# Evaluation of Explainable Artificial Intelligence for Predictive Process Mining in Education

DIJANA OREŠKI, MARIJA POKOS LUKINEC
Faculty of Organization and Informatics,
University of Zagreb,
Pavlinska 2, Varaždin,
CROATIA

*Abstract:* - Process mining leverages event log data to extract valuable knowledge and insights about the underlying processes. Education has embraced process mining, driven by the huge amounts of log data from student activities at the learning management systems (LMS) to enhance processes underlying the event logs of LMS. Educational predictive process mining supports predictions about the future of a running process instance. Predictive efforts are driven by machine learning (ML) and deep learning (DL) approaches. ML and DL approaches are characterized by a high level of efficiency and accuracy in prediction, but also increasing complexity and a low level of explainability. To overcome low explainability, various explainable artificial intelligence (XAI) methods emerged to explain the reasoning process. This study focuses on enhancing explainability in process outcome prediction by examining the properties of interpretability and faithfulness. We evaluate these properties across the primary dimensions of business process data: event attributes, case characteristics, and control flow patterns. Moodle events logs along with various ML and DL algorithms are used to validate the findings. The experiment is conducted to identify which xAI approach is best for educational predictive process mining. This is achieved through the application of key metrics: parsimony, functional complexity, importance ranking correlation, and level of disagreement. These metrics provide a structured approach to evaluating and enhancing the interpretability of predictive models in process mining. Research results in the form of guidelines assist practitioners and researchers in navigating the complex decision-making process by emphasizing the significance of explainability.

*Key-Words:* - explainable artificial intelligence, educational process mining, machine learning, deep learning, interpretability, faithfulness.

Received: March 9, 2024. Revised: December 11, 2024. Accepted: February 8, 2025. Published: April 1, 2025.

## 1 Introduction

Process mining is a discipline that enables the extraction of process-centric knowledge from time-ordered event logs collected by systems in a broad spectrum of domains. Process mining aims to discover process models, predict future activities, and recommend actions to mitigate possible risks. Process mining has been adopted across various fields, such as healthcare, finance, and education, [1]. Education has embraced process mining, driven by the huge amounts of log data from student activities at the learning management systems (LMS) and the growing interest in learning analytics and educational data mining. This has established educational process mining as a valuable tool in educational research.

In the educational process mining, event data reflects students' activities in LMSs which record time-stamped events of students' interactions with the tools. Educational process mining, among others, has been used to explore students' learning processes, and visualize students' paths and strategies, [2]. Such an approach allows us identification of most effective resources or challenging topics, or prediction of problems before they occur. Prediction is enabled by using predictive process mining (PPM), a use case of process mining, used for the providing future course of running business process instances. Predictive process mining experiments resulting from predictive models are developed by machine learning (ML) or deep learning (DL) approaches which provide superior results to alternative advanced statistical approaches, [3]. Most of the models that have been built with DL and ML are black boxes because their underlying structures are complex and difficult to explanation and interpretation, [4]. Various approaches have been proposed in ML and DL to explain obtained models, forming a new research area of explainable artificial intelligence (XAI), [5]. There are only a few studies so far that have attempted to apply XAI to interpret black box models in the process mining field. According to our knowledge, there is no study to apply or evaluate XAI methods for educational predictive process mining. Motivated by that, in this research paper, we apply and evaluate XAI measures for assessing explainable models based on ML and DL algorithms in the context of predictive educational process mining.

Accordingly, this paper aims to provide:

(i) Perceptions of explainable AI approaches used in machine learning and deep learning and evaluation metrics for explainability which are suitable for process mining,
(ii) comparative analysis of XAI metrics to explain and interpret results in educational predictive process mining,
(iii) evaluate machine learning and deep learning approaches for predictive process mining in education in terms of explainability and predictive performance.

Two research questions are set up:
RQ1: Which machine learning or deep learning algorithms give better results in terms of predictive performance in educational predictive process mining?

RQ2: Which machine learning or deep learning algorithms give better results in terms of explainability in educational predictive process mining?

The paper is structured as follows. Section two gives an overview of educational process mining and XAI, including state-of-the-art for explainable predictive process mining, indicating the gaps in the related literature that this research attempts to address. Section three explains the research methodology including XAI metrics, and approaches to their evaluation along with a description of the LMS data used in the empirical research. Section four provides findings drawn from the experiments. Section five concludes the paper by highlighting scientific contribution of the research and indicating guidelines for future research.

