Forecasting and Explaining Emergency Department Visits in a Public Hospital

Spyridon Petsis[1†], Areti Karamanou[1*†], Evangelos Kalampokis[1†] and Konstantinos Tarabanis[1†]

[1*]Information Systems Lab, University of Macedonia, Egnatia 156, Thessaloniki, 54636, Greece.

*Corresponding author(s). E-mail(s): akarm@uom.edu.gr; Contributing authors: bad20023@uom.edu.gr; ekal@uom.edu.gr; kat@uom.edu.gr;

†These authors contributed equally to this work.

Abstract

Emergency Departments (EDs) are the most overcrowded place of public hospitals. Using machine learning to accurately forecast the number of patients' visits at the ED could facilitate making more effective decisions on ED resource management and, hence, improve the quality of public healthcare. More importantly, using Explainable Artificial Intelligence (XAI) techniques to explain decisions could help deal with challenges like lack of trust caused towards decisions based on machine learning results. The objective of this paper is to use machine learning and XAI to forecast the number of patients' visits at the ED on the next on duty day in order to support policy makers in healthcare make accurate, transparent and trustworthy decisions. Towards this end, a case study is presented that uses the XGBoost algorithm to create a model that forecasts the number of patients' visits at the ED of the University Hospital of Ioannina, Greece based on data related to time, historical patients' visit, public holidays and events, and weather. SHAP framework is used to explain the model. The evaluation of the forecasting model resulted in MAE value of 18.37 revealing a much more accurate model than the baseline with MAE 29.38. According to the SHAP analysis, the day of the week, the mean number of visits the four previous days, and the daily maximum temperature are the three most effective variables in forecasting the number of patients' visits.

Keywords: forecasting emergency department visits, machine learning, XGBoost, SHAP

1 Introduction

Limited funding resources are one of the biggest bottlenecks that reaps the quality of public services. The public healthcare industry is no exception, albeit, according to OECD, public sources fund around 71% of health care [1]. Careful and responsible management and apportioning of funding resources would facilitate the improvement of the quality of public healthcare as well as of the operational efficiency in public hospitals. Digital technologies such as Artificial Intelligence (AI) and machine learning in public healthcare is a powerful tool that could be used towards this direction [2]. However, deploying AI in healthcare requires dealing with ethical challenges, such as the lack of trust towards AI-based decisions and the reinforcement of biases by algorithmic governance, organizational and managerial challenges, such as resistance to data sharing, and other types of challenges as well [3–5].

Emergency Departments (EDs) of hospitals are the place that patients visit in urgent situations during hospitals' on duty days and without prior appointment. They are the place where patients first declare their symptoms and be treated according to how critical their condition is. Especially in public hospitals, EDs are considered to be the busiest hospital entities that have to deal with high pressure due to the continuous high flows of patients. The large volume of patients visiting EDs results in overcrowding, a situation that "occurs when the identified need for emergency services exceeds available resources for patient care in the emergency department, hospital, or both" [6]. ED overcrowding is the main reason of the medical staff burnout and may result in high turnover rates, clinical errors, bad quality of healthcare services such as increased waiting times, increased mortality, and high hospital operation costs [7–10].

Forecasting the number of patients' visits at the ED could help make better decisions on how to effectively allocate daily resources and, hence, facilitate solving the problem of ED overcrowding in public hospitals. Towards this end, literature traditionally uses time series methods (e.g., [11–13]) and, only recently, powerful machine learning algorithms (such as eXtreme Gradient Boosting - XGBoost [14]) (e.g., [15–17]).

Although using machine learning is able to forecast the number of patients' visits with decent accuracy, explaining individual decisions of policy makers is also important especially when policy makers have to deal with challenges such as the the lack of trust towards AI-based decisions. Towards this end, Explainable Artificial Intelligence (XAI) techniques [18] that were recently introduced, can be used to help policy makers understand the forecasts of a model and, at the same time, increase transparency in decisions and trust. Literature, for example, has found that factors like weather (e.g., daily temperature and amount of rain), are highly connected to the number of patients' visits (e.g., [19, 20]). SHAP can be used to additionally explain which of these factors are the most effective in forecasting the number of visits, how (positively or negatively), and which specific values are more effective (e.g., daily temperature above a specific value).

The aim of this paper is to use machine learning and XAI to forecast the number of patients' visits at the ED on the next on duty day in order to support policy makers in healthcare to make more accurate, transparent and trustworthy decisions. Towards this end, a case study is presented that uses the XGBoost algorithm to create a model that forecasts the number of patients' visits at the ED of the University Hospital of Ioannina, Greece based on four categories of data; (i)time-based data, (ii) historical patients' visits data, (iii) holidays and event data, and (iv) weather data. Finally, SHAP framework is used to explain the model.

The structure of this research is organized as follows. Section 2 presents relevant works in literature that use machine learning to forecast the number of patients' visits at the ED. Section 3 presents in details the approach of this research. In addition, Section 4 specifies the problem solved in this work, Section 5 describes the results of data collection, Section 6 explores the collected data, and Section 7 creates the forecasting model. Section 8 explains the forecasting model. Finally, Sections 9 and 10 discuss the findings and conclude this paper respectively.

2 Literature Review

The issue of forecasting the number of patients that will visit a hospital's ED on an on duty day has been extensively explored in the past by many scientific works [11–13, 15–17, 19–44]. These works may be classified based on the method they use to achieve the forecast into three classes; (i) regression-based forecasting works, (ii) time series-based forecasting works, and (iii) machine learning-based forecasting works [45]. However, only a few of these works are classified into the third class.

Although machine learning-based forecasting works share the same objective, i.e., to forecast the number of patients' visits at the ED, they usually differ in the temporal level they base their forecasts (e.g., hourly, daily, etc.), the machine learning algorithm they employ, the variables they use to create the forecasting model, and the error metrics they use to evaluate the performance of their models.

Specifically, regarding the temporal level, most works forecast number of visits at the ED at the day level (i.e., one, two, three, etc., days into the future) (e.g., [15–17, 39]), at the hour level (e.g., [15–17, 40]), or at the week level (e.g., [41, 42]).

Regarding the machine learning method used to forecast ED patients' visits, most popular methods include Decision Tree [41], tree ensembles like

Random Forest [39] and Gradient Boosting ensemble with algorithms such as XGBoost [17, 39, 40], k-nearest neighbours [39], Generalized Linear Models like GLMNET with various types of regularizations (e.g., elastic net, lasso, and ridge) [39, 40], and deep learning Artificial Neural Network Algorithms (ANN) such as Recurrent Neural Network (RNN), Convolutional Neural Networks (CNN), and feedforward neural network [15–17, 41, 42, 44].

The variables used in literature to create the forecasting models include variables with historical data related to the number of patients' visits (such as the mean of ED visits per hour, year, month, day of the week, prior to one-year, etc.) [17, 39, 40], time-based variables (such as day of year, month, time of day, etc.) [39–41], holidays and special events related variables (such as whether it is a public holiday, weekend, Easter, Christmas, etc.) [17, 39–41], variables with weather data (such as the average, maximum, and minimum air temperature of the day) [39–41], variables related to socio-economic information (such as unemployment rate, and gross domestic product per capita) [41], and other variables like the number of flu hits on google [39].

Finally, common error metrics used to compare and evaluate the forecasting models include mean absolute error (MAE) [15, 17, 39, 40, 42], mean square error (MSE) [44], root mean squared error (RMSE) [15, 17, 40, 42], normalised mean squared error (NMSE) [41], mean absolute percentage error (MAPE) [39, 41, 44] and Symmetric mean absolute percentage error (sMAPE) [17], correlation coefficient (R) [42] and R^2 [15], and explained variance (EV) [15].

3 Research Approach

In this paper, we follow five broad steps of explainable machine leaning processes, i.e., (1) specify the problem, (2) collect data, (3) explore data, (4) create the forecasting model, and (5) explain the forecasting model.

