Combining Process Mining and Time Series Forecasting to Predict Hospital Bed Occupancy

Pieters, Annelore Jellemijn; Schlobach, Stefan

published in
Health Information Science
2022

DOI (link to publisher)
10.1007/978-3-031-20627-6_8

document version
Publisher's PDF, also known as Version of record

document license
Article 25fa Dutch Copyright Act

Combining Process Mining and Time Series Forecasting to Predict Hospital Bed Occupancy

Annelore Jellemijn Pieters and Stefan Schlobach(✉)

Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
k.s.schlobach@vu.nl

Abstract. This research investigates in how far AI methods can support the prediction of bed occupancy in hospital units based on individual patient data. We combine process mining and a Deep Spatial-Temporal Graph Modeling algorithm and show that this improves the performance of the prediction over existing approaches. To improve the model even more it is extended with knowledge available from patient records, like the day of the week, the time of the day, whether it is a vacation day or not and the amount of emergency cases per data point.

1 Introduction

Hospitals are more and more dealing with waits, delays and cancellations [8]. Sometimes hospitals can also get overflows of patients, which means that a hospital cannot handle the number of patients coming into that hospital anymore. Those problems used to be solved by adding more resources [10], such as beds, equipment or more staff. Nowadays those problems cannot be fixed that easily by adding resources for economic reasons. Moreover, research states that adding resources is not the best solution to the problem: instead literature shows that the focus should be on analysing how the patient moves through a hospital and optimising the capacity of bed occupancy.

When a patient goes to a hospital (s)he will follow a certain care pathway, e.g. starting the pathway with the emergency room for a broken leg. Hereafter the patient might need surgery, going to the operating rooms. After surgery (s)he will go to the post-anesthesia care unit to recover, and so the pathway grows, creating a so called *patient flow* through the hospital. According to research understanding this patient flow is crucial for reducing the overflows in a hospital [4,9,10,24].

Along with analysing the patient flows, the bed capacity management needs to be optimised as well. In a hospital there are a certain number of beds, though not always all the beds can be used. This can have several reasons, such as the absence of employees. For the hospital that we focused on, the beds are distributed over seventeen different departments, such as Day Care Unit for cancer, the General Care Unit hall B, the General Care Unit, the Post Anesthesia

A. Traina et al. (Eds.): HIS 2022, LNCS 13705, pp. 76–87, 2022.
https://doi.org/10.1007/978-3-031-20627-6_8

Care Unit level 3 or the Sleep Awake Unit, etc. The number of beds used on a day is determined by the number of employees working on that department that specific day. This means that each department has a number of beds available for patients, which varies per day.

To address this problem, this paper attempts to answer whether we can predict the bed occupancy of a hospital based on the previous and current occupancy? To answer this main research question, this study is divided into four sub-questions:

1. How to use Artificial Intelligence to find patient flows from patients in a hospital?
2. What methods perform best to predict bed occupancy of a hospital?
3. Can we improve existing bed occupancy perdition by combining process mining with a state-or-the-art forecasting algorithm?
4. Can we extend the forecasting model with domain knowledge in order to improve the performance of predicting bed occupancy?

Section 2 will present the relevant literature. Next, Sect. 3 describes the data that is used for this research, before the methodology of the different techniques is presented in Sect. 4. The experimental setup and the corresponding results are described in Sect. 5. The paper concludes in Sect. 6 with a discussion of the contributions and some recommendations, as well as the limitations of this research.

2 Theoretical Framework

Capacity Management. To get a grip on the occupation of the beds in a hospital, a hospital makes use of capacity management, which ensures that its capacity is used in an efficient and effective manner. There are three different kinds of capacity in a hospital; 1. Medical specialists, 2. Supporting personnel (e.g. nurses), 3. Resources (e.g. equipment or rooms). This paper focuses on resource capacity.