## 2 Related Work

This paper follows two main research paths: educational predictive process mining, as a use case of process mining in education, and explainable artificial intelligence.

Predictive process predictions refer to different tasks depending on the target to be predicted: continuous (regression task) or categorical (classification task). Thus, related work can be found for regression predictions of remaining cycle times, delays, and next timestamps [6], predictions of partial or final process outcomes [7], [8], anomaly classification [9], predictions regarding the next event [10] or a sequence of next events [11].

Recent studies also have addressed the predictive process prediction focusing on classification problems such as: (i) next event prediction, [12], [13], [14], [15]. (ii) business process outcome prediction, [16], [17]. (iii) remaining time prediction [15], and (iv) risk prediction [18]. From the perspective of predictive process models, deep learning approaches provide superior results to alternative statistical approaches [3]. Machine learning and deep learning techniques are powerful tools to solve a wide diversity of complex problems in various domains, including education, [19], [20]. Various deep learning approaches such as long-short term memory networks (LSTM), deep feedforward neural networks (FNN), or convolutional neural networks (CNN) have been successfully used for classification and regression tasks in predictive process modeling. While these advanced models yield more accurate results than traditional white-box methods, their implementation in process mining applications is hindered by their black box nature. The challenge in trusting these non-transparent models has created a gap between scientific research and its limited practical adoption. In recent years, there has been a growing focus on enhancing the interpretability of deep learning models across various research fields through various explanation techniques.

Although explaining deep learning-based black box approaches has received attention in many domains, there are only a few examples of an explainable artificial intelligence application in the domain of process mining, [21]. The study by [22] illustrated the need for explainable artificial intelligence in a manufacturing process mining. Authors in [23] also emphasized the importance of explainable artificial intelligence for business process management. In research [21] authors consider there is still much work on explainable predictive process mining and this gap needs to be filled, [21].

Explainable artificial intelligence strives to develop transparency of artificial intelligence-based models by ensuring insights into automated decision-making processes. In the field of XAI, few works have evaluated explanation models. In research [24] is presented a metric based on human-friendly properties, while in [25] authors defined a metric for model complexity. However, these metrics do not consider the process attributes. The study of [26] explores explainability models in the field of predictive process mining, but process data are not taken into account. Most importantly, these metrics did not evaluate the faithfulness of the model. Previous works emphasized challenges with the faithfulness of methods. In [27] authors demonstrated the relationship between the SHAP and the predictive performance and authors in [28] presented metrics that evaluated the relevance of the attribute value. However, research papers exploring the faithfulness of explainability in process mining are missing, [29], [30].

Previous papers identified metrics to evaluate the methods for explainability. Statistical approaches, such as linear regression, develop more transparent inferences than complex techniques, such as artificial neural networks, as they are intrinsically self-interpretable, [31]. However, simple approaches are mostly less accurate than complex ML and DL approaches. On the other hand, the interpretability and explainability of these models are low, [32]. Existing XAI metrics do not consider the different process-based characteristics: control flow, event, and case perspective. In this paper, explainability is defined through the interpretability of explanations and faithfulness of the explainability model suitable for process data.

Based on the aforementioned, there is a clear need for model-agnostic explainability metrics that are specific to process mining. Also, the faithfulness of explainability methods in process mining needs to be evaluated. These domain-specific metrics can guide professionals in selecting the optimal model. This paper tackles interpretability and faithfulness in domain-specific educational mining.

This paper examines the next event prediction process using previous activities based on data recorded in event logs. This task enables to proactively intervene and prevent undesired behavior in a timely manner, [10]. The task has received attention from researchers across different domains and is beneficial when introducing and comparing novel approaches to the field of PPM, [33].

All those examples come from the business community. Recently, process mining emerged in a variety of domains. In this paper, the focus is on educational process mining (EPM), process mining applied to educational data. EPM is focused on process-oriented knowledge discovery in educational systems, and it is indicated as an upcoming and emerging field of interest by [34] and [35].

This work builds upon the initially presented in [36] which focused on comparing explanations in the context of business predictive process mining. The current work applies their methodological approach to educational predictive process mining assuming differences in educational and business process data.

# 3 Research Methods

## 3.1 Educational Event Logs

Educational platforms collect huge amounts of data on every student's interaction within the learning environment. Such data is a valuable source to analyze and gain insights into the dynamics of the learning process. The dynamic nature of learning means that many learning processes unfold over time. This underscores the need for analytical methods that can capture and examine these temporal aspects. Techniques like process mining leverage the traces of student activity to understand how learning evolves over time. This is crucial for studying longitudinal processes, such as patterns of student engagement throughout an educational program. This is the main motivation behind domain selection in this research.