(1) Specify the problem. We conduct interviews with members of the administrative and medical staff of the University of Ioannina hospital's ED in order to understand and specify the topic and problem of this work. Specifically, the purpose of the interviews is to extract general information about the hospital (e.g., the location of the hospital, the size of the population it provides care to, etc.), details regarding the operation of the hospital during on duty days (e.g., which days and hours the hospital is on duty, etc.), and the factors that possibly affect the number of patients' visits. Finally, other details like the time period of interest are also discussed.

(2) Collect data. This work exploits three data sets; the first one includes data about patients' visits at the ED of the University Hospital of Ioannina (e.g., which day and what time the patient arrived at the ED, the gender and date of birth of the patient, etc.), the second one with actual weather data regarding the city of Ioannina where the hospital is located (e.g., the daily amount of rain, the minimum and maximum temperature of the city each day, etc.), and the third one with dates of international or local public and school holidays, as well as dates of special events in the city of Ioannina.

We collect ED patients' visit data from the ED registry of the University Hospital of Ioannina. In addition, we obtain actual weather data from meteo Search[1], a website that collects and archives Greek historical weather data from 520 weather stations in Greece. Data are published as open government data that are free for use and reuse. We develop an automated web scraping tool to collect data. The tool gets daily weather data for the years of interest including mean, minimum and maximum temperature, average amount of rain and average wind speed. For the public and school holidays, and special events dates we exploited "holidays"[2], a python library for getting dates of holidays for specific countries, provinces and/or states. The library includes 104 dates of Greek holidays. For the school holidays we also exploit the school calendar of the Greek Ministry of Education, Research and Religious Affairs. Finally, additional local holidays and special events are determined during the interview conducted in the previous step.

(3) Explore data. In this step we use various visualizations to explore and fully understand the three data sets. In addition, we use autocorrelation to examine the relationship between the number of patients' visits at the ED in a day and the same number in the past days (lags). Autocorrelation shows the correlation of data with itself lagged by some number of time units. For example, lag 1 is a version of the original data that is one period behind in time. In our case, autocorrelation checks how correlated is the number of patients' visits at the ED today to some days ago. For example, autocorrelation with lag 1 checks the correlation of the number of visits today with the number of visits on the previous on duty day. We use the Durbin-Watson test [46, 47] as well as the autocorrelation function (ACF) and plot [48] for time series data to examine our data for autocorrelation.

(4) Create the Forecasting model. In this work, we use the eXtreme Gradient Boosting (XGBoost) algorithm to create the forecasting model. XGBoost is a scalable tree boosting algorithm and that gives state-of-the-art results and has been applied on variable domains including health, finance, and energy [14]. Boosting refers to the combination of all hypotheses created in a single hypothesis to solve the general problem of boosting the performance of weak learning algorithms [49].

Furthermore, we use cross-validation to fine tune the model. Specifically, the method we use for the cross-validation is the TimeSeriesSplit[3], a variation of k-fold cross-validation, which uses the first k folds as train and k+1 fold as test ignoring the future values in each split. As TimeSeriesSplit deals with time series data, successive training sets are always super-sets of those that come before them. We use 5 splits to split the data set to 80% training data and 20% testing data.

We measure the performance of the forecasting models with three accuracy metrics; the mean absolute error (MAE), the root mean squared error (RMSE), and the mean absolute percentage error (MAPE).

---

[1] http://meteosearch.meteo.gr/
[2] https://pypi.org/project/holidays/
[3] https://scikit-learn.org/stable/modules/cross_validation.html

Finally, we create two more models that forecast the patients' visits at the ED after two and three on duty days. We measure the forecasting models' performance using the same accuracy metrics in order to compare the models with the respective metrics of the next on duty day model.

(5) Explain the Forecasting Model. Apart from forecasting the ED patients' visits with high accuracy, it is also very important to interpret the variables that affect the forecast. To this end, we apply the Shapley Additive eXplanations (SHAP) framework. SHAP was created by Shapley in 1953, based on game theory method that assigns payouts to players according to their contribution to the total payout [50]. On a predictive model, we can assume that each feature value of the model is a 'player' in a game, where the 'pay-out' is the prediction, and 'gain' is the difference between the prediction and a base value (i.e., the mean prediction) [51]. Shapley values help in the fair distribution of 'pay-out' to each feature, while also detecting the features that enhance the forecast higher than the features that push the forecast lower [52]. The Shapley value is the average marginal contribution of a feature value across all possible coalitions, can be calculated for any tree-based model and show how much each predictor variable contributes, either positively or negatively, to the target variable.

We also use the SHAP summary plot that classifies the variables used in the forecasting model based on their importance in predicting the number of ED patients' visits. More important variables are presented higher in the summary plot. Furthermore, we use SHAP dependence plots to show the impact of features' values on forecasting the number of ED patients' visits, i.e., to understand how the number of ED patients' visits change in relation to the value range of other variables, as well as which values of the variables affect the ED visits positively, negatively, or not at all. For example, positive Shapley values mean that there is a high probability for a large number of patient visits at the ED while negative values the opposite.

4 Specify the Problem

The University Hospital of Ioannina is one of the two public general hospitals located at Ioannina, a city at the Northwestern Greece, and provides care for a general population of 350,000. The hospital is on duty on odd dates, hence every second day, except for (i) months with 31 days and (ii) February of leap years, when the hospital is on duty for two straight days at the end of the month. For example, the hospital is on duty on the 31st of October as well as on the 1st of November. On duty days start at 8 a.m. and end at 8 a.m. of the next day. The hospital has 26 emergency departments (e.g., Pathology, Pediatrics, Cardiology, Orthopedics and others) and 24 outpatient clinics for appointments with the doctors. The shifts of the doctors and nurses for the on-duty days are monthly planned. Specifically, at the end of each month, a schedule is created with the next month's shifts of the medical and administrative staff.

Figure 1 presents the BPMN diagram that describes the steps and actors involved in the treatment process of a patient at the ED. Once a patient arrives at the emergency room during an on duty day a record is created based on the date and time of visit and her personal data. The patient is then examined by the emergency room's doctor in order to decide whether further examination or clinical test (e.g., blood tests, X-rays, etc.) are required and, hence, forward the patient to one of the emergency departments or discharge the patient from the ED to one of the outpatient clinics, to a hospital ward or to her home. In case of very urgent incidents, the patient is directly forwarded to the proper emergency department and not examined by the emergency room doctor. The patient is thereafter examined by the doctor of the emergency department she was forwarded to or, alternatively, conduct the clinical test, X-ray or scan ordered. Depending on the complexity of the incident, the patient may be further forwarded to other emergency departments, or even back to the emergency room until doctors decide to discharge her from the ED again to one of the outpatient clinics, to a hospital ward or to his/her home. Apart from when the patient first arrives at the ED, additional records are added for the same patient when, for example, the patient is forwarded to one of the emergency departments or when the patient is discharged from the ED. As a result, there are many records for the same patient and regarding the same ED visit.

In this work, we are interested in exploring the problem of forecasting the number of unique patients' visits at the ED of the University Hospital of Ioannina, Greece on the next on duty day. In addition, we explore the factors that contribute to this forecasting, that apart from historical patient visits (e.g., the mean number of visits three on duty days ago), include time-related factors (e.g., the month of patients' visits), weather related factors (e.g., the amount of rain during the day of visit at the ED), as well as factors related to public or school holidays and special events (e.g., whether the day of patient's visit is a public holiday or a special local event).

According to the interviews conducted with the administrative and medical staff, being aware of the number of patients' arrivals on the next on-duty day can be exploited to better manage the staff resources of the hospital by appropriately adapting the shift's schedule based on the forecasting. If the model, for example, foreacasts increased visits on the next on-duty day, shifts can be exceptionally re-arranged to include more medical staff. The forecasting can be also exploited to facilitate the management of other resources, such as opening additional beds in the hospital or discharging non-urgent incidents from the wards in order to ensure more space for possible new admissions because of increased arrivals at the ED.

