Hindawi
Advances in Operations Research
Volume 2023, Article ID 8867057, 10 pages
https://doi.org/10.1155/2023/8867057

Research Article

Decision-Making Based on Predictive Process Monitoring of Patient Treatment Processes: A Case Study of Emergency Patients

Agaraoli Aravazhi,[1] **Berit I. Helgheim**,[1] **and Petter Aadahl**[2]

[1] *Faculty of Logistics, Molde University College–Specialized University in Logistics, Molde 6410, Norway*
[2] *Department of Cardiothoracic Anesthesia and Intensive Care, Clinic of Anesthesia and Intensive Care, St. Olavs Hospital, Trondheim 7006, Norway*

Correspondence should be addressed to Agaraoli Aravazhi; agaraoli.aravazhi@himolde.no

Received 14 June 2023; Accepted 21 September 2023; Published 28 September 2023

Academic Editor: Hossein Moosaei

This paper investigates predictive process monitoring problems in emergency treatment by combining the fields of process management and artificial intelligence. The objective is to predict the next activity and its timestamp in the treatment of emergency patients who have undergone surgery at the gastroenterology or urology surgery units in a hospital in Norway. To achieve this goal, three models were developed using different algorithms, and the best performing model was identified using various performance metrics. The results demonstrate the potential of predictive process monitoring to accurately forecast the outcome of patient treatments. By leveraging the insights gained from predictive process monitoring, hospitals can make more informed decisions. The findings of this study suggest that predictive process monitoring holds significant promise as a tool for improving the efficiency and effectiveness of emergency patient treatment processes. This research has significant implications for the field of decision sciences, particularly regarding resource allocation, reducing waiting times, and improving patient outcomes. The ability to predict the outcomes of patient treatment processes has important implications for hospitals, allowing the streamlining and acceleration of the treatment process. Overall, this study provides a promising framework for predicting patient treatment processes by using the predictive process monitoring method. This could be expanded upon in future research, ultimately leading to improved patient outcomes and better decision-making in healthcare.

1. Introduction

Recent advances in medicine and technology have enabled hospitals to provide improved treatment for numerous diagnoses. The planning of care as well as the allocation of resources, however, involves prediction and hard decision-making. Further, patients now increasingly expect to receive the best care possible [1], and here, optimization and prediction of patient care processes may play an important role. With the growing number of patients and the need to provide the best care, regulating patient flow in a hospital has become increasingly complex. The traditional approach of reactive interventions adds to the complexity. Solving these complexities requires a high degree of coordination and decision-making from the healthcare providers at the hospital. Prescriptive analytics, based on machine learning algorithm, leads to optimized decision-making ahead of time [2].

Over the past two decades, significant research has been invested into the concept of process mining. Process mining encompasses the extraction of meaningful and previously unknown insights from historical event logs and additional process-related data [3]. This concept covers multiple fields such as process management and artificial intelligence (AI). Artificial intelligence covers numerous capabilities including the concept of prescriptive analytics. Using these two fields (process management and AI) in combination to predict the outcomes of future processes will facilitate decision-makers to take action.

Predicting future process outcomes is termed a “predictive process monitoring problem” [4]. For example, predicting the next activity in the treatment process will assist healthcare professionals to prepare the necessary steps to complete the event.

In predictive process monitoring, the input information is stored in the form of event logs in the information system. Event logs contain information about executed events with details such as case identifier (case ID), activity name, timestamp, and/or other relevant information such as resources used for and, in relation to this, the cost of each activity [5]. The output of the prediction problem is to predict a specific value for a new process instance, such as the activity name, the time-related details for the activity, and the resources used. The prediction output can be of any format, such as categorical, Boolean, or numeric, which defines the type of problem. For example, predicting the name of the next event is a categorical problem. Predicting whether a performance indicator exceeds the limit is a Boolean problem, and predicting the time at which a particular event will occur is a regression problem.

One of the main review studies on predictive process monitoring, conducted by Tama and Comuzzi [6], provides a benchmark on predictive process monitoring techniques and applications. In this study, 20 classification algorithms from five classifier families are benchmarked for predicting the name of the next event. Márquez-Chamorro et al. [7] presented an in-depth qualitative review on predictive process monitoring and the computational methods, predicted values, and quality evaluation metrics used.

There has, however, been less research focus on the prediction of the next activity and its time. Therefore, this paper focuses on the combined prediction values of the next activity and its start time. This is one of the principal use cases in predictive process monitoring. It is a combination of categorical prediction and regression prediction. An aspect to note is the “no free lunch” theorem [8], based on which no single model will have superior performance across datasets. Therefore, multiple models are built and compared against different evaluation metrics to select the best performing model for predicting the next event and the time of the start of the event.

The goal of this paper is to demonstrate how the predictive process monitoring method can be used by decision-makers in hospitals for better planning of the patient care process. This goal is twofold. First, we aim to predict the next activity and time for treatment in emergency patients at various triage levels. This can help healthcare professionals to prepare and has an impact on the planning processes. Second, the goal is to build multiple models to make predictions. The first model is built as a combination of a simple multinomial logistic regression model and a linear regression model. The second model is a random forest (RF) algorithm for classification combined with a neural network (NN) model for regression. The third model uses recurrent neural networks (RNNs) with long short-term memory (LSTM) architecture, separately for classification and regression. We also implemented k-fold cross-validation as the resampling strategy to avoid overfitting and underfitting issues. We used data from two surgical units, a gastroenterology (GA) and a urology (UR) surgical unit in a regional university hospital (herein called as hospital) in Norway.

