and to discover possible future research directions. To achieve this, a conceptual framework is created to evaluate the current state and to identify possible future directions [11].

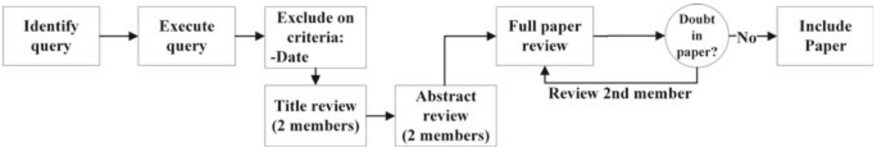

Fig. 1 Review protocol

Edmondson and Mcmanus [12] differentiate three archetypes of different levels of maturity: Nascent (e.g. Decision Mining), Intermediate (e.g. Process Mining), and Mature (e.g. Data Mining). The quality of research can be quantified with the H-index, created by Hirsch [13], where Decision mining has a H-index of 54, Process mining has a H-index of 116, and Data mining has a H-index of 305. With the rather immature state, "Nascent" [12], of the Decision mining research field and the low level of quality research (H-index of 54) a type of literature study is needed which utilizes existing empirical and conceptual studies for the creation of a conceptual framework together with possible future research directions. Therefore, a theoretical review is selected as the literature review method. Paré, Trudel, Jaana, and Kitsiou [11] state that a theoretical review draws on existing empirical and conceptual studies providing context for the identification, description, transformation into a higher theoretical structure for concepts, constructs, and their relations. The primary artefact of this type of literature study is the development of a conceptual framework with a set of research propositions [14, 15]. The contribution and value of the theoretical review lies in its ability to the development of novel conceptualizations or extend any current ones by the identification of knowledge gaps between the current knowledge and future directions [15]. To ensure quality, rigor, and transparency, the theoretical review was commenced by setting up a literature review protocol, as shown in Fig. 1. Google Scholar was used as main search database due to the fact that it has a higher coverage compared to other search engines or individual database searches [16–20].

2.1 Identify Query

Webster and Watson [15] state that Information Science is an interdisciplinary field and is based on research from other disciplines and therefore reviewing inside the IS field is not enough. To identify all concepts, constructs, and their relations concerning Decision mining, a wide range of search queries are used in order to achieve this. A difference is made between primary search queries focused on the research question (Decision Mining) and secondary search queries (Process Mining, Data Mining, Business Intelligence, Decision Management, Business Process Management, Decision Support Systems, Business Process Model and Notation, Decision Model and Notation) providing context for the primary search queries. The relations between the search queries are elaborated further in the result section.

2.2 Exclusion Criteria

Commonly used exclusion criteria for literature reviews are Date, Number of Citations, and High impact journals. Because of the low maturity and low H-index of the research field, no citation criterion is used during this study. The common use of the high impact factor journal inclusion and exclusion criterion [21–23] is not used for this specific study. The difference between maturity in research fields and the low research quality (H-index 54) is that the high impact factor journals do not publish relevant articles on Decision mining. The dissimilarity between search queries related to maturity and research quality is of that large a difference that the inclusion or exclusion based on a high impact factor source would exclude a large number of sources from the immature and lower research quality search queries. This review protocol includes only the Date criterion and is defined as everything between January 2000 and December 2018. The focus of this study lies on the current state of Decision mining and by implementing a date threshold the actual current state can be quantified in a date. Furthermore, Levy and Ellis [24] point out that by using a review protocol, the actual moment a literature review is completed becomes much clearer. An example of an element of a review protocol is a date range this clarifies which papers with the specific dates are included [25].

2.3 Paper Review

Two reviewers (R1 and R2) are involved during the abstract and paper review. R1 is a PhD-candidate with six years of practical and research experience in the field of Decision Management; R2 is a lecturer and researcher with seven years of practical and research experience in the field of Decision Management. The reviewers include or exclude the papers based on the title and based on abstract, see Fig. 3. When the two reviewers include the same paper based on the title, the paper is reviewed based on the abstract. If the situation occurs that a paper was marked with one included and one excluded, the two reviewers discuss the reviewing of the title. The same process of reviewing also applies to the abstract review. R1 decides during the review of the full paper on the relevancy of the paper. When any doubt on the relevancy exists by R1, R2 decides if the paper is excluded or included.

3 Search Results

The executed search query resulted in 810 potential useable articles, see Fig. 2. Duplicates and non-English papers were excluded, resulting in 74 papers being left out. After title reviews 656 papers were excluded, 80 papers were included. Reviewing the abstracts resulted in the exclusion of 24 papers, resulting in 56 papers that were

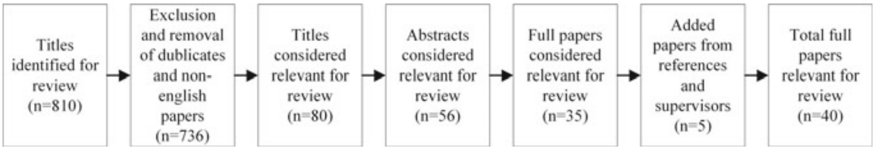

Fig. 2 Search results

included. After reviewing the full papers, 35 papers remained. Based on references and feedback from the supervisors 5 articles were included.

