



Predicting Remaining Time in Object-Centric Process Executions Using Machine Learning Models

Mohsen Mohammadi¹ 

1. Corresponding Author, Department of Computer engineering, Esfarayen University of Technology, Esfarayen, Iran.
Email: mohsen@esfarayen.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 12 June 2026
Received in revised form
12 July 2026
Accepted 17 September
2026
Available online 30
September 2026

Keywords:

Predictive process monitoring,
Object-centric process mining,
Machine learning,
Remaining time prediction

ABSTRACT

Predictive process monitoring is critical for effective business process management; however, current approaches relying on case-centric logs often fail to capture the dynamics of complex, multi-object interactions. This study proposes a framework for object-centric remaining time prediction, facilitating proactive operational decisions, including systematic delay reduction, optimized resource allocation, and enhanced service quality. We developed a prediction methodology leveraging object-centric event data from real-world logs. Our feature engineering captured multidimensional information, including temporal dynamics, elapsed execution time, sequential activity history, and intricate interactions between interrelated process objects. These features were used to train and compare three machine learning architectures: a Linear Regression baseline, Random Forest, and the Gradient Boosting algorithm. Empirical analysis reveals that ensemble tree-based methods consistently outperform the linear baseline. Specifically, the Gradient Boosting model demonstrated superior prediction accuracy, successfully mitigating the noise inherent in complex process data. Feature importance analysis confirmed that elapsed time, activity histories, and object-centric interaction features serve as the most influential predictors for remaining time. This research confirms that synthesizing object-centric feature extraction with ensemble learning provides a practical solution for time prediction in complex, multi-object environments.

Cite this article: Mohammadi, Mohsen.. Predicting remaining time in Object-centric process executions using machine learning models. *Journal of Data Analytics and Intelligent Decision-Making (JDAID)*, 2(3), 154-166.
<https://doi.org/10.22091/jdaid.2026.16354.1075>



© Author(s) retain the copyright.

Publisher: University of Qom.

DOI: <https://doi.org/10.22091/jdaid.2026.16354.1075>

Introduction

Process mining is a data-driven discipline that aims to analyze, monitor, and improve business processes based on event data recorded by information systems. As organizations increasingly rely on enterprise systems such as ERP, CRM, workflow management, and transaction-processing platforms, large volumes of event data are generated during process execution. These data provide objective evidence about how business processes are actually performed in practice. By analyzing event logs, process mining helps bridge the gap between real process behavior and prescribed process models, thereby supporting operational and managerial decision-making (Augusto et al., 2019; Carmona et al., 2018; de Leoni, 2022; Maggi et al., 2014; van der Aalst, 2016).

The main applications of process mining are commonly grouped into four categories. First, process discovery automatically derives process models from event logs to represent the observed behavior of process executions (van der Aalst, 2016). Second, conformance checking evaluates the alignment between recorded behavior and predefined models, rules, or regulations (Carmona et al., 2018). Third, process enhancement enriches process models with additional information related to performance, decisions, bottlenecks, or resource utilization (de Leoni, 2022). Fourth, predictive process monitoring uses historical event data to build predictive models that estimate future process states, such as the next activity, delay probability, risk level, or expected remaining execution time (Maggi et al., 2014). These perspectives show the importance of process mining as an analytical capability in management information systems. Figure 1 illustrates the general role of process mining in information systems and shows how event logs connect information systems, process models, and real-world process execution (van der Aalst, 2016).

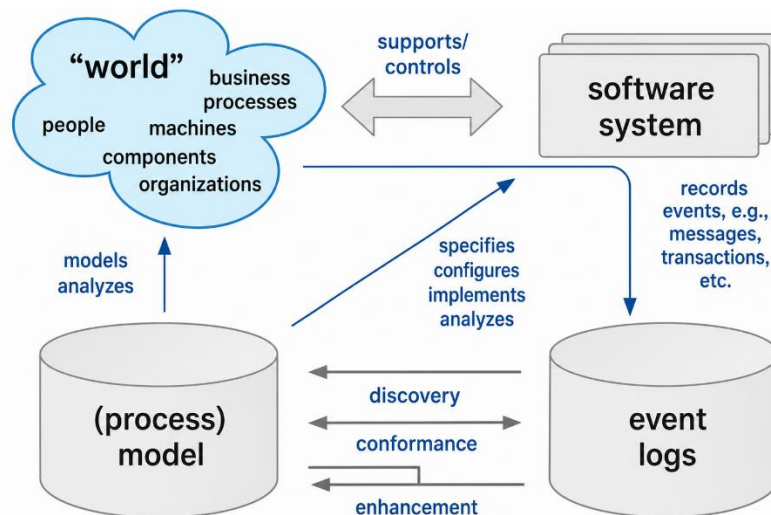


Figure 1. Process Mining in Management Information Systems (van der Aalst, 2016)