The basis for our analysis is an event log. An event log has three parts: case identifier, activity identifier, and timestamp. Based on the event log, multiple other metrics are calculated, e.g. cases per activity, number of activities, idle time, or throughput time.

In this research, we are using synthetic data generated from a real course at the University of Eastern Finland. The course has been described in [37] which used the original dataset. This synthetic data is explained here [1]

The course consisted of lectures including multiple practical sessions. Students submitted multiple assignments. The course had a project that accounted for 30% of the grade. Everything was available online, in the LMS Moodle. A description of the data is given in Table 1 (based on [1]).

Data refers to 95,580 timestamped Moodle logs for 130 distinct students. Activities include viewing lectures, discussing on forums, and working on individual assignments, among others. The logs were encoded putting together logs that typify the same action.

For example, actions associated with the group project were all coded as Groupwork, log activities related to feedback were coded as Feedback, logs of students' access to practical resources or assignments were coded as Practicals, etc., [1].

Table 1. Data description

| Variable | Description |
|---|---|
| Event.context | Learning management data / LMS data where the event occurred (e.g. Assignment: Literature review".) |
| user: | Username in the learning management system |
| timecreated: | Timestamp at which an event occurred |
| Component: | Sort of resource involved in the event. There are 13 different inputs (e.g. Forum, System, Assignment etc.) |
| Event.name: | Name of the event in Learning management system / LMS. There are 27 different inputs (e.g. Course module viewed, Course viewed, Discussion viewed etc. |
| Action: | A column created by combining event name and context. There are 12 different inputs (e.g. Group_work, Course_view, Practicals etc.) |

### 3.2 Explainable Artificial Intelligence

Term explainability is subjective, but many research papers have objectively quantified it, [38], [39]. These research papers mostly define it in terms of interpretability and faithfulness.

In our research, we have adopted definitions and formulas used in [36]. They define parsimony and functional complexity as properties of interpretability. Importance ranking correlation and level of disagreement as properties of faithfulness. All metrics are defined as model agnostic. Here, we give only brief definitions of the properties. For detailed descriptions and formulas see, [36].

Parsimony (P) is a property of interpretability that represents the complexity of a model. This is the non-zero weights provided by the post-hoc explainability model, [36].

Functional complexity (FC) serves as a measure of model complexity. The metric investigates how many altered predictions there would be when permuting the attribute values of an attribute type, and consequently measures how strongly the predictions depend on that attribute type, [36].

Importance ranking correlation (IRC) is a measure of the faithfulness of an explainability method and is quantified with the use of the non-parametric Spearman's correlation coefficient, [36].

Level of disagreement (LOD) is a metric of faithfulness that computes the percentage of similar predictions between the black box model and the explainability model, [36].

All four metrics will give us insights into the explainability of the models in educational predictive process mining.

## 4 Research Results and Discussion

We have applied model-agnostic explainability metrics that have been adapted for a process-oriented analysis. These metrics allow us to assess and compare the interpretability and faithfulness of explainability techniques in the field of educational process mining. Second, we applied four machine learning and deep learning algorithms. These models are complemented with post-hoc explainability techniques, allowing us to evaluate their performance in educational process mining settings from Moodle event logs. Specification of the event log used in the research is given in Table 2.

Table 2. Specification of Event Log

| Variable | Count |
|---|---|
| Number of events | 95626 |
| Number of cases | 9560 |
| Number of traces | 4076 |
| Number of distinct activities | 12 |
| Average trace length | 10.00272 |

Four different machine learning and deep learning algorithms are employed. Random Forrest (RF) and XGBoost (XGB) as machine learning approaches and Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN) as deep learning approaches. Predictive performance measured by AUC is presented in Table 3.

Table 3. Predictive Performance for Event Log per Classifier

| ***Classifier*** | ***AUC*** |
|---|---|
| RF | *87.57* |
| XGB | *87.12* |
| LSTM | *91.23* |
| CNN | *90.04* |

DL models outperform both ML models on educational event logs. This is opposite from results in [34] indicating differences in educational and business process data characteristics.

Next, we will investigate how four metrics evaluate the interpretability and faithfulness of different methods.

Table 4 presents the first metric, parsimony of each attribute type: control, case, and event.