To use the capacity in a hospital efficiently, one needs to take into account the capacity management triangle. This triangle considers the variability, the capacity and the service times [13]. The variability includes fluctuations in the arrival times of a patient, the availability of the capacity and the duration of an appointment. The capacity means that there is a certain capacity level and utilization rate. The service times include waiting times, cancellations etc. Besides the triangle, there are four elements that put pressure on the capacity management; 1. Financial, 2. Institutional and social, 3. Clinical, 4. Professional [23]. By taking all of the mentioned above into account, the capacity management critically contributes to the effectiveness of a hospital.

Patient Flow. Investigating patient flows in hospitals is an important challenge for capacity management of hospitals. According to Walley and Steyn, around 60/70% of the patients in a hospital bed is getting active treatment and 40/30% of the patients in a hospital bed are waiting to see a doctor or they are waiting to get a sign that they are allowed to go home [26]. This poor patient flow is accompanied by eight problems according to Villa, Barbieri and Lega [23]: 1. There are not enough supplies, 2. Patients must wait for a long time, 3. Bottlenecks, 4. Resources that are not used in an efficient way, 5. Patients need to stay for a long time in hospitals, 6. Not a lot of productivity, 7. Clinical settings that are used inappropriate, 8. The variability of the workload.

A successful tool developed by Domova et al. visualises the patient flows in order to help the hospital improve and optimise the services that they offer [5]. Analysing patient flows is also successfully used for hospital planning [11] to improve the quality of the healthcare network. Another research that uses patient flows is a simulation model that is used to analyse the patient flows in a hospital in Hong Kong [20].

Forecasting Methods. Most current methods to predict the number of occupied beds in hospital use Support Vector Machines (SVM) [3,7], claiming that SVM are among the best algorithms for bed occupancy prediction in hospitals. More recently, neural networks are used for predictions [6,19,25], outperforming older algorithms they used. However, in those articles the research focuses on all departments as individuals, ignoring the relations and flows between the different departments. In this study, we considered a different prediction model, called Deep Spatial-Temporal Graph Modeling algorithm, which is currently used to predict traffic flows [28], and applied it to predict bed occupancy.

3 Data

For this study, data from a Dutch hospital was used. For every patient the actions or treatments that (s)he underwent is saved in an Electronic Health Record (EHR) The EHR software that is used by this hospital is HiX (Healthcare Information eXchange).

The dataset is retrieved from the EHR system with an SQL query that states: which bed is occupied at which moment for the last six years and by which patient. From this view an overview was created with two timestamps a day, 10:00 and 16:00. For every occupied bed there is more information about that patient collected, such as whether the patient was an emergency case, the intake type (for a day or longer) and the patient origin (from home, another hospital or somewhere else). First we performed data preprocessing and data cleaning based on an exploratory data analysis.

Exploratory Data Analysis. In the original dataset the first timestamp in the data is January, 1st 2016 and the last timestamp February, 11, 2022. On each day at 10:00 and at 16:00 an entry is created of which beds are occupied. This results in 857418 observations for 2.234 days. Because of relocation, with new departments being created, and missing values (patients being still in the hospital) the time range of the dataset is restricted to a range between 29-01-2017 and 01-01-2022.

We can see that there are 115971 different hospitalizations, for 61261 different patients. This means that some patients are hospitalized multiple times in this period. After removing departments that are not of interest for this research from the dataset, we obtain 110236 different hospitalizations for 60431 different patients. Hereafter, we are deleting missing values and convert NaN values to zero values if there are no patients in the hospital at that time.

The next step is to analyse the time series data. First we analysed trends, or pattern that can be found in the data. The second component is seasonality, which means that there are (predictable) changes which occur every year. Figures 1, 2 and 3 show original data, the trend and the seasonality. In Fig. 2 we can, first, see that in March 2020, the COVID-19 pandemic started, because at this point large fluctuations in the number of beds that are occupied starts.[1] Moreover, holidays in the Netherlands are visible in the Figure, leading to a decreased number of occupied beds.[2] Figure 3 shows seasonality in the data. The seasonality in this data is that during weekdays the number of beds that are occupied are higher than during the weekend. This is the case every week. This can be explained since there are no appointments or surgeries scheduled during the weekends.