4 Results

The review process resulted in a longlist of papers distributed over the primary and secondary research queries. The identified concepts and their relationships are the pillars of the conceptual framework and this paper. The concepts and the relationships to each other are shown as a metamodel in Fig. 3. The following sections will create context to capture the current state of Decision mining.

4.1 Business Intelligence, Business Process Management and Decision Management

Business intelligence (BI) is described as a set of models and analysis methodologies that utilizes the available data to generate information and knowledge for the support of decision-making processes [26–28]. BI can be divided into four phases [29]: (1) identification of information needs, (2) information acquisition, (3) information analysis, and (4) storage and information utilization. Focusing more on information and data, BI can be defined as a subset of the two (information and data). Where data are facts or recorded measures of certain phenomena, information is structured

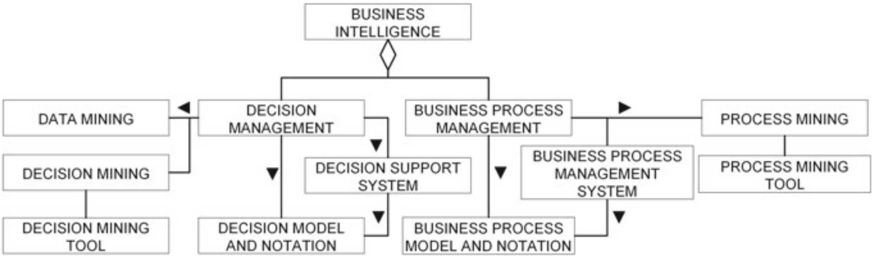

Fig. 3 Search query metamodel

data for the support of the decision-making process or to define any relationships between two facts [27]. Business process management (BPM) [30] and Decision management (DM) are examples of approaches utilizing available data to generate information and knowledge for the support of decision-making processes.

BPM is "*a collaboration of concepts, methods, and techniques to support the design, administration, configuration, enactment, and analysis of business processes*" [30]. The base of BPM is the explicit representation of business process containing activities with the execution constraints between these activities. After defining the business processes, they can be analyzed, improved, and enacted [30]. The industry standard for defining a business process is the Business Process Model and Notation (BPMN) [31].

DM is "*the practice of managing smart, agile decisions*" [32]. Decisions are amongst an organization most important assets [33] Therefore, adequately managing these decisions is vital. A decision is "*A conclusion that a business arrives at through business logic and which the business is interested in managing*" [34] and business logic is "*a collection of business rules, business decision tables, or executable analytic models to make individual business decisions*" [35]. Modelling the decisions and business logic to make them explicit for further analysis is an aspect of DM. Decision Model and Notation (DMN) [34] is used as an industry standard for the modelling of decisions.

4.2 Decision Support Systems

The first decision support systems (DSS) were developed for the support of decision makers by analyzing data using predefined models [36]. Due to recent advancements in technology and market demand, the DSSs are improved to serve future needs, implementing, i.e., analytical tools [37]. The current DSSs lack the transparency into how decisions are executed, especially for a group of decisions [5, 6]. DSSs play a major role in IS. Looking at the importance of DSS research, 12% of the articles published in IS journals are focused on DSSs [38] and citation-based analysis were DSSs is one of the three core subfields of IS [39].

4.3 Decision Mining, Process Mining, and Data Mining

The literature review showed that research published in the fields of Data mining, Process mining, and Decision mining have an overlap in used concepts The following meta models [40] are used to ground the literature of the Data mining, Process mining, and Decision mining fields.

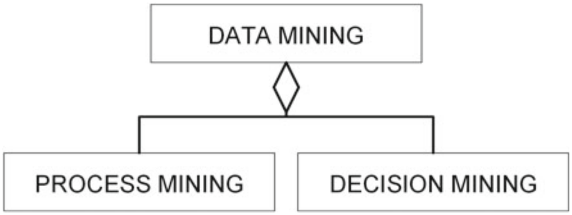

Fig. 4 Data mining literature relations

The Data mining field is consentient in its direction, the literature showed, on the other hand, no specifications towards Process mining and Decision mining, identifying these fields as an aggregation of Data mining focused on sequence (Process mining) and derivations (Decision mining), as shown in Fig. 4 [41–44].

The Process mining field is comparable to the Data mining field as it is consentient in its direction. Process mining utilizes Data mining techniques for mining of events from event logs for process discovery, conformance checking, and process improvement, as shown in Fig. 5 [2]. The Process mining field identifies Decision mining as a case perspective focus in Process mining, also known as decision point analysis, as shown in Fig. 5 [2, 3].

