6-18-2022

Outcome-Oriented Predictive Process Monitoring to Predict Unplanned ICU Readmission in MIMIC-IV Database

QIFAN CHEN
The university of Sydney, qche8411@uni.sydney.edu.au

Yang Lu
the University of Sydney, yalu8986@uni.sydney.edu.au

Charmaine Tam
TheUniversity of Sydney, charmaine.tam@sydney.edu.au

Simon Poon
The Sydney University of Sydney, simon.poon@sydney.edu.au

OUTCOME-ORIENTED PREDICTIVE PROCESS MONITORING TO PREDICT UNPLANNED ICU READMISSION IN MIMIC-IV DATABASE

Research in Progress

Qifan Chen, University of Sydney, Sydney, Australia, qifan.chen@sydney.edu.au

Yang Lu, University of Sydney, Sydney, Australia, yang.lu@sydney.edu.au

Charmaine S. Tam, University of Sydney, Sydney, Australia, charmaine.tam@sydney.edu.au

Simon K. Poon, University of Sydney, Sydney, Australia, simon.poon@sydney.edu.au

Abstract

Unplanned readmission entails unnecessary risks for patients and avoidable waste of medical resources, especially intensive care unit (ICU) readmissions, which increases likelihood of length of stay and more severely mortality. Identifying patients who are likely to suffer unplanned ICU readmission can benefit both patients and hospitals. Readmission is typically predicted by statistical features extracted from completed ICU stays. The development of predictive process monitoring (PPM) technique aims to learn from complete traces and predict the outcome of ongoing ones. In this paper, we adopt PPM to view ICU stay from electronic health record (EHR) as a continuous process trace to enable us to discover clinical and also process features to predict likelihood of readmission. Using events logs extracted from MIMIC-IV database, the results show that our approach can achieve up to 65% accuracy during the early stage of each ICU stay and demonstrate the feasibility of applying PPM to unplanned ICU readmission prediction.

Keywords: Readmission Prediction, Predictive Process Monitoring, MIMIC-IV

1 Introduction

Intensive care is expensive and its cost is increasing every year (Xue, Klabjan, and Y. Luo, 2019). Carefully timing the transfer of patients to low-level wards is thus essential to ensure the efficient allocation of finite resources (Desautels et al., 2017). Hospitals could benefit from cost reduction by discharging patients early from ICU. However, premature discharges may result in deterioration of patient health or adverse outcomes, and thus readmission (Xue, Klabjan, and Y. Luo, 2019). Readmitted patients tend to have higher mortality and longer length of stay compared to first admission patients (Durbin Jr and Kopel, 1993), which makes readmission rate an important factor when evaluating ICU quality (Xue, Klabjan, and Y. Luo, 2019). Thus, accurate predictions of patients with higher risk of readmission would benefit both patients and hospitals, and help doctors make appropriate decisions when discharging specific patients.

Different methods have been proposed to predict the likelihood of patients' readmission based on various features. Golmaei and X. Luo (2021) utilises NLP techniques to interpret features from clinical notes and patient network for predictions. Others rely on machine learning or deep learning algorithms combined with statistical features from chart events and demographic information to predict the risk of patients' readmission (Assaf and Jayousi, 2020; Barbieri et al., 2020; Desautels et al., 2017; Inan et al., 2018; Lin et al., 2019; Pakbin et al., 2018; Teo et al., 2020; Xue, Klabjan, and Y. Luo, 2019; Zebin and Chaussalet, 2019). Across all efforts made to address the predicting readmission task, few have considered process data. We propose an improvement by incorporating process monitoring to predict unplanned ICU readmissions.

Outcome-oriented PPM is a category of PPM that aims to analyse historical completed traces and predict outcomes of the current ongoing traces (Van der Aalst, Schonenberg, and Song, 2011). PPM plays a vital role in forecasting unwanted or abnormal outcomes. In this paper, we focus on predicting the readmission/non-readmission before the current discharge of a patient from ICU. Potential outcomes can be influenced by the sequence of activities and their corresponding attribute values (i.e. case-level and event-level) (Leontjeva et al., 2016). Previous studies mostly rely on machine learning to train models and predict outcomes, while deep learning achieves better performance when dealing with logs with high variants (Kratsch et al., 2020).