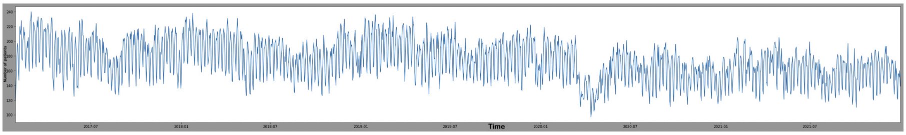

Fig. 1. Plot with the original data

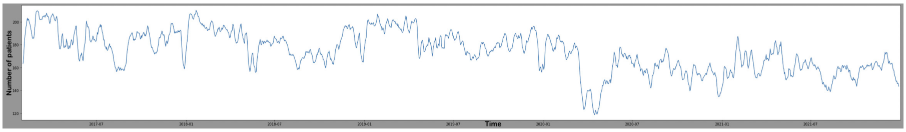

Fig. 2. Plot with the trend of the data

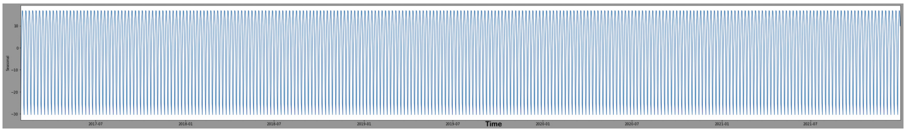

Fig. 3. Plot with the seasonality of the data

[1] Despite these fluctuations, we still chose to keep the data from when the pandemic started in the dataset, since COVID-19 is not completely gone yet, it is still possible that there are patients with COVID-19 in the hospital.

[2] We can see that during the Christmas holidays at the end of the year there is a decrease in occupied beds in the hospital. This is recurring every year.

4 Methods

The goal of this research is: Can we predict the bed occupancy of a hospital based on the currency occupancy? To do this, first the patient flow method is described, to solve the problem of finding the patient flows of the patients in a hospital. Then, we take a closer look at forecasting algorithms to find the bnest method to predict bed occupancy of a hospital.

4.1 Process Mining of Patient Flows

To determine the patient flow we apply process mining. Process mining is a method that is used to extract processes and insights from event logs [22]. Event logs are defined as: *"A set of process instances or traces, each of which contains a set of events. Events represent occurrences of process tasks or activities at a certain point in time"* [18]. The input for a process mining algorithm needs to have three different elements [17], a unique identifier, a timestamp and a name or description of the event.

In order to analyse the data using process mining, there are three different techniques which are comprised by process mining [21], 1) Process discovery, which discovers the processes behind the data, 2) Conformance checking which checks if the model conforms the event log data, and 3) Process enhancement, which tries to improve the process models. For this research we focus on process discovery, which takes event logs as input and creates a model without any a priori information. There are several different algorithms for process discovery [21].

1. Alpha Miner finds processes based on relations and causalities in a process,
2. the heuristic miner finds common processes and handles noises, and
3. the inductive miner, a sound algorithm that finds common processes.

The process models that are created using those algorithms from the event log data can be visualised in a directly follow graph (DFG) [14]. In the graph, each activity is visualised as a node and each relation is visualised as an edge between the nodes.

4.2 Forecasting Algorithm

Support Vector Machines. We use SVMs for regression in a variant called SVR, which is supervised machine learning algorithm based on finding optimal hyperplanes in a n-dimensional space, where n stands for the number of features or independent variables. By adapting the hyperplanes to a maximum margin, the error is reduced.

Deep Spatial-Temporal Graph Modeling. A forecasting method that takes into account the relations between the different hospital units is the Deep Spatial-Temporal Graph Modeling algorithm. Time series forecasting takes data of observations over time into account [2] and predicts values based on a series of historical observations. This algorithm takes a graph structure as input and takes spatial and temporal information into account [28]. The Deep Spatial-Temporal Graph Modeling model creates an adjacency matrix which will develop over time through training and node embedding. For an unweighted matrix the values are 1 or 0, where 1 means that the elements are connected and 0 means that the elements are not connected [27]. The weights in a matrix represent e.g. the strength of an edge (or 0 if there is no link).

