Association for Information Systems

AIS Electronic Library (AISeL)

International Conference on Information Systems 2020 Special Interest Group on Big Data Proceedings

Special Interest Group on Big Data Proceedings

12-14-2020

Forecasting Heterogeneous Patient Flow through Big Data Application in Medical Facilities for Rational Staffing

Elizaveta Prokofyeva

Maltseva Svetlana

Dmitry Tsiu-Zhen-Tsin

Forecasting Heterogeneous Patient Flow through Big Data Application in Medical Facilities for Rational Staffing

Short Paper

Prokofyeva, Elizaveta S., National Research University Higher School of Economics, company ANO Moi Raion, Department of Industry Projects Based on Big Data Technologies, Moscow, Russia, prokofyeva.liza@gmail.com

Prof. Maltseva, Svetlana V., National Research University Higher School of Economics, Moscow, Russia, smaltseva@hse.ru

Tsiu-zhen-tsin, Dmitry, National Research University Higher School of Economics, Moscow, Russia, dmitry.tsyuzhentsin@gmail.com

Abstract

An approach to automating resource management of a service organization based on simulation modeling integrating the predicted input flows of patients, is considered. Methods for predicting input flows based on SARIMA, Holt-Winters, LSTM, and controlled recurrent GRU models have been investigated. The results of a computational experiment on predicting patient flows in a medical organization are presented. Based on the results, a meta-algorithm for forecasting the input flow and its further integration into the simulation model of the service process of a multidisciplinary healthcare organization was developed.

Keywords: time series forecasting, patient flow, simulation and modelling, healthcare, organization structure

Introduction

Healthcare services are determined by the heterogeneous patient flow at different times of the day, as well as due to various emergencies, for example, epidemics, when the demand for medical services increases dramatically. Over the last months, thousands of people around the world have been affected by the Coronaviruses disease (COVID-19), which was declared a pandemic by the world health organization (Neher et al.,2020). Consequently, local healthcare systems in many countries had to adapt their organization of resource capacities to massive new patient flow. Rapidly increasing demands for the healthcare services require novel methods and approaches of capacity optimization to balance between meeting patient demand and the cost of providing an appropriate high level of service availability.

Optimizing resource utilization in modern healthcare facilities can lead to more effective work organization: data- driven decision about scheduling policies, patient flow characteristics, planning and distribution of resources, reducing queues and other tasks. Modern analytical technologies allow the development of decision-making tools based on empirical data in healthcare domain. For example, depersonalized data on actual movements of patients between medical units and specialists within these institutions allows management to plan the load of resources, ensure a high level of accessibility of services and optimize the work of the organization based on real demand for these services. Based on such data, patient flow management information systems and various services are being actively developed.

The task of resource management is fundamental for the effectiveness of service organizations. According to Guide to the Project Management Body of Knowledge, resource management is aimed at efficient and effective development of resources in organization. The role of automation in resource management is critical, as it provides meaningful analysis and distribute the load among of these resources optimally.

The active application of modern methods of big data processing and big data analysis significantly increases the quality of resource management automation in modern organizations. In addition, data collected by organization serves as a basis for predicting the required performance indicators. The use of predictive values, for instance, customer visits, in resource management is the basis for the formation of an adaptive organizational model and optimization of the resource usage.

The staffing strategy of service organizations is becoming more and more urgent as an important part of resource management. The problem of rational staffing is associated with the unevenness of the input flows of customers at different times of the day, in different periods of the year (the classic factor of the seasonality of demand), the peculiarities of the demographic situation of the area. Also, due to various emergency situations, for example, epidemics, when the demand for some services increases dramatically, decisions regarding the staffing of organizations must be made almost in real time.

By definition (Ahmady, Mehrpour and Nikooravesh, 2016), organizational structure is a way of dividing, organizing, and coordinating the activities of a company to control the productivity of the team. By modelling approach, the organizational structure might be represented as a set of interrelated elements aimed at ensuring the functioning and development of the organization (Kabanovskij, 2011).

Simulation modeling is used to solve the problems of choosing a rational organization structure when several criteria are applied. Simulation models enable one to analyzing the features of complex organizational structures (Akopov, 2017; Stanojevic et al., 2019; Zeltyn et al., 2011) and to define a rational staffing strategy based on experiments.