The remainder of this paper is organized as follows: First, the existing literature on predictive process monitoring is presented. Second, a description of the method used in the study is provided. Third, an overview of the results is presented. Fourth, we present a discussion of the results along with the research implications and limitations. Finally, we conclude the paper and provide details and suggestions for future research.

2. Literature Review

Since 2010, technological advancements have improved decision support in healthcare, and this has attracted considerable attention from researchers. In healthcare settings, patient flow management is important. Patient flow management involves coordinating the movement of patients through distinct stages of care, from admission to discharge [9, 10]. Therefore, making prompt and accurate decisions for patient flow management is critical. Failure to provide effective patient flow management can result in increased wait times and delayed diagnosis and treatment, which can have a negative impact on both the patient and healthcare professionals [11]. Therefore, effective use of decision support systems will have a positive impact on patient flow management.

Managing the flow of emergency patients through the hospital presents several challenges and issues including overcrowding, delays in care, and inefficient use of resources. Some studies have focused on this issue [12, 13] but have limited themselves to the study of the emergency department. Other studies have explored strategies to address these challenges, such as the use of Lean principles [14–16], the implementation of electronic health records to enhance communication and coordination [17, 18], and various other methods [19]. Recent advancements in the field of analytical techniques, such as machine learning, have been utilized to predict patient outcomes [20–23], including hospital readmissions [24] and mortality rates [25], for emergency patients in hospital. Despite such efforts to improve emergency patient flow in hospitals, challenges persist, and more research is needed to find effective solutions.

Recent developments in the field of process mining and machine learning have given rise to the technique of predictive process monitoring. Multiple machine learning algorithms have been applied for predictive process monitoring [26]. There are two main use cases concerned with the application of predictive process monitoring, such as predicting process outcomes and proactive process monitoring. Predicting process outcomes includes prediction of constraint values (such as costs) related to the process, while proactive process monitoring focuses on prediction of the next activity in, and the timestamp for, a case [6]. Since this study focuses only on proactive process monitoring, we are not presenting literature on predicting process outcomes. A summary of the existing literature on proactive process monitoring is presented in Table 1.

The literature summary shows that multiple studies have been conducted on proactive process monitoring. The

Table 1: Literature on proactive process monitoring.

Study	Industry/dataset	Activity	Time
Van der Aalst et al. [27]	Municipality		x
Rogge-Solti and Weske [28]	Logistics		x
Breuker et al. [29]	BPI Challenge	x	
Becker et al. [30]	BPI Challenge	x	
Polato et al. [31]	Road fines		x
Unuvar et al. [32]	Marketing campaign	x	
Verenich et al. [33]	Bondora: environmental permit	x	
Senderovich et al. [34]	Hospital		x
Tax et al. [35]	BPI Challenge	x	x
Evermann et al. [36]	BPI Challenge	x	
Mehdiyev et al. [37]	BPI Challenge	x	
Khan et al. [38]	BPI Challenge	x	x
Al-Jebrni et al. [39]	BPI Challenge	x	
Schönig et al. [40]	BPI Challenge	x	
Camargo et al. [41]	BPI Challenge	x	x
Lin et al. [42]	BPI Challenge	x	
Hinkka et al. [43]	BPI Challenge	x	
Theis and Darabi [44]	BPI Challenge	x	
Di Mauro et al. [45]	BPI Challenge	x	

majority of these have focused on activity-related predictions, i.e., on predicting the next activity, and a few studies have looked at time-related predictions. Only three studies have examined the prediction of both the activity and time. The majority of the research has focused on developing models to make predictions using deep learning algorithms such as autoencoder (AE), deep feedforward network (DFN), and LSTM. As can be seen in Table 1, these algorithms have been applied mostly to datasets available as part of the Business Process Intelligence (BPI) Challenge, while some have been applied to other real-life event logs for predictive process monitoring.

Despite the growing interest in predictive process monitoring, there are still several research gaps related to healthcare that need to be addressed. First, there is limited research on predicting both the next activity and its timestamp in patient flow. As mentioned, current research mainly focuses on predicting the next activity, which is essential for patient flow management, but timestamp prediction can provide valuable insights for optimizing workflow efficiency and reducing wait times. Second, most predictive analytics studies in healthcare rely on a limited dataset, which hinders their ability to generalize findings to other healthcare systems or settings. In addition, the majority of studies rely on deep learning algorithms, which can be computationally complex and require large amounts of computational time, limiting their practical application. Therefore, it is necessary to explore alternative techniques that can both achieve comparable predictive accuracy in a small amount of time and reduce computational complexity. Finally, there is a significant research gap in predicting treatment processes. While predicting patient outcomes and readmissions is essential, predicting the treatment process can provide healthcare professionals with valuable insights to make informed decisions about patient care.