Among the different areas of process mining, predictive process monitoring has gained increasing attention because it enables organizations to shift from retrospective analysis to proactive intervention (Maggi et al., 2014). Instead of evaluating process performance only after completion, predictive monitoring supports decisions while a process is still running. One important task in this field is predicting the remaining time of an ongoing process execution.

Accurate remaining time prediction can help organizations allocate resources more effectively, detect possible delays, prioritize critical cases, improve service-level compliance, and increase customer satisfaction. This capability is particularly valuable in domains such as procurement, logistics, healthcare, customer service, and supply chain management, where timely decisions can directly affect operational performance.

Despite these benefits, most existing predictive process monitoring approaches rely on the traditional case-centric view of event logs, in which each event belongs to one case and each case is represented as a linear sequence of activities (Maggi et al., 2014). This structure is useful for simple processes, but it is often insufficient for complex real-world processes involving multiple interacting entities. For example, one customer order may be connected to several items, invoices, shipments, and payments. Similarly, in procurement processes, purchase orders, goods receipts, invoices, and approvals may interact in different ways. Converting such processes into a single-case structure usually requires flattening the data, which can cause information loss, duplicated events, distorted dependencies, and reduced prediction quality.

Object-Centric Process Mining (OCPM) has been introduced to address these limitations by representing processes through events and multiple related objects instead of a single case identifier (van der Aalst & Berti, 2020). In OCPM, one event can be related to several objects, and each object may represent a different business entity involved in the process. This approach provides a more realistic view of complex process executions and preserves the interactions among objects. Recent studies have developed object-centric techniques for process discovery (van der Aalst & Berti, 2020), conformance checking (Adams & van der Aalst, 2021), performance analysis (Park et al., 2022), and feature extraction from object-centric event data (Adams et al., 2023). These developments create a foundation for extending predictive analytics to multi-object process environments.

However, predictive monitoring in object-centric settings is still less developed than traditional case-centric predictive monitoring. Recent studies have investigated object-centric predictive process monitoring (Gherissi et al., 2022), graph-based feature extraction from object-centric event logs (Berti et al., 2022), object-centric process predictive analytics (Galanti et al., 2023), and predictive object-centric monitoring frameworks (Adams et al., n.d.; Rohrer et al., 2022). Other works have examined comparative analyses between object-centric and classical predictive monitoring (Fioretto & Masciari, 2025), global predictive monitoring of object-centric processes (de Leoni & Volpato, 2025), and broader conceptual advances in object-centric process mining (Adams, 2025). Although these studies show the potential of object-centric approaches, the specific problem of predicting remaining time in object-centric process executions using machine learning models still requires further investigation.

This gap is important because remaining time prediction in object-centric environments is more complex than in traditional event logs. In case-centric prediction, the future behavior of a process is mainly estimated from the history of one case. In object-centric processes, however, the remaining time may depend on the states and interactions of several related objects. Therefore, effective prediction requires features that preserve object relationships rather than ignore them through flattening. There is a need to investigate how object-centric event data can

be transformed into meaningful predictive features and how machine learning models can use these features to estimate remaining execution time.

Accordingly, this study focuses on predicting the remaining time of object-centric process executions using machine learning models. The objective is to extract features from object-centric event data that reflect both activity histories and object interactions, and then use these features to train and evaluate predictive models. The novelty of this research lies in integrating object-centric process representation with machine learning-based remaining time prediction. Unlike traditional approaches that rely on flattened case-centric logs, this study preserves the multi-object structure of event data and uses it as the basis for prediction. Thus, the study contributes to object-centric predictive monitoring by addressing remaining time prediction, evaluating machine learning models on object-centric features, and demonstrating the managerial value of this approach for improving resource allocation, delay prevention, and operational decision-making.