Integrating PPM with readmission predictions is of value-adds. First, PPM views patients' journey as a process trace (i.e. different chart events with recorded timestamps and attributes), where various PPM technologies can be further integrated to predict readmissions. When EHR is utilised in PPM, time-series information is retrieved from event logs, including sequence of activities executed, the specific attribute value for each execution (i.e. the observation result for every chart event recorded) and the trend of results. Previous studies mostly relied on clinical information of chart events to develop statistical models to predict readmission and discarded executed sequences information of chart event and time-series values. The use of PPM enables us to perform predictions of current hospital inpatients (using their ongoing traces) before they have been cleared or scheduled for discharges. Nearly 33% of readmissions are due to premature discharge (Durbin Jr and Kopel, 1993). Process monitoring allows doctors to make appropriate adjustments to patients when they stay in ICU based on the prediction results to reduce the chance of unplanned readmission. Existing works' predictions are based on patients' completed records. However, patients are already discharged from ICU at that time, and so no further actions can be performed in ICU. To the best of our knowledge, this is one of the first attempts PPM is considered as an additional perspective to enhance outcome predictions for unplanned ICU readmission from a real-world EHR.

The aim of this paper is to present a novel approach to predict unplanned ICU readmission by applying PPM to the MIMIC-IV open access EHR database. We construct event logs based on historical ICU stays and encode selected features to train classifiers from proposed deep learning models. We benchmark three proposed models against a baseline approach under simulated real scenarios. The results demonstrate that our approach can efficiently predict readmissions for ongoing ICU stays, which provides doctors additional decision support and potentially reduce the chances of unplanned readmission for patients.

The paper is structured as follows. Section 2 discusses the background. In Section 3, we explain the main approach. In Section 4, we present the benchmarks. Section 5 concludes the paper.

2 Background

2.1 Readmission Prediction

Predicting readmission is a complicated task since the judgement is not made based on a single identifier (Lin et al., 2019). Instead, this involves considering more investigatory results and monitoring observations than at any other point in a patient's treatment process (Desautels et al., 2017). Barbieri et al. (2020) compare and benchmark the existing methods, which achieve an F-score of less than 38 %. Common methods involve extracting features from clinical notes or physiological measurements. Golmaei and X. Luo (2021) make predictions based on discovered features from unstructured clinical notes and patient network adopting NLP techniques. Statistical features (e.g. mean) are extracted from chart events and traditional machine learning algorithms (e.g. decision tree) are utilised for predictions in (Assaf and Jayousi, 2020; Inan et al., 2018). Desautels et al. (2017) propose a transfer learning based approach to transfer pre-learnt models to a new dataset, while Xue, Klabjan, and Y. Luo (2019) adopt grouped physiological and medication trends. Lin et al. (2019) and Zebin and Chaussalet (2019) utilise longitude data by grouping chart events based on an adaptive time window. However, none of them view the problem from a process perspective, which leads to the loss of informative features from EHR.

2.2 Outcome-oriented Predictive Process Monitoring

Outcome-oriented PPM seeks to predict the outcome of an ongoing trace as early as possible, given a set of historical completed traces (Maggi et al., 2014). Leontjeva et al. (2016) explore different methods to encode and represent features from event logs to achieve better results. Others apply various techniques to predict outcomes. Teinemaa et al. (2019) inspect different machine learning algorithms (e.g. decision tree and random forest) for predictions. Similar traces are first grouped and predictions are then made based on different clusters (Di Francescomarino et al., 2016), while Metzger et al. (2014) combine multiple techniques together (e.g. machine learning and constraint satisfaction). PPM has recently been widely applied to healthcare. Xu et al. (2020) adopt PPM to assist thrombolytic therapy decision-making for ischemic stroke patients. Huberts et al. (2021) use PPM to support healthcare workers in early identifying the risk of a mental health crisis in people diagnosed with schizophrenia. Pijnenborg et al. (2021) apply PPM to predict the life expectancy of stomach and esophageal cancer patients currently under treatment. Moreover, Kratsch et al. (2020) compare machine learning to deep learning for outcome predictions and points out that deep learning is more suitable for logs with high variant-to-instance ratio and event-to-activity ratio, which are the key features for EHR data (Munoz-Gama et al., 2022). Yet, existing research mainly focuses on applying deep learning to next event predictions instead of outcome predictions (Teinemaa et al., 2019). Therefore, we adopt several well-known deep learning algorithms, that have shown promising results in various prediction problems (Barbieri et al., 2020) to predict unplanned ICU readmissions.