Therefore, this study aims to address this research gap by focusing on predictive process monitoring for the treatment process. Moreover, in addition to predicting the next activity in patient flow, this study will also focus on predicting the timestamp of the next activity. To achieve this goal, multiple models will be explored to study the impact on prediction quality, which can help to identify the most effective approach to predictive process monitoring. Overall, this study aims to build on and contribute to the existing body of knowledge on patient flow management in healthcare settings by addressing the identified research gaps and providing new insights into the use of predictive analytics for treatment process monitoring.

3. Methods

This section presents the research framework comprising the following key aspects: data collection and preprocessing; models developed for the study; a thorough validation procedure; and the evaluation metrics used to access the performance of the developed models.

3.1. Data Collection and Preprocessing. The study setting was a regional hospital in Norway. Event log data was collected from the event records of the emergency patients who underwent surgery at the GA and UR surgical units over a period of approximately 7 years, from 2012 to 2018. The patients were categorized into four levels of triage marking: red, orange, yellow, and green. The number of events in the GA and UR surgical units for each triage level considered in this study is presented in Table 2.

An event log in the study recorded different instances of the treatment process for each patient. The sequence of activities for each case began with the start of a specific activity within the process. A sample of the traces in the event log, which includes a unique identifier for the case (case ID), the name of the activity, a timestamp showing when the activity started, and other meta-attributes (such as triage level and diagnosis code) related to the case are presented in Table 3.

Table 2: Number of events during the period 2012–2018.

Triage	GA (n)	UR (n)
Red	24 118	12 522
Orange	84 665	24 777
Yellow	142 919	52 447
Green	8 984	2 860
Total	260 686	92 606

GA = gastroenterology; UR = urology.

The event log data must be processed into a format that is compatible with machine learning algorithms. The transformed dataset consists of input attributes composed of the properties of different activities and the time differences between them. One class attribute and one numeric attribute are also generated, representing the next activity and the time difference between the activity and the previous activity within the window size (or prefix length) considered. Previous research by Márquez-Chamorro et al. [46], Tax et al. [35], and Tama and Comuzzi [6] considered a window size of three. Therefore, the transformed dataset includes seven nominal attributes (case ID, triage, DiagCategoryCode, DiagGroupCode, Activity 1, Activity 2, and Activity 3), two numeric attributes (duration between adjacent activities in the considered window size, namely, Time 12 and Time 23), and two target attributes (one class label attribute representing the next activity, and one numeric attribute representing the duration between the target activity and Activity 3). An example of the transformed event log data used in this study is presented in tabular format in Table 4.

As the aim of the study is to develop models separately for each surgical unit and triage level, the transformed data were filtered into separate datasets for each unit and triage level. To achieve this, an event-based sampling approach was implemented, whereby the case ID attribute is ignored for all datasets, and the triage attribute is ignored for individual triage datasets during the validation procedure to ensure that values are generated from the same case.

3.2. Model Development. In this study, we implemented three models in R, each consisting of a combination of a classification algorithm and a regression algorithm. The following is a brief description of each model, which is described in more detail in the remainder of this section.

3.2.1. Model 1: Multinomial Logistic Regression and Linear Regression. Multinomial logistic regression uses a logistic function to model the relationship between a set of independent variables and a categorical dependent variable. In this case, the categorical variable was used for classification purposes. The model predicts the probability of the dependent variable being in a particular class based on the independent variables. A cutoff of the predicted probability value determines which class is the predicted value [47]. This study uses the multinom function from the nnet package and implements it in R interface.

Linear regression, on the other hand, is a method of modeling the relationship between a dependent variable and one or more independent variables. In this model, the independent variables are used for regression purposes. The model assumes that the relationship between the dependent variable and the independent variables is linear and attempts to find the best fit line that describes this relationship [48]. For this, the lm function available in the stats package is implemented using R interface.

3.2.2. Model 2: Random Forest and Neural Network. Random forest is an ensemble of decision trees to make predictions. Each decision tree is trained on a subset of the training data and independently makes a prediction. The predictions of all the trees are then combined to make a final prediction [49]. In this study, a distributed RF framework in a random forest library was implemented through the R interface.

One of the most widely used algorithms is NN [50], which attempts to recognize patterns in data. It is designed to contain a set of artificial neurons that process the input data and make predictions [51]. To enable this, the Keras application programming interface (API) is used along with the TensorFlow package to build a simple NN in R interface.

3.2.3. Model 3: Long Short-Term Memory. Long short-term memory is one of the most powerful types of RNNs. An LSTM model uses memory cells to store information about past inputs and outputs, which is used to predict the future output values. In this study, two different LSTM models were developed, one for classification to predict the next activity and one for regression to predict the time of the activity. Keras API along with the TensorFlow package was used to build two different LSTM models in R interface.

3.3. Validation Procedure. To validate the prediction models developed in this study, which involve a combination of classification and regression algorithms, it is important to address common issues such as underfitting and overfitting. Several validation procedures have been developed over the years to achieve this, with the train/test data split being a commonly used approach. However, this method is not suitable when the data is imbalanced.

To address this issue, we implemented a stratified 80/20 train/test data split, ensuring that both datasets had an equal proportion of data. Furthermore, to ensure that the results were not due to chance, we implemented a tenfold cross-validation method. The training data was divided into ten stratified subsets, with no overlap between them. Nine of these were used to train the model; the remaining subset was used for validation. This process was repeated ten times, with each subset used as validation data once.