The Decision mining field has two main influences 1) Decision mining focused on mining Decision Points from business processes (decision-annotated) [3] and (2) a Decision mining approach where more implicit data involved in the decision-making process (decision-aware) is taken into account [10, 45]. Both directions have overlap and are a type of Decision mining with a Process mining focus and the utilization of Data mining techniques, as shown in Fig. 6.

Data Mining is defined as "*the automated or convenient extraction of patterns representing knowledge implicitly stored or captured in large databases, data warehouses, the Web, other massive information repositories, or data streams*" [41].

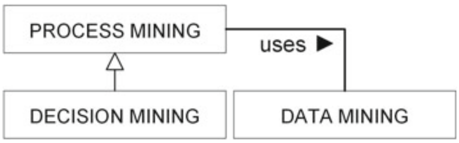

Fig. 5 Process mining literature relations

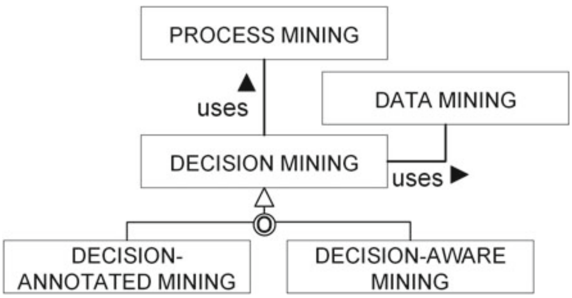

Fig. 6 Decision mining literature relations

Therefore, the term Data mining is a misnomer, as the goal of Data mining is the extraction of patterns of aggregations from large volumes of data, not the extraction of said data [41]. Data mining utilizes techniques for finding data patterns, these patterns enables accurate and fast decision making or provide new insights [42]. Data patterns could contain a specific sequence (processes), or a specific derivation (decisions). An example where Data mining can be performed is whether a specified class of customers will buy a combination of products, a business can utilize this data by predicting whether to increase the price on one of the articles or include an article in a sale, also known as a basket analysis [46]. Prediction, together with description, make up for the two Data mining categories [43]: (1) Predictive Data mining, which utilizes variables of existing data to predict any unknown or future values of other variables of value and (2) Descriptive Data mining, focusses on the discovery of patterns non trivially describing data which in turn can be interpreted by interested parties. Predictive and descriptive Data mining can be achieved by utilizing Data mining algorithms for the following Data mining classifications [43]: (1) Classification, a classification which discovers a predictive learning function that generalizes the known structure for the application of new data items. (2) Regression, a classification where a predictive learning function models the data, for the estimation of relations between data items. (3) Clustering, a classification where a descriptive task identifies a set of categories or clusters for the description of data. (4) Summarization, a classification where a descriptive task provides a compact representation of the data. (5) Dependency Modelling, a classification which searches for relations between multiple variables. (6) Change and Deviation Detection, a classification where changes or unusual records are discovered. The algorithms that are widely used for Data mining are C4.5 [47], *k*-means algorithm [48], Support vector machines [49], the Apriori algorithm [50], and CART [51], for a more extensive list see the work of Wu et al. [52]. The Data mining algorithms look for structural descriptions, these descriptions can become rather complex and are typically expressed as a set of rules or a decision tree. Rules and decision trees are easier understood by people and explain what has been learned, thereby serving as a base for future predictions [42]. Besides Data mining being in a mature state, major issues can be identified [41].

Process Mining is defined as "*the discovery, monitoring and improvement of real processes by extracting knowledge from event logs readily available in today's information systems*" [2], i.e., the mining of sequencing patterns. This technique is frequently used in BPM [30, 53]. Process Mining can be considered a twofold, on one side an analysis technique for Data mining and on the other side a technique for process modelling. The three types of Process mining are [2]; (1) process discovery, (2) conformance checking, and (3) process improvement. Process Discovery constructs a representation of an organization's business process and any major variations. Event logs are used as input to set up the process models [2] and serve as the starting point of Process mining [54]. Information systems utilized by organizations store detailed information about the specific sequence of activities performed by these information systems during the execution of a process [53]. The events in the event logs symbolize each one activity which in turn is part of a business process. Detailed information on events is stored in these event logs concerning: e.g. finishing time of an activity. The algorithms that are widely used for process discovery include Alpha Miner, Heuristic Miner, Trace Clustering, and Fuzzy Miner [54, 55]. During conformance checking, the discovered and created process model is analyzed and it is checked on any discrepancy between the event logs and the process model [2, 56]. The main purpose of conformance checking is to identify any problem areas which can be improved by utilizing this knowledge. The modification of the process model to comply with the event logs is done during process improvement. Process improvement aims to extend or improve the process model using information from the event logs about the actual process [2]. Two improvement types exists (1) repair, the modification of the process model to better reflect the actual process [57], and (2) extension, adding a new perspective after cross-referencing this with the event log [2]. In relation to the three types of Process mining, four perspectives are characterized [2] control-flow, organizational, case, and time. The four perspectives are all related to each other. The control-flow perspective focuses on the ordering of activities. The organizational perspective focuses on resource information hidden in the event logs e.g. involved actors, and their relations. The case perspective focuses on case properties e.g. characterization of a case by the process path. The time perspective focuses on the time and frequency on events e.g. prediction of processing time utilizing timestamps.

