Received March 11, 2018, accepted April 13, 2018, date of publication April 25, 2018, date of current version May 24, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2828652

Predicting the Urgency Demand of COPD Patients From Environmental Sensors Within Smart Cities With High-Environmental Sensitivity

JAVIER MEDINA QUERO[1], MIGUEL ÁNGEL LÓPEZ MEDINA[2], ALBERTO SALGUERO HIDALGO[3], AND MACARENA ESPINILLA[1]

[1]Department of Computer Science, University of Jaén, Campus Las Lagunillas, 23071 Jaén, Spain
[2]Council of Health, Andalusian Health Service, 41071 Sevilla, Spain
[3]Department of Computer Science, University of Cádiz, 11001 Cádiz, Spain

Corresponding author: Javier Medina Quero (jmquero@ujaen.es)

This work was supported in part by the REMIND Project Marie Sklodowska-Curie EU Framework for Research and Innovation Horizon 2020 under Grant 734355and in part by the Council of Health for the Andalusian Health Service, Spain, together the research from the Spanish Government under Project PI-0203- 2016 and Project TIN2015-66524-P.

ABSTRACT Predicting the urgency demand of patients at health centers in smart cities supposes a challenge for adapting emergency service in advance. In this paper, we propose a methodology to predict the number of cases of chronic obstructive pulmonary disease (COPD) from environmental sensors located in the city of Jaén (Spain). The approach presents a general methodology to predict events from environmental sensors within smart cities based on four stages: 1) summarize and expand features by means of temporal aggregations; 2) evaluate the correlation for selecting relevant features; 3) integrate straightforwardly expert knowledge under a fuzzy linguistic approach; and 4) predict the target event with the sequence-based classifier long short-term memory under a sliding window approach. The results show an encouraging performance of the methodology over the COPD patients of the city of Jaén based on a quantitative regression analysis and qualitative categorization of data.

INDEX TERMS Predicting urgency demand, long short-term memory, temporal aggregation, fuzzy linguistic approach.

I. INTRODUCTION

Allocating patients at Emergency Services in smart cities is a key factor in decision making of Health Centers [1]. Predicting this demand enables adapting policies and distributing dedicated resources [2] in order to improve the Emergency Service, especially in crisis scenarios of seasonal stages.

At the same time, from the last years, the number and type of environmental sensors [3] installed in smart cities [4] is increasing due to their low-power, low-cost, high-capacity, and miniaturized [5]. Frequently, these sensors provide different data types and collecting rate, which hinder the data fusion requiring spatiotemporal processing and an *ad hoc* development guided by expert knowledge [6].

Based on these challenges, in this work we propose a methodology which has been implemented to predict the number of cases of Chronic Obstructive Pulmonary Disease (COPD) [7] within the emergency service of Jaén (Spain) from heterogeneous environmental sensors located in the city. The study of COPD patients in the city of Jaén is encouraged by two aims. On the first hand, the incoming from COPD generates a high assistance and economic impact on Andalusian Health Services [8]. On the other hand, Jaén is well-known by leading olive oil-based agriculture [9] being immersed in a agricultural territory surrounded mainly by olive oil trees, which develops a high-environmental sensitivity within the city with a deep repercussion in COPD patients [10].

The remainder of the paper is structured as follows: in Section I-A, previous related works are presented and the approach is introduced; in Section II the methodology to predict events from environmental sensors within Smart Cities is formally defined; in Section III, experiments based on the proposed methodology for COPD patients within the Emergency Service of Jaén (Spain) are performed. Results are discussed in Section IV. Finally, in Section V, conclusions and ongoing adaptations are introduced.

A. RELATED WORKS

On the first hand, forecasting the demand in Emergency Services has been widely studied in previous works. A reference work with a further analysis on the demand of emergency care is [11]. This work is centered on forecasting the number of daily emergency admissions (without analyzing an specific disease) in Bromley Hospitals in United Kingdom. Here, several key factors were identified: (i) the influence of seasonality, (ii) the influence of temporal aggregated weather, (iii) the pre-analysis of features by correlation coefficients, and (iv) the impact of context information from health services, such as number of calls to nurse telephone advice lines.

In other previous works related to the urgency demand, we find [12], where monthly COPD hospitalization were analyzed by a time-series analysis where some external variables, such as, environmental or contextual data, were related to the demand prediction. Similarly, the work [13] details the prediction of daily patient attendances at the pediatric emergency department in Lille Regional Hospital Centre (France), which serves four million inhabitants (7% of the French population). The forecasting of these works is based on the classical and the extension of ARIMA without integrating external variables.

Moreover, the technological advances in smart cities provide a fine granularity in small population nuclei. In this way, in the recent work [14], the monitoring within intra-urban scales enables forecasting events on a neighborhood level during warm weather episodes. In general, the potential of the integration of heterogeneous sources in smart cities is being highly valuable for many fields, such as, i) in [15] utilizes temperature and humidity data to inform both household-level and city-wide prediction of electricity demand; ii) in [16], forecasting and monitoring of energy efficiency developed in public buildings, and iii) in [17] where social aspects (crime, safety and employment) were analyzed from the perception of citizens.

On the second hand, in order to deal with the prediction of temporal events within a smart city from environmental sensors, we propose a general methodology. For that, several stages are defined to process information from sensor data streams to identify features which can be integrated in machine learning algorithms. This methodology has been described in an open and general way in order to be suitable in the integration within future or current implementations based on the persistence of environmental sensors from ambient stations in Smart Cities [18].

Firstly, aggregating and sampling the sensor data streams from environmental sensors are required [19] to provide an homogeneous temporal granularity. Secondly, when the number of environmental sensors in the smart city arises, we aim to detect the relevant sources by means of statistical analysis methods, such as *correlation coefficient* [20], which has been properly proposed for analyzing factors in forecasting demand of emergency care [11]. Thirdly, we normalize the features from the sensor data streams of environmental sensors and identify the most relevant.