It should also be noted that the development of online patient communities, various recommendation services, healthcare platforms and ecosystems bring elements of self-organization to the overall health system. Patients can implement the course of examination and treatment, relying, among other things, on independent or community-recommended decisions when choosing a specialist, medical facility, methods of examination and treatment.

The emergence of critical events caused by external factors also makes it relevant to use models based on predicting patient flows. This paper is focused on input patient flow integration of a multidisciplinary medical facility to simulation model when choosing the rational staffing structure in organization.

Patient Flow Description

Two types of input flows are investigated. First type of patient flow is based on historical data of an organization, and second type is based on forecast. Both types of patient flow are sequentially integrated to simulation input. In the beginning, the series of experiments is carried out based on statistics of the simulated medical organization. Further, the input flow of patients is represented by the predicted values of attendance. The incoming patient flow is divided into subgroups, $g_n \in G$, depending on the primary department, where patient arrives:

- g_1 – patients admitted to the CCU (Coronary Care Unit),
- g_2 – patients admitted to the MICU (Medical Intensive Care Unit),
- g_3 – patients admitted to NICU (Neonatal Intensive Care Unit),
- g_4 – patients admitted to TSICU (Trauma Surgical Intensive Care Unit),
- g_5 – patients admitted to SICU (Surgical Intensive Care Unit),
- g_6 – patients admitted to the CSRU (Clinical Recovery and Support Unit).

For each group of patients, an attendance forecast for future periods is separately constructed based on autoregressive and neural network models for predicting time series. One of the popular models for predicting attendance is Seasonal Autoregressive Integrated Moving Average (SARIMA) model (Schweigler, 2009), which was chosen for the kind of high accuracy demonstrated in studies to solve similar problems, in particular in (Kam et al., 2010; Luo et al., 2017).

Holt-Winters model (Andrysiak et al.,2018) was also constructed and evaluated, considering the exponential trend and additive seasonality. The controlled recurrent unit GRU (Afanasieva, Platov, 2019) and the LSTM model (Siami-Namini et al., 2018) were chosen as deep learning models.

Computational Experiments

Dataset for prediction was from MIMIC-III relational database. The latest version of the database is MIMIC-III v1.4 dated September 2, 2016, which includes 61,532 ICU stays: 53,432 for adults and 8,100 for newborns. The dataset covers the period from June 2001 to October 2012. For the experimental part, the ICUSTAYS table is utilized. All dates in MIMIC-III have been shifted to protect patient privacy, therefore the modelling is performed within one year.

Predictive models were trained in the Google Colaboratory cloud platform and optimized using the hyperopt library1. Table 1 shows the results of the RMSE metric for each model and for each group. Since this metric is minimized, the models with the lowest metrics have the best results. In this case, the best results are observed in the LSTM model by groups g_2,g_4, g_5, g_6, however, in some cases SARIMA and GRU perform better in groups g_1 and g_3, respectively.

Table 1. Results of Predicting Attendance in a Medical Facility

	g_1	g_2	g_3	g_4	g_5	g_6
SARIMA	4.6180	7.6787	4.7832	4.5814	5.0447	4.6760
HWES	4.6507	7.6230	4.9137	4.5863	5.0551	4.6828
GRU	4.5376	7.7650	4.8439	4.4483	5.0295	4.6324
LSTM	4.5380	7.6721	4.8566	4.4444	4.9595	4.6205

To improve the forecasting results, a meta-algorithm using the example of linear regression was considered, which showed an improvement in forecasting accuracy for groups g_3,g_4, g_5:

$$y_t = w_1 y_t^{EXP} + w_2 y_t^{SARIMA} + w_3 y_t^{LSTM} + w_4 y_t^{GRU} \quad (1)$$

where y_t^{EXP} – prediction for the t^{th}period by HWES model,
y_t^{SARIMA} – prediction for the t^{th}period by SARIMA model,
y_t^{LSTM} – prediction for the t^{th}period by LSTM model,
y_t^{GRU} – prediction for the t^{th}period by GRU model,
w_1, w_2, w_3 и w_4 – prediction weights.