An often-used definition of Decision mining, and also referred to as decision point analysis, is "*aims at the detection of data dependencies that affect the routing of a case*" [3], i.e., the mining of derivation patterns. One of the major tasks of Decision mining is the estimation of data quality and interpretation of their semantics, interpretation of data whether it is relevant, the actual meaning, and unit of measurement [1]. The second major task is the classification of decisions which allow the discovery of correspondence between decision makers and their roles through the development of decision models and (semi) automatic decision analysis techniques [1]. The work of Rozinat and van der Aalst [3] assumes the mining of decisions as the mining of decision points from a business process. This method is comparable to the three types of Process mining [2]: discovery, conformance checking, and improvement. This approach uses a version of the C4.5 algorithm [47] for the construction of decision trees which allow the analysis of the choice in the decision points of a process.

Rozinat and van der Aalst [3] proposes the use of Petri Nets [58] for this approach, thereby being able to determine the specific points where a choice is made and which branches are followed. After the identification of a decision point, the authors try to determine if certain cases with certain properties follow each a specific route. However, this approach has many limitations. For example, this approach cannot deal with event logs containing deviating behavior and with more complex control-flow constructs [59]. Many variations on the decision point analysis variant of Decision mining have been published, mostly focusing on refining the way to retrieve the decision information [60–62].

Another approach of Decision mining is the focus on implicit knowledge of the decision maker. The actions of the decision maker and the identification of the decision-making strategy, and determining what data and information is used, how this is used, and what knowledge is employed by the decision maker for the choice between two alternatives [63]. The implicit knowledge is captured using the Decision Model and Notation standard (DMN) [34], together with the clear intention to use it in the context of business processes using Business Process Model and Notation (BPMN) [31]. Hybrids also surfaced such as Product Data Model (PDM) [64]. Showing the influence that context plays on decision making is an important task because [1]: (1) finds and models typical scenarios of interactions between users, (2) discover reoccurring situations within large volumes of raw data, and (3) cluster decision makers, allowing the reduction of supported user models and increasing data presentation. The current Decision mining technique has a limited applicability and is only usable for workflows that uses representations in Petri Nets. This current approach utilizes control flow-based discovery techniques which lack in-depth analysis of the specific action that is performed. Future Decision mining techniques should be driven by the construction of a decision model than by a control flow containing decision points [10, 65]. The Data mining techniques used for Decision mining, which do focus on temporal aspects, often leave out the importance of decisions being the driver of the analyzed result. More can be gained to know which attributes are important to a decision. Furthermore, it is paramount to understand the importance of the decision-making process. The use of Data mining techniques can be improved by implementing decision models [10].

Organizations using algorithms in decision-making processes have to adhere to European legislation which stipulates that automated decision making, without the interaction of a person, is in principle not allowed. However, an exception is made for automated decision making based on a legal basis, provided that it provides for appropriate measures to protect the rights and freedoms and legitimate interests of the data subject [66]. Therefore, the importance of transparent algorithms is clear. Another example of mandatory transparency is the safety critical systems in autonomous cars [67] which need to adhere to the ISO 26262 standard [68].

Furthermore, by reviewing the literature on Data mining, Process mining, and Decision mining, the segregation of these concepts seems incorrect. Reviewing the focus of each concept Data mining is data pattern focused, Process mining is event log focused, and Decision mining can be depicted into two directions, calling it a decision point analysis focused on annotating decisions, or focused on being decision-aware. Therefore, Data mining seems a more general concept where Process mining and Decision mining are included as a type of Data mining, as shown in Fig. 4.

5 Conceptual Framework

The conceptual framework depicts Decision mining and its related concepts, as shown with annotations[x] in Fig. 7, and is the result of the theoretical review. *Data mining*[1], *Process mining*[2], and *Decision mining*[3] are identified as elements of BI, were BI is described as a set of models and analysis methodologies that utilizes the available