5 Collect data

In order to forecast the number of patients' visits at the ED of the University Hospital of Ioannina on the next on duty day we collect three data sets

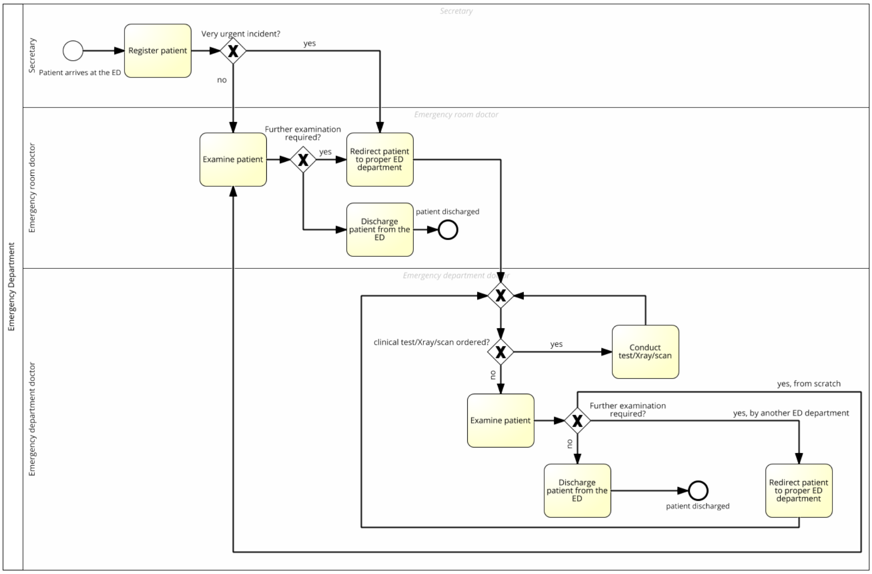


Fig. 1: BPMN diagram for the process of ED treatment

including (1) data regarding patients' visit data from the hospital's registry, (2) weather data, and (3) dates of public and school holidays and other special events.

We collect 893046 records with data about patients' visits at the ED from the ED registry of the University Hospital of Ioannina. Data regard a seven year time period, i.e., from the 1st of March 2013 to the 31st of December 2019. Each record includes a unique patient's identifier, the gender (1 = male, 2 = female) and date of birth of the patient, the precise date and time the record was created, and the ED department the record is related to. The unique patient's identifier is not the real patient identifier but a number randomly generated from it. As a result, someone can not go backwards to the initial identifier and, hence, personal data are protected. We adapt date values in order to refer to the related on duty day and not to the calendar date. As a result, records with time from 8 a.m. of an on duty day to 8 a.m. of the next day are all considered records of the on duty day. For example, patients that arrive at the ED on the 31st of October 9 a.m. or at the 1st of November 7 a.m. are both considered to be arrived at the 31st of October. Since data include sensitive information, the hospital gave us the consent to use it for research purposes.

In addition, as described in section 4, a patient's visit at the ED may be related to multiple records in the registry. In order to correctly calculate the number of visits at the ED per on duty day, we count the number of unique patient visits. We result in a total of 349898 unique visits at the ED during 1272 on duty days.

We also collect weather data for the time period from the 1st of March 2013 to the 31st of December 2019. The weather data include the mean, minimum, and maximum daily temperature (in Celsius scale), the amount of rain per day (in mm), and the average wind speed per day (in km per hour) for the city of Ioannina.

Finally, we collect the dates of public and local holidays, school holidays, and other special events of the city of Ioannina for the time period from the 1st of March 2013 to the 31st of December 2019. For example, the 1st of January is a public holiday in Greece, while the 17th of January is a local holiday in the city of Ioannina. School holidays include dates that are considered as holidays according to the school calendar provided by the Ministry of Education, Research and Religious affairs of Greece every year. Special events are topic and nationwide dates with special importance to the city of Ioannina or Greece like, for example, the Carnival date when people in Ioannina traditionally light big fires. Some of the previous days are movable, i.e., they do not always occur the same day or date each year. For example, the 1st of January is a movable day, while Carnival is celebrated in a movable date.

6 Explore data

In this study, we utilize three data sets including (i) historical patients' visits at the ED, (ii) weather data about the city of Ioannina, and (iii) data with dates of public and local holidays, school holidays and special events during 1272 on duty days. Table 1 reports the descriptive statistics of the number of patients' visits with an annual granularity. Start date for the collection of data is the 1st of March 2013, hence 2013 has fewer patients' visits at the ED. The number of visits at the ED is increased over the years. However, there is not a significant variation in the mean number of visits, hence our data is stationary.

Table 1: Descriptive statistics of the number of patients' visits

	All	2013	2104	2015	2016	2017	2018	2019
Total	**349898**	**41688**	**51009**	**50530**	**51134**	**49881**	**53255**	**52401**
Mean	**274.86**	267.23	274.24	271.67	273.44	268.18	286.32	281.72
Median	**275.5**	266.0	273.5	271.0	272.0	267.0	287.5	281.5
Maximum	**392**	345	357	343	348	364	392	373
Minimum	**115**	176	200	190	185	115	210	196

In addition, Fig. 2 shows the daily variation of patients' visits during all on duty days. The figure doesn't show any strong underlying trend throughout the days.

Regarding the monthly variation in the daily number of patients' visits (Fig. 3), August has the most visits per day than the rest of months and December the fewer visits. However, there is not a significant variation in the

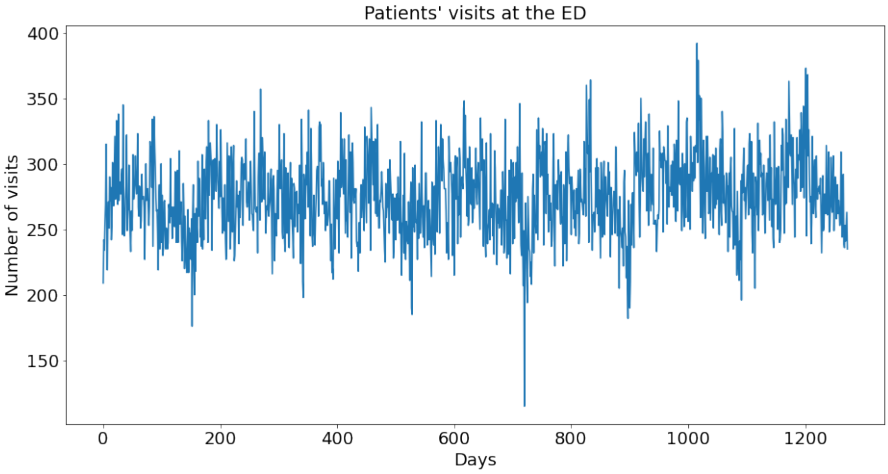


Fig. 2: Daily number of visits

number of visits throughout the months of the year nor the boxplot reveals the existence of seasonality in data.

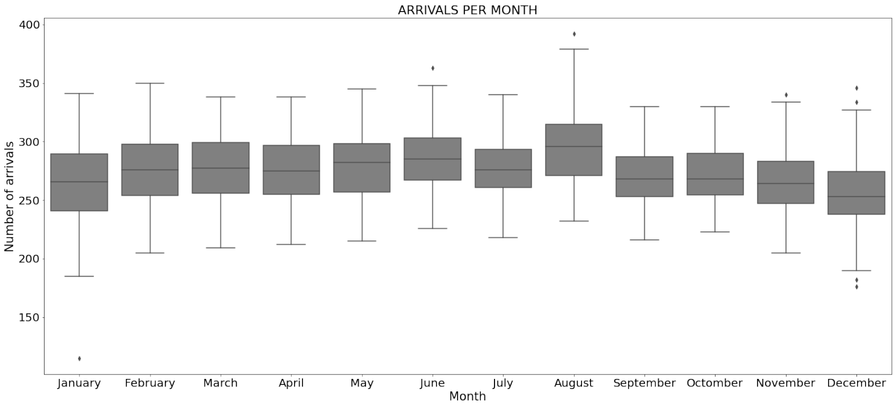


Fig. 3: Variation in the number of patients' visits at the ED on different months

Furthermore, we explore the existence of daily patterns in the visits of patients at the ED (Fig. 4). According to the plot, Monday is the day with the highest number of patients' visits at the ED. Moreover, it seems that the average number of visits is lower during weekend that in the rest of the days. The variation in the number of visits is similar in all days.

