IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 15, 2021, accepted March 10, 2021, date of publication March 17, 2021, date of current version March 26, 2021.
Digital Object Identifier 10.1109/ACCESS.2021.3066562

Predicting Length of Stay Across Hospital Departments

JESÚS MANUEL PUENTES GUTIÉRREZ, MIGUEL-ÁNGEL SICILIA, SALVADOR SÁNCHEZ-ALONSO, AND ELENA GARCÍA-BARRIOCANAL
Departamento Ciencias de la Computación, Universidad de Alcalá, 28054 Madrid, Spain
Corresponding author: Jesús Manuel Puentes Gutiérrez (jesus.puentes@edu.uah.es)

ABSTRACT The length of hospital stay and its implications have a significant economic and human impact. As a consequence, the prediction of that key parameter has been subject to previous research in recent years. Most previous work has analysed length of stay in particular hospital departments within specific study groups, which has resulted in successful prediction rates, but only occasionally reporting predictive patterns. In this work we report a predictive model for length of stay (LOS) together with a study of trends and patterns that support a better understanding on how LOS varies across different hospital departments and specialties. We also analyse in which hospital departments the prediction of LOS from patient data is more insightful. After estimating predictions rates, several patterns were found; those patterns allowed, for instance, to determine how to increase prediction accuracy in women admitted to the emergency room for enteritis problems. Overall, concerning these recognised patterns, the results are up to 21.61% better than the results with baseline machine learning algorithms in terms of error rate calculation, and up to 23.83% in terms of success rate in the number of predicted which is useful to guide the decision on where to focus attention in predicting LOS.

INDEX TERMS Length of stay, hospital department, machine learning, decision tree, random forest.

I. INTRODUCTION

Extracting knowledge from databases is essential for organizations, in both private enterprises and in government agencies. If enterprises are able to recognise patterns or trends in recurrent processes, then they will be able to direct resources where they are needed, allowing a more efficient management of those available. Besides, the ability to predict events or specific behaviours within some level of confidence confers additional benefits in terms of savings both at economic and human levels or managing resources.

According to a study [9], healthcare systems generate large amounts of administrative data about patients, departments, medical material costs, bed availability, diseases, etc. This study departs from readily available administrative data to assess resource use in hospital systems. Concretely, a substantial amount of data stored in computer databases which, after an adequate analysis, can be helpful to improve the management of internal resources, to reduce costs savings, improve patients care among other tasks. Besides, as claimed by [22], to make a sustainable and successful integration of healthcare systems and, consequently, improve not only management but the overall system, some of the main factors to consider are patients' needs, information systems data collection and performance management. Therefore, innovation should be a priority in health care, patients care and hospital management.

A prolonged stay of patients in hospitals implies considerable costs and discomfort for patients. It also entails the need for efficient use of resources and facilities for better planning at forthcoming resources demands. These reasons motivate in-depth studies that attempt to reduce the length of stay (LOS) in hospitals, as pursued elsewhere by [6], [18] or [19]. Previous works have used hospital datasets to analyse LOS in particular departments with specific cases, and focus on the predictive effectiveness of the resulting models, but do not take into account how the context can be used to improve the models or gather additional insights. This research provides a comprehensive approach to the problem using data from all hospital departments in a large Spanish hospital located in Madrid. The cohort consists of hospitalised patients in a period that starts on June 1st, 2010 and ends on September 29th, 2015. Here we approach the problem in global terms, analysing all hospital departments to get an overall

The associate editor coordinating the review of this manuscript and approving it for publication was Shagufta Henna.

idea of which departments are more appropriate to assign more resources. So, predicting LOS in these departments may support more effective management and enables a more successful vision while searching for patterns about specific or generic cases in the patient's health history.

In our research, machine learning techniques are applied to hospital management in an attempt to optimise hospital resources more efficiently within the departments, providing an extra advantage in favour of patients and hospital entities. Thus, we report our results on predicting LOS in all departments from a hospital and present additional patterns that may complement the models with relevant insights.

Results show that the most effective hospital departments to predict LOS are Obstetrics, Ophthalmology and Endocrinology. It can also be observed that the best results in predictions are obtained by those supervised classification algorithms based in tree-like structures or those based in Radial Basis Function kernels (RBF). As said, some additional patterns were discovered, patterns related to specific discoveries improving results inside some departments, like the ones in the Cardiology department, where the results were up to 23,83% better than in regular tests. These results were obtained with the Decision Tree algorithm, in the same way as [10] article. In this article, the best results were also obtained in the Cardiology department with Decision Tree and Support Vector Machines (as RBF) algorithms. Other authors, as [14], also obtained the best results with tree-like algorithms, Decision Tree C4.5 in this case.

The rest of this paper is structured as follows. In Section 2, we provide a survey of previous related research. Section 3 describes the dataset analysed and the methods used to evaluate the initial situation, then discussing the supervised machine learning techniques used and the approach to compare the different hospital departments. In Section 4, results are discussed, describing which hospital departments are more appropriated to predict LOS. Finally, Section 5 provides conclusions and outlook.

II. BACKGROUND

Adjusting LOS can save costs by reducing the resources earmarked for the purpose. The calculation of the number of days that patients stay overnight in the hospital is usually made as described elsewhere in [25]. In [9], they describe the factors (e.g. age, type of admission, hospital type) with the most substantial influence in predicting LOS. Both the research of Freitas *et al.* and the one reported in this article share the concern of drawing consequences to improve management resources. However, Freitas *et al.* focus on the factors that affect prediction, while this research focuses instead on the specificities of hospital departments and attempts to determine which departments are the more promising to predict more accurately the length of stay.

Similar to how other previous studies have been conducted, we use different machine learning techniques to predict LOS than found elsewhere. In other previous studies, some authors – [9], [24] or [6] – use regression, while others –[3], [10] or [1] – apply classification techniques. Although researchers seem to use classification techniques predominantly, methodologies using decision trees with generic hospital data have yielded better results than others in previous studies, as it is stated in [1], a study where decision trees also outstand. An algorithm often used is Random Forest, which offers similar results to Decision Trees in those cases that data maintains the appropriate balance between bias and variance. For example, see [5]. In [12] also compare the Random Forest algorithm and neural networks, obtaining better results with the first one. Other outstanding classification models are Support Vector Machine algorithms, which offer remarkable results ([10]) or KNN ([11]).

Regression algorithms also yield good results as shown in [8]. These algorithms are commonly used to search for behavioural patterns. In [3], it is verified that previous treatments affect the number of days that patients must remain in hospital. Instead, our study uses classification algorithms to predict results and, depending on them, via analysis and data mining, to obtain patterns based on previous results.

Supervised classification algorithms often try to figure out the answer to a yes/no question, such as in the survey made in [18], where they predict if a patient will remain in the hospital more than a week. Our research could have yielded better results by making groups of days instead of using unique days, but the focus and the effort have been made to search a group of medical specialties with best results. Hence, in [8] determined if the LOS would be within a specific range of days or not. Tables 1, 2 and 3 provide a summary of all the referenced articles.