A graph neural network architecture is used to model the different node-levels which are dynamic with the help of the relations between the different nodes. The model is heavily used in forecasting traffic speed flows [15]. To capture the relation between the spatial variables and the temporal variable they formally used graph convolution networks (GCN) and transformed it into recurrent neural networks (RNN).

For our research a different approach has been used, consisting of multiple spatial-temporal layers and one output layer [28]. Before the data is entering a spatial-temporal layer, the input data is transformed by a linear layer. After the linear layer the data goes on to the Spatial-Temporal layers, in which the first step is the Gated Temporal Convolution Module (Gated TCN), containing gating mechanisms where complex temporal dependencies in the data are learned. The Graph Convolution Layer (GCN) works together with the previous layer, the TCN layer, to keep the spatial-temporal dependencies of the data. Each of the Spatial-Temporal layers is connected with the output layer by skip connections [16].

Another component that is used is a stacked dilated 1D convolution component, this makes it possible to create a receptive field which can keep track of the amount of layers and increase exponentially with these number of layers. By using this algorithm the relation between the spatial and the temporal components can be observed.

As a baseline we use a majority baseline, which always predicts the most frequent class label in the dataset.

5 Experiments and Results

5.1 Modeling Hospital Patient Flows

In order to find the patient flows using process mining we used the Process Mining for Python (PM4Py) framework. The heuristic miner algorithm of this package is used acting on the Directly-Follows Graph and in this way the most common processes are found. For this part of the experimental setup the heuristic miner will be used to obtain the patient flows and visualise them. This algorithm is chosen since this algorithm has the goal to obtain the most common flows, which is what we want to obtain.

Results. Figure 4 shows the most common 0.0001% of the patient flows from the (whole) hospital. This 0.0001% is chosen as otherwise the graph would be too big to analyse, based on parameter tuning to keep overview small enough, while still providing an idea of how many different patient flows there are.

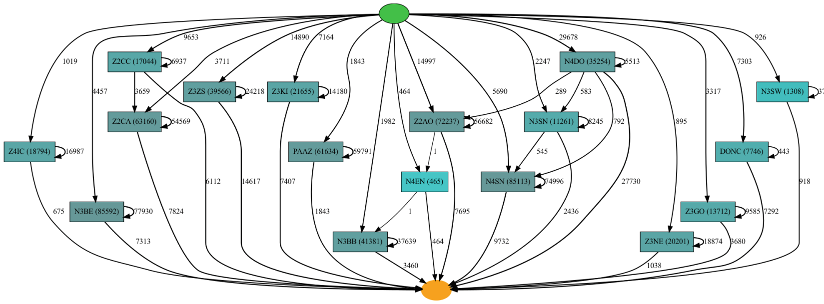

Fig. 4. 0.0001% most common patient flows in a hospital

The most common patient flows, are when patients are moving to another bed on the same department. A second movement that frequently occurs is the flow from the Cardiac Care Unit (Z3CC) to the Cardiology (Z2CA) unit. Another patient flow that occurs often is patients that move from the Day Care Unit (N4DO) to the Post Anesthesia Care Unit (N4SN).

Next to visualising the data we can also get a table with the most common patient flows in a hospital. An observation is that the most frequent flows in the department are from a department to itself, as visualised in Table 1. However also a movement from the Cardiac Care Unit to the Cardiology unit happens often in this timeframe.

Table 1. The 5 most common patient flows between two units

Number	Patient flows
1	Day Care Unit –> Day Care Unit
2	Care Suites Unit –> Care Suites Unit
3	Acute Medical Unit –> Acute Medical Unit
4	Children Ward –> Children Ward
5	Post Anesthesia Care Unit level 4 –> Post Anesthesia Care Unit level 4

Table 1 shows the place in the top 5 and the respective patient flow in the chosen time period. We can see that the patient flow is from Day Care Unit to Day Care Unit, which means that the patient is laying on the Day Care Unit and is moving to another bed in the time-frame. This happens 5062 times in almost five years, which is the highest number of all the patient flows.

5.2 Forecasting Bed Occupancy

In order to investigate which forecasting algorithm is performing best in predicting bed occupancy in a hospital compared Support Vector Regression (SVR) with the Deep Spatial-Temporal Graph Modeling algorithm and the baseline.