The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 presents the research methodology and proposed prediction approach. Section 4 reports and discusses the experimental results. Finally, Section 5 concludes the paper and outlines future research directions.

Literature Review

Predictive Process Monitoring (PPM) is an important branch of process mining that aims to predict the future behavior of running process instances before they are completed (Maggi et al., 2014). Common prediction tasks include next-activity prediction, outcome prediction, risk prediction, delay prediction, and remaining time prediction. Among these, remaining time prediction is especially valuable because it helps organizations estimate how much time is still needed to complete an ongoing process. This supports resource allocation, workload management, service-level monitoring, and early intervention in delayed or risky executions. Most traditional PPM approaches are based on case-centric event logs, where each event belongs to one case and each case is represented as a sequence of activities (Maggi et al., 2014). Although this representation is useful for many simple processes, it is limited for complex real-world processes involving multiple interacting objects. In processes such as order-to-cash, procure-to-pay, logistics, production, and healthcare, one event may be related to several objects, such as orders, items, invoices, shipments, and payments. When such processes are transformed into a single case-centric log, important relations between objects may be lost, duplicated, or distorted.

Object-Centric Process Mining (OCPM) has been introduced to overcome this limitation by representing processes through events and multiple related objects rather than a single case identifier (van der Aalst & Berti, 2020). In object-centric event logs, one event can be connected to different object types and object instances, allowing the process model to preserve interactions among entities. Recent studies have extended OCPM to process discovery (van der Aalst & Berti, 2020), conformance checking (Adams & van der Aalst, 2021), performance analysis (Park et al., 2022), and feature extraction from object-centric event data (Adams et al.,

2023). These developments provide a foundation for applying predictive monitoring to object-centric processes.

However, PPM for object-centric processes is still a relatively recent research area (Gherissi et al., 2022; Berti et al., 2022; Adams et al., n.d.). A major challenge is how to transform object-centric event data into a suitable representation for machine learning models. The traditional solution is to flatten object-centric logs into case-centric logs by selecting one object type as the case notion. Although this allows the use of conventional predictive models, flattening may cause several data quality problems, including event duplication, loss of object interaction information, artificial ordering of events, and biased process representation.

To reduce these problems, some studies have enriched flattened logs with object-related features. Galanti et al. (2023) and Gherissi et al. (2022) suggested using features such as the number of objects of specific types, the percentage of objects passing through certain activities, and aggregated object-level values. These features help preserve part of the object-centric information and can improve predictive performance compared with purely case-centric representations. Nevertheless, these methods still depend on flattened logs and therefore cannot fully exploit the native structure of object-centric event data.

Further developments introduced graph-based and relation-aware representations. Park et al. (2022) proposed performance-related features based on graph structures of process executions. Similarly, Berti et al. (2022) studied feature-space innovations based on statistical relationships between objects. These approaches show that object relations contain useful predictive information and can improve the representation of complex process behavior. However, some of them still require transformation or partial flattening, which limits their ability to fully capture multi-object dependencies.

To overcome the limitations of flattening, Adams et al. (n.d.) proposed a native object-centric predictive monitoring framework that works directly on object-centric event logs without converting them into case-centric logs. This framework uses a graph-based concept of process execution and directly considers relations among events and objects. By avoiding flattening, it reduces data quality problems and enables predictive models to use richer object-centric structures.

Recent studies have further emphasized the importance of moving beyond classical case-centric predictive monitoring. Comparative research has shown that object-centric representations may provide advantages when process behavior depends on multiple interacting objects (Fioretto & Masciari, 2025). Other studies have proposed global predictive monitoring approaches that consider the broader object-centric execution context instead of focusing on only one selected object type (de Leoni & Volpato, 2025). In addition, recent conceptual developments in OCPM highlight the importance of preserving event-object relationships, multi-entity dependencies, and object interactions for advanced process analytics (Adams, 2025). To summarize these developments, Table 1 compares recent methods in object-centric predictive monitoring in terms of their main approaches, feature representations, and limitations.