Another aspect common to many of LOS studies is that they are limited to a specialization area only. Thus, the case described in [20] targets patients admitted in the Intensive Care Unit; [4] focus on LOS of women diagnosed with uterine fibroid, a relatively common benign tumour found in female's reproductive organs; while in [15], LOS with inpatients in psychiatric hospitals. We also noted that it is unusual to predict the exact number of days when calculating LOS. Instead, it is more common to provide a range of days as the final result (frequently classified as a short, medium or long stay), defining these stays as specific intervals of days, e.g. in the way reported in [21] or [7].

Other studies first analyse the hospital department with the best results, to later identify behavioural patterns or parameters that would interfere with the LOS prediction. Thus, in [10] the authors determine the variations in their results when the diastolic blood pressure varies or when other secondary diseases exist. Also, in [23] it is identified higher prediction errors for patients admitted recently.

The main contribution of the present article is that the LOS prediction is carried out for all hospital departments rather than just for one, further determining which departments are more suitable to predict LOS. Cost reduction estimations can be made at the hospital level and not only in an individual department, thus obtaining a comprehensive view of business and finances. Besides, the method makes it possible

TABLE 1. Summary of articles about supervised learning techniques.

Article	SVM	SVC	KNN	DT	RF	NN	n	Department
Aghajani & Kargari (2016)		X	X	X			327	General Surgery
Chang, & Lu (2016)	X					X	320	Obstetrics, Gynaecology
Chuang et al. (2015)					X		897	General Surgery
Combes C., Kadri F., & Chaabane S. (2014)	X			X	X		12,498	Emergencies (Pediatrics)
Hachesu, Ahmadi, Alizadeh, & Sadoughi (2013)	X			X		X	4,948	Cardiology
Houthooft et al. (2015)	X		X		X	X	14,480	Intensive Care Unit
LaFaro et al. (2015)		X		X	X	X	185	Intensive Care Unit
Liu et al. (2004)		X		X			4,722	Geriatric Medicine
Liu et al. (2006)		X		X			4,722	Geriatric Medicine
Lowell et al. (1997)						X	-	Psychiatry
Panchami, & Radhika (2014)	X	X				X	-	-
Parag C. P., & Hitesh K. (2014)				X			-	Psychiatry
Stoean et al. (2015)	X			X		X	368	General and Digestive Surgery
Tanuja, Acharya, & Shailesh (2011)		X	X	X		X	401	-
Turgeman, May, & Sciulli (2017)	X			X		X	20,321	Admissions
Yasinski E. et al. (2017)				X			-	-

Caption:

SVM: *Support Vector Machine Classification (SVM - C)*

SVC: *Support Vector Classifiers (SVC)*

KNN: *KNN*

DT: *Decision Trees*

RF: *Random Forest*

NN: *Neural Network*

n: *Number of samples*

to decide which departments are more suitable to cost reductions.

III. MATERIAL AND METHODS

A. DATA PREPARATION

All data was collected from the data management system and anonymised removing all personal details and quasi-identifiers to comply with data protection regulations. The study population were the 63,932 patients admitted to the hospital during the five years under review –from June 2010 to September 2015–. If the number of samples is compared to other authors, it may seem a similar number in an analysis of this kind, but in this article, the research is performed across all hospital departments, so the number of samples is significantly lower. Thus, in Anesthesiology and Resuscitation,

TABLE 2. Summary of articles about supervised learning techniques – regression.

Article	MT	LgR	LnR	SVM	NS	n	Department
Baylis (2009)					X	-	-
Combes C., Kadri F., & Chaabane S. (2014)		X				12,498	Emergencies (Pediatrics)
Ferrão, Duarte, Janela,& Martins (2015)		X				5,089	-
Freitas et al. (2012)		X				9,253,087	-
Panchami, & Radhika (2014)		X				-	-
Parag C. P., &Hitesh K.(2014)			X			-	Psychiatry
Rouzbahman, Jovicic, & Chignell (2017)			X			23,000	Intensive Care Unit
Turgeman, May, & Sciulli (2017)	X					20,321	Admissions
Yang, Wei, Yuan, & Schoung (2010)	X			X		1,080	-
Yasinski E. et al. (2017)				X		-	-

Caption:

MT: *Model-tree-based Regression*

LgR: *Logistic Regression*

LnR: *Linear Regression*

SVM *Support Vector Machine Regression (SVM - R)*

NS: *Non-specified Regression model*

n: *Number of samples*

TABLE 3. Summary of articles about supervised learning techniques –clustering.

Article	KM	H	GM	n	Department
El-Darzi et al. (2009)	X	X	X	105,77	-

Caption:

KM: *K - Means (Clustering)*

H: *Hierarchical (Clustering)*

GM: *Gaussian Mixture Model (Density estimation, Dimensionality Reduction)*

n: *Number of samples*

Neonatology and Dermatology departments, the number of samples is insufficient to obtain good predictive results. The dataset initially contained 45 attributes per patient. However, only 12 were selected after filtering with a procedure called SelectKBest, which is part of the scikit-learn machine learning library (version 0.20), that belongs to feature_selection class. The parameters used with this procedure were the chi2 score function, the best one for classification tasks, and 12 for the number of attributes, partially due to a better efficiency but also because adding more attributes did not improve the results significantly. This feature selection technique consists of picking up only those features that contributed most to the target variable –number of days in our case–. The chi-squared statistical test for non-negative features and classification problems was used in this data selection because it fits better to the available dataset. The final result is what we call the ''feature importance attribute'', i.e. a score for each feature, where the higher the score, the more relevant is the feature towards the prediction of the LOS.

Table 4 shows the distribution of the attributes in the hospital dataset after the data cleansing.

Data preparation and analysis was carried out by creating an ad hoc Python program, making use of open source libraries such as Scipy (mathematics, science and engineering), Numpy (N-dimensional arrays, vectorisation, indexing, etc.) and Matplotlib (library for creating static, animated and interactive visualisations).

The data set was therefore organised around 12 numerical features. One of the features derived in a new field from the existing data: the number of days that the patient remained in the hospital, obtained by subtracting the exit date from the entrance date. The other features obtained were the patient's unique identifier, surgery date, the disease that causes hospital admission (codified as diagnosis-2), diagnosis-4, diagnosis-6, diagnosis-7, diagnosis-8, diagnosis-9, diagnosis-10, diagnosis-11, and diagnosis-12. The number after each diagnosis indicates the order of all the diseases diagnosed in the patient, but not necessarily all the patients will have all the diagnosis features filled. For example, a patient could only have diagnosis-2 and diagnosis-4 filled (i.e. ''atrial and flutter fibrillation'' – code 427.32 – and ''non specified cardiac insufficiency'' – code 428.9 –, both diagnosis in International Classification of Diseases) with the rest diagnosis empty, while another patient could have an additional diagnosis filled. Other features from the original dataset were taken into account to extract additional information, such as gender and age. The selection algorithm chose 12 features out of 45, but two of them – gender and age– were taken out of the selection as they were used to identify and analyse patterns instead.