3 Methodology

In this paper, we aim to integrate PPM to predict ICU readmissions, i.e. predict patient's outcome during the time they are in-ICU stay. The proposed PPM approach (Figure 1) follows the framework suggested by Teinemaa et al. (2019). This approach consists of two stages: offline, to train a prediction model based on the historical event log, and online, to make predictions on running traces. To this end, we first present how to construct the event log based on completed ICU stays from the MIMIC-IV database. Then, we describe our method to extract and encode features in the event log. The encoded event log is inputted to train a prediction model, where several deep learning models are proposed and implemented.

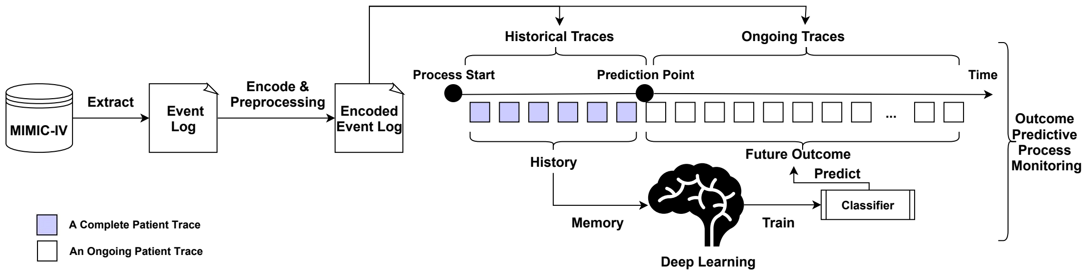

Figure 1. Overview of the proposed approach.

3.1 Event Log Construction

The readmission event log is constructed from the publicly available MIMIC-IV database [1], which contains real hospital stays for patients admitted to a medical centre in the US from 2008 to 2019 (A. Johnson

[1] `https://mimic.mit.edu/`

et al., 2021). First, we filter out patients under 18 years old and patients died in the ICU. This results in a total of 47,642 patients with 67,727 ICU stays. Each ICU stay is treated as a process trace since a patient can have multiple ICU stays. Then, we take ICU patients' chart events as activities and their observation results as attributes from the *chartevents* table in database. It must be noted that we only take the last 48-hour data of each ICU stay for the offline stage, since they are found to be the most informative data for the prediction of readmission (Barbieri et al., 2020). Moreover, we extract several demographic features along with ICD diagnoses from *patients*, *admissions* and *diagnoses icd* tables as case-level attributes. The structure of the event log before encoding is shown in Figure 2.

3.2 Feature Extraction and Encoding

As mentioned, we treat ICU patients' chart events as activities and their observation results as event-level attributes in the event log. We further filter in 12 chart events as activities (e.g. Glucose and Blood pressure) since they are proven to be valuable contributing factors to readmission (Xue, Klabjan, and Y. Luo, 2019; Zebin and Chaussalet, 2019). Details of selected activities are presented in Table 1. We first propose a 12-dim chart event features to encode each activity's observation results, as shown in Figure 2. One important implication worth noting is that many chart events recorded simultaneously in EHR with the same timestamp, i.e. the actual sequence of events could have been lost and problematic for PPM. To work around this issue, we aggregate the same timestamp events within each trace together. First, their attribute values are easily merged together by copying each value to the same line in the event log, since each activity has its own dimension to record the attribute values, as shown in Figure 2. We then add a 12-dim binary indicator feature to track the events that occurred at this timestamp, as shown in Figure 2. For example, if three events happen together, namely "pH", "Glucose" and "Weight", they are merged into the same line in the event log, with their observation results recorded in the corresponding attribute dimension. In addition, for the 12-dim binary indicator feature, the three corresponding activities are encoded with "1" and the rest with "0". In this way, we avoid using the inaccuracy relations between events with the same timestamp.