Our approach includes a fuzzy linguistic processing in computing the features in order to: (i) include interpretable linguistic information from expert human knowledge [21], (ii) extend the features of environmental sensors by means of linguistic terms [22]–[24], and (iii) normalize the aggregated measurements with linguistic terms, with membership degree is defined between [0, 1]. Our linguistic approach is based on fuzzy logic [25], which has provided successful results in developing intelligent systems from sensor data streams [26]–[29].

Fourthly, a sequence-based classifier learns temporal features from environmental sensors under a sliding window approach [30]. The proposed sequence-based classifier in this work is the Long Short-Term Memory (LSTM). LSTM is a Recurrent Neural Networks which is made up of a chain of repeated modules, called memory cell. A memory cell is composed of an input gate, a self-recurrent connection, a forget gate and an output gate. The cell states of LSTMs can be controlled in order to remove or add information based on the learning of the gates. LSTM has shown promising performance in multivariate time series of observations to recognize patterns of clinical measurements [30], as well as, moderating the forgetting and consolidation of illness memory in readmission prediction [31].

II. METHODOLOGY

In this section we describe the proposed methodology to predict temporal events within a smart city from environmental sensors. This is based on the next following stages:

- Integrating information from heterogeneous sources, chiefly environmental sensors of the city, to collect and summarize the environmental information within a homogeneous temporal granularity.
- Selecting relevant features from environmental sensors based on the correlation regarding the target class to predict. This stage is key to reduce the candidate number of environmental sensors which describe the status of smart cities.
- Applying a fuzzy linguistic processing to sensor data streams by means of expert knowledge to develop linguistic features which describe the environmental sensors in a linguistic way.
- Including a sequence based classifier, such as LSTM [32], to learn a sequence of features under a sliding window approach.

In Figure 1, we show a scheme of the proposed approach, which is widely described in the next Sections.

A. TEMPORAL AGGREGATION FROM HETEROGENEOUS ENVIRONMENTAL SENSORS

In this section we describe how to provide an homogeneous temporal granularity of sensor data streams from environmental sensors by means of aggregation operators.

In a formal way, an environmental sensor s^i provides a sensor data stream $S^i = \{v^i_{t*}, v^i_{t*+\Delta t^i}, v^i_{t*+\Delta t^i \cdot j}\}$ in a current

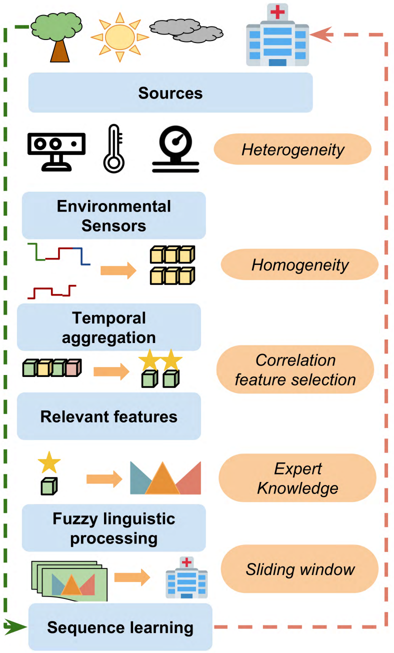

FIGURE 1. Scheme of the approach. From environmental sensors the information processing provides (i) temporal aggregation, (ii) selection of relevant features, (iii) integration of expert knowledge, and (iv) a sequence learning of events.

time t^* within a collecting rate Δt^i, where v_t^i represents a measurement of the sensor s^i.

In order to homogenize the data streams based on the temporal granularity Δt of the target class T, we aggregate those environmental sensors of higher collection rate $\Delta t^i < \Delta t$, which granularity is finer. For that, we propose summarizing the data streams using different aggregation operators, such as *max*, *min*, *sum* or *average*, which are proposed by experts. These aggregated sensor data streams configure the *candidate attributes* to predict events from environmental sensors.

Thereby, several aggregation operators can be defined for a given environmental sensor: (i) enlarging the number of candidate attributes in regard to the original number of environmental sensors and (ii) keeping a homogeneous temporal granularity Δt for all environmental sensors.

In a formal way, a temporal aggregation operator $O^k_{\Delta t}$ summarizes a data stream S^i within a homogeneous temporal granularity defined by Δt obtaining a candidate attribute S_k^i:

$$\begin{aligned} S_k^i &= O^k_{\Delta t}(S^i) \\ O^k_{\Delta t}(S^i) &= \{O^k_{t^*}(S^i), \ldots, O^k_{t^*+\Delta t \cdot j}(S^i)\} \end{aligned} \tag{1}$$

where each measurement obtained by $O_t^k(S^i)$ is included in the candidate attribute S_k^i with an aggregation operation $\cup^k$, which describes the semantic of the operator O^k, from the original measurement $v^i_{t'}$:

$$O_t^k(S^i) = \bigcup^{k}_{t' \in [t, t-\Delta t]} v^i_{t'} \tag{2}$$

At this point, the data stream from environmental sensors have been aligned and aggregated to provide a set of *candidate attributes* with homogeneous temporal granularity.

B. CORRELATION FEATURE SELECTION FROM ENVIRONMENTAL SENSORS

In this section, we describe how to select the most *relevant features* from candidate attributes based on the correlation regarding the target class which defines the event to predict. The aim of this stage is (i) to reduce the number of attributes involved, which increases the performance and avoid the over-fitting [33] and (ii) to obtain a value of relevance from features regarding the target class, which provides an interpretable metric for experts [34].