The final dataset was organised according to patients' medical diagnosis. In this way, patients were grouped in hospital departments teams to study the length of their stay. Fig. 1 shows a dataset overview of the maximum and the average length of stay per department. It reveals the high variability in the predicted feature shown by some departments, which could explain the low accuracy in the predictions. An example of this can be observed in Neonatology department, where the standard deviation is very high, bigger than the average value for that department, and the prediction rate is very low, a maximum of 19,51% with Naïve Bayes algorithm.

TABLE 4. Hospital features distribution.

Variable		Number of encounters	% of the population
Gender			
	Female	32,920	51.64%
	Male	30,832	48.36%
Admission date			
	Min.:	June 1st, 2010	
	Max.:	September 29th, 2015	
Specialty of Admission			
	Different departments (see Fig. 1)		
Discharge disposition			
	Discharge to home	61,059	95.78%
	Discharge to other hospital	1,018	1.60%
	Voluntary discharge	194	0.30%
	Discharge due to death	1,481	2.32%
	Discharge due to transfer to other social-sanitary centre	-	0.00%
	Discharge due to escape	-	0.00%
	Discharge due to hospitalization at home	-	0.00%
Surgery			
	yes:	18,487	29.00%
	no:	45,265	71.00%
Primary diagnosis			
	Congestive heart failure	2,349	3.68%
	Atherosclerosis coronary in native coronary artery	1,403	2.20%
	Acute exacerbation	1,312	2.06%
	Pneumonia	1,124	1.76%
	Atrial fibrillation	833	1.31%
	Breathing disorder	795	1.25%
	Subendocardial infarction	674	1.06%
	Left heart failure	648	1.02%
	Precordial pain	609	0.96%
	Acute respiratory failure	552	0.87%
	Others (lower than 0,85%)	53,453	83.85%
Primary procedures			
	Karyotype	5,694	8.93%
	Electrocardiogram	3,534	5.54%
	Chest x-ray	2,928	4.59%
	Coronary angioplasty	2,426	3.81%
	Head computerized axial tomography	1,653	2.59%
	Doppler echocardiogram	1,398	2.19%
	Mental state evaluation	1,298	2.04%
	Cardiac tomography	1,134	1.78%
	Abdominal ultrasound	1,045	1.64%
	Antibiotic shot	1,026	1.61%
	Heart tissue excision	949	1.49%
	Others (lower than 1,36%)	40,667	63.79%
Age			
	30 years old or younger	4,576	7.18%
	31 - 60 years old	18,800	29.49%
	Older than 60	40,376	63.33%
Cost weight			
	below mean value	15,417	24.18%
	over mean value	48,335	75.82%
Groups related with diagnosis			
	Childbirth with sterilization	2,756	4.32%
	Renal procedures	2,413	3.78%
	Dilation and curettage	2,401	3.77%
	Respiratory system	2,019	3.17%
	Upper airways	1,331	2.09%
	Other respiratory procedures	953	1.49%
	Rehabilitation procedures	907	1.42%
	Other rehabilitation procedures	873	1.37%

TABLE 4. ***(Continued.)*** **Hospital features distribution.**

Variable	Number of encounters	% of the population
Others (lower than 1,35%)	50,099	78.58%
Days between admission and discharge (Length of Stay)		
0 - 7 days	37,425	58.70%
8 - 14 days	15,342	24.07%
15 - 21 days	5,555	8.71%
22 - 28 days	2,355	3.69%
More the 28 days	3,075	4.82%

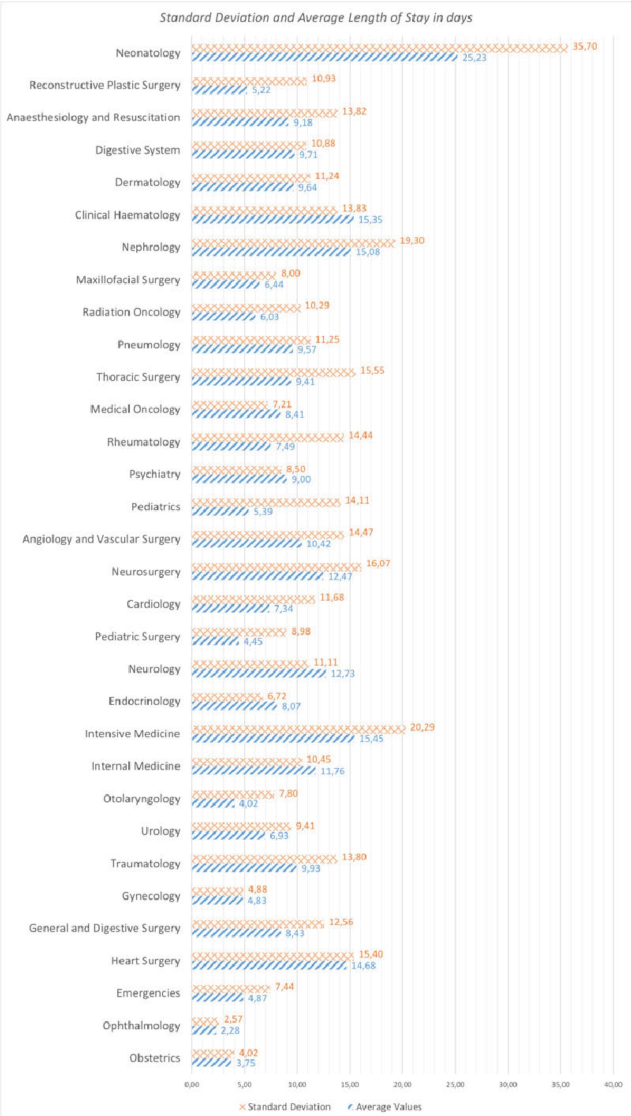

FIGURE 1. **Hospital departments LOS overview.**

In this kind of studies, data cleansing and pre-processing are essential to have optimal results. Regarding this, we removed all registers with either unknown or missing hospital department, as well as those whose fields had spelling errors, missing values and other irregularities. Specifically, records without relevant information were removed, such as those missing diagnosis values, or sex information or information on the procedures or all of them. Therefore, those records –a total of 180– were discarded as it was not viable to infer missing values for a diagnosis or a procedure.

Given that our primary interest was to predict the length of stay in each hospital department, we created an attribute that contains the number of days that patients remained in hospital, calculated with both admission and discharge dates.

B. TRAINING AND TESTING DATA SETUP

The original dataset was partitioned into a training set and a testing set, without any stratification. The procedure is simple: training records were randomly selected until they reached the 67% threshold, then, the remaining 33% were assigned to the testing dataset. This decision was based on the most common percentages used in similar studies. Thus, in [13], they proceed with two thirds for training and one third for testing, while in [1], it is used 70% of the data for training and 30% for testing. These partition ratios appear to be the most appropriate to ensure that datasets about population statistics are marginally different from that of the overall data. Afterwards, we tested another data partitioning procedure – using k-fold cross-validation with five partitions randomly selected as the splitting strategy– to improve results, but the results were only slightly better. This method, obtaining the average of the five independent pieces of training, provided the best option for data selection and was finally selected (others have successfully done the same in previous studies). For instance, in [2], it is used cross-validation when partitioning their datasets to better balance variance and performance.