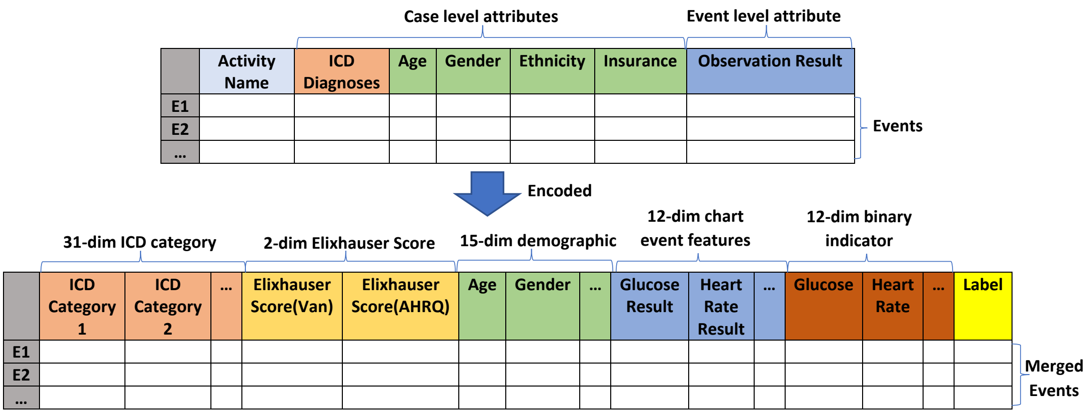

Figure 2. Overview of the event log before and after encoding.

We also extract several demographic features, including gender, age, ethnicity and insurance, as case-level attribute. It is widely known that demographic features are factors that may influence unplanned ICU readmission. For instance, it has been shown that, elderly patients have a higher chance of readmission to ICU than young patients (Assaf and Jayousi, 2020). For categorical demographic features (i.e. all except age), they are encoded using one-hot encoding, which leads to a 15-dim feature (see Figure 2).

The other common issue for existing approaches is that they can only handle a single version of the ICD (i.e. International Classification of Diseases) code (i.e. ICD-9) and require large feature dimensions to

encode ICD codes (i.e. 300 dimensions for each patient's diagnosis). However, in the latest MIMIC-IV database, ICD codes are mixed with version 9 and 10 (A. Johnson et al., 2021). In this context, information would be lost if we simply transfer one version to the other, since ICD-10 codes are much more specific than ICD-9 codes (i.e. 69,000 ICD-10 diagnosis codes compared to 14,000 ICD-9 diagnosis codes). It is unlikely that event logs have so many dimensions for attributes in PPM. Therefore, this paper treats each patient's diagnoses as case-level attributes. We also propose a method to encode diagnoses with relatively small dimensions and suitable for ICD-9 and ICD-10 by utilising the Elixhauser comorbidity score as a feature to provide a composite measure of health conditions. The Elixhauser comorbidity score, analogously as the Charlson comorbidity score, which is an approach of categorizing comorbidities of patients based on the ICD diagnosis (Elixhauser et al., 1998). A primary reason that we selected Elixhauser comorbidity score is that it develops 31 comorbidity categories, that are more specific than the Charlson comorbidity score (i.e. 19 categories). We first group patients' all diagnoses into a 31-dim binary indicator feature based on Elixhauser comorbidity score, where “1" signifies that the patient has comorbidity in the specific category. Then, we compute the final score based on two well-known weight settings (van Walraven Algorithm (Walraven et al., 2009) and AHRQ Algorithm (Moore et al., 2017)) and append the 2-dim score to the event log as case-level attributes, as shown in Figure 2.
It is essential to label the input data for any supervised learning tasks. In this paper, we categorised all historical ICU stays into positive (further readmission of the patient is recorded in the database) and negative (no further readmission recorded within the next 30 days) cases. This leads to a single dim feature in Figure 2. We adopt the below criteria in (Zebin and Chaussalet, 2019) to define readmitted patients:

- Patients were transferred to low-level wards from ICU but returned to ICU again.
- Patients were discharged, but returned to ICU within the next 30 days.
- Patients were discharged from ICU and died within the next 30 days.

Using these criteria, 9,535 traces are identified as readmission and the rest are non-readmission. We adopt stratified random down-sampling to achieve a balanced event log to train classifiers (Shahrokh Esfahani and Dougherty, 2014).