Table 4. Parsimony for Event Log per Classifier

| Classifier | Control | Case | Event |
|---|---|---|---|
| RF | *35* | *28* | *30* |
| XGB | *20* | *18* | *25* |
| LSTM | *3* | *3* | *8* |
| CNN | *4* | *4* | *9* |

The parsimony of the total model is the sum of the parsimony for attribute types. A simple model is preferred. Accordingly, LSTM and CNN are the most parsimonious models. The second element of the model interpretability is functional complexity. Table 5 presents Functional Complexity per attribute type.

Table 5. Functional complexity for Event Log per Classifier

| Classifier | Control | Case | Event |
|---|---|---|---|
| RF | *93.32* | *5,44* | *19,98* |
| XGB | *91,11* | *4,18* | *16,73* |
| LSTM | *0.05* | *0.02* | *0.03* |
| CNN | *0.05* | *0.04* | *0.02* |

Functional complexity is used to identify the dependency of the attribute types on the explanations and to see if certain attribute types are important for explanation. If a change in all the attributes of a certain type does not change prediction, then these attribute values are not important. It is like sensitivity analysis in machine learning-based models. Therefore, interpretability is interpreted together with the metric parsimony, as this allows us to see whether the most prominent attribute types are also the most influential on the predictions.

The second property of explainability is faithfulness consisting of two properties, Importance Ranking Correlation and Level of Disagreement.

Table 6 presents Importance Ranking Correlation (IRC) values for each algorithm. This metric does not distinguish between the different attribute types, as the focus is on the relative ranking of the attributes in general.

Table 6. Importance ranking correlation for Event Log per Classifier

| Classifier | IRC |
|---|---|
| RF | *0.41* |
| XGB | *0.31* |
| LSTM | *0.18* |
| CNN | *0.23* |

The positive correlation coefficient near 1 indicates a faithful model, while a decrease in faithfulness is indicated by a value closer to 0. Here, machine-learning approaches yielded better results than deep-learning approaches.

The second metric of faithfulness is the Level of Disagreement. Table 7 presents the Level of Disagreement for ML and DL algorithms employed in the study.

Table 7. Level of disagreement for Event Log per Classifier

| Classifier | LOD |
|---|---|
| RF | *1.24* |
| XGB | *3.54* |
| LSTM | *1.22* |
| CNN | *3.44* |

The level of Disagreement investigates whether the attribute importance and the explainability model focus on the same attribute type. In our research results, machine learning and deep learning approaches yielded similar results.

Our results prove that the trade-off between performance and explainability measured through interpretability and faithfulness is difficult in the field of educational predictive process mining.

# 5 Conclusion

This paper evaluated the explainability of educational predictive models developed by deep learning in process mining. Explainability was defined through characteristics of interpretability of explanations and faithfulness of the explainability model. Interpretability was evaluated through the metrics of parsimony and functional complexity. Faithfulness was evaluated through metrics of importance ranking correlation and level of disagreement.

Our research focused on events, case, and control perspectives to capture process data. The approach was applied to the educational event log and applying four machine and deep learning algorithms.

Empirical analysis led us to the following answers to defined research questions:

RQ1: Which machine learning or deep learning algorithms give better results in terms of predictive performance in educational predictive process mining?
Deep learning algorithms provided better results in terms of predictive performance.

RQ2: Which machine learning or deep learning algorithms give better results in terms of explainability in educational predictive process mining?

Deep learning and machine learning algorithms provided the same results in terms of explainability.

We demonstrated that each model has its advantages and disadvantages. Although our research makes a scientific contribution to educational predictive process mining, it has several limitations. First, only one data set is used in empirical research. Second, there are various machine and deep learning algorithms which were not included in the research. Finally, other metrics for the explainability of the models could be employed.

Our research has practical implications for enhancing trust in complex machine learning and deep learning models working on their explainability.

**Declaration of Generative AI and AI-assisted Technologies in the Writing Process**

The authors wrote, reviewed and edited the content as needed and they have not utilised artificial intelligence (AI) tools. The authors take full responsibility for the content of the publication.

*References:*


[1] M. Saqr, S. López-Pernas, *Learning Analytics Methods and Tutorials: A Practical Guide Using R*, Springer, 2024.

[2] H. AlQaheri, M. Panda, An Education Process Mining Framework: Unveiling Meaningful Information for Understanding Students' Learning Behavior and Improving Teaching Quality, *Informaion*, Vol.13, No.1, 2022, pp. 1-19. https://doi.org/10.3390/info13010029

[3] D. A. Neu, J. Lahann, P. Fettke, A systematic literature review on state-of-the-art deep learning methods for process prediction, *Artificial Intelligence Review,* Vol.55, No.2, 2022, pp. 801-827.