Once the training and testing datasets were ready, the data from each hospital department was analysed with all the machine learning methods selected. Once the model of each machine learning algorithm was fixed, and all departments and methods were evaluated, two other rounds of tests were run to verify the performance's stability and the correctness of the results. The datasets composition did not vary until all the algorithms were tested with all the departments, but given that three groups of tests were run, re-selecting different samples and re-partitioning occurred. The three groups of tests were performed to confirm results and to check that over-fitting or any other undesirable effect was taking place. Overall, the training dataset was used to adjust the parameters of the different models, while the testing dataset was used to evaluate their predictive capability. Besides, cross-validation was used in another group of tests to see if the results improved, although no significant differences were found.

C. EVALUATION CRITERIA OF THE METHODS

Machine learning methods used in the classification works referenced in Table 1 were compared to identify the most appropriate approach to predict LOS and identify the best-ranked hospital departments in this prediction. To evaluate predictions, we built a model of each algorithm for each department to later proceed to analyse every department with the most common classification algorithms used in the article

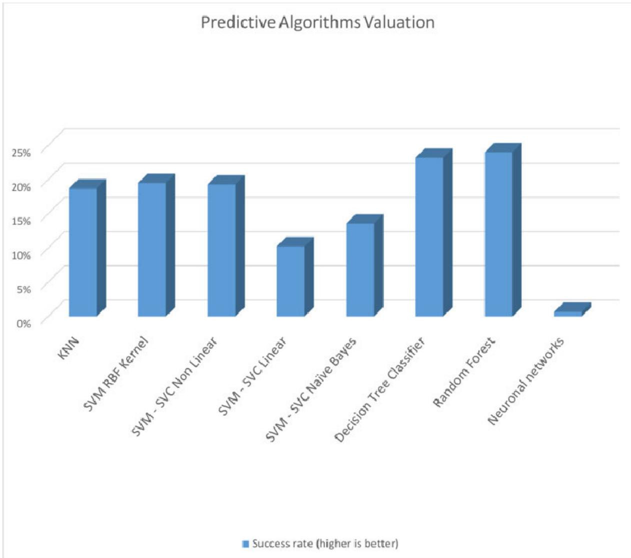

FIGURE 2. Success rate algorithms results overview.

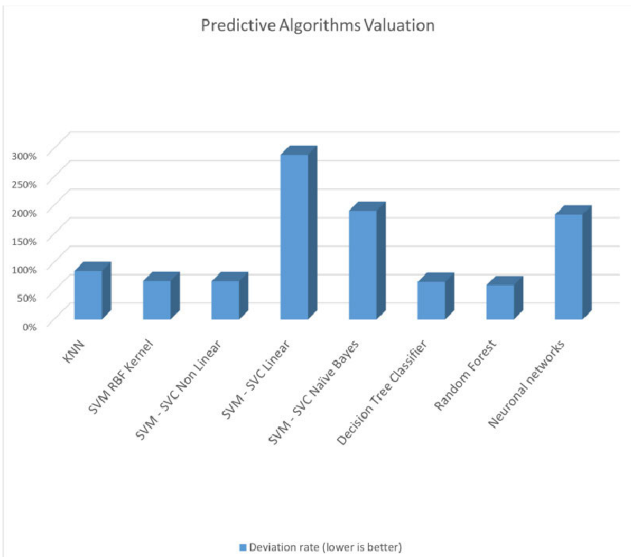

FIGURE 3. Deviation rate algorithms results overview.

review. Similarly, in this research, we used several methods to obtain the best results, as tables show in Fig. 3 and Fig. 4.

After reviewing the output feature, i.e. the number of predicted days that patients remain in hospital, we observed that the results were a small group of discrete values. These results lead us to decide using classification algorithms instead of regression algorithms as initially planned. Such a small amount of discrete values can be treated as labels that can be later transformed into integers, or even into groups of integers if necessary.

Depending on particular characteristics of every method, a specific procedure was applied to search the hyper-parameters for the best cross-validation score, as it is proposed in [17], where this procedure is used to obtain the optimal parameters. The different methods scanned the data

			Average deviation in days							
Hospital Department	Length of Stay average value	Length of Stay maximum value	KNN	SVM RBF Kernel	SVM - SVC Non Linear	SVM - SVC Linear	SVM - SVC Naïve Bayes	Decision Trees	Random Forest	Neural network
Ophthalmology	2,28	29	0,84	0,84	0,90	1,74	3,00	0,70	0,78	2,26
Obstetrics	3,75	69	1,55	1,48	1,42	4,03	3,23	1,21	1,27	3,15
Endocrinology	8,07	46	4,63	4,56	4,49	12,25	10,00	4,73	3,41	7,10
Radiation Oncology	6,03	102	2,99	3,73	4,19	15,25	12,08	3,22	2,68	30,05
Gynecology	4,83	57	3,35	3,80	4,34	5,32	5,53	2,33	2,20	4,33
Urology	6,93	159	3,83	3,62	3,62	13,22	18,53	3,36	3,42	7,53
Heart Surgery	14,68	153	11,89	10,99	11,57	30,74	24,20	7,23	7,11	9,00
Otolaryngology	4,02	114	2,07	1,95	1,98	2,76	4,59	2,18	1,99	7,07
Emergencies	4,87	103	2,69	2,49	2,48	25,72	12,11	2,49	2,37	10,97
Pediatric Surgery	4,45	151	2,55	2,66	2,95	58,95	3,84	2,32	2,29	2,98
Rheumatology	7,49	281	5,90	6,22	6,31	5,92	6,31	4,49	3,90	25,96
Thoracic Surgery	9,41	150	7,97	6,64	6,53	10,01	14,48	5,62	4,93	14,75
General and Digestive Surgery	8,43	295	5,69	5,06	5,03	52,25	17,83	4,98	4,43	11,52
Traumatology	9,93	301	7,72	7,01	6,99	8,29	20,27	5,76	5,37	51,61
Reconstructive Plastic Surgery	5,22	156	3,30	3,06	2,93	12,13	11,91	2,88	2,94	15,00
Intensive Medicine	15,45	304	13,41	13,99	14,20	41,78	16,83	8,64	9,07	10,54
Maxillofacial Surgery	6,44	78	3,83	3,64	3,77	5,38	8,50	4,18	3,76	15,43
Pediatrics	5,39	340	3,38	3,05	3,07	13,73	11,17	3,15	3,17	3,48
Internal Medicine	11,76	188	7,70	7,03	7,02	17,80	10,34	7,03	6,94	7,16
Medical Oncology	8,41	56	5,62	5,39	5,46	13,80	9,47	5,35	5,00	14,08
Neurology	12,73	107	7,78	8,01	7,69	18,60	28,10	8,07	8,14	13,01
Neurosurgery	12,47	172	9,02	9,23	8,62	32,56	15,58	8,02	7,53	14,97
Angiology and Vascular Surgery	10,42	169	7,86	7,25	7,26	24,17	42,87	7,48	6,37	12,53
Cardiology	7,34	275	5,73	4,67	4,77	30,22	20,19	4,54	4,57	7,21
Psychiatry	9,00	78	9,40	6,24	6,57	33,31	72,20	5,60	5,81	25,94
Neonatology	25,23	174	133,06	18,64	18,70	381,51	15,95	62,30	29,16	131,76
Digestive System	9,71	120	8,11	6,80	6,66	12,43	23,98	6,37	6,42	17,34
Clinical Haematology	15,35	88	13,79	10,14	10,92	16,02	20,26	10,43	10,64	23,05
Anaesthesiology and Resuscitation	9,18	83	7,11	7,15	7,09	24,07	11,77	6,07	7,85	29,17
Nephrology	15,08	334	15,60	14,30	14,03	45,77	44,29	13,44	11,50	20,33
Pneumology	9,57	236	8,02	8,72	8,65	14,50	20,20	8,06	7,33	12,76
Dermatology	9,64	128	9,79	11,38	7,75	29,03	12,79	9,33	11,33	17,52