Weather data include the maximum, minimum, and mean daily temperature in Celsius degrees, the amount of daily rain (in mm), and the average

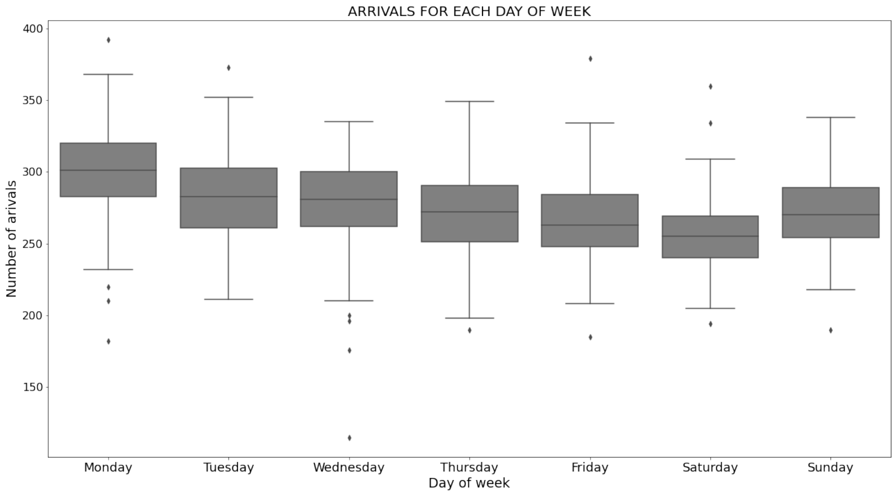


Fig. 4: Variation in the number of patients' visits at the ED on different days of week

wind speed in km/h in the city of Ioannina. According to Table 2, the mean daily temperature of the city of Ioannina throughout all years is 13.69°C.

Table 2: Descriptive statistics of weather data

	Maximum temp (°C)	Minimum temp (°C)	Mean temp (°C)	Rain (mm)	Average wind speed (km/h)
Mean	21.42	6.9	13.69	3.61	2.48
Median	21.7	7.3	13.7	0.2	1.8
Maximum	38.1	22.4	29.2	83.6	22.7
Minimum	-1.7	-11.5	-5.3	0.0	0.0

The holidays and events data set consists of dates related to public holidays, school holidays, and special events. The final data set includes 91 dates that are public holidays, 868 dates that are school holidays, and 35 dates that are special events in the period from the 1st of March 2013 to the 31st of December 2019.

In order to further explore data sets, we join them based on the date. Fig. 5 presents daily patients' visits as well as the variations of the maximum temperature in the city of Ioannina during all on duty days. There is a connection between the maximum daily temperature and the number of daily visits at the ED.

Finally, we check data for autocorrelation. We first perform a Durbin-Watson Test and the test statistic is 0.0178 meaning that there is autocorrelation in data. We also create the autocorrelation plot (Fig. 6). The x-axis in the

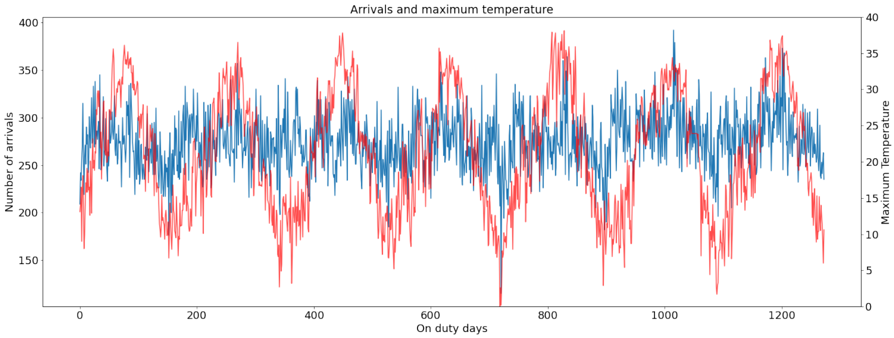


Fig. 5: Number of patients' visits and maximum temperature

plot shows the lag values (1 = previous on duty day, 2 = two on duty days ago, etc.) while the y-axis shows the coresponding autocorrelation value. According to the plot, the number of visits 1, 3, 4, and 7 on duty days ago (lags 1, 3, 4 and 7 respectively) has a strong impact on the number of visits today.

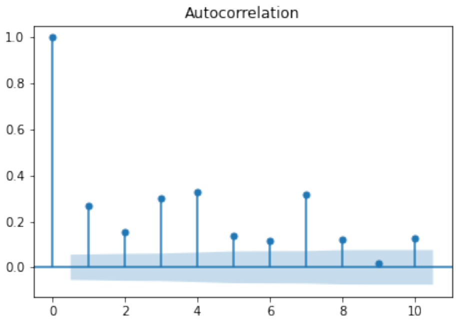


Fig. 6: Number of daily visits autocorrelation plot

7 Create the Forecasting Model

We use four types of variables to forecast the number of patients' visits at the ED; (i) time-based variables, (ii) historical patient visits variables, (iii) holidays and event based variables, and (iv) weather related variables. The final number of variables is 38.

Time-based variables (Table 3) are numeric variables that indicate the day of year, month, day of month, day of week, and week of year.

Historical patient visits variables (Table 4) are numeric values that are derived from patients' visits at the ED in the past. Some of them (numbers 1-13 in the Table) refer to minimum, maximum, and mean calculations of

Table 3: Time-based variables

	Variable	Description
1	day of year	The day of the year the patient visited the ED. Values in range [1, 366].
2	month	The month of the year the patient visited the ED. Values in range [1, 12].
3	day of month	The day of the month the patient visited the ED. Values in range [1, 31].
4	day of week	The day of the week the patient visited the ED. Values in range [1, 7] (1: Monday, 7:Sunday).
5	week of year	The week of the year the patient visited the ED. Values in range [1, 53].

previous on duty days with windows 1 to 5 respectively. For example, "max of previous 4 days" (window 4) is the maximum number of visits of previous 4 on duty days. The rest of the variables (numbers 14-20 in the Table) refer to minimum, maximum, and mean calculations of last year's on duty days with windows 1, 3, and 5. For example, "mean of 3 days last year" (window 3) calculates the mean number of visits of the same on duty day last year, one on duty day prior, and one on duty after it.

Table 4: Historical patient visits variables

	Variable	Description
1	previous day	The number of patients' visits at the ED the previous on duty day.
2 -13	mean/min/max of previous n days	The mean/min/max number of patients' visits at the ED of the n previous on duty days. n is in range [2, 5].
14	one year ago	Last year's number of patients' visits at the ED the same on duty day.
15 - 20	mean/min/max of n days last year	Last year's mean/minimum/maximum number of patients' visits at the ED with window n. The value of n is 3 and 5. The mean/minimum/maximum is calculated using the number of visits the same on duty day last year as well as the one/two next, prior on duty days.

Holidays and event based variables (Table 5) are boolean variables that indicate if it is weekend, public holiday, special event, or school holiday (school holidays are sometimes public holidays too). Additional variables of this category indicate if it is a day before or after a public holiday or special event.

Finally, weather related variables (Table 6) are numeric data that provide information about the temperature (minimum, maximum or mean) amount of rain and average wind speed per day in the city of Ioannina where the hospital is located.

Table 5: Holidays and event based variables

	Variable	Description
1	weekend	If the day of visit was during weekend. Values in range [0,1].
2	public holiday	If the day of visit was a public holiday. Values in range [0,1].
3	before public holiday	If the day of visit was one day before a public holiday. Values in range [0,1].
4	after public holiday	If the day of visit was one day after a public holiday. Values in range [0,1].
5	event	If the day of visit was a special event day. Values in range [0,1].
6	before event	If the day of visit was a day before a special event. Values in range [0,1].
7	after event	If the day of visit was a day after a special event. Values in range [0,1].
8	school holiday	If the day of visit was a school holiday. Values in range [0,1].