Item ID	Activity Name	Normal Value	Frequency
220045	Heart rate	81.0 bpm	28.6%
220210	Respiratory rate	14.0 insp/min	28.3%
224639	Weight	81.3 kg	0.5%
220739	GSW eye opening	4 Spontaneously	7.1%
223900	GSW verbal response	5 Oriented	7.1%
223901	GSW motor response	6 Obeys Commands	7.1%
220051	Diastolic blood pressure	75.0 mmHg	5.9%
220052	Mean blood pressure	103.0 mmHg	5.9%
220621	Glucose	85.0	1.7%
220227	Oxygen saturation	96.8 %	0.2%
223761	Body temperature	98.6 °F	6.9%
223830	pH	7.4	0.7%

Table 1. Activities selected from the MIMIC-IV database

3.3 Preprocessing

Some essential preprocessing operations are performed to ensure the quality of the encoded event log. One major issue is the existence of missing values in the event log, especially in the 12-dim chart event features, since each activity has its own attribute to store observation results and not all activities occur at every timestamp. For instance, “Glucose" is the activity recorded at a specific timestamp and is the only attribute that has the value, while others are blank. In this context, we fill in the missing values with the latest recorded attribute value for this activity in the trace. If such results are not available, the mean of this activity’s attribute value is calculated and filled in the event log. The attribute is filled with –1 to indicate that this activity does not happen in this trace. Therefore, all dimensions are filled in values, which is an essential requirement for any supervised learning algorithm. We also remove several events with extreme attribute values since they have significant effects when performing min-max normalisation for each dimension. Since process monitoring utilises supervised learning algorithms, it is necessary to ensure that the input has the same size, which means the same length for each trace. We first remove traces with extremely small lengths (i.e. traces with a single event). Instead of using zero padding for short traces, we use Last-Event-Carried-Forward (LECF) imputation method inspired by (Zebin and Chaussalet, 2019), which keeps duplicating the last event until the trace reaches the certain length. Therefore, each trace is of the same length.

3.4 Deep Learning Models

In this paper, we propose three well-known models for PPM. We first propose a CNN model consists of a series of pairs of convolutional and max-pooling layers, since it has achieved satisfactory results in outcome-oriented PPM (Weytjens and De Weerdt, 2020). We first build a convolutional layer to filter the encoded event log into feature maps, followed by a max-pooling layer to summarise feature maps. The structure is used again to further extract features and reduce the number of network parameters. An increasing number of filters (16, 128) is applied to each convolutional layer. The computed feature maps are concatenated and fully connected to a dense decision layer with one output neuron.

LSTM has been widely applied in outcome-oriented PPM (Wang et al., 2019). LSTM has the ability to process sequences of data instead of a single data point, which makes it more suitable to work with time-series data compared to CNN. The bidirectional LSTM can preserve information from both the past and future, which allows it to learn from two directions. We build a single bidirectional LSTM layer with 64 neurons connected to a dense decision layer.

We also implement a structure that combines the above two models. This model has shown promising results in traditional readmission prediction (Zebin and Chaussalet, 2019). However, it has not previously been applied to PPM. The convolutional layer (i.e. with 16 filters) receives the output from the bidirectional LSTM layer (i.e. with 64 neurons), which transfers the sequence data to spatial feature maps to further understand the relationships between the different events in the log. A dense decision layer with a single output neuron is connected at the end to output the final result.

4 Benchmark

In this section, we benchmark the three models and a baseline approach using the encoded event log described in the previous section. To simulate the outcome prediction in real scenarios where prediction models are trained using historic traces and applied to ongoing traces, all traces are sorted based on start time (Teinemaa et al., 2019). A prediction point is selected, where the first 80 % of the event log are utilised for selecting and training the best model parameters and the remaining 20% are used for testing. In other words, the training utilises complete traces before a given time and the testing is performed afterwards on traces with different prefix lengths. All three deep learning models follow the same settings. The maximum number of epochs is 100 (Early stop is set to 10 epochs), the batch size is set to 16, and the learning rate is set to 0.01.

4.1 Evaluation Metrics

As the aim of PPM is to predict the outcome as early as possible, earliness is a measure to assess how early (in terms of length of the ongoing trace) the prediction performance can reach a certain level (Teinemaa et al., 2019). We apply the trained model to different prefixes in the testing set to explore how early we can predict unplanned ICU readmission. We set the prefix limitation to 21, the justification is that this prefix limitation is half of the mean trace length in the testing set and the majority of traces (i.e. over 83 %) are longer than 21 in the testing set. The accuracy calculates how many correct predictions are made out of all predictions. We also compute the area under the ROC curve (AUC) metrics that expresses the probability of a given classifier will rank a positive case higher than a negative one (Teinemaa et al., 2019).Both accuracy and AUC range from 0 to 1, where 1 means perfect predictions. We perform five-fold cross-validation to avoid bias and the average accuracy and AUC are reported for each prefix length.