All ten validation results were then compared, and the best performing cross-validated model was selected for each developed model. Employing such validation procedures ensures that the models developed are better equipped to avoid common issues such as underfitting and overfitting and that the results are reliable and robust.

3.4. Evaluation Metrics. In this study, we have developed prediction models that include both classification and regression models. To evaluate their performance, multiple metrics were considered for each part of the model.

Table 3: Sample event log for predicting the next event.

Case ID	Triage	DiagCategoryCode	DiagGroupCode	Activity	Timestamp
1_1	Yellow	K00–K95	K55–K64	A	04-03-2017 10 : 06
1_1	Yellow	K00–K95	K55–K64	B	04-03-2017 10 : 12
. . .	. . .	. . .	. . .	. . .	. . .
1_1	Yellow	K00–K95	K55–K64	M	04-03-2017 19 : 35

For multiclass classification problems, four performance measures are considered, namely, accuracy, Matthews correlation coefficient (MCC), confusion entropy, and area under the receiver operating characteristic (AUC-ROC) curve. While accuracy is a simple metric that provides a rough goodness of the model, the confusion entropy value is difficult to interpret [52]. Therefore, this study considers MCC and the AUC-ROC curve as crucial performance measures. The MCC value ranges from −1 to 1, with 1 representing perfect classification and −1 representing extreme misclassification [53]. The AUC-ROC value ranges from 0 to 1, with 1 representing perfect classification and 0 representing complete classification inaccuracy [54]. The MCC value was calculated using the following equation, and the AUC-ROC value was evaluated using the multiclass.roc function of the pROC package in R [55].

$$\mathrm{MCC} = \frac{c * s - \sum_{k}^{K} p_k * t_k}{\sqrt{\left(s^2 - \sum_{k}^{K} p_k^2\right) * \left(s^2 - \sum_{k}^{K} t_k^2\right)}}, \tag{1}$$

where K is the number of classes, k is the class from 1 to K, $t_k = \sum_{i}^{K} C_{ik}$ is the number of times class k truly occurred, $p_k = \sum_{i}^{K} C_{ki}$ is the number of times class k was predicted, $c = \sum_{k}^{K} C_{kk}$ is the total number of correct predictions, and $s = \sum_{i}^{K} \sum_{j}^{k} C_{ij}$ is the total number of predictions.

For regression evaluation, two widely used indices, mean absolute error (MAE) and root mean square error (RMSE), are considered. The first, MAE, treats all errors equally, while RMSE penalizes large errors. The lower values of both metrics indicate better model performance. These metrics are calculated by using the following equations, respectively:

$$\mathrm{RMSE} = \frac{1}{n} \sqrt{\sum_{t=1}^{n} \left(y_t - \hat{y}_t\right)^2}, \tag{2}$$

$$\mathrm{MAE} = \frac{1}{n} \sum_{t=1}^{n} \left|y_t - \hat{y}_t\right|, \tag{3}$$

where n is the number of data points, y_t is the actual value, and $\hat{y}_t$ is the predicted value.

4. Results

The focus of this study is on predictive process monitoring in which predictions are made for determining the next activity and its timestamp for emergency patients who have undergone surgery either at the GA or at the UR surgical unit. For this study, three different models were developed, and the performance of these models was compared to identify the best performing model. A comparison of the consolidated results is presented in Table 5.

The first model built for the study was a combination of the MLR algorithm for classification problems and the LR algorithm for regression problems. For the MLR algorithm, a maximum iteration of 5000 was used to build the model for predicting the next activity. This model uses the Akaike Information Criterion (AIC) for fitting. For the LR algorithm, the default QR decomposition method was used to fit the model. The performance results of the combined model show that the MLR model performed best for the UR surgical unit datasets, while the LR model outperformed the other methods for most of the GA surgical unit datasets.

The second model, a combination of the RF algorithm and NN, addresses the classification problem and the regression problem, respectively. The RF is built using a maximum of 500 trees to predict the next activity. The NN model is built using one input layer with seven or eight input dimensions based on the dataset, three hidden layers (with 128, 64, and 32 neurons, respectively), and one output layer with one neuron. The model was compiled using the “adam” optimizer and “mean absolute error” as a loss function. For the RF model, the evaluation metric MCC had the best performance for UR datasets, while in most cases, the AUC-ROC values were better for GA datasets. For the NN model, the results were mixed and only with the total triage dataset and the yellow triage dataset both the MAE and RMSE metrics did have better performance for UR.

The final model was a combination of two different LSTM models, one each for the classification problem and the regression problem. The first model for classification was designed with one input layer with seven or eight input dimensions based on the dataset, three hidden layers (with the first two layers having 128 and 64 neurons, and with the third layer being a dropout layer with a rate of 0.25), and one output layer with 23 output neurons for each activity. The model was compiled using the “adam” optimizer and “sparse categorical cross entropy” as a loss function.