For the SVR algorithm we use default settings, including an rbf kernel, the gamma is scale and the probability is False. With the help of feature selection, we use features: the hour of the day, emergency or not, intake type (e.g. is it for a day or longer), and origin (e.g. from home or the emergency room) and the timestamp.

For the Deep Spatial-Temporal Graph Modeling algorithm we also used the default settings, i.e. only knowledge about the hour of the day. The sequence length is set to seven[3]. The algorithm is trained for 100 iterations.

The data is divided into a training set (70%) and a test set of 30% of the data. This experiment were evaluated by comparing Mean Absolute Error (MAE) [1] and and Root Mean Squared Error (RMSE) [12][4].

For all values that are predicted for each datapoint of the test set, the MAE and the RMSE are calculated. The mean of all those error scores is calculated in order to find the error rate for each different algorithm. Each experiment is repeated 20 times and the mean of the MAE and RMSE is calculated.

Results. In this subsection we show the results of the comparison between the SVR algorithm, the baseline and the Deep Spatial-Temporal Graph Modeling algorithms. In Table 2 the results of the forecasting with the different algorithms are shown. Table 2 shows that the baseline, non surprisingly, that always predicts random values has the highest MAE and RMSE scores. The second highest scores for the MAE and RMSE are obtained by the SVR algorithm. This means that the SVR algorithm is performing the second worst. Furthermore can we see that the values for the MAE and RMSE scores for the majority baseline are better than SVR, but worse than the mean Deep Spatial-Temporal Graph Modeling. This means that the Deep Spatial-Temporal Graph Modeling algorithm has the lowest error values for the MAE and RMSE and is performing the best.

Table 2. Different forecasting algorithms predicting the bed occupancy

	Mean MAE	Mean RMSE
SVR	2.7186	3.3218
Deep Spatial-Temporal Graph Modeling	**1.9189**	**2.5711**
Baseline majority	2.5133	2.9445

[3] A prediction looks at the number of patients of the last week (seven days).

[4]

$$MAE = (\frac{1}{n})\sum_{i=1}^{n}|F_t - A_i| \qquad RMSE = \sqrt{(\frac{1}{n})\sum_{i=1}^{n}(F_i - A_i)^2} \quad (1)$$

Where n is the number of observation, F_t is the predicted output value, A_t is the actual output value and t is the time point the value is predicted for.

5.3 Combining Patient Flows with A Forecasting Model

We now create a model that combines process mining and the Deep Spatial-Temporal Graph Modeling algorithm. To investigate this model different inputs for the Deep Spatial-Temporal Graph Modeling algorithm will be used, adjusting the adjacency matrix based on what was learned during process mining.

We compared four different inputs for the model, an empty adjacency matrix, where all values in the matrix are set to zero. The second input is an adjacency matrix with all values set to zero except for the diagonal row, those values are filled with ones. This diagonal input is chosen because of frequent flows within departments. The third input is an adjacency matrix filled with random values differing between zero and one. The fourth and last input is an adjacency matrix filled with the input retrieved from the process mining method. To find the correlations between the different departments, we divided the amount of occurrences of a flow by the total number of occurrences of all flows.

Results. Table 3 shows the comparison between the combination of process mining and Deep Spatial-Temporal Graph Modeling and models with other input.

Table 3. Comparison of various deep spatial-temporal graph modeling algorithm

	MAE	RMSE
Empty matrix	2.0121	2.6852
Diagonal matrix	1.9841	2.6548
Random matrix	1.9441	2.5967
Process mining matrix	**1.9189**	**2.5711**

Table 3 shows that when the process mining values are used as input for the model, the model has the lowest error values for the MAE and RMSE evaluation metrics and thus performs the best.

5.4 Model Extension with Knowledge

To extend the model created with the combination of process mining and Deep Spatial-Temporal Graph Modeling, we investigated if adding knowledge to the model improves the performance of the model. A set of interviews with hospital employees provided the following factors influencing the number of patients: 1. The time of the day, 2. The day in the week, 3. Vacation periods and 4. The amount of emergency cases on a day. Each of these statements was added to the model individually and in every possible combination.