Among several existing methods, the baseline method (Assaf and Jayousi, 2020) has shown to be of relatively high accuracy. It takes the minimum, maximum and mean of each chart event as features and utilises the pre-trained 300-dimension ICD embeddings from (Mikolov et al., 2013) to encode diagnoses. It is worth mentioning that the baseline approach uses all information in the testing set, while our approach makes predictions based on partial information in the testing set.

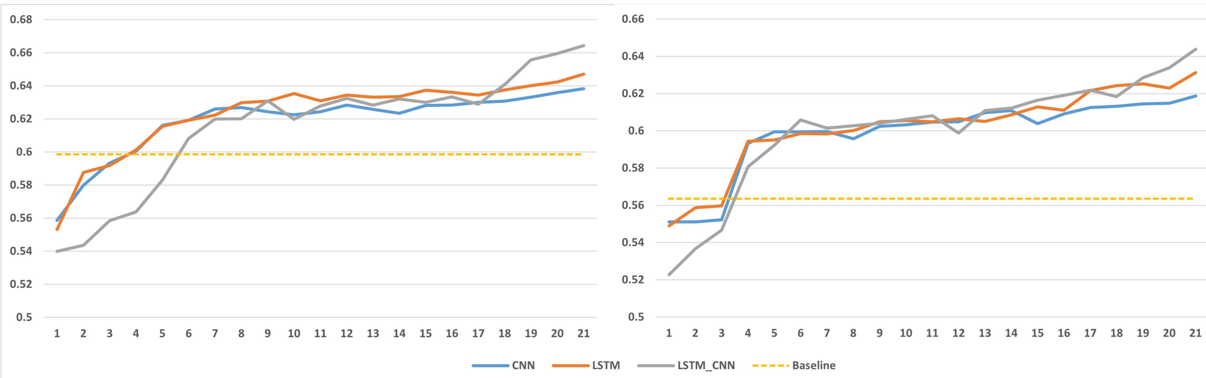

Figure 3. Accuracy (left) and AUC (right) comparisons with different deep learning models on different prefix lengths.

We report the accuracy and AUC comparisons of ICU readmission prediction across different implementations and the baseline approach on different prefix lengths in Figure 3. We evaluate the baseline approach using the completed testing set instead of prefixes since the baseline approach is not specifically designed for PPM, which displayed as constant lines in Figure 3. Our approach achieves worse performance than the baseline initially when the prefix lengths are short. This is consistent with our expectation that prediction performance will grow as the number of chart events increases. Moreover, unplanned ICU readmission is a complicated problem, given that the average trace length is 43 in the event log. The results also show that the LSTM+CNN model performs slightly worse when dealing with short prefixes. It is again likely due to insufficient time-series data for the convolutional layer to filter. The predictive performance progressive improves as prefix length increases and surpasses the baseline (i.e. after six events being considered for accuracy and four events for AUC). This demonstrates that our PPM method can make accurate predictions as early as the first four events in the trace. The LSTM model performs slightly better than the other two models when the prefix length is between 7 and 16, while the LSTM+CNN model outperforms towards the end. This is expected with increasing prefix lengths, since the LSTM+CNN model has better capacity when dealing with relatively large time-series data that contain many features. It also demonstrates the potential to apply the LSTM+CNN model to other outcome prediction tasks in process monitoring. If we adopt accuracy as the measurement to evaluate earliness, the earliness can achieve as early as four if the threshold is set to 0.6 (i.e. CNN and LSTM models). If the threshold is increased to 0.65, the earliness is delayed to 19 (i.e. LSTM+CNN model), which is still relatively ahead of time considering the average trace length.

5 Discussion and Conclusion

Although the prediction accuracy of our approach as demonstrated requires improvement, there is potential to apply process monitoring to predict unplanned ICU readmission. By viewing ICU stays as process traces, rich time-series information can be retrieved from the event log (such as the sequence of chart events and trends of observation results), which provides opportunities to predict readmissions with only short prefix lengths. The baseline approach uses statistical features of chart events from each ICU stay, while ignoring process information. Compared with the baseline approach, our approach can make predictions in the early stage of each ICU stay, which would give doctors additional perspectives to consider in their evidence-based decision adjustments before patients discharge from ICU and may reduce the likelihood of readmission. Hence, medical resources can be better planned and personalised to target patients predicted to be readmitted. The key idea of the proposed approach is to allow predictions can be progressively made during the patient stay in ICU. The underlying principle of the methods is to enable decision making at a particular time point with the most recent historical information (time length of previous 48 hours in our case) and let doctors to reassess the patient's chance of readmission based on predictions. For patients with less than 48 hours of historical data, we apply LECF imputation method to make the required length of the trace.