The second LSTM model was designed with one layer with seven or eight input dimensions based on the dataset, two hidden layers (consisting of 128 and 64 neurons, respectively), and one output layer with one output neuron. The model was compiled using the “adam” optimizer and “mean absolute error” as a loss function. The LSTM models produced better performance results for the UR datasets than for the GA datasets.

4.1. Model Selection. The next step was to select the best performing models for each dataset. To select the model for this study, the performance values of all the models presented in Table 5 were compared. Based on the comparison, the RF and NN ensembles provided the best performance results in most cases, followed by the LSTM model performances.

Table 4: Transformed format of the event log sample shown in Table 3.

Case ID	Triage	. . .	Activity 1	Activity 2	Activity 3	Time 12	Time 23	Target activity	Target time
1_1	Yellow	. . .	A	B	C	6.0	3.0	D	9.0
1_1	Yellow	. . .	B	C	D	8.0	7.0	E	10.0
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
1_1	Yellow	. . .	K	L	M	3.0	5.0	N	15.0

Table 5: Evaluation metrics results of prediction models both for the gastroenterology (GA) and for the urology (UR) surgical units.

	Models	GA				UR			
		MCC	AUC-ROC	MAE	RMSE	MCC	AUC-ROC	MAE	RMSE
Total triage	MLR + LR	0.622	0.758	8.961	12.631	0.804	0.784	9.227	12.763
	RF + NN	0.850	0.840	5.061	10.948	0.920	0.827	4.851	10.549
	LSTM	0.837	0.822	5.068	10.900	0.909	0.836	4.835	10.373
Red triage	MLR + LR	0.647	0.735	8.652	12.370	0.869	0.877	8.153	12.384
	RF + NN	0.841	0.836	4.899	10.258	0.935	0.889	4.765	11.235
	LSTM	0.815	0.793	4.953	10.585	0.928	0.895	4.415	10.500
Orange triage	MLR + LR	0.677	0.765	8.699	12.380	0.775	0.787	9.194	12.635
	RF + NN	0.865	0.841	4.882	10.543	0.908	0.824	5.055	10.244
	LSTM	0.851	0.813	4.842	10.695	0.888	0.824	5.071	10.249
Yellow triage	MLR + LR	0.627	0.763	9.089	12.778	0.826	0.792	9.199	12.768
	RF + NN	0.849	0.845	5.220	11.081	0.934	0.838	4.889	10.633
	LSTM	0.833	0.820	5.282	11.179	0.922	0.818	4.968	10.745
Green triage	MLR + LR	0.628	0.770	9.710	13.693	0.711	0.797	9.390	12.589
	RF + NN	0.825	0.845	6.653	12.834	0.879	0.840	6.750	12.530
	LSTM	0.785	0.838	6.663	13.080	0.755	0.821	7.134	12.409

AUC-ROC = area under the receiver operating characteristic; LR = linear regression; LSTM = long short-term memory; MAE = mean absolute error; MCC = Matthew correlation coefficient; MLR = multinomial logistic regression; NN = neural network; RF = random forest; RMSE = root mean square error.

For patients with red, yellow, or green triage who underwent surgery at the GA surgical unit, the ensemble of the RF and NN models provided the most accurate predictions. Similarly, for orange and yellow triage patients who underwent surgery at the UR surgical unit, the same ensemble model performed the best.

However, for all emergency patients at the GA surgical unit and for patients with green triage who underwent surgery at the UR surgical unit, the ensemble of the RF and the NN models was better than other models except for the RMSE value. In these cases, RMSE values were low for the LSTM model. Emergency patients with orange triage at the GA surgical unit had better performance results with the RF and NN models in all metrics except for MAE. The MAE value was lower when the LSTM model was implemented.

The LSTM model outperformed other models in all metrics except for the MCC value for all emergency patients and red triage patients at the UR surgical unit. The MCC value was higher with the RF and NN ensembles.

5. Discussion

The objective of this study is to address the predictive process monitoring problem by combining the fields of process management and AI. Specifically, we aim to predict the next activity and its timestamp in the treatment process of emergency patients who underwent surgery at the GA or UR surgical unit in a hospital in Norway. This research highlights the possibility of using the predictive process monitoring method in emergency patient care to predict the next activity and its timestamp.

To achieve this goal, we propose a novel approach that combines two models, one for predicting the next activity and the other for predicting the timestamp. We developed three models using different algorithms, namely, (i) MLR and linear regression; (ii) RF and NN; and (iii) two separate LSTM models. To address the challenge of unbalanced data, we employed stratified sampling to split the data and implemented k-fold cross-validation to avoid underfitting and overfitting issues.

The study found that the RF and NN models provided the best performance results in most cases, followed by the LSTM model. This combination accurately predicted the next activity and its timestamp for emergency patients who underwent surgery at the GA and UR surgical units, with low error values for most cases. The LSTM model outperformed other models in most metrics, except for the MCC value for red triage patients at the UR surgical unit, where the RF and NN models performed better.

Based on the results of this study, it appears that the RF and NN ensembles outperformed the LSTM model in most cases. One factor that may have contributed to this result is the difference in computational complexity between the two models. Random forest and NN models are relatively less computationally intensive compared to the LSTM model, which requires more computation resources due to its complex architecture and training process [56]. This combination of algorithms provides an effective and efficient approach for predictive process monitoring in emergency patient treatment processes with less computational time, making it a preferable choice over the more computationally complex LSTM model. In addition, the RF and NN models will be able to provide results in real time, which is particularly important in emergency situations where prompt decision-making is critical.