For that, we compute the correlation from candidate features, which are composed of each data sensor stream S_k^i from previous stage, to the target class T which defines the event to predict. Correlation feature selection is a well-know method based on the principle that relevant features are those with are highly correlated with the target classification [35]. Additionally, for selecting relevant features, additional metrics [36], [37] or models based on decision trees [33] have been effectively proposed and they can be taken into consideration in other works and contexts.

In this work we apply the *Pearson correlation coefficient* [20], which provides a measure of the linear correlation between two features in the range $[-1, 1]$:

$$\rho_{S_k^i,T} = \frac{cov(S_k^i, T)}{\sigma S_k^i \cdot \sigma T} \tag{3}$$

where $cov(S_k^i, T)$ represents the co-variance between a given feature S_k^i and the target class T, and $\sigma S_k^i, \sigma T$ their standard deviations, respectively.

Finally, an α-cut is applied to the absolute value of coefficient for selecting the most relevant features from environmental sensors $S^*_{k,i}/|\rho_{S_k^i,T}| > \alpha$. We note positive and negative correlation is also properly described by the sign of the correlation coefficient.

For sake of simplicity, we can write S_*^i for the *relevant features* from the environmental sensor s^i instead of $S^*_{k,i}$.

C. COMPUTING LINGUISTIC FEATURES FROM DATA STREAMS OF ENVIRONMENTAL SENSORS

In this section we describe how to compute *linguistic features* from the relevant features S_*^i obtained in previous stages.

For each relevant aggregated data stream S_*^i, we associate a fuzzy linguistic terms V_r^i to describe the measures of the sensor $O_t^k(S^i)$. V_r^i represents a fuzzy linguistic set characterized by a membership function $\mu_{\tilde{V}_r^i}(O_t^k(S^i))$, which computes the degree of membership of a measurement $O_t^k(S^i)$ in the fuzzy linguistic set V_r^i. For sake of simplicity, we write V_r^i instead of $\mu_{\tilde{V}_r^i}(O_t^k(S^i))$ and $V_r^i(t)$ instead of $O_t^k(S^i)$.

Thereby, the linguistic terms V_r^i, which describe the environmental sensors S_*^i, configures the features by means a numerical representation between [0, 1] defined by the degree of membership.

Finally, we note step provides also a normalization of data, which has been demonstrated to improve the Machine Learning capabilities [38]. Other relevant statistical approaches, which are non based on expert knowledge, for this purpose could be included in ongoing works, such as Kernel Fisher Discriminant Analysis [39] or Discriminative Linear Transforms [40].

1) CONFIGURING LINGUISTIC TERMS WITH FUZZY SCALES

In order to provide a straightforward methodology to describe the linguistic features using fuzzy linguistic scales, which provide high interpretability with minimal expert knowledge. In concrete, we propose:

- To model the relevant features by means of Computing with Words, where fuzzy linguistic multi-granular modelling [21] enables experts expressing their preferences using a linguistic notation.
- To propose a fuzzy linguistic scale $|\bar{L}^i|$ of granularity g to describe each relevant environmental sensor S^i_*, in where each term $\bar{A}^i_l$ is characterized by using a triangular membership function $\mu_{\bar{A}^i_l}(x)$ [41]. The terms within the fuzzy linguistic scale (i) fit naturally and equally ordered within the domain of discourse of the sensor data stream S^i_* from the interval values $[L_1, \dots L_g]$, (ii) fulfill the principle of overlapping to ensure a smooth transition [42].

$$|\bar{L}^i| = \{\bar{A}^i_1, \dots, \bar{A}^i_l, \dots, \bar{A}^i_g\}, \tag{4}$$

- To enable experts aggregating $A^i_{k\cup k+1}$ terms to describe hesitant fuzzy linguistic terms [43], which configure trapezoidal fuzzy sets [44] within fuzzy scales [45].

In this way, the experts select those *linguistic features* V^i_r from the terms $V^i_r \in \bar{A}^{i\,*}_r$ within the fuzzy scale $|\bar{L}^i|$ which better describe the relevant features S^i_* based on the expert criteria with minimal configuration using a intuitionistic representation.

D. LEARNING A SEQUENCE OF FEATURES UNDER A SLIDING WINDOW APPROACH

In this section, we describe how to develop a *sequence features* from environmental sensors under a sliding window approach.

Previously, following the proposed methodology, we have computed the linguistic features V^i_r, which describe a temporal aggregation $V^i_r(t)$ from the relevant environmental sensor S^i_* in a given time-stamp t .

In order to describe the evolution of the membership degrees from linguistic terms within a previous time interval, we include a sliding window of size W to compose a sequence of features. In this way, from a time-stamp t_k, which represents a point of time $t_k = t^* - \Delta t \cdot k$ from the current time t^* and the temporal granularity Δt of the target class T, we obtain:

- The target value $T(t_k)$ of the class T to learn.
- For each relevant environmental sensor S^i_*, a *sequence of features* $V^i_r(t_k)^*$ defined by the sliding window $t^*_k = \{t_k, t_{k-1}, \dots, t_{k-W}\}$ from the current time t^* :

$$\begin{aligned} V^i_r(t_k)^* &= \{V^i_r(t^*_k)\} \\ k^* &\in \{0, \dots, W\} \end{aligned} \tag{5}$$

Finally, $T(t_k)$ and $V^i_r(t_k)^*$ represent respectively the target class (output) and *sequence features* (input) to be learned by a sequence-based classifier. Here, we propose LSTM as sequence-based classifier modeled by a Recurrent Neural Network for multivariate time series. Other approaches, such as Hidden Markov Models [46] could be similarly integrated in other contexts, although LSTM has been demonstrated to overcome the performance in sequence-based classifiers in several domains [47].

III. EXPERIMENTAL SETUP

In this section we describe the experimental setup applied over a real scene of prediction the urgency demand of COPD patients from environmental sensors within smart cities with high-environmental sensitivity.