It is important to note that predicting readmission is a complex problem, as mentioned in Section 2. This paper presents a preliminary attempt to apply PPM to the readmission prediction problem based on the MIMIC-IV database. Certain improvements can further improve the performance. The first is the event log construction, where additional relevant features can be included to enhance predictive performance. For instance, MIMIC-IV contains other information relating to continuous infusions and intermittent administrations during ICU stays, which are potential indicators for readmission. Given that diagnoses are billed on hospital discharge and determined by trained persons who read signed clinical notes in the MIMIC-IV database (A. Johnson et al., 2021). Clinical notes are generated continuously during the patient's hospital stay. Certain differences may exist in diagnoses between admission and discharge for patients. The robustness of this study would have been increased if we could obtain the admission diagnoses from the MIMIC-IV database. In addition, we treat each ICU stay independently. However, some inter-trace features could contribute to the predictions. For example, the previous readmission times and intervals for each patient. Finally, model selection can be improved. We only utilised basic deep learning models for this preliminary attempt; more advanced and complex models may improve performance, such as the LSTM model with attention mechanism (Wang et al., 2019).

In this paper, we proposed an approach to predict unplanned ICU readmission through PPM and implemented three deep learning models. Compared with the previous studies, process monitoring views each ICU stay as a process trace and utilises the rich time-series information in event logs. Additionally, process monitoring is now shown to be used to provide decision support to doctors when patients are still in ICU, as process monitoring learns from historical complete traces and makes predictions for ongoing, incomplete traces. This paper presents interesting opportunities for future research on predicting unplanned ICU readmission using PPM technologies. Anticipating accuracy could be further improved by extracting more features in event logs and proposing more advanced deep learning algorithms. Future work might explore what features are more significant in predicting unplanned ICU readmission, as data richness and quality of EHR improves, these features can be of increased significance in predicting unplanned ICU admissions.

References

Assaf, R. and R. Jayousi (2020). "30-day Hospital Readmission Prediction using MIMIC Data." In: *2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT)*. IEEE, pp. 1–6.

Barbieri, S., J. Kemp, O. Perez-Concha, S. Kotwal, M. Gallagher, A. Ritchie, and L. Jorm (2020). "Benchmarking deep learning architectures for predicting readmission to the ICU and describing patients-at-risk." *Scientific reports* 10 (1), 1–10.

Desautels, T., R. Das, J. Calvert, M. Trivedi, C. Summers, D. J. Wales, and A. Ercole (2017). "Prediction of early unplanned intensive care unit readmission in a UK tertiary care hospital: a cross-sectional machine learning approach." *BMJ open* 7 (9), e017199.

Di Francescomarino, C., M. Dumas, F. M. Maggi, and I. Teinemaa (2016). "Clustering-based predictive process monitoring." *IEEE transactions on services computing* 12 (6), 896–909.

Durbin Jr, C. and R. F. Kopel (1993). "A case-control study of patients readmitted to the intensive care unit." *Critical care medicine* 21 (10), 1547–1553.

Elixhauser, A., C. Steiner, D. R. Harris, and R. M. Coffey (1998). "Comorbidity measures for use with administrative data." *Medical care*, 8–27.

Golmaei, S. N. and X. Luo (2021). "DeepNote-GNN: predicting hospital readmission using clinical notes and patient network." In: *Proceedings of the 12th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*, pp. 1–9.

Huberts, L. C., R. J. Does, B. Ravesteijn, and J. Lokkerbol (2021). "Predictive monitoring using machine learning algorithms and a real-life example on schizophrenia." *Quality and Reliability Engineering International*.

Inan, T. T., M. B. R. Samia, I. T. Tulin, and M. N. Islam (2018). "A Decision Support Model to Predict ICU Readmission through Data Mining Approach." In: *PACIS*, p. 218.

Johnson, A., L. Bulgarelli, T. Pollard, S. Horng, L. A. Celi, and R. Mark (Mar. 2021). *MIMIC-IV*. URL: `https://doi.org/10.13026/s6n6-xd98`.

Kratsch, W., J. Manderscheid, M. Röglinger, and J. Seyfried (2020). "Machine learning in business process monitoring: a comparison of deep learning and classical approaches used for outcome prediction." *Business & Information Systems Engineering*, 1–16.