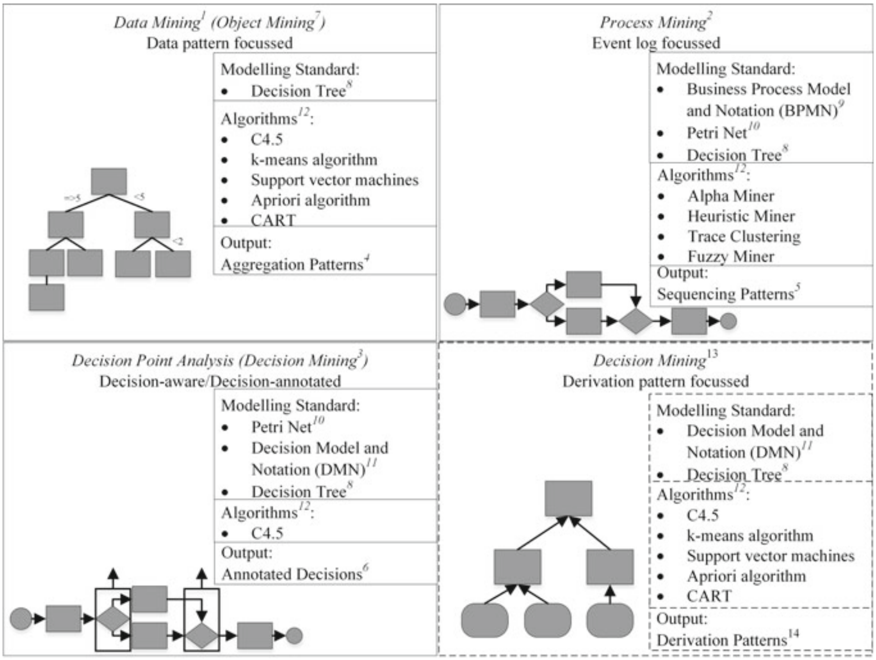

Fig. 7 Conceptual framework

data to generate information and knowledge for the support of decision-making processes [26–28]. Data mining is the mining of *aggregation patterns*[4], Process mining is the mining of *sequencing patterns*[5], and Decision point analysis is utilizing *annotated decisions*[6], thereby utilizing data to generate information and knowledge for the support of decision-making processes. Additionally, reviewing the literature of the context of Decision mining showed that Data mining is not the mining of data, a so-called misnomer [41]. Changing this to *Object mining*[7] would be more appropriate. Specifying and designing focused on data, processes, or decisions each has its standards, where *decision trees*[8] are widely used, and for the more specific models (processes and decisions) *BPMN*[9] [31], *Petri Nets*[10] [69], and *DMN*[11] [34] are used. Each mining technology uses each their own set of *algorithms*[12] to identify patterns based on statistical analysis.

6 Conclusion

The goal of this research is to review the current state of Decision mining. To do so, the following research question was addressed: "*What is the current state of the Decision mining research field?*" In order to answer this question, we conducted a theoretical review consisting of 810 papers which were cut down to a total of 40 papers which were deemed relevant in covering the topic of Decision mining. This study depicted the term Decision mining and where Decision mining falls under (BI, BPM, and DM) or is directly related to (Data mining/Object mining and Process mining). The theoretical review resulted into a conceptual framework were all the discovered concepts were depicted. The conceptual framework revealed several gaps in the existing research. Therefore, several research directions could be considered for researchers to target and extend this scarcely covered topic. Additionally, the conceptual framework revealed the inconsistency in the involved research fields covering the same subjects and using the same concepts but not using a clear definition, or the same name.

7 Discussion and Future Research

The theoretical review resulted into a conceptual framework were all the discovered concepts are depicted. The conceptual framework revealed several gaps in the existing research Therefore, several research directions could be considered for researchers to target.

The main future research focus should be directed to improve the level of research maturity of the Decision mining research field. Therefore, the level of maturity should mature out of the "Nascent" phase using the classification of research domain maturity of Edmondson and Mcmanus [12]. Other authors support the notion of a more direct focus for improving the maturity of the Decision mining research field [10, 65].

In relation to the previous research focus, the current research on *Decision mining*[3] is focused on the perspective from a business process point of view, and a more standalone viewpoint is needed where *Decision mining*[13] is focusing on the decision viewpoint with *derivation patterns*[14] as output, as shown in Fig. 7. The two main Decision mining influences lack the capacity to deal with event logs containing deviating behavior and dealing with more complex control-flow constructs [59] and lacks a holistic overview of the decision model [10]. Recent research focusses on a more holistic discovery of decisions [70], which still has a business process point of view with event logs as its data input. Additionally, with the focus on the viewpoint shift, the data input for Decision mining should shift from an event log to a decision log [10], as shown in Fig. 8. Supporting the previous notion the following Decision mining definition is created to cover the new research directions: *"the method of extracting and analysing decision logs with the aim to extract information from such decision logs for the creation of business rules, to check compliance to business rules and regulations, and to present performance information*".