Table 1. Comparison of Previous Methods in Object-Centric Predictive Monitoring

Reference	Main approach	Features representation	Main limitation
Galanti et al. (2023)	Enriched flattened logs	Number of objects, percentage of objects passing activities, aggregated values	Depends on flattening and only partially preserves object interactions
Gherissi et al. (2022)	Feature-based prediction on flattened logs	Number of specific object types and object-related attributes	Keeps limited object information and loses complex dependencies
Park et al. (2022)	Graph-based process representation	Performance-related and structure-based features	Requires specialized representation and additional preprocessing
Berti et al. (2022)	Object-relation feature innovation	Statistical relationships between objects	Uses object relations but does not fully solve native prediction
Adams et al. (2025)	Native object-centric predictive monitoring	Graph-based processing without flattening	Promising but requires advanced tools and more validation
Fioretto and Masciari (2025)	Classical vs. object-centric comparison	Case-centric and object-centric feature representations	Needs broader evaluation across datasets and tasks
de Leoni and Volpato (2025)	Global object-centric predictive monitoring	Broader object-centric execution context	Still an emerging research direction
Adams (2025)	Conceptual OCPM development	Event-object relations and multi-object dependencies	Provides foundation but needs more task-specific studies

Overall, previous research shows a clear transition from flattened case-centric adaptations toward graph-based and native object-centric predictive monitoring. Early methods improved prediction by adding object-related features to flattened logs (Galanti et al., 2023; Gherissi et al., 2022), while later studies focused on graph-based structures, object relationships, and native object-centric frameworks (Park et al., 2022; Berti et al., 2022; Adams et al., n.d.). More recent research has compared object-centric and traditional predictive monitoring and confirmed the potential of object-centric representations for complex processes (Adams, 2025; de Leoni & Volpato, 2025; Fioretto & Masciari, 2025). Nevertheless, the specific task of predicting the remaining time of object-centric process executions using machine learning models remains insufficiently explored. This study addresses this gap by extracting object-centric features and evaluating their usefulness for remaining time prediction.

Methodology

The proposed method for predicting the remaining time of object-centric process executions consists of five main stages: data preparation, feature extraction, model development, hyperparameter optimization, and evaluation, as presented in Figure 2.

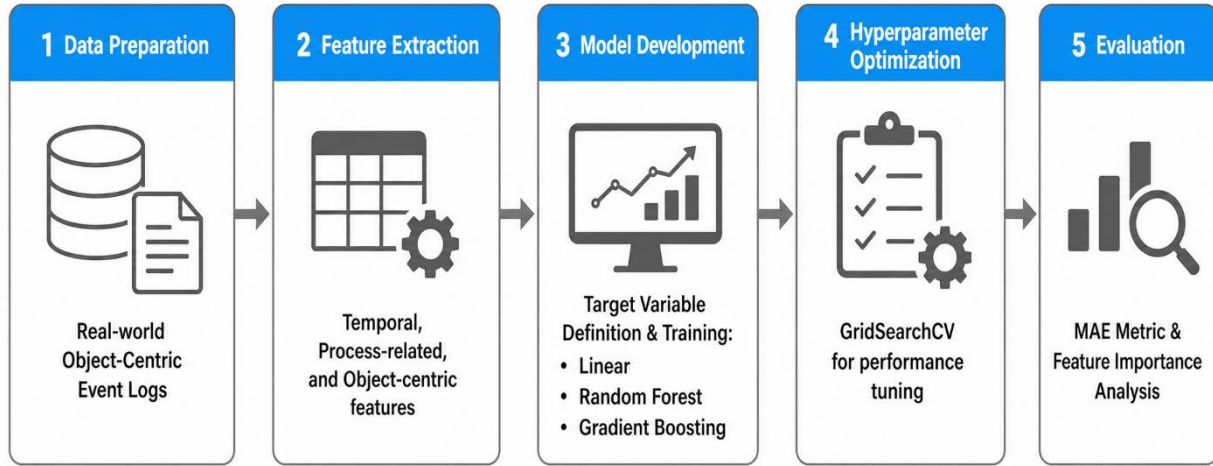


Figure 2. The proposed method

In the first stage, real-world object-centric event logs were selected from the IEEE Task Force on Process Mining repository (Task Force on Process Mining, n.d.), namely BPI2017-Final, order_process, and p2p-normal. These datasets were chosen because they represent complex business processes involving multiple interacting objects, which makes them appropriate for evaluating object-centric predictive monitoring approaches (Fioretto & Masciari, 2025; Maggi et al., 2014; van der Aalst & Berti, 2020). The selected logs cover different operational contexts, including loan application handling, order management, and purchase-to-pay execution.