Leontjeva, A., R. Conforti, C. Di Francescomarino, M. Dumas, and F. M. Maggi (2016). "Complex symbolic sequence encodings for predictive monitoring of business processes." In: *International Conference on Business Process Management*. Springer, pp. 297–313.

Lin, Y.-W., Y. Zhou, F. Faghri, M. J. Shaw, and R. H. Campbell (2019). "Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long short-term memory." *PloS one* 14 (7), e0218942.

Maggi, F. M., C. Di Francescomarino, M. Dumas, and C. Ghidini (2014). "Predictive monitoring of business processes." In: *International conference on advanced information systems engineering*. Springer, pp. 457–472.

Metzger, A., P. Leitner, D. Ivanović, E. Schmieders, R. Franklin, M. Carro, S. Dustdar, and K. Pohl (2014). "Comparing and combining predictive business process monitoring techniques." *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 45 (2), 276–290.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean (2013). "Distributed representations of words and phrases and their compositionality." In: *Advances in neural information processing systems*, pp. 3111–3119.

Moore, B. J., S. White, R. Washington, N. Coenen, and A. Elixhauser (2017). "Identifying increased risk of readmission and in-hospital mortality using hospital administrative data." *Medical care* 55 (7), 698–705.

Munoz-Gama, J., N. Martin, C. Fernandez-Llatas, O. A. Johnson, M. Sepúlveda, E. Helm, V. Galvez-Yanjari, E. Rojas, A. Martinez-Millana, D. Aloini, et al. (2022). "Process mining for healthcare: Characteristics and challenges." *Journal of Biomedical Informatics* 127, 103994.

Pakbin, A., P. Rafi, N. Hurley, W. Schulz, M. H. Krumholz, and J. B. Mortazavi (2018). "Prediction of ICU readmissions using data at patient discharge." In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, pp. 4932–4935.

Pijnenborg, P., R. Verhoeven, M. Firat, H. van Laarhoven, and L. Genga (2021). "Towards Evidence-Based Analysis of Palliative Treatments for Stomach and Esophageal Cancer Patients: a Process Mining Approach." In: *2021 3rd International Conference on Process Mining (ICPM)*. IEEE, pp. 136–143.

Shahrokh Esfahani, M. and E. R. Dougherty (2014). "Effect of separate sampling on classification accuracy." *Bioinformatics* 30 (2), 242–250.

Teinemaa, I., M. Dumas, M. L. Rosa, and F. M. Maggi (2019). "Outcome-oriented predictive process monitoring: Review and benchmark." *ACM Transactions on Knowledge Discovery from Data (TKDD)* 13 (2), 1–57.

Teo, K., K. W. Lai, C. W. Yong, B. Pingguan-Murphy, J. H. Chuah, and C. A. T. Tee (2020). "Prediction of Hospital Readmission Combining Rule-based and Machine Learning Model." In: *2020 International Computer Symposium (ICS)*. IEEE, pp. 352–355.

Van der Aalst, W. M., M. H. Schonenberg, and M. Song (2011). "Time prediction based on process mining." *Information systems* 36 (2), 450–475.

Walraven, C. van, P. C. Austin, A. Jennings, H. Quan, and A. J. Forster (2009). "A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data." *Medical care*, 626–633.

Wang, J., D. Yu, C. Liu, and X. Sun (2019). "Outcome-oriented predictive process monitoring with attention-based bidirectional LSTM neural networks." In: *2019 IEEE International Conference on Web Services (ICWS)*. IEEE, pp. 360–367.

Weytjens, H. and J. De Weerdt (2020). "Process outcome prediction: CNN vs. LSTM (with attention)." In: *International Conference on Business Process Management*. Springer, pp. 321–333.

Xu, H., J. Pang, X. Yang, M. Li, and D. Zhao (2020). "Using predictive process monitoring to assist thrombolytic therapy decision-making for ischemic stroke patients." *BMC Medical Informatics and Decision Making* 20 (3), 1–10.

Xue, Y., D. Klabjan, and Y. Luo (2019). "Predicting ICU readmission using grouped physiological and medication trends." *Artificial intelligence in medicine* 95, 27–37.

Zebin, T. and T. J. Chaussalet (2019). "Design and implementation of a deep recurrent model for prediction of readmission in urgent care using electronic health records." In: *2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE, pp. 1–5.