Overall, the result of this study demonstrates the use of predictive process monitoring to accurately forecast the values of patient treatment processes. The prediction of these values has significant implications for hospitals, as it allows for the streamlining and acceleration of the treatment process. In addition, these predictions can serve as a valuable tool for decision-making in hospital operations, particularly in regard to resource allocation, reducing waiting times, and improving patient outcomes [57]. By leveraging the insights gained from predictive process monitoring, hospitals can make more informed decisions about staffing levels, bed availability, and other critical factors that impact the quality of care they provide. Ultimately, the findings of this study suggest that predictive process monitoring holds significant promise as a tool for improving the efficiency and effectiveness of emergency patient treatment processes.

In this study, two limitations should be acknowledged. First, the study examined patient treatment processes at a single hospital in Norway. Therefore, the results may not be generalizable to other healthcare settings. The hospital's specific patient population, care protocols, and data collection processes may have influenced the model's performance. In addition, the study only focused on emergency patients who underwent surgery at the GA or UR unit. This is a representative study which needs to be extended to other patient populations and medical conditions. Despite these limitations, this study presents the prediction of patient treatment processes, which can be expanded upon in future research.

6. Conclusion

In conclusion, this study proposes an approach that combines two models for predicting the next activity and its timestamp in the treatment process of emergency patients who underwent surgery at the GA or UR surgical units in a hospital in Norway. The study found that the RF and NN models outperformed the LSTM model in most cases, providing an effective and efficient approach for predictive process monitoring in emergency patient treatment processes with less computational time.

Predictive process monitoring has significant implications for hospitals, as it allows for the streamlining and acceleration of the treatment process and can serve as a valuable tool for decision-making in hospital operations. By using the insights gained from predictive process monitoring, hospitals can make more informed decisions about staffing levels, bed availability, and other critical factors that impact the quality of care they provide. Although the study limited itself to examining patient treatment processes at a single hospital in Norway and, moreover, focused on emergency patients who underwent surgery at the GA or UR unit. The findings of this study suggest that predictive process monitoring has the potential to enhance the efficiency and efficacy of emergency patient treatment processes.

This study contributes to the growing body of literature on AI-driven decision sciences, demonstrating the potential for combining the fields of process management and AI to address predictive process monitoring problems. The proposed approach can be applied to other medical conditions and patient populations and can be further developed and refined to improve its performance. Overall, the results of this study provide valuable insights for hospitals and researchers seeking to improve the quality of care provided to emergency patients.

Data Availability

The hospital event log data used to support the findings of this study have not been made available because of confidentiality reasons.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] X. Peng, B. Flynn, A. Narayanan, and R. Fan, "Alternative information processing mechanisms in hospital supply chains: impact on cost, quality, and patient satisfaction," *Decision Sciences*, pp. 1–20, 2022.

[2] K. Lepenioti, A. Bousdekis, D. Apostolou, and G. Mentzas, "Prescriptive analytics: literature review and research challenges," *International Journal of Information Management*, vol. 50, pp. 57–70, 2020.

[3] W. Van Der Aalst, *Process Mining: Data Science in Action*, Springer, Berlin, Germany, 2016.

[4] I. Teinemaa, M. Dumas, F. Maria Maggi, and C. D. Francescomarino, *Predictive Business Process Monitoring with Structured and Unstructured Data*, Springer International Publishing, Cham, Switzerland, 2016.

[5] R. S. Mans, W. M. Van der Aalst, and R. J. Vanwersch, *Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes*, Springer, Berlin, Germany, 2015.

[6] B. A. Tama and M. Comuzzi, "An empirical comparison of classification techniques for next event prediction using business process event logs," *Expert Systems with Applications*, vol. 129, pp. 233–245, 2019.

[7] A. E. Márquez-Chamorro, M. Resinas, and A. Ruiz-Cortés, "Predictive monitoring of business processes: a survey," *IEEE Transactions on Services Computing*, vol. 11, no. 6, pp. 962–977, 2018.

[8] D. H. Wolpert, "The lack of A priori distinctions between learning algorithms," *Neural Computation*, vol. 8, no. 7, pp. 1341–1390, 1996.

[9] P. Nugus, A. Holdgate, M. Fry, R. Forero, S. McCarthy, and J. Braithwaite, "Work pressure and patient flow management in the emergency department: findings from an ethnographic study," *Academic Emergency Medicine*, vol. 18, no. 10, pp. 1045–1052, 2011.

[10] S. Villa, A. Prenestini, and I. Giusepi, "A framework to analyze hospital-wide patient flow logistics: evidence from an Italian comparative study," *Health Policy*, vol. 115, no. 2-3, pp. 196–205, 2014.

[11] R. Hall, D. Belson, P. Murali, and M. Dessouky, "Modeling patient flows through the healthcare system," in *Patient Flow: Reducing Delay in Healthcare Delivery*, R. W. Hall, Ed., pp. 1–44,Springer US, Boston, MA, USA, 2006.