[4] R. Galanti, B. Coma-Puig, M. de Leoni, J. Carmona and N. Navarin, Explainable predictive process monitoring, *In 2020 2nd International Conference on Process Mining (ICPM),* pp. 1-8, IEEE.

[5] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A survey of methods for explaining black box models, *ACM Computing Surveys,* Vol.51, 2018. https://doi.org/10.1145/3236009

[6] G. Park, M. Song, Predicting performances in business processes using deep neural networks, *Decision Support Systems,* Vol. 129, 2019, p. 113191.

[7] Kratsch, W., Manderscheid, J., Röglinger, M., Seyfried, J., Machine learning in business process monitoring: a comparison of deep learning and classical approaches used for outcome prediction, *Business & Information Systems Engineering*, Vol. 63, 2021, pp. 261-276.

[8] M. Hinkka, M., Lehto, T., Heljanko, K., Jung, A. (2019). Classifying process instances using recurrent neural networks. *In Business Process Management Workshops: BPM 2018 International Workshops, Sydney, NSW, Australia,* September 9-14, 2018, Revised Papers 16 (pp. 313-324). Springer International Publishing.

[9] Nolle, T., Luettgen, S., Seeliger, A., Mühlhäuser, M., Binet: Multi-perspective business process anomaly classification, *Information Systems,* Vol. 103, 2022, p. 101458.

[10] Mehdiyev, N., Evermann, J., Fettke, P., A novel business process prediction model using a deep learning method, *Business & information systems engineering,* Vol. 62, 2020, pp. 143-157.

[11] Di Francescomarino, C., Ghidini, C., Maggi, F. M., Petrucci, G., Yeshchenko, A., (2017). An eye into the future: leveraging a-priori knowledge in predictive business process monitoring. *In Business Process Management: 15th International Conference, BPM 2017, Barcelona, Spain,* September 10–15, 2017, Proceedings 15 (pp. 252-268). Springer International Publishing.

[12] Evermann, J., Rehse, J.-R., Fettke, P., Predicting process behaviour using deep learning, *Decis. Support Syst,* Vol. 100, 2017, pp. 129-140.

[13] Le, M., Gabrys, B., Nauck, D., A hybrid model for business process event and outcome prediction, *Expert Syst,* Vol. 34, 2017, p. e12079.

[14] Márquez-Chamorro, A.E., Resinas, M., Ruiz-Cortés, A., Toro, M., Run-time prediction of business process indicators using evolutionary decision rules, *Expert Syst,* Vol. 87, 2017, pp. 1-14.

[15] Polato, M., Sperduti, A., Burattin, A., de Leoni, M., Time and activity sequence prediction of business process instances, *Computing*, Vol. 100, 2018, pp. 1005-1031.

[16] Leontjeva, A., Conforti, R., Di Francescomarino, C., Dumas, M., & Maggi, F. M., Complex symbolic sequence encodings for predictive monitoring of business processes. In Business Process Management: 13th International Conference, BPM 2015, Innsbruck, Austria, August 31--September 3, 2015, Proceedings 13 (pp. 297-313). Springer International Publishing.

[17] De Leoni, M., van der Aalst, W.M.P., Dees, M., A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs, *Inf Syst*, Vol. 56, 2016, pp. 235-257.

[18] Conforti, R., Fink, S., Manderscheid, J., Röglinger, M., PRISM–a predictive risk monitoring approach for business processes, *In: International Conference on Business Process Management, Springer,* 2016, pp. 383-400

[19] Karypidis, E., Mouslech, S. G., Skoulariki, K., & Gazis, A., Comparison Analysis of Traditional Machine Learning and Deep Learning Techniques for Data and Image Classification, *WSEAS TRANSACTIONS on MATHEMATICS,* 2022, pp. 122-130. doi: 10.37394/23206.2022.21.19

[20] Sabrigiriraj M and K. Manoharan, Teaching Machine Learning and Deep Learning Introduction: An Innovative Tutorial-Based Practical Approach, *WSEAS TRANSACTIONS ON ADVANCES in ENGINEERING EDUCATION,* Vol.21 No.6, 2024, pp. 54-61. https://doi.org/10.37394/232010.2024.21.8

[21] Mehdiyev, N., & Fettke, P., 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 granular computing, Springer*, 2021, pp. 1-28.

[22] Rehse, J.-R., Mehdiyev, N., Fettke, P., Towards Explainable Process Predictions for Industry 4.0 in the DFKI-Smart-Lego-Factory, *KI-Künstliche Intelligenz*, 2019, pp. 1-7.