In the second stage, the logs, originally stored in JSON-OCCEL format, were preprocessed and transformed into a structured representation suitable for machine learning analysis. Particular attention was given to preserving the object-centric structure of the data in order to avoid the limitations of traditional case-centric flattening, which may lead to duplicated events and the loss of important interaction information (Adams et al., n.d.; Galanti et al., 2023; Gherissi et al., 2022). Based on the preprocessed data, three categories of features were extracted. The first category includes temporal features, such as elapsed time and time intervals between events. The second category includes process-related features, such as the current activity, previous activities, and activity sequence information. The third category includes object-centric features, such as the number of related objects, object types, and interaction patterns among objects. These feature groups were selected because prior studies have demonstrated the importance of combining behavioral and object-related information in predictive monitoring tasks (Adams et al., n.d.; Berti et al., 2022; Park et al., 2022).

In the third stage, the target variable was defined as the remaining execution time, calculated as the difference between the timestamp of the current event and the completion time of the corresponding execution. The dataset was then divided into training and testing subsets. Several machine learning algorithms were applied, including Linear Regression, Random Forest, and Gradient Boosting. These models were implemented using Scikit-learn, while preprocessing

and process mining tasks were supported by PM4Py and object-centric process mining tools (Adams et al., 2023; Adams et al., n.d.; van der Aalst & Berti, 2020).

In the fourth stage, hyperparameter tuning was performed using GridSearchCV to improve predictive performance and reduce overfitting. Finally, model performance was evaluated using Mean Absolute Error (MAE), and feature importance analysis was conducted to identify the variables with the greatest predictive influence. The results indicate that elapsed time is the most influential predictor, while process history and object-centric interaction features also contribute significantly to prediction accuracy.

Finding Discussion

This section discusses the performance of three regression models—Linear Regression (LR), Random Forest (RF), and Gradient Boosting (GB)—for predicting the remaining time of object-centric process executions. Model performance was evaluated using Mean Absolute Error (MAE), where lower values indicate better prediction accuracy. The comparative results are shown in Figure 3.

As shown in Figure 3, Gradient Boosting achieved the best overall performance among the evaluated models. Linear Regression produced MAE values of 0.64, 1.20, and 0.07 across the three datasets. Random Forest improved the results, reducing the MAE values to 0.61, 0.90, and 0.04. Gradient Boosting achieved the lowest errors, with MAE values of 0.59, 0.85, and 0.04, respectively. These results indicate that nonlinear ensemble models are more effective than a simple linear baseline for remaining time prediction in object-centric processes.

The weaker performance of Linear Regression can be explained by its assumption of a linear relationship between the input features and the target variable. This assumption is often insufficient for object-centric processes, where remaining time depends on multiple interacting factors, including elapsed time, previous activities, object types, and relations among objects. In contrast, Random Forest and Gradient Boosting can capture nonlinear dependencies between temporal, behavioral, and object-centric features. Gradient Boosting performs best because it sequentially learns from previous prediction errors and gradually improves the model. Random Forest also shows strong performance, indicating that tree-based ensemble methods are suitable for complex object-centric event data.