FIGURE 4. Average deviation. Summary of results measured in days.

to configure optimal parameters with GridSearch, RandomizedSearch and HyperOpt Bayesian techniques. We must take into account that not all methods can select the same parameters and that even though several options are passed as parameters, only the one with the best results is used in each algorithm. The methods, which were selected because they were the most used according to state of the art, are shown below. All the methods used in this study, including feature selection and parameters optimisation, are included in the open-source machine learning library Scikit-learn, that supports supervised and unsupervised learning. Some examples are [1] that work with Naïve Bayes, KNN and Decision Tree Classifier, obtaining the best results with the last one, [4] that work with Back-propagation Neural Networks and Support Vector Machines, where the best one was the second algorithm (SVM), [12] that use Neural Networks for a primary analysis, and Decision Trees and Random Forest in secondary analyses, being the best one Random Forest, or [5] that work with Support Vector Machines, Decision Trees and Random Forest, obtaining the best results with the last one.

- *KNN with K nearest neighbours.* The hyper-parameters chosen in this method were the best estimators of weight options (uniform and distance) and the number of neighbours (between 1 and 31).
- *Support Vector Machines (SVM).* Depending on the type of kernel, the methods used from this group are described below.
  - *Polynomial Kernel.* The results were not suitable with the hospital dataset used in this article due to the significant deviation from real values in its predictions.
  - *RBF Kernel.* The hyper-parameters were the best estimators of parameter C (which trades off correct classification of training samples against maximisation of the decision function's margin: 0.001, 0.01, 0.1, 1 and 10) and parameter gamma (which defines how far the influence of a single training example reaches: 0.001, 0.01, 0.1 and 1).
  - *Support Vector Classifiers (SVC).*

- *Non-Linear Kernel.*
- *Linear Kernel.* The hyper-parameters were the best estimators of parameter C (which can define the degree of correct classification that the algorithm has to meet: 0.001, 0.01, 0.1, 1 and 10) and the maximum number of iterations (100, 1,000 and 10,000).

- *Naïve Bayes.* Based on the assumption of independence between predictors, no parameters are needed.
- *Decision Tree Classifier.* The hyper-parameters were the best estimators of maximum depth (between 1 and 120, stepping 2 by 2), minimum number of samples splits (between 2 and 500, stepping 10 by 10), minimum samples leaf (1, 5 and 10) and the impurity measures such as Gini and entropy, which indicate either how to maximise the purity of each split or the extent to which the groupings are homogenised in the trees.
- *Random Forest.* The hyper-parameters were the best estimators of maximum depth (none and 3), maximum features (between 1 and 11), the minimum number of samples split (between 2 and 11), the selection of bootstrap (False and True) and the impurity measures such as Gini and entropy.
- *Neural Networks.* A recurrent neural network was used with 50 previous results in each sequence to obtain the final target output. The network had 30 nodes, an Adam type optimiser and a quadratic loss function.

To evaluate and assess, in an objective manner, which departments are more appropriate to predict LOS, we have used two measurement variables: the error made in the calculation of the days and the success percentage in the calculation of the exact days. The first variable will allow us to decide more accurately which are the more appropriate departments to predict the LOS, while the second is a complementary way to confirm or to determine what department is the most appropriate finally. So, two specific functions were created to calculate the data quality measurement variables, which allowed to choose the more suitable ones. The variables are defined down below.

- *Mean Error.* It measures the percentual deviation between the predicted results and the real values expected. Once we know this percentage (error), we calculate the mean number of days that the predicted value deviates from the real value, which is a metric of the proximity to the real value.
- *Exact Success percentage (accuracy).* It measures the success percentage that an algorithm achieves when predicting the length of stay in a hospital, success meaning that the predicted value matches the real value exactly.

IV. RESULTS

In this study, hospital departments were compared according to the algorithms used to predict results, taking into account those which offer better results in comparative tests. Methods based in decision trees stand out from the rest, as in Decision Tree Classifier and Random Forest algorithms. The same applies to those using RBF kernel, such as the Support Vector Machines (RBF Kernel) and Support Vector Classifiers (Non-Linear Kernel) algorithms. This behaviour can be observed in Fig. 2. Other studies – e.g. [5] – confirm our choice, as their results indicate that the random forest method in surgery patients is the most accurate and stable prediction model among all the methods analysed.

In Fig. 2, the success rate was obtained by calculating the average value of all departments in each algorithm, as can be observed in the following formulas:

$$accuracy_{department} = \frac{TP}{TP + FP} \tag{4.1}$$

where TP are true positive values and FP are false positive values. Finally, the average value was obtained applying:

$$\bar{X}_{algorithm} = \frac{\sum_{i=1}^{n} accuracy_{department_i}}{N} \tag{4.2}$$

where N is the number of departments. In Fig. 3, the deviation rate was obtained by calculating the average value of all departments in each algorithm, as observed in the following formulas:

$$deviation_{department} = \frac{\sum_{i=1}^{n} |X_{p_i} - X_{r_i}|}{S} \tag{4.3}$$

where X_p is the predicted value, X_r is the real value and S is the number of samples. Finally, the average value per department:

$$\bar{D}_{algorithm} = \frac{\sum_{i=1}^{n} deviation_{department_i}}{N} \tag{4.4}$$

where N is the number of departments.

Once the prediction method has been described, we now proceed to detail the results. The departments General and Digestive Surgery obtained very similar results, where the best ones were obtained using the random forest method. Other surgeries in dataset departments (Angiology and Vascular Surgery, Thoracic Surgery, Maxillofacial Surgery, Heart Surgery and Neurosurgery) also confirm random forest as the best method. It is important to remark that the SVM with Polynomial Kernel algorithm was omitted because the results were not appropriate to predict LOS.