Another example of a meta-algorithm is a fully connected neural network optimized with the hyperopt library. The predictions of the models y_t^{EXP}, y_t^{SARIMA}, y_t^{LSTM}, y_t^{GRU} with n neurons on the first layer arrive at the input of such a neural network, and the final prediction y_t is calculated at the output. The neural network is trained using backpropagation. According to the results of a computational experiment, a three-layer neural network shows a significant improvement in the quality of prediction, especially for groups g_1, g_2, g_3, g_4 и g_5.

[1]http://hyperopt.github.io/hyperopt/

Concluding Remarks

This work considers methods for predicting an input patient flow of a medical facility and proposes an approach to integrating a predicted input flow into a simulation model. The series of experiments carried out shows that the creation of a meta-algorithm based on the results obtained from SARIMA, Holt-Winters, GRU and LSTM models significantly improves the quality of prediction in almost all patient groups.

The integration of patient flow prediction into simulation models allows one to assess the expected flow of applications to a service organization, consider the demand of the population in services and adaptively form the rational staffing structure of the organization.

The presented approaches serve as the basement for further development of simulation model of urban medical facility and recommendation service for patients. The system will offer an optimal choice of different specialists and the best location of medical facility to provide the required service. Therefore, the shared economy principles are applied by the implementation of such service that would reduce the load of appointments to therapists by matching patients to needed doctors, taking into account the schedule, symptoms, and other relevant factors.

This work was supported by the Russian Foundation for Basic Research (Project 20-07-00651 A - Investigation into the stability of self-organizing distributed digital systems based on models of the social networks dynamics).

References

Neher, R.A., Dyrdak, R., Druelle, V., Hodcroft, E.B., Alber, J. 2020. "Potential impact of seasonal forcing on a SARS-CoV-2 pandemic, " Swiss Med Wkly.

A Guide to the Project Management Body of Knowledge, Third Edition. 2004. Newtown Square, Pennsylvania: Project Management Institute (PMI).

Ahmady, G.A., Mehrpour, M., Nikooravesh A. 2016 “Organizational Structure”. Procedia - Social and Behavioral Sciences, (230).

Kabanovskij, D.S. 2011. "Fundamentals of Organizational Formation Service Management Structures," Izvestiya SPbGEU (4).

Akopov, A.S. 2017. "Simulation Modeling: Textbook and Workshop for Academic Bachelor's Degree," Yurayt Publishing House.

Stanojevic, P., Misita, M., Zoran, J. et al. 2019. "Organizational design based on simulation modeling, " The International Journal of Advanced Manufacturing Technology.

Zeltyn S., Marmor Y., Mandelbaum A. et al. 2011. “Simulation-Based Models of Emergency Departments: Operational, Tactical, and Strategic Staffing”. ACM Transactions on Modeling and Computer Simulation.

Schweigler, L.M., Desmond, J.S., McCarthy, M.L., Bukowski, K.J., Ionides, E.L., Younger, J.G. 2009. "Forecasting models of emergency department crowding". Academic Emergency Medicine, 16 (4), pp. 301-308.

Kam, H. J., Sung, J. O., & Park, R. W. 2010. "Prediction of Daily Patient Numbers for a Regional Emergency Medical Center using Time Series Analysis". Healthcare informatics research, 16(3), pp. 158–165.

Luo, L., Luo, L., Zhang, X., He, X. 2017. "Hospital daily outpatient visits forecasting using a combinatorial model based on ARIMA and SES models," BMC health services research, 17(1), pp. 469.

Andrysiak T., Saganowski Ł., Maszewski M. 2018. "Time Series Forecasting Using Holt-Winters Model Applied to Anomaly Detection in Network Traffic," Pérez García H., Alfonso-Cendón J., Sánchez González L., Quintián H., Corchado E. (eds) International Joint Conference SOCO’17-CISIS’17-ICEUTE’17 León, Spain, Proceeding.

Afanasieva T., Platov P. 2019. "The study of recurrent neuron networks based on GRU and LSTM in time series forecasting," Proceedings of International Conference on Time Series and Forecasting (ITISE 2019), 25–27 Sept., Granada (Spain), Godel Impresions Digitales, vol. 1, pp. 190–201.

Siami-Namini, Sima et al. 2018. "A Comparison of ARIMA and LSTM in Forecasting Time Series," 17th IEEE International Conference on Machine Learning and Applications (ICMLA): pp. 1394-1401.

MIMIC-III, a freely accessible critical care database. 2016. Johnson AEW, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, and Mark RG. Scientific Data.