Table 4 shows that the values are lower, when the knowledge about the time of day is added to the model. Besides this, we can see that the error is the lowest if the knowledge about the time in the day, the day in the week and vacation is added to the model.

Table 4. Model extension with knowledge

	MAE	RMSE
Time of the day	1.9189	2.5711
Day in the week	2.2165	2.8973
Vacation	2.3159	3.0507
Emergency	2.2215	2.9045
Time of the day - Day in the week - Vacation	**1.8564**	**2.4812**
Time of the day - Day in the week - Emergency	1.8928	2.5197
Day in the week - Vacation - Emergency cases	2.2342	2.8990
Time of the day - Day in the week - Vacation	1.9009	2.5282

6 Discussion

Conclusion. In this paper we have shown that Process Mining can be used in order to find the patient flows from the patients. By using a heuristic miner the patient flows can be analysed and visualised. We tested four different algorithms for predicting the bed occupancy of a hospital, showing that a Deep Spatial-Temporal Graph Modeling algorithm is performing best to predict the bed occupancy. Moreover, a model is made that combines process mining with Deep Spatial-Temporal Graph Modeling forecasting. Table 3 shows that this model has the lowest error values and thus is performing the best. Finally, we extended the model with domain knowledge and showed that adding time of the day always improves the model, and if this statement is missing, the results are getting worse. There is also found that overall the more statements added to the model, the better the model is performing. However, the combination of time of the day, day in the week and vacation day or not, is performing the best. We can conclude that the best way to predict hospital occupancy is by combining the patient flow algorithm with the Deep Spatial-Temporal Graph Modeling algorithm and extend this with knowledge about the time of the day, the day in the week and if it is a vacation day or not.

Limitations and Recommendations. While the literature suggests that SVR is the best performing method for bed prediction in hospitals, there has been no approach using Deep Spatial-Temporal Graph Modeling algorithm as of yet. However, there are limitations to this research, such as the time range of the data, which included the start of COVID-19. This results into big fluctuations the bed occupancy. However, we still chose to keep this data in the dataset, since COVID-19 is still in our lives and may still be there for a long period of time.

Another limitation is that this research is only performed on one dataset. For future research one can investigate this research on multiple datasets of different hospitals to see if the model created in this research, can also be utilized on different datasets.

Initially within this research we wanted to extend the model with knowledge about the diagnosis of each patient. However, access to more structured data (such as diagnostic knowledge using, e.g., ICD10 was not available at the time of this research). For future research this might be very interesting to focus on.