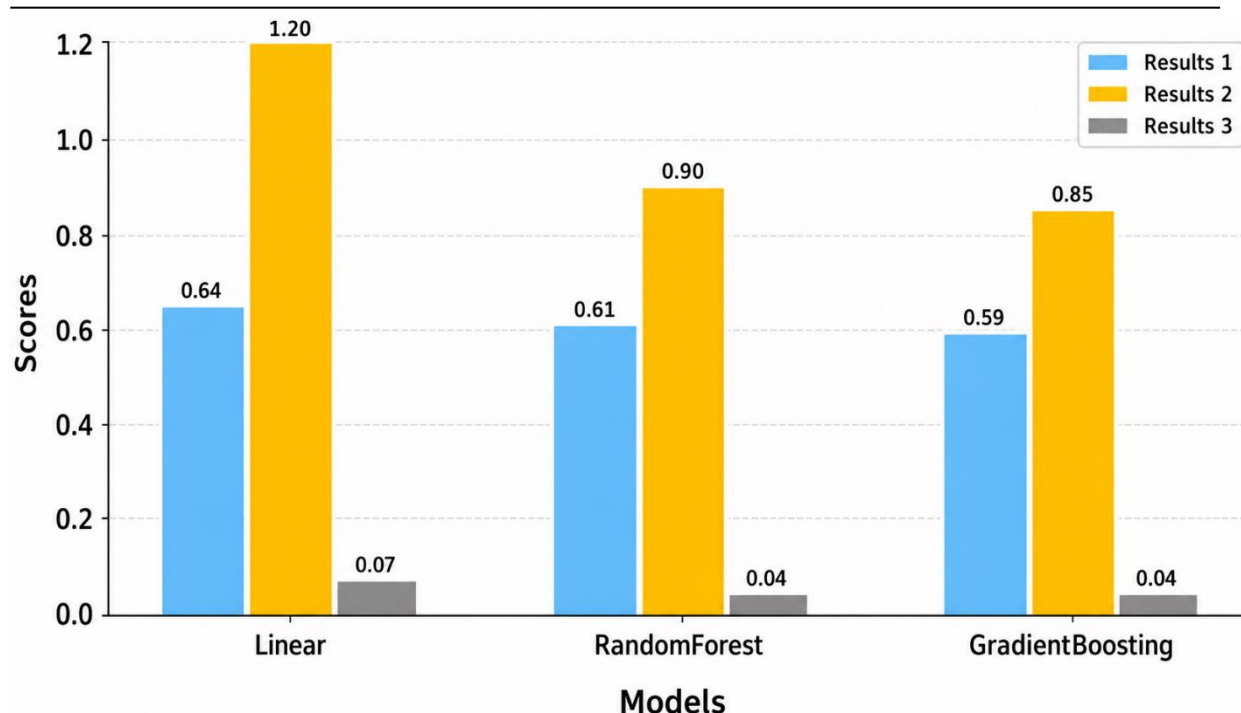


Figure 3. Comparison of machine learning models for predicting remaining time

The results are consistent with recent studies in predictive process monitoring and object-centric process mining. Earlier research has shown that remaining time prediction is influenced by elapsed time, event history, activity sequences, and contextual process attributes (Maggi et al., 2014; Fioretto & Masciari, 2025). However, many traditional predictive monitoring approaches are based on case-centric event logs, where each event is assigned to a single case. This assumption may be inappropriate for real business processes involving several interacting objects, such as orders, items, invoices, customers, and resources (Galanti et al., 2023; Gherissi et al., 2022; Park et al., 2022; Berti et al., 2022; Adams et al., 2025).

Compared with Galanti et al. (2023), which highlights the limitations of traditional case-centric process mining for multi-object processes, this study provides empirical evidence that object-centric information can be used for remaining time prediction. In relation to Gherissi et al. (2022), which discusses the problems caused by flattening object-centric logs into a single case notion, the proposed approach preserves multi-object relationships during preprocessing and feature extraction. The improved performance of RF and GB supports the idea that preserving object interactions provides useful predictive information that may be lost in flattened representations.

Compared with van der Aalst and Berti (2020), which focuses on representing complex business processes through object-centric event data, this study extends the discussion from process representation to predictive performance. The findings show that object-related attributes, such as the number of related objects, object types, and interaction patterns, contribute to more accurate predictions. Similarly, while Adams et al. (2023) emphasizes object-centric process mining for analyzing processes with multiple related entities, this study adds a quantitative comparison of machine learning models and shows that ensemble-based methods outperform Linear Regression.

The present work is also complementary to Adams et al. (n.d.), which introduces or applies object-centric process mining tools and techniques. While that study supports the extraction and analysis of object-centric event data, this study demonstrates how such data can be transformed into predictive features for estimating remaining execution time. In relation to Fioretto and Masciari (2025), the findings confirm the importance of temporal and behavioral features for remaining time prediction, but extend this perspective by incorporating object-centric features. Compared with (Nasseri and Mohseni 2024), which reports the effectiveness of machine learning methods for process prediction, this study confirms that nonlinear models are more suitable than linear models when process behavior contains complex dependencies. Finally, in line with (Adams et al. 2025), which highlights recent advances in predictive monitoring and advanced learning techniques, the results show that ensemble tree-based models remain highly competitive for object-centric remaining time prediction.

Overall, the comparison with recent studies shows that this research is aligned with current developments in predictive monitoring, while extending them to object-centric event logs. The results confirm that combining object-centric feature representation with nonlinear ensemble learning provides a more appropriate solution for complex multi-object business processes.