[12] J. Furmaga and S. A. McDonald, "Impact of rapid medical evaluation on patient flow in an urban emergency department," *Journal of Medical Systems*, vol. 45, no. 6, p. 63, 2021.

[13] A. Sayah, L. Rogers, K. Devarajan, L. Kingsley-Rocker, and L. F. Lobon, "Minimizing ED waiting times and improving patient flow and experience of care," *Emergency Medicine International*, vol. 2014, Article ID 981472, 8 pages, 2014.

[14] A. Alowad, P. Samaranayake, K. Ahsan, H. Alidrisi, and A. Karim, "Enhancing patient flow in emergency department (ED) using lean strategies–an integrated voice of customer and voice of process perspective," *Business Process Management Journal*, vol. 27, no. 1, pp. 75–105, 2021.

[15] H. Chan, S. Lo, L. Lee et al., "Lean techniques for the improvement of patients' flow in emergency department," *World Journal of Emergency Medicine*, vol. 5, no. 1, pp. 24–28, 2014.

[16] M. Sánchez, M. Suárez, M. Asenjo, and E. Bragulat, "Improvement of emergency department patient flow using lean thinking," *International Journal for Quality in Health Care*, vol. 30, no. 4, pp. 250–256, 2018.

[17] S. S. Kennebeck, N. Timm, M. K. Farrell, and S. A. Spooner, "Impact of electronic health record implementation on patient flow metrics in a pediatric emergency department," *Journal of the American Medical Informatics Association*, vol. 19, no. 3, pp. 443–447, 2012.

[18] H. Salmasian, A. B. Landman, and C. Morris, "An electronic notification system for improving patient flow in the emergency department," *AMIA Jt Summits Transl Sci Proc.*, vol. 2019, pp. 242–247, 2019.

[19] H. Vass and Z. K. Szabo, "Application of queuing model to patient flow in emergency department. Case study," *Procedia Economics and Finance*, vol. 32, pp. 479–487, 2015.

[20] G. Bunney, S. Tran, S. Han et al., "Using machine learning to predict hospital disposition with geriatric emergency department innovation intervention," *Annals of Emergency Medicine*, vol. 81, no. 3, pp. 353–363, 2023.

[21] M.-C. Chen, T. Y. Huang, T. Y. Chen, P. Boonyarat, and Y. C. Chang, "Clinical narrative-aware deep neural network for emergency department critical outcome prediction," *Journal of Biomedical Informatics*, vol. 138, Article ID 104284, 2023.

[22] M. Greco, P. F. Caruso, S. Spano et al., "Machine learning for early outcome prediction in septic patients in the emergency department," *Algorithms*, vol. 16, no. 2, p. 76, 2023.

[23] Z. Zhou, C. Huang, P. Fu et al., "Prediction of in-hospital hypokalemia using machine learning and first hospitalization day records in patients with traumatic brain injury," *CNS Neuroscience and Therapeutics*, vol. 29, no. 1, pp. 181–191, 2023.

[24] H. Jung, H. W. Park, Y. Kim, and Y. Hwangbo, "Machine learning-based prediction for 30-day unplanned readmission in all-types cancer patients," in *Proceedings of the 2023 IEEE International Conference on Big Data and Smart Computing*, BigComp, Jeju, Republic of Korea, February 2023.

[25] T. Huang, D. Le, L. Yuan, S. Xu, and X. Peng, "Machine learning for prediction of in-hospital mortality in lung cancer patients admitted to intensive care unit," *PLoS One*, vol. 18, no. 1, Article ID e0280606, 2023.

[26] W. Deng, Y. Li, K. Huang, D. Wu, C. Yang, and W. Gui, "LSTMED: an uneven dynamic process monitoring method based on LSTM and Autoencoder neural network," *Neural Networks*, vol. 158, pp. 30–41, 2023.

[27] W. M. P. van der Aalst, M. H. Schonenberg, and M. Song, "Time prediction based on process mining," *Information Systems*, vol. 36, no. 2, pp. 450–475, 2011.

[28] A. Rogge-Solti and M. Weske, *Prediction of Remaining Service Execution Time Using Stochastic Petri Nets with Arbitrary Firing Delays*, Springer Berlin Heidelberg, Berlin, Germany, 2013.

[29] D. Breuker, P. Delfmann, M. Matzner, and J. Becker, *Designing and Evaluating an Interpretable Predictive Modeling Technique for Business Processes*, Springer International Publishing, Cham, Switzerland, 2015.

[30] J. Becker, D. Breuker, P. Delfmann, and M. Matzner, "Designing and implementing a framework for event-based predictive modelling of business processes," in *Proceedings of the International Workshop on Enterprise Modeling and Information Systems Architectures (EMISA 2014)*, Luxembourg, September 2014.

[31] M. Polato, A. Sperduti, A. Burattin, and M. de Leoni, "Data-aware remaining time prediction of business process instances," in *Proceedings of the 2014 International Joint Conference on Neural Networks*, IJCNN, Beijing, China, July 2014.

[32] M. Unuvar, G. T. Lakshmanan, and Y. N. Doganata, "Leveraging path information to generate predictions for parallel business processes," *Knowledge and Information Systems*, vol. 47, no. 2, pp. 433–461, 2016.