The data included in this section correspond to the city of Jaén (Andalucia, Spain), where several heterogeneous environmental sensors are installed to collect the information of the locality and the number of cases of (COPD) within the Emergency Service within the Andalusian Health Service.

A. DESCRIPTION OF DATA

The data source collected in this work is related to the environmental sensors from the Andalusia Environment Council[1] and the Spanish Society of Allergology and Clinical Immunology.[2] The environmental sensors located in the city of Jaén are:

- ***Pollutant***. The specific pollutants studied in relation to COPD are particulate matter (PM10), carbon monoxide (CO), sulfur dioxide (SO2), nitrogen dioxide (NO2) [48].
- ***Ambient airborne pollen***. It is associated to respiratory hospital admissions [49]. Moreover, the pollen species reveal significant effects in a separated way in hospital admissions [10].
- ***Atmospheric phenomena***. It is described as set of environmental factors which affect exacerbation of COPD visits within Emergency Service [50].
- ***The day of the week***. It is included as additional source, which is non based on ambient sensors, due to the fact that the behavior of an Emergency Service in Spain has been demonstrated to be affected by the day of the week [51]. In concrete, Mondays and weekends were identified as the days of most visits with regards to ending working days.

In Table 1 we show the acronym, magnitude and collecting rate.

[1] http://www.juntadeandalucia.es/medioambiente/site/portalweb/

[2] http://www.seaic.org/

TABLE 1. Environmental sensors located in the city of Jaén. The acronyms for ambient stations are GP) gaseous pollutants, WD) weather data and PL) pollen level.

Acronym	Sensor	Magnitude	Station	Collecting rate
SO2	Sulfur dioxide	μ_g/m^3	GP	Minutes
CO	Carbon monoxide	μ_g/m^3	GP	Minutes
NO2	Nitrous oxide	μ_g/m^3	GP	Minutes
O3	Ozone	μ_g/m^3	GP	Minutes
PM10	PM10	μ_g/m^3	GP	Minutes
TEMP	Temperature	°C	WD	Hour
RAIN	Rain fall	l/m^2	WD	Hour
HUM	Relative humidity	%	WD	Hour
WIND	Wind	Km/h	WD	Hour
OLEA	Olea	u/m^3	PL	Day
AMA	Amaranta	u/m^3	PL	Day
CUP	Cupra	u/m^3	PL	Day
GRA	Graminea	u/m^3	PL	Day

TABLE 2. Aggregation operators for environmental sensors.

Sensor	Aggregation operators
SO2	avg,max,min
CO	avg,max,min
O3	avg,max,min
NO2	avg,max,min
MP10	avg,max,min
TEMP	avg,max,min
RAIN	sum
HUM	avg
WIND	max
OLEA	max
AMA	max
CUP	max
GRA	max

TABLE 3. Correlation coefficient $\rho_{S^i_k,T}$ of relevant features: Feature I) further stationary and seasonal stage; Feature II) seasonal stage.

$\rho_{S^i_k,T}$	Feature I
0.73	max(OLEA)
0.49	max(GRA)
0.34	max(AMA)
0.11	max(O3)
-0.15	avg(CO)
$\rho_{S^i_k,T}$	Feature II
-0.31	avg(TEMP)
0.27	avg(HUM)

Secondly, the number of COPD patients (NCP) registered in the Hospital Emergency represents the target class. Data have been collected from the Council of Health for the Andalusian Health Service[3] within a confidential protocol in accordance with the ethical standards of the 1964 Declaration of Helsinkwhich which guarantees the anonymity of the data being summarized the number of cases for each day. The NCP has been obtained from a recent codification process which categorizes a case of urgency to a specific label. The data from this system includes two recent years (2015 and 2016).

We note, the target class describes an imbalanced dataset [52], [53] where the distribution of NPC is not equally represented. The average and standard deviation of NPC are $\mu = 2.70, \sigma = 4.43$, respectively. Moreover, two main stages are clearly identified within NCP: (i) *seasonal stage* where $NCP > 10$ and the distribution parameters are $\mu = 14.60$, $\sigma = 10.31$; and (ii) *stationary stage* where $NCP <= 10$ and the distribution parameters are $\mu = 1.83, \sigma = 1.48$.

B. TEMPORAL AGGREGATION

The temporal granularity of the measurements of the target class (NCP), which determine the the number of cases of COPD patients, is $\Delta t = 24h$. However, some collecting rates from the environmental sensors varies between minutes and hours.

In order to provide a temporal aggregation from heterogeneous environmental sensors, which was discussed in Section II-A, we include the aggregation operators to obtain the *candidate attributes*, which are described in Table 2.

C. SELECTION OF RELEVANT FEATURES

This section is focused on selecting the *relevant features* from once of the candidate. For that, as we described en Section II-A, we compute the *Pearson correlation coefficient* [20] between the candidate features and the target class NCP $S^*_{k,i} = |\rho_{S^i_k,T}| > \alpha$ with $\alpha = 0.1$ obtaining the results shown in Table 3:

Moreover, as the correlation is affected by the imbalanced distribution within the *seasonal stage*, we have also computed the *Pearson correlation coefficient* only within the *stationary stage* to extract relevant features for it, obtaining the results shown in Table 3.

D. COMPUTING LINGUISTIC FEATURES

Next, as we detail in Section II-C, we describe the relevant features using linguistic approach with a minimal expert knowledge.

For that, we generate for each relevant feature a normalized linguistic scale L^i with granularity $g = 5$, where proposed linguistic terms fit naturally ordered within the domain of discourse of the environmental sensor. A granularity 5 in scales increases the response rate and response quality, and it is readily comprehensible to enabling experts to express their view [54]. In order to expand the number of features of based on their correlation coefficients, we set the number of linguistic features based on the criteria:

$$|\rho_{S^i_k,T}| = \begin{cases} 1, & \rho_{S^i_k,T} \leq 1/3 \\ 2, & 1/3 < \rho_{S^i_k,T} < 2/3 \\ 3, & \rho_{S^i_k,T} \geq 2/3 \end{cases}$$

Thereby, the selected linguistic terms, which describe and expand the relevant features, are shown in Figure 2 and Table 4. They compose the final features to be learned by the sequence-based classifier.