The findings have important practical implications for organizations managing complex processes such as loan processing, order management, logistics, procurement, and purchase-to-pay workflows. In these environments, delays are often caused not by a single activity but by interactions among multiple objects. Therefore, object-centric remaining time prediction can provide more realistic and actionable information than traditional case-centric prediction.

Accurate remaining time prediction can help managers identify possible delays, monitor service-level agreements, allocate resources more effectively, and support proactive decision-making. For example, in loan processing, the model can help detect applications that are likely to be delayed. In order management, it can support coordination among orders, items, deliveries, and invoices. In purchase-to-pay processes, it can help identify bottlenecks between purchase orders, goods receipts, invoices, and payments.

From a modeling perspective, Gradient Boosting is the most suitable option when prediction accuracy is the main objective, as it achieved the lowest MAE. Random Forest is also a strong choice for practical implementation because it provides competitive accuracy, robustness, and relatively efficient computation. Linear Regression remains useful as a transparent baseline, allowing the improvements of more advanced models to be clearly demonstrated.

In addition to accuracy, tree-based ensemble models offer interpretability through feature importance analysis. This is valuable for process improvement because it helps identify which features most strongly influence remaining time. In this study, elapsed time was one of the most influential predictors, while previous activities and object-centric interaction features also contributed to prediction accuracy. This confirms that remaining time prediction in object-centric processes depends on both temporal progress and multi-object behavior.

Overall, the results demonstrate that ensemble tree-based models are more reliable than linear models for predicting remaining time in object-centric processes. By combining object-centric feature extraction with nonlinear machine learning, the proposed approach improves prediction accuracy and provides useful insights for operational decision-making and process optimization.

(Galanti et al., 2023; Gherissi et al., 2022; Park et al., 2022; Berti et al., 2022; Adams et al., 2025; Fioretto & Masciari, 2025).

Conclusions

This study investigated the prediction of remaining execution time in object-centric process executions using machine learning models. In contrast to traditional case-centric predictive monitoring, which relies on a single case identifier, the proposed approach used object-centric event data to preserve relationships among multiple interacting objects within the same process. Temporal features, previous activity information, and object-centric interaction features were extracted from real-world object-centric event logs and used to train Linear Regression, Random Forest, and Gradient Boosting models.

The results showed that ensemble tree-based models outperformed the linear baseline. Gradient Boosting achieved the lowest MAE, demonstrating its ability to capture complex and nonlinear relationships between process features and remaining time. Random Forest also achieved strong performance, confirming the suitability of tree-based ensemble methods for object-centric predictive monitoring. Linear Regression was less effective because it could not adequately model nonlinear dependencies and multi-object interactions.

Feature importance analysis showed that elapsed time, previous activities, and object-centric interaction features were among the most influential predictors of remaining execution time. This confirms that object-centric information provides valuable predictive signals that may be lost when complex processes are transformed into traditional case-centric event logs. From a practical perspective, the proposed approach can help organizations identify potential delays, improve resource allocation, monitor service-level agreements, and support proactive operational decision-making.

The main limitation of this study is that the evaluation was conducted on selected object-centric event logs and a limited set of machine learning models, which may restrict the generalizability of the findings. Future work should evaluate the approach on more diverse real-world datasets and explore advanced methods such as deep learning, hybrid models, online prediction, and explainable artificial intelligence. Overall, this study demonstrates that combining object-centric process mining with ensemble machine learning, especially Gradient Boosting, is an effective approach for improving remaining time prediction in complex business processes.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRedit authorship contribution statement: Mohsen Mohammadi: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing original draft, Writing– review & editing.

Declaration of Generative AI and AI-assisted technologies in the writing process: During the preparation of this manuscript, the author used ChatGPT (OpenAI) to assist with language editing and to improve the clarity and readability of selected textual passages. The tool was used solely for writing support and was not used to generate research data, perform analysis, design the methodology, or draw scientific conclusions. All AI-assisted content

was reviewed, revised, and verified by the author, who takes full responsibility for the final content of the manuscript.

Conflict of interest: The authors declare no conflict of interest.

Ethical considerations: The authors avoided data fabrication, falsification, and plagiarism, and any form of misconduct.

Data availability statement: The data that support the findings of this study are openly available in 4TU.ResearchData as two separate datasets: BPI Challenge 2017, order_process, and p2p-normal. These datasets are freely available at <https://data.4tu.nl/>

Acknowledgements: The authors would like to thank anonymous reviewers for their valuable suggestions in manuscript revision.