Table 6: Weather related variables

	Variable	Description
1	Minimum temperature	The minimum temperature (in Celsius scale) of the day.
2	Maximum temperature	The maximum temperature (in Celsius scale) of the day.
3	Mean temperature	The mean temperature (in Celsius scale) of the day.
4	Rain	The amount of rain (in mm) during the day.
5	Average wind speed	The average wind speed (in km per hour) of the day.

We use XGboost to create the model that forecasts the number of patients' visits on the next on duty day. We use the persistence model as a baseline. The persistence model assumes that the number of patients' visits on the next on duty day is the same as the number of current on duty day's visits.

In order to be able to calculate historical variables no 14 - 20 (see Table 4) in our forecasting model, i.e., variables calculated using previous year's data, we remove data that include first 185 on duty days, i.e., on duty days of the first year of collected data. We result in 1087 on duty days. We thereafter split the resulting data set into two data sets, i.e., training data that include the first 870 on duty days (80%), and test data that include the rest 217 on duty days (20%).

We create the forecasting model using XGBoost. We also use cross-validation to fine tune the model. Table 7 shows the optimal hyperparameter values.

We evaluate the performance of the forecasting model using mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE). After testing the accuracy of the model's predictions MAE value is 18.37, RMSE value is 22.96, and MAPE value is 6.5% (Table 8). Our model's value of MAE shows that the model forecasts patients' visits at the ED much more accurately than the baseline whose MAE value is 29.38.

Table 7: Optimal hyperparameter values

	Parameter	Optimal value
1	n_estimators	998
2	learning_rate	0.01
3	max_depth	2
4	min_child_weight	3
5	gamma	0
6	subsample	1
7	colsample_bytree	0.77

Table 8: Forecasting model's error metrics

	MAE	RMSE	MAPE (%)
Train	14.99	19.04	5.5
Test	18.37	22.96	6.5

Finally, Figure 7 shows the difference between actual and forecast values in the test set. According to the plot, the forecasting model is able to catch the visits behaviour except for peak and trough points of visits.

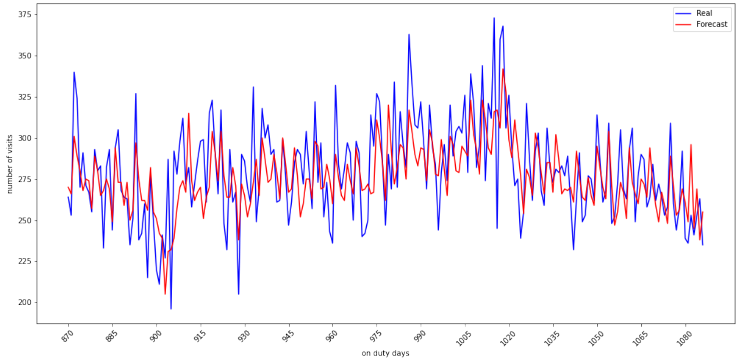


Fig. 7: Actual and forecast values of patients' visits at the ED

In addition, we forecast the number of visits for the second and third on duty days, i.e., after four and six calendar days respectively. Table 9 shows the optimal hyperparameter values for the two forecasting models.

We also evaluate the performance of the two forecasting models using mean absolute error (MAE), root mean squared error (RMSE), and mean absolute

Table 9: Optimal hyperparameter values for the models that predict the number of visits after two and three on duty days

	Parameter	Optimal value (2nd on duty day)	Optimal value (3rd on duty day)
1	n_estimators	840	650
2	learning_rate	0.01	0.01
3	max_depth	2	7
4	min_child_weight	4	3
5	gamma	0	0
6	subsample	1	1
7	colsample_bytree	0.6	0.8

percentage error (MAPE) (Tables 10 and 11). For the 2nd on duty day model, MAE value is 19.41, RMSE value is 23.9, and MAPE value is 6.91%. For the 3rd on duty day model, MAE value is 21.36, RMSE value is 26.59, and MAPE value is 7.5%. Both models' values of MAE shows that the models forecast patients' visits at the ED much more accurately than the baseline with MAE value is 29.38. At the same time, the forecasts of both models are less accurate worse than the next on duty day's forecasts.

Table 10: Forecasting model's error metrics (2nd on duty day)

	MAE	RMSE	MAPE (%)
Train	16.85	21	6.27
Test	19.41	23.9	6.91

Table 11: Forecasting model's error metrics (3rd on duty day)

	MAE	RMSE	MAPE (%)
Train	6.83	8.86	2.5
Test	21.36	26.59	7.5

8 Explain the model

We use SHAP summary and dependence plots to explain the created forecasting model. Figure 8 displays the SHAP summary plot. Each dot in the plot represents one observation, in our case one on duty day of the ED. Y-axis presents the variables used to create the model in order of importance from top to bottom. In our plot, variable "day of week" has the higher impact on the number of patients' visits at the ED, followed by the mean number of visits the four previous days and the maximum daily temperature. In addition, x-axis shows the Shapley values which indicate the probability of success; higher Shapley values mean higher probability for increased number of patients' visits at the ED. The color of the dots demonstrate whether a variable has high (red) or low (blue) value for the specific observation. According to the plot, the first days of the week, i.e., Monday, Tuesday, (lower values of the variable represented with the blue color) have high Shapley values, meaning that in these days the probability for a large number of ED visits is high. In the same way, in weekends (greater values of the variable represented in the plot with the red color), the same probability is low (low Shapley values). In addition, the mean number of visits of the four previous on duty days is also important for the forecast, with higher four days means having a greater probability of increased number of visits in the current day. Finally, patients' visits will be probably increased on duty days when maximum daily temperature in the city of Ioannina is high.

SHAP dependence plots show how Shapley values change as the values of a particular variable evolve. When a Shapley value is positive, then the probability for increased numbers of patient visits at the ED is high as opposed to negative values which are connected with decreased numbers of patient visits.

Fig. 9a shows the dependence plot of the variable "day of week" (x-axis) vs its Shapley value (y-axis). Each dot in the plot represents an on duty day. The variable "day of week" has discrete values (0, 1, 2, 3, etc.), hence all dots are grouped above these seven values. For example, all dots above day of week with value 1 represent the number of patient's visits in each Tuesday. The plot illustrates that Mondays (day of week 0) have the highest positive Shapley value. This means that Monday is the day with the highest probability to have increased numbers of patients' visits at the ED related to the rest of the days. In addition, Tuesday, Wednesday, and Wednesday have Shapley values (either positive or negative) very close to 0, meaning that they have little effect on the number of patients' visits. Finally, Saturday and Sunday have negative Shapley values, meaning that the probability of great number of patients' visits in these days decreases. This claim agrees with other works in literature that find Monday as the busiest day of the week and weekends as the days with decreased number of visits at the EDs [17, 19, 22, 23, 43, 44].