In the case of the non-ambient feature *day of the week*, we included a straight forward fuzzyfication based on the fact that the days closer to Monday and weekend have been demonstrated to have a higher impact in urgency demand [51]. So, the domain of the days of the week is related

[3] http://www.juntadeandalucia.es/servicioandaluzdesalud

TABLE 4. Relevant features and linguistic terms selected from normalized fuzzy scale. The columns represent (from left to right): F) relevant feature, R) positive or negative correlation, $|\rho_{S^i_k,T}|$) number of linguistic features, $|S^i_*|$) domain of discourse, and V^i_r) selected linguistic terms by expert.

F	R	$\lvert\rho_{S^i_k,T}\rvert$	$\lvert S^i_*\rvert$	V^i_r
avg(HUM)	$+$	1	$[20, 100]$	$A_{4\cup5}$
avg(TEMP)	$-$	1	$[0, 35]$	$A_{1\cup2}$
avg(CO)	$-$	1	$[50, 700]$	$A_{1\cup2}$
max(O3)	$+$	1	$[40, 175]$	$A_{4\cup5}$
max(AMA)	$+$	2	$[0, 60]$	$A_{3\cup4}$ $A_{4\cup5}$
max(GRA)	$+$	2	$[0, 600]$	$A_{3\cup4}$ $A_{4\cup5}$
max(OLEA)	$+$	3	$[0, 20k]$	$A_{2\cup3}$ $A_{3\cup4}$ $A_{4\cup5}$

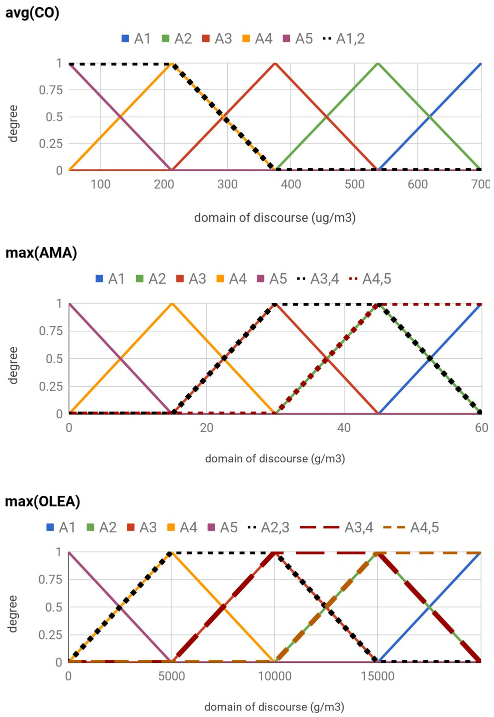

FIGURE 2. Triangular fuzzy scales (in regular line) and proposed linguistic terms (in dotted lines) for the environmental sensors: max(Olea), max(AMA) and avg(CO), which describe an example of selection for 1, 2 and 3 relevant linguistic features.

to a degree of relevance $\{MO, TU, WE, TH, FR, SA, SU\} \rightarrow \{1, 0.75, 0.5, 0.25, 0.25, 0.5, 0.75\}$.

E. LEARNING A SEQUENCE FEATURES UNDER A SLIDING WINDOW APPROACH WITH LSTM

Following the methodology described in Section II-D, we compose a *sequence features* from the linguistic features which describe the environmental sensors under a sliding window approach. The size of the sliding window has set to $W = 30$ that represents a sequence features of thirty days.

TABLE 5. RMSD and MAE in regression on NCP.

Day ahead	Measure	Value
0-day	MAE	1.29
0-day	RMSD	2.11
1-day	MAE	1.27
1-day	RMSD	2.04

The sequence-based classifier proposed in this work is LSTM due to the encourageing performance developed in Machine Learning [55]. The configuration of LSTM was:

- Learning rate $= 0.0001$, to work well as standard parameterization [56], [57].
- Number of neurons $= 64$, as a minimal reference value in learning patterns in RNN [58]
- Number of layers $= 3$, because a great number of layer increases exponentially the learning time without significance in accuracy [59].
- Batch size $= 150$ and training epochs $= 30$, in order to provide an adequate learning within LSTM from 9000 samples.

IV. RESULTS

In this section we present the results obtained under the experimental setup described in Section III and two main analysis:

- **Regression** of the NPC from target class, which is represented by a continuous quantitative variable.
- **Classification**, which is represented by a qualitative categorization from the value of NPC.

In both analysis, the results are evaluated within the approaches of:

- **1-day ahead**, which determines the NPC based on the sequence of features from previous days without taking into account the information from the current day.
- **0-day ahead**, which determines the NPC based on the sequence of features from the current day to the previous day.

In all cases, the evaluation has been configured under a leave-one-out cross validation [60] within further timeline.

A. REGRESSION OF NCP

In the regression analysis, the number of COPD patients (NCP) registered in the Hospital Emergency defines the continuous quantitative variable to be predicted in the target class. Here, LSTM has been configured within an Adam Optimizer and the cost function as mean squared error [61].

The Figure 3 shows the NCP in the two years dataset together with 1-day and 0-day ahead predictions. Root-mean-square deviation (RMSD) and mean absolute error (MAE) are detailed in Table 5.