References

- Adams, J. N. (2025). *Object-centric process mining: Concepts and advancements* [Doctoral dissertation, RWTH Aachen University].
- Adams, J. N., Drescher, H., Swoboda, A., Günnemann, N., Park, G., & van der Aalst, W. M. P. (n.d.). *Improving predictive process monitoring using object-centric process mining*.
- Adams, J. N., Park, G., Levich, S., Schuster, D., & van der Aalst, W. M. P. (2023). *A framework for extracting and encoding features from object-centric event data*. *Expert Systems with Applications*, 213, Article 119173. <https://doi.org/10.1016/j.eswa.2022.119173>
- Adams, J. N., & van der Aalst, W. M. P. (2021). *Precision and fitness in object-centric process mining*. In *Proceedings of the International Conference on Process Mining (ICPM 2021)* (pp. 128–135). IEEE. <https://doi.org/10.1109/ICPM53251.2021.9576886>
- Augusto, A., Conforti, R., Dumas, M., La Rosa, M., Maggi, F. M., Marrella, A., Mecella, M., & Soo, A. (2019). *Automated discovery of process models from event logs: Review and benchmark*. *IEEE Transactions on Knowledge and Data Engineering*, 31(4), 686–705. <https://doi.org/10.1109/TKDE.2018.2841877>
- Berti, A., Herforth, J., Qafari, M. S., & van der Aalst, W. M. P. (2022). *Graph-based feature extraction on object-centric event logs* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2207.03904>
- Carmona, J., van Dongen, B. F., Solti, A., & Weidlich, M. (2018). *Conformance checking: Relating processes and models*. Springer. <https://doi.org/10.1007/978-3-319-99414-7>
- de Leoni, M. (2022). *Foundations of process enhancement*. In W. M. P. van der Aalst & J. Carmona (Eds.), *Process mining handbook: Lecture notes in business information processing* (Vol. 448, pp. 243–273). Springer. https://doi.org/10.1007/978-3-031-08848-3_8
- de Leoni, M., & Volpato, P. (2025). *Global predictive monitoring of object-centric processes*. In *International Conference on Business Process Management* (pp. 255–272). Springer.
- Fioretto, S., & Masciari, E. (2025). *A comparative analysis of predictive process monitoring: Object-centric versus classical event logs*. *Knowledge and Information Systems*, 67(9), 7355–7398. <https://doi.org/10.1007/s10115-025-02384-y>
- Galanti, R., de Leoni, M., Navarin, N., & Marazzi, A. (2023). *Object-centric process predictive analytics*. In *Proceedings of the 5th International Conference on Process Mining*.
- Gherissi, W., El Haddad, J., & Grigori, D. (2022). *Object-centric predictive process monitoring*. In *Service-oriented computing: ICSOC 2022 workshops* (pp. 27–39). Springer. https://doi.org/10.1007/978-3-031-26507-5_3
- Maggi, F. M., Di Francescomarino, C., Dumas, M., & Ghidini, C. (2014). *Predictive monitoring of business processes*. In *Advanced information systems engineering: CAiSE 2014: Lecture notes in computer science* (Vol. 8484, pp. 1–16). Springer. <https://doi.org/10.1007/978-3-319-07881-6>
- Park, G., Adams, J. N., & van der Aalst, W. M. P. (2022). *OPerA: Object-centric performance analysis*. In *Conceptual modeling: ER 2022* (pp. 281–292). Springer. https://doi.org/10.1007/978-3-031-17995-2_20
- Rohrer, T., Ghahfarokhi, A. F., Behery, M., Lakemeyer, G., & van der Aalst, W. M. P. (2022). *Predictive object-centric process monitoring* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2207.10017>
- Task Force on Process Mining. (n.d.). *Resources: Logs*. <https://www.tf-pm.org/resources/logs>

- van der Aalst, W. M. P. (2016). *Process mining: Data science in action* (2nd ed.). Springer.
<https://doi.org/10.1007/978-3-662-49851-4>
- van der Aalst, W. M. P., & Berti, A. (2020). *Discovering object-centric Petri nets*. *Fundamenta Informaticae*, 175(1–4), 1–40. <https://doi.org/10.3233/FI-2020-1946>