In addition, Fig. 9b shows the dependence plot of the daily maximum temperature in Ioannina (x-axis) vs. its Shapley value (y-axis). The plot illustrates that there is an increased probability for greater numbers of patients' visits when the daily maximum temperature in Ioannina is above 24 Celsius degrees (positive Shapley values). Interestingly, in days with lower maximum temperature, i.e., temperature below 8 Celsius degrees (negative Shapley values),

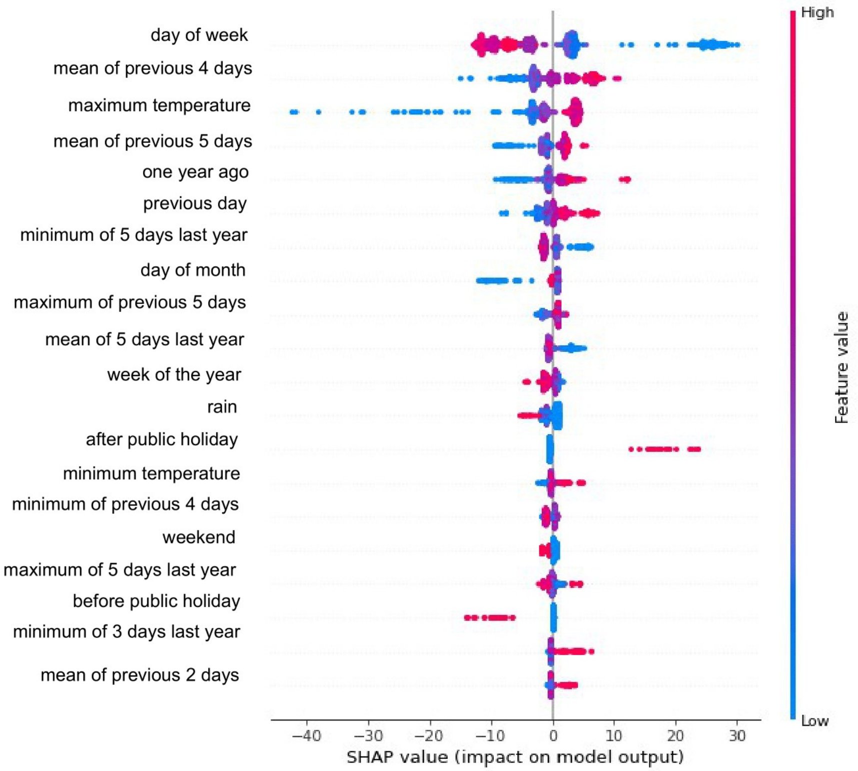
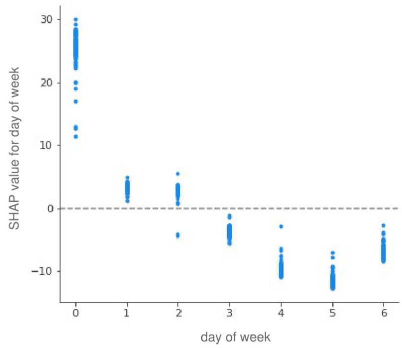


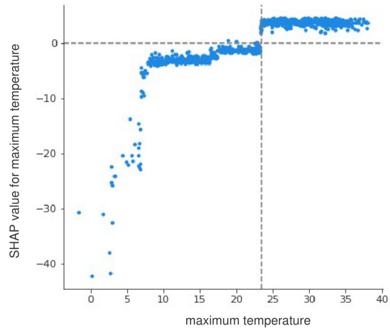
Fig. 8: SHAP summary plot

there is a decreased probability for great number of patients' visits, especially in temperatures close to 0 Celsius degrees. Referring back to Table 2, we can see that the mean temperature at the city of Ioannina is 13.69 and the mean maximum temperature is 21.42 Celsius degrees, hence we may conclude that hottest days in the city of Ioannina are highly connected with greater numbers of patient visits at the ED and vice versa. Our findings about temperature are in general in line with literature that outlines the significance of temperature in the prediction of ED visits [13, 19, 20, 44] although some works found that it doesn't add value to the forecasting model (e.g., [23]). However, to the best of our knowledge, none of the previous works was able to explain in what way (positively or negatively) and which specific values of temperature are significant in the forecasting.

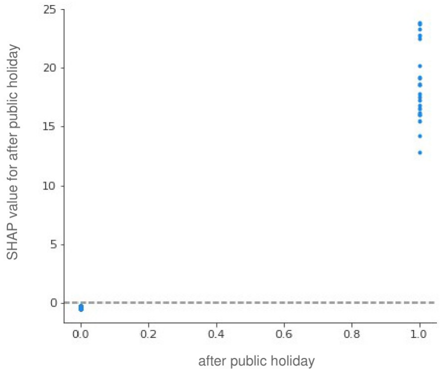
Furthermore, Fig. 9c shows the effect of the variable "after public holiday" on the number of patients' visits at the ED. The type of the variable "after public holiday" is boolean and its values indicate whether an on duty day is after a public holiday (value 1) or not (value 0). As a result, all dots in the



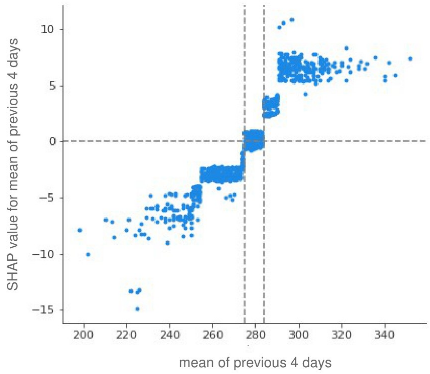
(a) Day of week vs. its Shapley value



(b) Max daily temperature vs. its Shapley value



(c) After holiday vs. its Shapley value



(d) Mean of previous 4 days vs. its Shapley value

Fig. 9: SHAP dependence plots

dependence plot are above these two values. The plot indicates that the number of patients' visits at the ED after public holidays will be probably large.

Finally, Fig. 9d shows the Shapley values for the mean number of patients' visits of four previous on duty days. In general, the plot shows a clear upward trend in the Shapley values when the mean of four previous on duty days increases. In particular, when the mean of number of visits is above a point around 284, there is a high probability for increased number of visits the next on duty day. On the contrary, when values of mean are below 274, then the probability of increased number of visits is low. Referring back to Table 1, we can see that the mean number of daily patients' visits is 274.86. We may, hence, result in the claim that, when the mean number of visits of the four previous on duty days is above the mean value of total visits, then there is a high chance for increased number of visits on the next on duty day.

9 Discussion

ED overcrowding during on duty days is a common problem of hospitals and, according to literature, dealing with it requires achieving a balance between input, throughput, and output of patients [53]. Accurately forecasting the patients' visits is, hence, only part of an overall effort to ensure properly managing resources and making decisions that could help solving ED overcrowding. Other factors that affect overcrowding include the efficiency of the triage and room placement procedures, the cohesiveness of patient care teams, the physical layout of the ED, the efficiency and efficient use of diagnostic testing (e.g., laboratory), the accessibility of medical information, the quality of documentation and communications systems, and the availability of timely specialty consultation [54].

According to the medical and administrative staff of the University of Ioannina, the created forecasting model can be used to ensure more efficient allocation of staff and other resources, such as the number of beds in the patients' wards on the next on duty day, i.e., after two calendar days. Literature suggests that the number of patients' visits can additionally be useful for other purposes as well, such as communicating with admitting physicians about alternative logistical options for their patients, and admitting appropriate patients to units outside the hospital [55].

10 Conclusion

Dealing with overcrowding is a difficult task since public hospitals usually have limited resources and forecasting the number of patients' visits can be useful to deal with it. The objective of this paper is to use an advanced tree-based ensemble algorithm and XAI to forecast the number of patients' visits at the ED of a public hospital. Towards this end, we presented a case study that uses the XGBoost algorithm to develop a model that forecasts the number of patients' visits at the ED of the Greek University of Ioannina hospital on the next on duty day. The creation of the model is based on 38 variables classified in four categories; (i) time-based variables, (ii) historical patient visits variables, (iii) holidays and event based variables, and (iv) weather related variables. Data about historical patients' visits were collected from the ED registry of the University of Ioannina making this work original and unique. Finally, SHAP framework was used to find which variables affect the forecasts and how.

The evaluation of the forecasting model resulted in a MAE value of 18.37, an RMSE value of 22.96, and a MAPE value of 6.5%. The model performed much better than the persistence model used as a baseline whose MAE value is 29.38. We also created two more models to forecast the patients' visits at the ED after two and three on duty days. Both models' accuracy metrics we worse than the next on duty day's and baselines' metrics.