[33] I. Verenich, M. Dumas, M. L. Rosa, F. M. Maggi, and C. D. Francescomarino, *Minimizing Overprocessing Waste in Business Processes via Predictive Activity Ordering*, Springer International Publishing, Cham, Switzerland, 2016.

[34] A. Senderovich, A. Shleyfman, M. Weidlich, A. Gal, and A. Mandelbaum, *P3-Folder: Optimal Model Simplification for Improving Accuracy in Process Performance Prediction*, Springer International Publishing, Cham, Switzerland, 2016.

[35] N. Tax, I. Verenich, M. La Rosa, and M. Dumas, *Predictive Business Process Monitoring with LSTM Neural Networks*, Springer International Publishing, Cham, Switzerland, 2017.

[36] J. Evermann, J.-R. Rehse, and P. Fettke, "Predicting process behaviour using deep learning," *Decision Support Systems*, vol. 100, pp. 129–140, 2017.

[37] N. Mehdiyev, J. Evermann, and P. Fettke, "A multi-stage deep learning approach for business process event prediction," in *Proceedings of the 2017 IEEE 19th Conference on Business Informatics (CBI)*, Thessaloniki, Greece, July 2017.

[38] A. Khan, H. Le, K. Do et al., "Memory-augmented neural networks for predictive process analytics," 2018, https://www.arxiv-vanity.com/papers/1802.00938/.

[39] A. Al-Jebrni, H. Cai, and L. Jiang, "Predicting the next process event using convolutional neural networks," in *Proceedings of the 2018 IEEE International Conference on Progress in Informatics and Computing (PIC)*, Suzhou, China, December 2018.

[40] S. Schönig, R. Jasinski, L. Ackermann, and S. Jablonski, "Deep learning process prediction with discrete and continuous data features," in *Proceedings of the 13th International Conference on Evaluation of Novel Approaches to Software Engineering*, pp. 314–319, Madeira, Portugal, March 2018.

[41] M. Camargo, M. Dumas, and O. González-Rojas, *Learning Accurate LSTM Models of Business Processes*, Springer International Publishing, Cham, Switzerland, 2019.

[42] L. Lin, L. Wen, and J. Wang, "MM-pred: a deep predictive model for multi-attribute event sequence," in *Proceedings of the 2019 SIAM International Conference on Data Mining (SDM)*, pp. 118–126, Alberta, Canada, May 2019.

[43] M. Hinkka, T. Lehto, and K. Heljanko, *Exploiting Event Log Event Attributes in RNN Based Prediction*, Springer International Publishing, Cham, Switzerland, 2020.

[44] J. Theis and H. Darabi, "Decay replay mining to predict next process events," *IEEE Access*, vol. 7, pp. 119787–119803, 2019.

[45] N. Di Mauro, A. Appice, and T. M. A. Basile, "Activity prediction of business process instances with inception CNN models," in *Proceedings of the XVIIIth International Conference of the Italian Association for Artificial Intelligence*-Springer International Publishing, Cham, Switzerland, November 2019.

[46] A. E. Márquez-Chamorro, M. Resinas, A. Ruiz-Cortés, and M. Toro, "Run-time prediction of business process indicators using evolutionary decision rules," *Expert Systems with Applications*, vol. 87, pp. 1–14, 2017.

[47] C. Kwak and A. Clayton-Matthews, "Multinomial logistic regression," *Nursing Research*, vol. 51, no. 6, pp. 404–410, 2002.

[48] X. Su, X. Yan, and C.-L. Tsai, “Linear regression,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 3, pp. 275–294, 2012.

[49] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[50] Y. Yu, T. Wang, and Y. Shi, “Analytics for multiperiod risk-averse newsvendor under nonstationary demands,” *Decision Sciences*, 2021.

[51] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, “State-of-the-art in artificial neural network applications: a survey,” *Heliyon*, vol. 4, no. 11, Article ID e00938, 2018.

[52] G. Jurman and C. Furlanello, “A unifying view for performance measures in multi-class prediction,” 2010.

[53] M. Grandini, E. Bagli, and G. Visani, “Metrics for multi-class classification: an overview,” 2020, https://arxiv.org/abs/2008.05756.

[54] L. Lau, Y. Kankanige, B. Rubinstein et al., “Machine-learning algorithms predict graft failure after liver transplantation,” *Transplantation*, vol. 101, no. 4, pp. e125–e132, 2017.

[55] X. Robin, N. Turck, A. Hainard et al., “pROC: an open-source package for R and S+ to analyze and compare ROC curves,” *BMC Bioinformatics*, vol. 12, no. 1, p. 77, 2011.

[56] O. Schumann, C. Wöhler, M. Hahn, and J. Dickmann, “Comparison of random forest and long short-term memory network performances in classification tasks using radar,” in *Proceedings of the 2017 Sensor Data Fusion: Trends, Solutions, Applications*, SDF, Bonn, Germany, October 2017.

[57] I. Teinemaa, M. Dumas, M. L. Rosa, and F. Maria Maggi, “Outcome-oriented predictive process monitoring: review and benchmark,” *ACM Transactions on Knowledge Discovery from Data*, vol. 13, no. 2, pp. 1–57, 2019.