The third focus for future research is conducting research on algorithms in the Decision mining field. Current Decision mining algorithm research is focused on the assumption that the algorithms in ProM [3] are the best fit for mining decisions. Future research should focus on creating an overview of useable algorithms for Decision mining, and if needed, create a new algorithm if existing algorithms lack the capacity for mining decisions. Recent technologies as Reinforcement learning

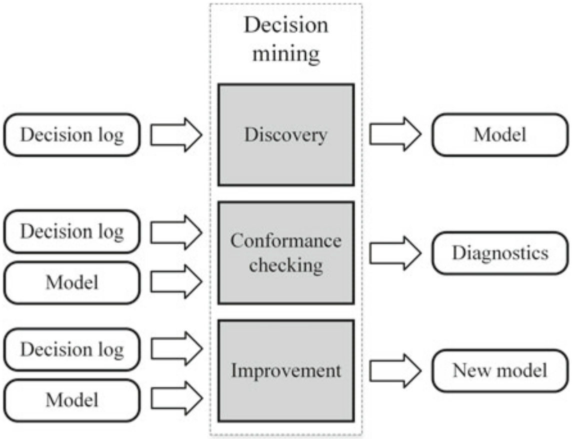

Fig. 8 Decision mining activities

[71, 72] and Deep learning [73, 74] algorithms could create added value for Decision mining. Lastly, the previous mentioned future research should be focused on the elements centered around the activities of Decision mining: Discovery, Conformance checking, and Improvement, as shown in Fig. 8.

References

1. Smirnov, A., Pashkin, M., Levashova, T., Kashevnik, A., Shilov, N. (2009). Context-driven decision mining. In *Encyclopedia of data warehousing and mining* (pp. 320–327). Information Science Reference, Hershey, NY.
2. van der Aalst, W. M. P.: *Process mining: Discovery, conformance and enhancement of business processes*. Springer Science & Business Media (2011).
3. Rozinat, A., van der Aalst, W. M. P.: Decision mining in ProM. In: Dustdar, S., Fiadeiro, J. L., & Sheth, A. P. (Eds.), *Business process management: 4th international conference, BPM 2006, Vienna, Austria*, 5–7 September 2006. Proceedings (pp. 420–425). Heidelberg, Berlin: Springer. https://doi.org/10.1007/11841760_33.
4. Arnott, D., & Pervan, G. (2005). A critical analysis of decision support systems research. *Journal of Information Technology, 20,* 67–87.
5. Horita, F. E. A., de Albuquerque, J. P., Marchezini, V., & Mendiondo, E. M. (2017). Bridging the gap between decision-making and emerging big data sources: An application of a model-based framework to disaster management in Brazil. *Decision Support Systems, 97,* 12–22. https://doi.org/10.1016/j.dss.2017.03.001.
6. Mohemad, R., Hamdan, A. R., Othman, Z. A., & Noor, N. M. M. (2010). Decision support systems (DSS) in construction tendering processes. *International Journal of Computer Science Issues, 7,* 35–45. https://doi.org/10.1109/ICSSSM.2008.4598482.
7. Chiang, W. Y. K., Zhang, D., & Zhou, L. (2006). Predicting and explaining patronage behavior toward web and traditional stores using neural networks: A comparative analysis with logistic regression. *Decision Support Systems, 41,* 514–531. https://doi.org/10.1016/j.dss.2004.08.016.
8. Li, X.-B. (2005). A scalable decision tree system and its application in pattern recognition and intrusion detection. *Decision Support Systems, 41,* 112–130. https://doi.org/10.1016/j.dss.2004.06.016.