More importantly, we found which variables are more effective in forecasting the number of visits and also which values of these variables affect the forecast and how. According to the analysis based on Shapley values, the three most effective variables in forecasting the number of patients' visits on the next on duty day come from three out of the four categories of variables used to create the forecasting model; (i) the day of week (time-based variables), (ii) the mean number of visits of the four previous on duty days (historical patient visits variables), and (iii) the daily maximum temperature of the city of Ioannina (weather variables). Specifically, among the most interesting insights revealed from the Shapley values' analysis is that Monday is the busiest day for the ED, a claim that agrees with relevant works in literature. In addition, the maximum daily temperature is also important, especially when this temperature is above 24 Celsius degrees when we may expect large numbers of visits, as opposed to maximum temperature below 8 Celsius degrees that are associated with fewer patients' visits. Although temperature in general has been found in other works in literature as significant for forecasting ED visits, to the best of our knowledge, none of these works was able to explain in what way and which values of temperature are more significant. Increased visits should also be expected on days after holidays. Finally, a mean above 284 number of visits in the four previous on duty days (which is actually very close to the mean value of total visits in our data) means that probably there would be a lot of visits on the next on duty day.

Although the subject of forecasting the number of patients' visits at the ED is not new in literature, to the best of our knowledge, this is the first attempt to understand which variables play the most important role in the forecasting as well as which specific values of these variables are more effective and how. This piece of information is very helpful for policy makers because, apart from facilitating taking accurate and evidence-based decisions, it can increase their confidence about their decisions, decrease their hesitancy for adopting Artificial Intelligence, and help them make more transparent decisions.

Declarations

Availability of data and materials

The datasets generated during and/or analysed during the current study are not publicly available due to confidentiality but are available from the corresponding author on reasonable request.

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] OECD: Heath at a Glance 2019: OECD Indicators (2019). https://doi.org/10.1787/4dd50c09-en

[2] Lee, D.: Effects of key value co-creation elements in the healthcare system: focusing on technology applications. Service Business **13**(2), 389–417 (2019). https://doi.org/10.1007/s11628-018-00388-9

[3] Sun, T.Q., Medaglia, R.: Mapping the challenges of artificial intelligence in the public sector: Evidence from public healthcare. Government Information Quarterly **36**(2), 368–383 (2019). https://doi.org/10.1016/j.giq.2018.09.008

[4] Janssen, M., Kuk, G.: The challenges and limits of big data algorithms in technocratic governance. Government Information Quarterly **33**(3), 371–377 (2016). https://doi.org/10.1016/j.giq.2016.08.011. Open and Smart Governments: Strategies, Tools, and Experiences

[5] World Bank: Artificial intelligence in the public sector: Maximizing opportunities, managing risks (2020)

[6] Statement, P.: Crowding. Annals of Emergency Medicine **74**(3), 31 (2019). https://doi.org/10.1016/j.annemergmed.2019.05.003

[7] Cremonesi, P., di Bella, E., Montefiori, M., Persico, L.: The robustness and effectiveness of the triage system at times of overcrowding and the extra costs due to inappropriate use of emergency departments. Applied health economics and health policy **13**(5), 507–514 (2015). https://doi.org/10.1007/s40258-015-0166-5

[8] Moukarzel, A., Michelet, P., Durand, A.-C., Sebbane, M., Bourgeois, S., Markarian, T., Bompard, C., Gentile, S.: Burnout syndrome among emergency department staff: prevalence and associated factors. BioMed research international **2019** (2019). https://doi.org/10.1155/2019/6462472

[9] Pines, J.M., Hilton, J.A., Weber, E.J., Alkemade, A.J., Al Shabanah, H., Anderson, P.D., Bernhard, M., Bertini, A., Gries, A., Ferrandiz, S., *et al.*: International perspectives on emergency department crowding. Academic Emergency Medicine **18**(12), 1358–1370 (2011). https://doi.org/10.1111/j.1553-2712.2011.01235.x

[10] Galanis, P., Siskou, O., Charalambous, G., Konstantakopoulou, O., Liarigkovinou, A., Karagkouni, I., Tsolas, G., Chatzi, P., Giotaki, A., Keramidas, G., *et al.*: Inappropriate Use of Public Hospitals Emergency Departments in Greece: Magnitude and Associated Factors. Stud Health Technol Inform **262**, 224–227 (2019). https://doi.org/10.3233/SHTI190059

[11] Butler, M., Gu, H., Kenney, T., Campbell, S.: P017: Does a busy day predict another busy day? a time-series analysis of multi-centre emergency

department volumes. Canadian Journal of Emergency Medicine **18**(S1), 83–84 (2016). https://doi.org/10.1017/cem.2016.193

[12] Calegari, R., Fogliatto, F.S., Lucini, F.R., Neyeloff, J., Kuchenbecker, R.S., Schaan, B.D.: Forecasting daily volume and acuity of patients in the emergency department. Computational and mathematical methods in medicine **2016** (2016). https://doi.org/10.1155/2016/3863268

[13] Kam, H.J., Sung, J.O., Park, R.W.: Prediction of daily patient numbers for a regional emergency medical center using time series analysis. Healthcare informatics research **16**(3), 158–165 (2010). https://doi.org/10.4258/hir.2010.16.3.158

[14] Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, pp. 785–794 (2016). https://doi.org/10.1145/2939672.2939785

[15] Harrou, F., Dairi, A., Kadri, F., Sun, Y.: Forecasting emergency department overcrowding: A deep learning framework. Chaos, Solitons & Fractals **139**, 110247 (2020). https://doi.org/10.1016/j.chaos.2020.110247

[16] Kadri, F., Abdennbi, K.: Rnn-based deep-learning approach to forecasting hospital system demands: application to an emergency department. International Journal of Data Science **5**(1), 1–25 (2020). https://doi.org/10.1016/j.mlwa.2021.100200

[17] Rocha, C.N., Rodrigues, F.: Forecasting emergency department admissions. Journal of Intelligent Information Systems **56**(3), 509–528 (2021). https://doi.org/10.1007/s10844-021-00638-9

[18] Deeks, A.: The judicial demand for explainable artificial intelligence. Columbia Law Review **119**(7), 1829–1850 (2019)

[19] Diehl, A.K., Morris, M.D., Mannis, S.A.: Use of calendar and weather data to predict walk-in attendance. Southern medical journal **74**(6), 709–712 (1981). https://doi.org/10.1097/00007611-198106000-00020

[20] Duwalage, K.I., Burkett, E., White, G., Wong, A., Thompson, M.H.: Forecasting daily counts of patient presentations in australian emergency departments using statistical models with time-varying predictors. Emergency Medicine Australasia **32**(4), 618–625 (2020). https://doi.org/10.1111/1742-6723.13481

[21] Jones, S.S., Thomas, A., Evans, R.S., Welch, S.J., Haug, P.J., Snow, G.L.: Forecasting daily patient volumes in the emergency department. Academic Emergency Medicine **15**(2), 159–170 (2008). https://doi.org/10.

1111/j.1553-2712.2007.00032.x

[22] Cheng, Q., Argon, N.T., Evans, C.S., Liu, Y., Platts-Mills, T.F., Ziya, S.: Forecasting emergency department hourly occupancy using time series analysis. The American Journal of Emergency Medicine **48**, 177–182 (2021). https://doi.org/10.1016/j.ajem.2021.04.075

[23] Erkamp, N.S., van Dalen, D.H., de Vries, E.: Predicting emergency department visits in a large teaching hospital. International journal of emergency medicine **14**(1), 1–11 (2021). https://doi.org/10.1186/s12245-021-00357-6

[24] Bergs, J., Heerinckx, P., Verelst, S.: Knowing what to expect, forecasting monthly emergency department visits: A time-series analysis. International emergency nursing **22**(2), 112–115 (2014). https://doi.org/10.1016/j.ienj.2013.08.001

[25] Davis, R.E., Markle, E.S., Windoloski, S., Houck, M.E., Enfield, K.B., Kang, H., Balling Jr, R.C., Kuehl, D.R., Burton, J.H., Farthing, W., *et al.*: A comparison of the effect of weather and climate on emergency department visitation in roanoke and charlottesville, virginia. Environmental Research **191**, 110065 (2020). https://doi.org/10.1016/j.envres.2020.110065