B. CLASSIFICATION OF NCP

In the classification analysis, the quantitative value has been translated to a qualitative categorization of NPC with three variables:

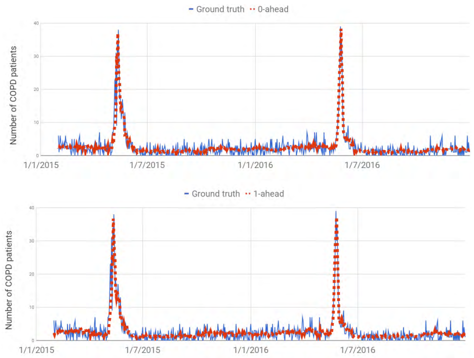

FIGURE 3. Regression on NCP for 0-day and 1-day ahead.

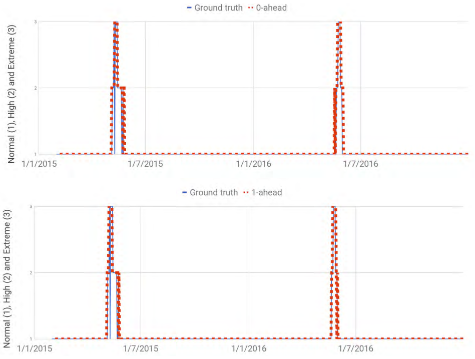

FIGURE 4. Classification on NCP (Normal (1), High (2) and Extreme (3)) for 0-day and 1-day ahead.

- *Normal*, if the number of cases is ordinary $NCP < 10$, which mainly describes the stationary stage.
- *High*, if the number of cases is unusually high $10 <= NCP <= 25$.
- *Extreme*, if the number of cases represents an alarming state for the resources and planing within the Health Center $NCP > 25$.

The activation function for classification over LSTM is *softmax* [62] trained under *cross entropy* [63] as cost function.

We have included two metrics for evaluating the classification provided by the approach: (i) *accuracy* within the timeline, which represents the percentage correctly classified, being TP, true positives, TN, true negatives, FP, false positives and, finally, FN, fase negatives: $acc = \frac{TP+TN}{TP+TN+FP+FN}$; and (ii) *averaged F1-score* from discrete variables, which provides an insight into the balance between classes considering $precision = \frac{TP}{TP+FP}$, and $recall = \frac{TP}{TP+FN}$.

The Figure 4 shows the qualitative discrete variables *Normal (1), High (2) and Extreme (3)* of NCP in the two years dataset together with 1-day and 0-day ahead predictions. *Accuracy* and *averaged F1-score* are detailed in Table 6.

C. DISCUSSION

Overall, the performance of the approach for predicting the urgency demand of COPD patients in the city of Jaén from environmental sensors is encouraging.

TABLE 6. Accuracy (Acc) and averaged F1-score (Avg F1) in regression on NCP.

Day ahead	Measure	Value
0-day	Acc	98.16%
0-day	Avg F1	82.49%
1-day	Acc	98.50%
1-day	Avg F1	81.39%

On the first hand, the classification of LSTM on NCP has described a good performance in the accuracy within the timeline (up to 98%). For solving imbalance between classes in metric, we evaluate the *averaged F1-score* obtaining a notable result of 82%. More prominently, the evaluation of predictions 1-day ahead has presented similar results with 0-day ahead, which enables the anticipation of the resources in the Emergency Service.

On the second hand, the regression of LSTM on NCP provides a suitable estimation of COPD patients with $MAE = 1.2$ and $RMSD = 2.0$, with a slight improvement 1-day ahead. LSTM has learned the critical impact of *seasonal stage* ($\mu = 14.60$, $\sigma = 10.31$) on the number of cases in urgency demand within the Health Service. The main lack is predicting NPC within the *stationary stage* ($\mu = 1.83$, $\sigma = 1.48$), where slight fluctuations have not be related to the temporal evolution of environmental sensors within the city.

In this way, we suggest other sources not related in this work (due to current impossibility of technical integration) could develop an impact in urgency demand of COPD patients, highlighting:

- Erroneous codification or diagnosis. There is a high percentage of patients with COPD erroneously diagnosed, especially in the field of primary care [64].
- Impact of environmental context from neighboring localities, due to the fact that 42.24% of current COPD patients are not living in Jaén city where environmental sensors are located, but in bordering villages within the province.
- Influenza infection [65] with coexisting diseases, such as *seasonal flu* [66].
- Imprecision of Particulate Matter (PM10), which granularity to detect consistent data with the daily cycle of gaseous pollutants emitted by traffic, requires higher precision sensors, such as, PM1 or PM2.5 [67].

V. CONCLUSIONS AND ONGOING WORKS

In this work, we present a general methodology to predict events in Smart Cities from environmental sensors located within it. The main proposed stages are introduced in reference works, but we mainly include: (i) a linguistic approach to integrate expert knowledge with minimal configuration based on the intuitionistic representation; (ii) LSTM as sequence based classifier, which shows promising performance to recognize patterns in multivariate time series.

The results describe a further evaluation to predict urgency demand of COPD patients within the city of Jaén.

A quantitative regression and qualitative classification are analyzed showing a good performance in predicting the evolution of incidences between *seasonal stage* and *stationary stage*. In the case of *seasonal stage*, the slight fluctuations in number of COPD patients has not been related to data collected from environmental sensors, which is identified in similar problems in literature to the low precision of PM10 measures. In this way, we agree that the quality and cost of ambient stations are currently identified as key factors in the deployment of Smart Cities [68].

In ongoing works, other external factors, which could impact in urgency demand of COPD patients, are expected to be further analyzed. Moreover, we will focus on relating the stational flu within the environmental sensors and Health Centers from Andalusian Health Service. For that next purposes, several environmental stations located in different cities and Health Centers will be involved, which geolocation will be key to be integrated and weighted in order to provide a spatial forecasting in larger territories. In this onoging work, the current methodology will present the advantage of identifying and describe the relevant features from environmental sensors for each citiy, which will provide an interesting and descriptive analysis of environmental factors based on the geolocation.