The methods producing best results, shown in Fig. 3, offer either a lower average deviation in days or a lower error rate compared to the actual values. For example, the average deviation predicting LOS in Obstetrics was 1.21 days, while it was 0.70 days in Ophthalmology (as shown in Fig. 4). Regarding accuracy, best-ranked departments in the list of more accurate predictions also obtain less average deviation in the projections. In most cases, average success rates are 37.85% and 64.05% in Obstetrics Ophthalmology, respectively (see Fig. 4). Although results with neural networks could have been improved by obtaining more samples and spending more time supporting algorithm development, we decided not to carry out these improvements at the light of the other algorithms' results. The results obtained using several types

Hospital Department	KNN	SVM RBF Kernel	SVM - SVC Non Linear	SVM - SVC Linear	SVM - SVC Naïve Bayes	Decision Trees	Random Forest	Neural network
Ophthalmology	58,69%	59,83%	57,62%	57,07%	9,17%	64,05%	61,34%	1,29%
Radiation Oncology	49,28%	38,57%	30,53%	19,99%	25,94%	49,62%	51,15%	0,26%
Rheumatology	22,15%	19,39%	17,88%	8,60%	22,93%	36,20%	38,24%	0,18%
Obstetrics	30,41%	32,99%	33,94%	2,46%	31,55%	37,85%	37,46%	0,51%
Otolaryngology	32,62%	33,80%	32,20%	20,66%	25,11%	34,72%	35,82%	0,91%
Emergencies	33,33%	35,11%	35,60%	20,78%	23,24%	35,11%	35,54%	0,53%
Pediatrics	23,01%	25,96%	25,64%	4,85%	25,19%	30,17%	33,61%	2,07%
Pediatric Surgery	28,49%	26,31%	27,05%	26,45%	21,76%	33,35%	32,40%	1,08%
Anaesthesiology and Resuscitation	20,95%	22,12%	22,77%	4,22%	24,33%	33,34%	32,18%	0,00%
Reconstructive Plastic Surgery	24,98%	30,57%	32,34%	15,73%	11,56%	30,86%	30,36%	0,51%
Cardiology	20,58%	21,69%	21,38%	11,19%	14,47%	27,41%	28,52%	2,66%
Endocrinology	22,14%	22,70%	24,74%	17,70%	8,12%	24,59%	26,07%	0,00%
Gynecology	16,01%	14,83%	15,08%	10,09%	13,41%	25,88%	25,98%	0,42%
Urology	17,82%	19,32%	18,68%	8,27%	15,47%	24,67%	23,59%	0,70%
General and Digestive Surgery	17,07%	17,77%	17,82%	14,62%	12,21%	21,45%	22,83%	1,57%
Angiology and Vascular Surgery	16,45%	15,97%	15,71%	6,49%	10,20%	19,93%	21,74%	0,82%
Thoracic Surgery	9,13%	12,97%	15,93%	5,53%	11,33%	18,20%	20,86%	0,16%
Psychiatry	13,37%	16,75%	15,40%	5,38%	1,59%	20,68%	20,71%	0,11%
Traumatology	13,61%	14,08%	14,50%	8,38%	9,69%	19,93%	20,25%	0,12%
Neonatology	6,00%	18,43%	18,43%	1,39%	19,51%	8,22%	14,37%	0,00%
Maxillofacial Surgery	16,54%	17,80%	18,26%	3,71%	16,32%	15,39%	19,26%	0,60%
Medical Oncology	14,06%	12,87%	11,73%	3,86%	10,62%	16,12%	17,60%	0,25%
Digestive System	12,44%	12,83%	14,30%	7,20%	9,14%	14,82%	16,96%	0,70%
Heart Surgery	12,68%	13,20%	11,19%	5,62%	5,93%	14,90%	16,13%	0,87%
Intensive Medicine	11,85%	9,42%	9,18%	4,21%	11,72%	14,64%	14,13%	1,17%
Neurosurgery	8,89%	9,91%	9,49%	3,72%	7,32%	12,70%	12,94%	0,72%
Dermatology	9,49%	6,65%	6,49%	3,07%	6,39%	11,84%	9,26%	0,00%
Pneumology	7,50%	8,67%	8,77%	5,58%	7,44%	9,55%	11,24%	1,04%
Nephrology	8,14%	7,98%	7,19%	2,60%	5,23%	10,50%	10,88%	0,66%
Neurology	8,48%	8,87%	9,85%	8,88%	2,67%	8,48%	8,94%	0,40%
Internal Medicine	7,65%	9,06%	8,99%	4,59%	8,55%	9,09%	9,09%	1,10%
Clinical Haematology	4,99%	8,31%	8,84%	5,43%	6,01%	7,88%	8,89%	0,37%

FIGURE 5. **Summary of results showing the average exact success rates (accuracy).**

of neural networks with different configurations were similar between them and for all cases worse than those obtained with the other algorithms. Figures 4 and 5 show the results per departments sorted from more appropriate to less appropriate. Fig. 4 uses the deviation as sorting criterion (from the lowest to highest deviation), while Fig. 5 sorts departments from lower to higher accuracy. Both cases take into account the maximum value of all machine learning methods per department. The selected methods in the sorting criterion were those with decision trees and RBF kernels in their algorithms, due to the achievement of better results.

Fig. 4 shows the average prediction error measured in days, with all methods named above and in every hospital department with sufficient and appropriate data. Thus a list of departments ordered is obtained, where departments on top are those more suitable to predict LOS. As it can be observed, the top list departments are Obstetrics, Ophthalmology and Endocrinology. In Fig. 5, showing the exact success rate in each department, higher prediction rates remain on top, and one can notice almost the same order observed in Fig. 4: Ophthalmology, Radiation Oncology and Rheumatology are listed as the most appropriate departments.

As previously mentioned, every hospital department dataset was subjected to three different series of tests, one per each method described in this article. Regarding the data for these tests, three different sample selections were chosen both for testing and training. The same sample selection was used with all methods in each series of tests to have equal status with other predictions. From these series of tests were obtained the average results shown in Fig. 4 and 5. These tables summarise the results of predicting the LOS. The first one shows the average deviation in days, while the second shows the average accuracy rates. Notice that the departments ''Rehabilitation'' and ''Pain Control Unit in Emergencies'' could not be evaluated due to lack of data.

V. DISCUSSION

Considering the results obtained by other authors, for example in [1], the results of this research are consistent with them, that is, best predictions were given by Decision Tree algorithm, as can be observed in Fig. 5. Other authors obtain better results with Random Forest algorithm as stated in [12] and [5]. In our study, the Random Forest algorithm obtained the best mean predicted value in all hospital departments which is also in concordance with these two articles. This can be seen in Fig. 2.

From the hospital departments perspective, both [1] and [5] analyse General Surgery department getting the best results with Decision Tree and Random Forest algorithms, respectively. Our research results align with them because all surgery departments also obtained the best results with these two algorithms. It must be taken into account that the hospital dataset analysed does not have a unique generic surgery department. The surgery departments that obtained better results with the Decision Tree algorithm were Pediatric Surgery and Reconstructive Plastic Surgery. The best results for the Random Forest algorithm were General and Digestive Surgery, Angiology and Vascular Surgery, Thoracic Surgery, Maxillofacial Surgery, Heart Surgery, and Neurosurgery. By contrast, in [10], in the Cardiology department, the best results were reached with Support Vector Machines algorithms. In contrast, the best results were obtained, once again, by the Random Forest algorithm. Please note that the results also depend on each particular dataset.