[23] Jan, S.T.K., Ishakian, V., Muthusamy, V., AI Trust in business processes: The need for process-aware explanations, arXiv Prepr. arXiv2001.07537. (2020).

[24] Islam, S. R., Eberle, W., & Ghafoor, S. K., Towards quantification of explainability in explainable artificial intelligence methods. In R. Bart´ak, & E. Bell (Eds.), Proceedings of the thirty-third international florida artificial intelligence research society conference, May 17–20, 2020 (pp. 75–81). AAAI Press.

[25] Molnar, C., Casalicchio, G., & Bischl, B., Quantifying model complexity via functional decomposition for better post-hoc interpretability, In Communications in Computer and Information Science: vol. 1167. PKDD/ECML workshops (1) (pp.193–204). Springer

[26] El-khawaga, G., Abu-Elkheir, M., & Reichert, M., Xai in the context of predictive process monitoring: An empirical analysis framework, *Algorithms,* Vol.15, No.6, 2022, p. 199

[27] Ma, S., & Tourani, R., Predictive and causal implications of using Shapley value for model interpretation, *Proceedings of the 2020 KDD workshop on causal discovery, PMLR,* 2020, pp. 23-38

[28] Velmurugan, M., Ouyang, C., Moreira, C., & Sindhgatta, R., Evaluating fidelity of explainable methods for predictive process analytics, *International conference on advanced information systems engineering, Springer,* 2021, pp. 64-72

[29] Stevens, A., Smedt, J. D., Peeperkorn, J., Munoz-Gama, J., & Lu, X. Quantifying explainability in outcome-oriented predictive process monitoring, *In Lecture notes in business information processing: vol. 433. Process mining workshops – ICPM, Springer,* 2021, pp. 194-206.

[30] Rudin, C., Stop explaining black box machine learning models for high stake decisions and use interpretable models instead, *Nature Machine Intelligence,* Vol.1, No.5, 2019, pp. 206-215.

[31] H. Khanh Dam, T. Tran, A. Ghose, Explainable software analytics in: Proceedings of the 40th International Conference on Software Engineering: New Ideas and Emerging Results, ACM, Gothenburg, Sweden, 2018, pp. 53–56, http://dx.doi.org/10.1145/3183399.3183424.

[32] Vilone, G., & Longo, L., Notions of explainability and evaluation approaches for explainable artificial intelligence, *Information Fusion*, Vol.76, 2021, pp. 89-106.

[33] Márquez-Chamorro, A. E., Resinas, M., Ruiz-Cortés, A., Predictive monitoring of business processes: a survey IEEE Transactions on Services Computing, Vol.11, No.6, 2017, pp. 962-977.

[34] Bogarín, A., Cerezo, R., Romero, C., A survey on educational process mining, *Wiley Interdisciplinary Reviews: Data Mining and*

*Knowledge Discovery*, Vol.8, No.1, 2018, p. e1230.

[35] Liu, F., Zhao, L., Zhao, J., Dai, Q., Fan, C., & Shen, J., Educational process mining for discovering students' problem-solving ability in computer programming education, *IEEE Transactions on Learning Technologies,* Vol.15, No,6, 2022, pp. 709-719.

[36] Stevens, A., & De Smedt, J., Explainability in process outcome prediction: Guidelines to obtain interpretable and faithful models, *European Journal of Operational Research,* Vol.317, No.2, 2024, pp. 317-329.

[37] Saqr, M., & López-Pernas, S., Why learning and teaching learning analytics is hard: An experience from a real-life LA course using LA methods, *In International conference on technological ecosystems for enhancing multiculturality, Singapore: Springer Nature Singapore,* 2023, pp. 781-789.

[38] Markus, A. F., Kors, J. A., & Rijnbeek, P. R., The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies, *Journal of Biomedical Informatics,* Vol.113, 2021. P. 103655

[39] Zhou, J., Gandomi, A. H., Chen, F., Holzinger, A., Evaluating the quality of machine learning explanations: A survey on methods and metrics, *Electronics*, Vol.10, No.5, 2021, pp. 593

**Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)**

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

**Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself**

This paper is supported by the Croatian Science Foundation under the project SIMON: Intelligent System for automatic selection of machine learning algorithms in Social Sciences, UIP-2020-02-6312.

**Conflict of Interest**

The authors have no conflicts of interest to declare.

**Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0
https://creativecommons.org/licenses/by/4.0/deed.en_US