REFERENCES

[1] M. Williams, "Hospitals and clinical facilities, processes and design for patient flow," in *Patient Flow: Reducing Delay in Healthcare Delivery*. Springer, 2006, pp. 45–77.

[2] G. G. van Merode, S. Groothuis, and A. Hasman, "Enterprise resource planning for hospitals," *Int. J. Med. Informat.*, vol. 73, no. 6, pp. 493–501, 2004.

[3] J. Gabrys, "Automatic sensation: Environmental sensors in the digital city," *Senses Soc.*, vol. 2, no. 2, pp. 189–200, 2007.

[4] L. G. Anthopoulos, "Understanding the smart city domain: A literature review," in *Transforming City Governments for Successful Smart Cities*. Granada, Spain: Univ. Granada, 2015, pp. 9–21.

[5] B. P. L. Lau, N. Wijerathne, B. K. K. Ng, and C. Yuen, "Sensor fusion for public space utilization monitoring in a smart city," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 473–481, Apr. 2017.

[6] J. L. Castro, M. Delgado, J. Medina, and M. D. Ruiz-Lozano, "Intelligent surveillance system with integration of heterogeneous information for intrusion detection," *Expert Syst. Appl.*, vol. 38, no. 9, pp. 11182–11192, 2011.

[7] T. F. Murphy and S. Sethi, "Chronic obstructive pulmonary disease," *Drugs Aging*, vol. 19, no. 10, pp. 761–775, 2002.

[8] J. L.-C. Bodineau *et al.*, "Análisis de los ingresos por enfermedad pulmonar obstructiva crónica en andalucýa, año 2000," *Arch. de Bronconeumol.*, vol. 38, no. 10, pp. 473–478, 2002.

[9] L. Viladomiu and J. Rosell, "12. Olive oil production and the rural economy of Spain," in *Sustaining Agriculture and the Rural Environment: Governance, Policy, and Multifunctionality*. London, U.K.: Edward Elgar Publishing, 2004, p. 223.

[10] J. Díaz, C. Linares, and A. Tobías, "Short-term effects of pollen species on hospital admissions in the city of Madrid in terms of specific causes and age," *Aerobiologia*, vol. 23, no. 4, pp. 231–238, 2007.

[11] S. A. Jones, M. P. Joy, and J. Pearson, "Forecasting demand of emergency care," *Health Care Manage. Sci.*, vol. 5, no. 4, pp. 297–305, 2002.

[12] A. Gershon, D. Thiruchelvam, R. Moineddin, X. Y. Zhao, J. Hwee, and T. To, "Forecasting hospitalization and emergency department visit rates for chronic obstructive pulmonary disease. A time-series analysis," *Ann. Amer. Thoracic Soc.*, vol. 14, no. 6, pp. 867–873, 2017.

[13] F. Kadri, F. Harrou, S. Chaabane, and C. Tahon, "Time series modelling and forecasting of emergency department overcrowding," *J. Med. Syst.*, vol. 38, no. 9, p. 107, 2014.

[14] R. J. Ronda, G. J. Steeneveld, B. G. Heusinkveld, J. Attema, and A. Holtslag, "Urban finescale forecasting reveals weather conditions with unprecedented detail," *Bull. Amer. Meteorol. Soc.*, vol. 98, no. 12, pp. 2675–2688, 2017.

[15] S. Suffian, D. P. de Leon Barido, and P. Singh, "Temperature and humidity dependence for household- and city-wide electricity demand prediction in Managua, Nicaragua," in *Proc. Int. Conf. Ubiquitous Comput. Ambient Intell.*, 2017, pp. 721–727.

[16] J. Massana, C. Pous, L. Burgas, J. Melendez, and J. Colomer, "Identifying services for short-term load forecasting using data driven models in a smart city platform," *Sustain. Cities Soc.*, vol. 28, pp. 108–117, Jan. 2017.

[17] Z. Khan, A. Anjum, K. Soomro, and M. A. Tahir, "Towards cloud based big data analytics for smart future cities," *J. Cloud Comput.*, vol. 4, p. 2, Dec. 2015.

[18] J. H. Lee, M. G. Hancock, and M.-C. Hu, "Towards an effective framework for building smart cities: Lessons from Seoul and San Francisco," *Technol. Forecasting Social Change*, vol. 89, pp. 80–99, Nov. 2014.

[19] S. Gebbert and E. Pebesma, "A temporal GIS for field based environmental modeling," *Environ. Model. Softw.*, vol. 53, pp. 1–12, Mar. 2014.

[20] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise Reduction in Speech Processing*. Berlin, Germany: Springer-Verlag, 2009, pp. 1–4.

[21] J. A. Morente-Molinera, I. J. Pérez, R. Ureña, and E. Herrera-Viedma, "On multi-granular fuzzy linguistic modelling in decision making," *Procedia Comput. Sci.*, vol. 74, pp. 49–60, 2015.

[22] M. Espinilla, J. Medina, Á.-L. García-Fernández, S. Campaña, and J. Londoño, "Fuzzy intelligent system for patients with preeclampsia in wearable devices," *Mobile Inf. Syst.*, vol. 2017, Oct. 2017, Art. no. 7838464.

[23] J. Medina, L. Martínez, and M. Espinilla, "Subscribing to fuzzy temporal aggregation of heterogeneous sensor streams in real-time distributed environments," *Int. J. Commun. Syst.*, vol. 30, no. 5, p. e3238, 2017.

[24] J. Medina, M. Espinilla, Á. L. García-Fernández, and L. Martínez, "Intelligent multi-dose medication controller for fever: From wearable devices to remote dispensers," *Comput. Elect. Eng.*, vol. 65, pp. 400–412, Jan. 2018.