After examining the results, we found two main reasons why some departments are more predictable than others. The first reason is the number of records, given that the more registers for a department, the better are the results. The second one refers to having samples (patients) with similar (or the same) diseases; if those samples are either sufficiently equal or sufficiently different, then the prediction for such department is more favorable.

It can be noticed that not the same departments are at the top of both the lists shown in tables of Fig. 4 and 5. This does not necessarily have to be inconsistent, as they are two factors that help to decide better on how to choose the best department. As an example of this, Rheumatology was not on top of the list in Fig. 3, although it was on top of Fig. 4. This fact indicates that this department is also appropriate to predict LOS, but taking into account that, in those cases where prediction fails, the deviation from the real value is high.

The use of descriptive statistics and visual inspection tools such as histograms and scatter plots, as well as the inspection of the experimental results related to the algorithms' accuracy, made visible behaviour patterns that help interpret the final model. Also, clustering database tools and statistical information about the dataset helped in the detection.

The samples that describe those patterns were prepared into special datasets for each corresponding department, and then processed with the algorithms used with the generic dataset, improving their results.

The outcomes were improved in Obstetrics, Ophthalmology, Emergencies and Cardiology departments. It is observed that patients with specific conditions, following a particular pattern, can improve prediction rates. In such a way, these patterns achieve their respective departments as the most appropriate ones to predict LOS.

The observed patterns in each hospital department are listed below:

- **Ophthalmology**: When patients are younger or equal to 75 years of age, the error rate in predictions reduces by 12.01% (in absolute terms, it decreases from 0.91 days to an average of 0.63). Our results demonstrate that the predictions of the exact days of LOS are up to 8.53% better for patients under 75.
- **Obstetrics**: Patients with a prolonged (post-term) high-risk pregnancy, exceeding 42 weeks of gestation, the risk due to higher perinatal mortality and morbidity, the predictions reduce the error rate by up to 10.11%, reducing this rate from 1.21 days to an average of 0.83 days. In the case of predicting the exact days, the results are improved by up to 9.56%.
- **Emergencies**: For the Emergencies department, the results also improve when the inpatients are women, and the reason for their admission is enteritis –an inflammation of the intestines–. The error rate in the predictions decreases by 13.65%, involving a reduction from 2.37 days to an average of 1.71 days. For the projection of the exact days, the results are improved by up to 16.11%.
- **Cardiology**: When the reason for admission is either fibrillation or heart palpitation, the error rate in the predictions is reduced by as much as 20.57%. The variance between the real and the predicted value decreases from 4.54 days to an average of 3.03 days. But if inpatients are men, then the error rate reduces by up to 21.61% and the variance to an average of 2.95 days. The prediction of the exact days, the results are improved by up to 23.83%.

VI. CONCLUSION

This research shows that reductions in hospital costs and an improvement in quality patient care are possible. We prove that estimating the length of stay is possible at the hospital level, contrarily to other department-dependent studies. This means that it is possible to assess which departments are a better choice to save costs. The limitation on the Rehabilitation and Pain Control Unit in Emergencies previously commented should also be considered, as those departments could not be evaluated due to the lack of data. Although the results predicting the exact number of days were not as expected, they were good enough as to provide a ranked list of hospital departments which was fundamental to detect behavioural patterns. Besides, we consider a low accuracy with low deviation in predictions a good result, although it is not realistic to expect getting good results for all departments.

Other practical implications of our findings are of considerable interest. For instance, hospital managers can –according to the number of patients that would be admitted – either reserve a particular number of beds according to the predictions or most accurately decide on the allocation of resources to particular departments. Managers could make these decisions just consulting the ordered lists of departments, both with success rate list (Fig. 5) and average deviation list (Fig. 4).

The improvement in the success rates when predicting the LOS in some hospital departments is interesting, especially the improvement in the days of LOS variance within the same department. These achievements were possible thanks to patients with particular conditions in specific departments. The improvement, for example, in women admitted to the emergency department suffering from enteritis or in the other cases described in the paper is especially relevant. The study carried out here allows observing several behaviour patterns under certain conditions, which allows suggesting which cases are likely to use a particular machine learning method to increase the success rate.

Our research also offers directions to future research in those hospital departments which are the most appropriate ones to predict LOS. This helps avoid unnecessary effort and provides a contextualised vision to hospitals about which departments are susceptible to cost and resource-saving options.

As a final reflection, we think it would be interesting to repeat the analysis of predictions in those departments where results were not favorable when there will be greater availability of new samples. Also, it would be interesting to analyse the length of stay after suffering a big pandemic as COVID-19 to check how it affects to this feature (LOS) and to find out how this fact affects to hospital departments. Recent studies have detected an important reduction, as stated in [16], where LOS have been reduced as much as 20%.

REFERENCES

[1] S. Aghajani and M. Kargari, "Determining factors influencing length of stay and predicting length of stay using data mining in the general surgery department," *Hospital Practices Res.*, vol. 1, no. 2, pp. 51–56, May 2016.

[2] S. Barnes, E. Hamrock, M. Toerper, S. Siddiqui, and S. Levin, "Real-time prediction of inpatient length of stay for discharge prioritization," *J. Amer. Med. Inform. Assoc.*, vol. 23, no. e1, pp. e2–e10, Apr. 2016.

[3] P. Baylis, "Better health care with data mining," SPSS, Shared Med. Syst. Ltd., London, U.K., 2009.

[4] C.-L. Chang and P.-Y. Lu, "The study on evaluating length of hospital stay for myomectomy," *Int. J. Sci. Eng. Invest.*, vol. 5, no. 59, pp. 157–162, 2016.

[5] M.-T. Chuang, Y.-H. Hu, C.-F. Tsai, C.-L. Lo, and W.-C. Lin, "The identification of prolonged length of stay for surgery patients," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2015, pp. 3000–3003.

[6] C. Combes, F. Kadri, and S. Chaabane, "Predicting hospital length of stay using regression models: Application to emergency department," in *Proc. 10th Conférence Francophone de Modélisation, Optimisation et Simulation (MOSIM)*, 2014. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01081557

[7] E. El-Darzi, R. Abbi, C. Vasilakis, F. Gorunescu, M. Gorunescu, and P. Millard, "Length of stay-based clustering methods for patient grouping," in *Intelligent Patient Management* (Series Studies), vol. 189. Berlin, Germany: Springer, 2009, pp. 39–56.

[8] J. C. Ferrão, M. Duarte, F. Janela, and H. Martins, "Predicting length of stay and assignment of diagnosis codes during hospital inpatient episodes," in *Proc. 1st Karlsruhe Service Summit Workshop Adv. Service Res.* Karlsruhe, Germany: KIT Scientific, 2015, pp. 65–72.