References

1. Akdemir, B., Çetinkaya, N.: Long-term load forecasting based on adaptive neural fuzzy inference system using real energy data. Energy Procedia **14**, 794–799 (2012)
2. Chatfield, C.: Time-Series Forecasting. Chapman and Hall/CRC (2000)
3. Daghistani, T.A., Elshawi, R., Sakr, S., Ahmed, A.M., Al-Thwayee, A., Al-Mallah, M.H.: Predictors of in-hospital length of stay among cardiac patients: a machine learning approach. Int. J. Cardiol. **288**, 140–147 (2019)
4. Dart, T., Cui, Y., Chatellier, G., Degoulet, P.: Analysis of hospitalised patient flows using data-mining. In: Studies in Health Technology and Informatics, pp. 263–268 (2003)
5. Domova, V., Sander-Tavallaey, S.: Visualization for quality healthcare: patient flow exploration. In: 2019 IEEE International Conference on Big Data (Big Data), pp. 1072–1079. IEEE (2019)
6. Gentimis, T., Ala'J, A., Durante, A., Cook, K., Steele, R.: Predicting hospital length of stay using neural networks on MIMIC III data. In: 2017 IEEE 15th International Conference on Dependable, Autonomic and Secure Computing, 15th International Conference on Pervasive Intelligence and Computing, 3rd International Conference on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), pp. 1194–1201. IEEE (2017)
7. Hachesu, P.R., Ahmadi, M., Alizadeh, S., Sadoughi, F.: Use of data mining techniques to determine and predict length of stay of cardiac patients. Healthc. Inform. Res. **19**(2), 121–129 (2013)
8. Hall, R.: Patient flow. AMC **10**, 12 (2013)
9. Hanne, T., Melo, T., Nickel, S.: Bringing robustness to patient flow management through optimized patient transports in hospitals. Interfaces **39**(3), 241–255 (2009)
10. Haraden, C., Resar, R.: Patient flow in hospitals: understanding and controlling it better. Front. Health Serv. Manag. **20**(4), 3 (2004)
11. Jay, N., Kohler, F., Napoli, A.: Using formal concept analysis for mining and interpreting patient flows within a healthcare network. In: Yahia, S.B., Nguifo, E.M., Belohlavek, R. (eds.) CLA 2006. LNCS (LNAI), vol. 4923, pp. 263–268. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-78921-5_19
12. Kalogirou, S.A.: Solar thermal systems: components and applications-introduction. Elsevier Ltd. (2012)
13. Klassen, R.D., Menor, L.J.: The process management triangle: an empirical investigation of process trade-offs. J. Oper. Manag. **25**(5), 1015–1034 (2007)
14. Leemans, S.J., Poppe, E., Wynn, M.T.: Directly follows-based process mining: exploration and a case study. In: 2019 International Conference on Process Mining (ICPM), pp. 25–32. IEEE (2019)
15. Li, Y., Yu, R., Shahabi, C., Liu, Y.: Diffusion convolutional recurrent neural network: data-driven traffic forecasting. arXiv preprint arXiv:1707.01926 (2017)
16. Liu, F., Ren, X., Zhang, Z., Sun, X., Zou, Y.: Rethinking skip connection with layer normalization. In: Proceedings of the 28th International Conference on Computational Linguistics, pp. 3586–3598 (2020)

17. Moeke, D.: De meerwaarde van process mining bij het optimaliseren van de patient flow (2021)
18. de Murillas, E., Reijers, H.A., van der Aalst, W.M.: Case notion discovery and recommendation: automated event log building on databases. Knowl. Inf. Syst. **62**(7), 2539–2575 (2020). https://doi.org/10.1007/s10115-019-01430-6
19. Pofahl, W.E., Walczak, S.M., Rhone, E., Izenberg, S.D.: Use of an artificial neural network to predict length of stay in acute pancreatitis. Am. Surg. **64**(9), 868 (1998)
20. Rado, O., Lupia, B., Leung, J.M.Y., Kuo, Y.-H., Graham, C.A.: Using simulation to analyze patient flows in a hospital emergency department in Hong Kong. In: Matta, A., Li, J., Sahin, E., Lanzarone, E., Fowler, J. (eds.) Proceedings of the International Conference on Health Care Systems Engineering. SPMS, vol. 61, pp. 289–301. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-01848-5_23
21. Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Process mining in healthcare: a literature review. J. Biomed. Inform. **61**, 224–236 (2016)
22. Van Der Aalst, W.: Process mining. Commun. ACM **55**(8), 76–83 (2012)
23. Villa, S., Barbieri, M., Lega, F.: Restructuring patient flow logistics around patient care needs: implications and practicalities from three critical cases. Health Care Manag. Sci. **12**(2), 155–165 (2009). https://doi.org/10.1007/s10729-008-9091-6
24. Villa, S., Prenestini, A., Giusepi, I.: A framework to analyze hospital-wide patient flow logistics: evidence from an Italian comparative study. Health Policy **115**(2–3), 196–205 (2014)
25. Walczak, S., et al.: Predicting hospital length of stay with neural networks. In: FLAIRS Conference, pp. 333–337 (1998)
26. Walley, P., Silvester, K., Steyn, R., Conway, J.B.: Managing variation in demand: lessons from the UK national health service/practitioner application. J. Healthc. Manag. **51**(5), 309 (2006)
27. Weisstein, E.W.: Adjacency matrix (2007). https://mathworld.wolfram.com/
28. Wu, Z., Pan, S., Long, G., Jiang, J., Zhang, C.: Graph WaveNet for deep spatial-temporal graph modeling. arXiv preprint arXiv:1906.00121 (2019)