9. Thomassey, S., & Fiordaliso, A. (2006). A hybrid sales forecasting system based on clustering and decision trees. *Decision Support Systems, 42,* 408–421. https://doi.org/10.1016/j.dss.2005.01.008.
10. De Smedt, J., Vanden Broucke, S. K. L. M., Obregon, J., Kim, A., Jung, J. Y., Vanthienen, J. (2017). Decision mining in a broader context: An overview of the current landscape and future directions. In *Lecture notes in business information processing* (pp. 197–207). Springer International Publishing. https://doi.org/10.1007/978-3-319-58457-7_15.
11. Paré, G., Trudel, M. C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: a typology of literature reviews. *Information & Management, 52,* 183–199. https://doi.org/10.1016/j.im.2014.08.008.
12. Edmondson, A. C., & Mcmanus, S. E. (2007). Methodological fit in management field research. *Academy of Management Review, 32,* 1155–1179.
13. Hirsch, J. E. (2010). An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship. *Scientometrics, 85,* 741–754. https://doi.org/10.1007/s11192-010-0193-9.
14. Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology, 1,* 311–320. https://doi.org/10.1037/1089-2680.1.3.311.
15. Webster, J., Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly, 26*, xiii–xxiii. https://doi.org/10.1.1.104.6570.
16. Harzing, A. -W., Alakangas, S. (2016). Google scholar, scopus and the web of science: A longitudinal and cross-disciplinary comparison. *Scientometrics 106*, 787–804 (2016). https://doi.org/10.1007/s11192-015-1798-9.
17. Wildgaard, L. (2015). A comparison of 17 author-level bibliometric indicators for researchers in astronomy, environmental science, philosophy and public health in web of science and google scholar. *Scientometrics, 104,* 873–906. https://doi.org/10.1007/s11192-015-1608-4.
18. Franceschet, M. (2010). A comparison of bibliometric indicators for computer science scholars and journals on Web of Science and Google Scholar. *Scientometrics, 83,* 243–258. https://doi.org/10.1007/s11192-009-0021-2.
19. Amara, N., & Landry, R. (2012). Counting citations in the field of business and management: Why use Google scholar rather than the web of science. *Scientometrics, 93,* 553–581. https://doi.org/10.1007/s11192-012-0729-2.
20. Gehanno, J.-F., Rollin, L., & Darmoni, S. (2013). Is the coverage of google scholar enough to be used alone for systematic reviews. *BMC medical informatics and decision making, 13,* 7. https://doi.org/10.1186/1472-6947-13-7.
21. Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature, 421,* 37–42. https://doi.org/10.1038/nature01286.
22. Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature, 536,* 285–291. https://doi.org/10.1038/nature19057.
23. Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., et al. (2012). Biodiversity loss and its impact on humanity. *Nature, 486,* 59–67. https://doi.org/10.1038/nature11148.
24. Levy, Y., Ellis, T.J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science 9*, 181–211. https://doi.org/10.28945/479.
25. Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems, 37,* 879–910. https://doi.org/10.2139/ssrn.1954824.
26. Vercellis, C.: *Business intelligence: Data mining and optimization for decision making*. New York: Wiley.
27. Zikmund, W., Babin, B., Carr, J., Griffin, M. (2009). *Business research methods.* South-Western.
28. Chen, H., Chiang, R. H. L., Storey, V. C.: Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 1165–1188. https://doi.org/10.2307/41703503.

29. Loennqvist, A., & Pirttim, V. (2006). The measurement of business intelligence. *Information Systems Management, 23,* 32–40. https://doi.org/10.1080/07366980903446611.
30. Weske, M. (2012). *Business process management*. Berlin: Springer. https://doi.org/10.1007/978-3-642-28616-2.
31. Object Management Group (OMG) (2011) Business process model and notation (BPMN) Version 2.0. *Business, 50*, 508. https://doi.org/10.1007/s11576-008-0096-z.
32. Von Halle, B., & Goldberg, L. (2009). *The decision model: A business logic framework linking business and technology*. New York, NY: Taylor and Francis Group, LLC.
33. Blenko, M.W., Mankins, M.C., Rogers, P.: The Decision-Driven Organization. Harv. Bus. Rev. 10 (2010).
34. Object Management Group. (2016). *Decision model and notation.*
35. Object Management Group. (2016). *ArchiMate® 3.0 specification.*
36. Shim, J., Merrill, W., Courtney, J., Power, D., Sharda, R., & Carlsson, C. (2002). Past, present and future of decision support system. *Decision Support Systems, 33,* 111–126.
37. Chugh, R., & Grandhi, S. (2013). Why Business Intelligence? Significance of business intelligence tools and integrating BI governance with corporate governance. *International Journal of Entrepreneurship and Innovation, 4,* 1–14. https://doi.org/10.4018/ijeei.2013040101.
38. Arnott, D., & Pervan, G. (2014). A critical analysis of decision support systems research revisited: The rise of design science. *Journal of Information Technology, 29,* 269–293. https://doi.org/10.1057/jit.2014.16.
39. Taylor, Dillon, & Wingen, Van. (2010). Focus and diversity in information systems research: Meeting the dual demands of a healthy applied discipline. *MIS Quarterly, 34,* 647. https://doi.org/10.2307/25750699.
40. van de Weerd, I., Brinkkemper, S. (2008). Meta-modeling for situational analysis and design methods. In: Handbook of research on modern systems analysis and design technologies and applications (vol. 35).
41. Han, J., Kamber, M., Pei, J. (2011). *Data mining: Concepts and techniques*. Morgan Kaufmann Publishers, Burlington, MA. https://doi.org/10.1016/B978-0-12-381479-1.00001-0.
42. Witten, I. H., Frank, E., Hall, M. A., & Pal, C. (2016). *Data mining: Practical machine learning tools and techniques*. Burlington, MA: Morgan Kaufmann Publishers.
43. Kantardzic, M. (2011). *Data mining : Concepts, models, methods, and algorithms*. Wiley Online Library.
44. Rokach, L., Maimon, O. (2015). *Data mining with decision trees: Theory and application.*
45. Petrusel, R., Vanderfeesten, I., Dolean, C. C., Mican, D. (2011). Making decision process knowledge explicit using the decision data model. In: Business Information Systems, 340.
46. Berry, M. J. A., Linoff, G. S. (2004). *Data mining techniques: for marketing, sales, and customer support*. New York: Wiley.
47. Ross, Q. J. (1993).*C4.5: Programs for machine learn*. Morgan Kaufmann Publishers.
48. Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE transactions on information theory, 28,* 129–137. https://doi.org/10.1109/TIT.1982.1056489.
49. Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York: Springer.
50. Agrawal, R., Srikant, R. (1994). Fast algorithms for mining association rules in large databases. In *Proceedings of the 20th International Conference on Very Large Data Bases* (pp. 487–499). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
51. Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees. *Chapman & Hall/CRC*. https://doi.org/10.1201/9781315139470.
52. Wu, X., Kumar, V., Ross, Q. J., Ghosh, J., Yang, Q., Motoda, H., et al. (2008). Top 10 algorithms in data mining. In: *Knowledge and information systems* (pp. 1–37). https://doi.org/10.1007/s10115-007-0114-2.
53. Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2018). *Fundamentals of business process management*. Berlin: Springer.
54. van der Aalst, W. M. P., Adriansyah, A., De Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., et al. (2012). *Process mining manifesto. Lecture notes in business information processing*. 99 LNBIP (pp. 169–194). https://doi.org/10.1007/978-3-642-28108-2_19.

55. Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D. (2016). Process mining in healthcare: A literature review. *Journal of Biomedical Informatics, 61*, 224–236. https://doi.org/10.1016/j.jbi.2016.04.007.
56. Rozinat, A., & van der Aalst, W. M. P. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems, 33,* 64–95. https://doi.org/10.1016/j.is.2007.07.001.
57. Rovani, M., Maggi, F. M., de Leoni, M., & van der Aalst, W. M. P. (2015). Declarative process mining in healthcare. *Expert Systems with Applications, 42,* 9236–9251. https://doi.org/10.1016/j.eswa.2015.07.040.
58. Petri, C. A. (1966). Communication with Automata. *Application Data Research, 15*, 357–62. https://doi.org/AD0630125.
59. de Leoni, M., van der Aalst, W. M. P. (2013). Data-aware process mining: Discovering decisions in processes using alignments. In: *Proceedings of the 28th Annual ACM Symposium on Applied Computing* (pp. 1454–1461). dl.acm.org.
60. Kim, A., Obregon, J., Jung, J.-Y. (2014). Constructing decision trees from process logs for performer recommendation. In *International Conference on Business Process Management* (pp. 224–236) (2014). https://doi.org/10.1007/978-3-319-06257-0_18.
61. Mannhardt, F., de Leoni, M., Reijers, H. A., van der Aalst, W. M. P.: Decision mining revisited-discovering overlapping rules. In *International Conference on Advanced Information Systems Engineering* (pp. 377–392).
62. de Leoni, M., Dumas, M., Garcķa-Bańuelos, L.: Discovering branching conditions from business process execution logs. In *International Conference on Fundamental Approaches to Software Engineering* (pp. 114–129). Berlin: Springer. https://doi.org/10.1007/978-3-642-37057-1.
63. Petrusel, R. (2010). *Decision mining and modeling in a virtual collaborative decision environment*. Rijeka: In-Tech.
64. Vanderfeesten, I., Reijers, H. A., van der Aalst, W. M. P. (2008). Product based workflow support: A recommendation service for dynamic workflow execution. In *Proceedings of 20th International Conference on Advance Information System Engineering* (pp. 571–574). https://doi.org/10.1007/978-3-540-69534-9_42.
65. Sarno, R., Sari, P. L. I., Ginardi, H., Sunaryono, D., Mukhlash, I. (2013). Decision mining for multi choice workflow patterns. In *Proceeding—2013 International Conference on Computer, Control, Informatics and Its Applications. Recent Challenges Computer, Control Informatics, IC3INA* (pp. 337–342). https://doi.org/10.1109/IC3INA.2013.6819197.
66. European Union. (2016). *General data protection regulation. Official Journal of European Union, L119*, 1–88 (2016).
67. Borg, M., Englund, C., Durán, B. (2017). Traceability and deep learning—safety-critical systems with traces ending in deep neural networks. In: Grand challenges of traceability: The next ten years (pp. 48–49). https://arxiv.org/abs/1710.03129.
68. ISO. (2018). ISO 26262-2:2018.
69. Petri, C. A. (1962). Kommunikation mit Automaten.
70. De Smedt, J., Hasić, F., vanden Broucke, S. K. L. M., Vanthienen, J. (2017). Towards a holistic discovery of decisions in process-aware information systems. In *International Conference on Business Processing Management* (pp. 183–199). https://doi.org/10.1007/978-3-319-65000-5_11.
71. Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., et al. (2017). Mastering the game of go without human knowledge. *Nature, 550,* 354.
72. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction.* Cambridge: MIT Press.
73. De Fauw, J., Ledsam, J. R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., et al. (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine*. https://doi.org/10.1038/s41591-018-0107-6.
74. Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521,* 436–444. https://doi.org/10.1038/nature14539.