[9] A. Freitas, T. Silva-Costa, F. Lopes, I. Garcia-Lema, A. Teixeira-Pinto, P. Brazdil, and A. Costa-Pereira, "Factors influencing hospital high length of stay outliers," *BMC Health Services Res.*, vol. 12, no. 1, p. 265, Dec. 2012.

[10] P. R. Hachesu, M. Ahmadi, S. Alizadeh, and F. Sadoughi, "Use of data mining techniques to determine and predict length of stay of cardiac patients," *Healthcare Informat. Res.*, vol. 19, no. 2, pp. 121–129, 2013.

[11] R. Houthooft, J. Ruyssinck, J. van der Herten, S. Stijven, I. Couckuyt, B. Gadeyne, F. Ongenae, K. Colpaert, J. Decruyenaere, T. Dhaene, and F. De Turck, "Predictive modelling of survival and length of stay in critically ill patients using sequential organ failure scores," *Artif. Intell. Med.*, vol. 63, no. 3, pp. 191–207, Mar. 2015.

[12] R. J. LaFaro, S. Pothula, K. P. Kubal, M. E. Inchiosa, V. M. Pothula, S. C. Yuan, and M. A. Inchiosa, "Neural network prediction of ICU length of stay following cardiac surgery based on pre-incision variables," *PLoS ONE, Public Library Sci.*, vol. 10, pp. 1–19, Dec. 2015.

[13] P. Liu, E. El-Darzi, C. Vasilakis, P. Chountas, W. Huang, and L. Lei, "Comparative analysis of data mining algorithms for predicting inpatient length of stay," in *Proc. PACIS*, vol. 86, 2004, pp. 1087–1097.

[14] P. Liu, L. Lei, J. Yin, W. Zhang, W. Naijun, and E. El-Darzi, "Healthcare data mining: Prediction inpatient length of stay," in *Proc. 3rd Int. IEEE Conf. Intell. Syst.*, Sep. 2006, pp. 832–837.

[15] W. Lowell, G. Davis, W. Lajousky, S. Stieffel, G. Davis, M. Breau, S. Harvey, and H. Shirazi, "A field trial using artificial neural networks to predict psychiatric in-patient length-of-stay," *Neural Comput. Appl.*, vol. 5, no. 3, pp. 184–193, Sep. 1997.

[16] M. Z. AlShehery, B. Duraisamy, A. R. Z. Zaidi, N. F. AlShehry, F. Z. Zaidi, A. A. Rababah, A. A. Assiri, M. S. AlGhamdi, A. Al Mutair, and A. Al-Omari, "COVID-19 and palliative care services: Comparative patterns of inpatient, outpatient, and consultation services in a tertiary care center in riyadh," *Cureus*, vol. 12, no. 12, pp. 1–11, Dec. 2020.

[17] R. M. Mekhaldi, P. Caulier, S. Chaabane, A. Chraibi, and S. Piechowiak, *Using Machine Learning to Predict the Length of Stay in a Hospital Setting Trends and Innovations in Information Systems and Technologies*, vol. 1159. Cham, Switzerland: Springer, 2020, pp. 202–211.

[18] V. U. Panchami and N. Radhika, "A novel approach for predicting the length of hospital stay with DBSCAN and supervised classification algorithms," in *Proc. 5th Int. Conf. Appl. Digit. Inf. Web Technol. (ICADIWT)*, Feb. 2014, pp. 207–212.

[19] C. P. Parag and K. K. Hitesh, "Machine learning techniques for predicting hospital length of stay in pennsylvania federal and specialty hospitals," *Int. J. Comput. Sci. Appl.*, vol. 11, no. 3, pp. 45–56, 2014.

[20] M. Rouzbahman, A. Jovicic, and M. Chignell, "Can cluster-boosted regression improve prediction of death and length of stay in the ICU?" *IEEE J. Biomed. Health Informat.*, vol. 21, no. 3, pp. 851–858, May 2017.

[21] R. Stoean, C. Stoean, A. Sandita, D. Ciobanu, and C. Mesina, "Ensemble of classifiers for length of stay prediction in colorectal cancer," in *Advances in Computational Intelligence*, vol. 9094. Cham, Switzerland: Springer, 2015, pp. 444–457.

[22] E. Suter, N. D. Oelke, C. E. Adair, and G. D. Armitage, "Ten key principles for successful health systems integration," *Healthcare Quart., Toronto, Ont.*, vol. 13, pp. 16–23, Dec. 2009.

[23] L. Turgeman, J. H. May, and R. Sciulli, "Insights from a machine learning model for predicting the hospital length of stay (LOS) at the time of admission," *Expert Syst. Appl.*, vol. 78, pp. 376–385, Jul. 2017.

[24] C.-S. Yang, C.-P. Wei, C.-C. Yuan, and J.-Y. Schoung, "Predicting the length of hospital stay of burn patients: Comparisons of prediction accuracy among different clinical stages," *Decis. Support Syst.*, vol. 50, no. 1, pp. 325–335, Dec. 2010.

[25] E. Yasinski, D. Reilly, N. Duggal, B. S. Walker, E. Carpintero, S. Nag, and C. Hentz, "Understanding & predicting length of stay (LOS) using machine learning," in *Proc. Dexur. 15th Floor, 575 5th Avenue*, New York, NY, USA, vol. 10017, 2017. [Online]. Available: https://dexur.com/a/ml-research-los/6/

JESÚS MANUEL PUENTES GUTIÉRREZ received the B.S. degree in computer systems technical engineering and the M.S. degree in free software from the Universitat Oberta de Catalunya, in 2010 and 2012, respectively, and the B.S. degree in computer engineering from the Universidad de Alcalá, Madrid, Spain, in 2014, where he is currently pursuing the Ph.D. degree in acknowledgement and information engineering. Since 2014, he has been researching with machine learning algorithms and artificial intelligence and the team leader of development and maintenance teams in the management chart of a medical insurances company.

MIGUEL-ÁNGEL SICILIA received the degree in computer science from the Pontifical University of Salamanca and the Ph.D. degree from the Carlos III University of Madrid. Before joining academia, he was an E-Commerce Architect. He is currently a Professor with the Department of Computer Science, University of Alcalá. He has developed his research activity in artificial intelligence, machine learning, and analytics applied to different areas, including learning, health, computational science, command and control systems, and information security. He is also the Editor-in-Chief of *Data Technologies and Applications* Journal (Emerald).

SALVADOR SÁNCHEZ-ALONSO is currently an Associate Professor and a Senior Member of the Information Engineering Group, a research unit of the Computer Science Department, University of Alcalá, Spain. He previously worked as a software engineer with a software solutions company. With a good number of high impact factor publications in the last ten years, his current research interests include social network analysis applications, and Web science and computer science education.

ELENA GARCÍA-BARRIOCANAL received the M.Sc. degree in computer science from the Pontifical University of Salamanca and the Ph.D. degree from the Department of Computer Science, University of Alcalá, Spain. She is currently a Professor. She is also the director of various training programs related to big data and business analytics within UAH. Her research interests include knowledge representations, semantic interoperability, and the storage, mining, and analysis of large volumes of data, areas in which she has been actively working in research projects, publications, and Ph.D. supervisions.