[26] Ghada, W., Estrella, N., Pfoerringer, D., Kanz, K.-G., Bogner-Flatz, V., Ankerst, D.P., Menzel, A.: Effects of weather, air pollution and oktoberfest on ambulance-transported emergency department admissions in munich, germany. Science of The Total Environment **755**, 143772 (2021). https://doi.org/10.1016/j.scitotenv.2020.143772

[27] Rosychuk, R.J., Youngson, E., Rowe, B.H.: Presentations to emergency departments for copd: A time series analysis. Canadian respiratory journal **2016** (2016). https://doi.org/10.1155/2016/1382434

[28] Xu, Q., Tsui, K.-L., Jiang, W., Guo, H.: A hybrid approach for forecasting patient visits in emergency department. Quality and Reliability Engineering International **32**(8), 2751–2759 (2016). https://doi.org/10.1002/qre.2095

[29] Afilal, M., Yalaoui, F., Dugardin, F., Amodeo, L., Laplanche, D., Blua, P.: Forecasting the emergency department patients flow. Journal of medical systems **40**(7), 1–18 (2016). https://doi.org/10.1007/s10916-016-0527-0

[30] Gul, M., Guneri, A.F.: Planning the future of emergency departments: Forecasting ed patient arrivals by using regression and neural network models. International Journal of Industrial Engineering **23**(2) (2016)

[31] Rosychuk, R.J., Youngson, E., Rowe, B.H.: Presentations to alberta emergency departments for asthma: a time series analysis. Academic Emergency Medicine **22**(8), 942–949 (2015). https://doi.org/10.1111/acem.12725

[32] Mai, Q., Aboagye-Sarfo, P., Sanfilippo, F.M., Preen, D.B., Fatovich, D.M.: Predicting the number of emergency department presentations in western australia: A population-based time series analysis. Emergency Medicine Australasia **27**(1), 16–21 (2015). https://doi.org/10.1111/1742-6723.12344

[33] Aboagye-Sarfo, P., Mai, Q., Sanfilippo, F.M., Preen, D.B., Stewart, L.M., Fatovich, D.M.: A comparison of multivariate and univariate time series approaches to modelling and forecasting emergency department demand in western australia. Journal of biomedical informatics **57**, 62–73 (2015). https://doi.org/10.1016/j.jbi.2015.06.022

[34] Petrou, P.: An interrupted time-series analysis to assess impact of introduction of co-payment on emergency room visits in cyprus. Applied health economics and health policy **13**(5), 515–523 (2015). https://doi.org/10.1007/s40258-015-0169-2

[35] Zlotnik, A., Gallardo-Antolin, A., Alfaro, M.C., Pérez, M.C.P., Martínez, J.M.M., *et al.*: Emergency department visit forecasting and dynamic nursing staff allocation using machine learning techniques with readily available open-source software. CIN: Computers, Informatics, Nursing **33**(8), 368–377 (2015). https://doi.org/10.1097/cin.0000000000000173

[36] Kang, S.W., Park, H.S.: Emergency department visit volume variability. Clinical and experimental emergency medicine **2**(3), 150 (2015). https://doi.org/10.15441/ceem.14.044

[37] Aroua, A., Abdul-Nour, G.: Forecast emergency room visits–a major diagnostic categories based approach. International Journal of Metrology and Quality Engineering **6**(2), 204 (2015). https://doi.org/10.1051/ijmqe/2015011

[38] Ekström, A., Kurland, L., Farrokhnia, N., Castrén, M., Nordberg, M.: Forecasting emergency department visits using internet data. Annals of emergency medicine **65**(4), 436–442 (2015). https://doi.org/10.1016/j.annemergmed.2014.10.008

[39] Vollmer, M.A., Glampson, B., Mellan, T., Mishra, S., Mercuri, L., Costello, C., Klaber, R., Cooke, G., Flaxman, S., Bhatt, S.: A unified machine learning approach to time series forecasting applied to demand at emergency departments. BMC Emergency Medicine **21**(1), 1–14 (2021). https://doi.org/10.1186/s12873-020-00395-y

[40] Padthe, K.K., Kumar, V., Eckert Carly M. MD MPH, M.N.M., Zahid, A., Ahmad, M. Aurangzeb, Teredesai, A.: Emergency department optimization and load prediction in hospitals. Proceedings of the AAAI Fall Symposium on AI for Social Good **2884** (2020)

[41] López Ibáñez, B., Torrent-Fontbona, F., Roman, J., Inoriza, J.M.: Forecasting of emergency department attendances in a tourist region with an operational time horizon. Information Systems (2021)

[42] Khaldi, R., Afia, A.E., Chiheb, R.: Forecasting of weekly patient visits to emergency department: real case study. Procedia Computer Science **148**, 532–541 (2019). https://doi.org/10.1016/j.procs.2019.01.026. The Second International Conference on Intelligent Computing in Data Sciences, ICDS2018

[43] Sun, Y., Heng, B.H., Seow, Y.T., Seow, E.: Forecasting daily attendances at an emergency department to aid resource planning. BMC emergency medicine **9**(1), 1–9 (2009). https://doi.org/10.1186/1471-227x-9-1

[44] Sudarshan, V.K., Brabrand, M., Range, T.M., Wiil, U.K.: Performance evaluation of emergency department patient arrivals forecasting models by including meteorological and calendar information: A comparative study. Computers in Biology and Medicine, 104541 (2021). https://doi.org/10.1016/j.compbiomed.2021.104541

[45] Gul, M., Celik, E.: An exhaustive review and analysis on applications of statistical forecasting in hospital emergency departments. Health Systems **9**, 263–284 (2018). https://doi.org/10.1080/20476965.2018.1547348

[46] Durbin, J., Watson, G.S.: Testing for serial correlation in least squares regression: I. Biometrika **37**(3/4), 409–428 (1950). https://doi.org/10.2307/2332391

[47] Durbin, J., Watson, G.S.: Testing for serial correlation in least squares regression: Ii. Biometrika **38**(1/2), 159–177 (1951). https://doi.org/10.2307/2332325

[48] Box, G.E.P., Jenkins, G.M., Reinsel, G.C., Ljung, G.M.: Time Series Analysis: Forecasting and Control. Wiley Series in Probability and Statistics. Wiley, New York (2015). https://doi.org/10.1111/jtsa.12194

[49] Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. Journal of computer and system sciences **55**(1), 119–139 (1997). https://doi.org/10.1006/jcss.1997.1504

[50] Shapley, L.S.: A value for n-person games. Contributions to the Theory of Games, Volume II, 307–318 (1953). https://doi.org/10.1515/

9781400881970-018

[51] Slack, D., Hilgard, S., Jia, E., Singh, S., Lakkaraju, H.: Fooling lime and shap: Adversarial attacks on post hoc explanation methods. In: Proceedings of the AAAI ACM Conference on AI, Ethics, and Society, pp. 180–186 (2020)

[52] Kalampokis, E., Karamanou, A., Tarabanis, K.: Applying explainable artificial intelligence techniques on linked open government data. In: Scholl, H.J., Gil-Garcia, J.R., Janssen, M., Kalampokis, E., Lindgren, I., Rodríguez Bolívar, M.P. (eds.) Electronic Government, pp. 247–258. Springer, Cham (2021)

[53] Morris, Z.S., Boyle, A., Beniuk, K., Robinson, S.: Emergency department crowding: towards an agenda for evidence-based intervention. Emergency Medicine Journal **29**(6), 460–466 (2012). https://doi.org/10.1136/emj.2010.107078

[54] Asplin, B.R., Magid, D.J., Rhodes, K.V., Solberg, L.I., Lurie, N., Camargo Jr, C.A.: A conceptual model of emergency department crowding. Annals of emergency medicine **42**(2), 173–180 (2003)

[55] Asheim, A., Bache-Wiig Bjørnsen, L.P., Næss-Pleym, L.E., Uleberg, O., Dale, J., Nilsen, S.M.: Real-time forecasting of emergency department arrivals using prehospital data. BMC emergency medicine **19**(1), 1–6 (2019). https://doi.org/10.1186/s12873-019-0256-z