[25] L. A. Zadeh, "Fuzzy sets," in *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A Zadeh*. Singapore: World Scientific, 1996, pp. 394–432.

[26] M. Espinilla and C. Nugent, "Computational intelligence for smart environments," *Int. J. Comput. Intell. Syst.*, vol. 10, no. 1, pp. 1250–1251, 2017.

[27] J. Medina, M. Espinilla, and C. Nugent, "Real-time fuzzy linguistic analysis of anomalies from medical monitoring devices on data streams," in *Proc. 10th EAI Int. Conf. Pervasive Comput. Technol. Healthcare*, 2016, pp. 300–303.

[28] J. Medina, M. Espinilla, D. Zafra, and L. Martínez, and C. Nugent, "Fuzzy fog computing: A linguistic approach for knowledge inference in wearable devices," in *Proc. Int. Conf. Ubiquitous Comput. Ambient Intell.*, 2017, pp. 473–485.

[29] J. M. Quero, M. R. F. Olmo, M. D. P. Aguilera, and M. E. Estévez, "Real-time monitoring in home-based cardiac rehabilitation using wrist-worn heart rate devices," *Sensors*, vol. 17, no. 12, p. 2892, 2017.

[30] Z. C. Lipton, D. C. Kale, C. Elkan, and R. Wetzel. (2015). "Learning to diagnose with LSTM recurrent neural networks." [Online]. Available: https://arxiv.org/abs/1511.03677

[31] T. Pham, T. Tran, D. Phung, and S. Venkatesh, "DeepCare: A deep dynamic memory model for predictive medicine," in *Proc. Pacific–Asia Conf. Knowl. Discovery Data Mining*, 2016, pp. 30–41.

[32] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[33] J. R. Quinlan, *C4. 5: Programs for Machine Learning*. New York, NY, USA: Elsevier, 2014.

[34] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, vol. 112. New York, NY, USA: Springer, 2013.

[35] M. A. Hall, "Correlation-based feature selection for machine learning," Ph.D. dissertation, Dept. Comput. Sci., Univ. Waikato Hamilton, Hamilton, New Zealand, 1999.

[36] A. Bellet, A. Habrard, and M. Sebban. (2013). "A survey on metric learning for feature vectors and structured data." [Online]. Available: https://arxiv.org/abs/1306.6709

[37] J. T. Kent, "Information gain and a general measure of correlation," *Biometrika*, vol. 70, no. 1, pp. 163–173, 1983.

[38] S. Ioffe and C. Szegedy. (2015). "Batch normalization: Accelerating deep network training by reducing internal covariate shift." [Online]. Available: https://arxiv.org/abs/1502.03167

[39] L. Bo, L. Wang, and L. Jiao, "Feature scaling for kernel Fisher discriminant analysis using leave-one-out cross validation," *Neural Comput.*, vol. 18, no. 4, pp. 961–978, 2006.

[40] S. Tsakalidis, V. Doumpiotis, and W. Byrne, "Discriminative linear transforms for feature normalization and speaker adaptation in HMM estimation," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 3, pp. 367–376, May 2005.

[41] D.-Y. Chang, "Applications of the extent analysis method on fuzzy AHP," *Eur. J. Oper. Res.*, vol. 95, no. 3, pp. 649–655, 1996.

[42] A. S. Markowski, M. S. Mannan, and A. Bigoszewska, "Fuzzy logic for process safety analysis," *J. Loss Prevention Process Ind.*, vol. 22, no. 6, pp. 695–702, 2009.

[43] R. M. Rodriguez, L. MartiÌĄnez, and F. Herrera, "Hesitant fuzzy linguistic term sets for decision making," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 1, pp. 109–119, Feb. 2012.

[44] S.-H. Chen, "Ranking fuzzy numbers with maximizing set and minimizing set," *Fuzzy Sets Syst.*, vol. 17, no. 2, pp. 113–129, 1985.

[45] S.-M. Chen and J.-A. Hong, "Multicriteria linguistic decision making based on hesitant fuzzy linguistic term sets and the aggregation of fuzzy sets," *Inf. Sci.*, vol. 286, pp. 63–74, Dec. 2014.

[46] S. R. Eddy, "Hidden Markov models," *Current Opinion Struct. Biol.*, vol. 6, no. 6, pp. 361–365, 1996.

[47] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," in *Proc. 15th Annu. Conf. Int. Speech Commun. Assoc.*, 2014, pp. 338–342.

[48] S. S. Salvi and P. J. Barnes, "Chronic obstructive pulmonary disease in non-smokers," *Lancet*, vol. 374, no. 9691, pp. 733–743, 2009.

[49] I. C. Hanigan and F. H. Johnston, "Respiratory hospital admissions were associated with ambient airborne pollen in Darwin, Australia, 2004–2005," *Clin. Experim. Allergy*, vol. 37, no. 10, pp. 1556–1565, 2007.

[50] O. E. Brzezińska-Pawłowska, A. Rydzewska, M. Łuczyńska, B. Majkowska-Wojciechowska, M. L. Kowalski, and J. S. Makowska, "Environmental factors affecting seasonality of ambulance emergency service visits for exacerbations of asthma and COPD," *J. Asthma*, vol. 53, no. 2, pp. 139–145, 2016.

[51] M. Sánchez and A. Smally, "Comportamiento de un servicio de urgencias según el dýa de la semana y el número de visitas," *Emergencias, Revista de la Sociedad Española de Med. de Urgencias y Emergencias*, vol. 19, no. 6, pp. 319–322, 2007.

[52] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002.

[53] N. Japkowicz and S. Stephen, "The class imbalance problem: A systematic study," *Intell. Data Anal.*, vol. 6, no. 5, pp. 429–449, Oct. 2002.

[54] F. Saleh and C. Ryan, "Analysing service quality in the hospitality industry using the SERVQUAL model," *Service Ind. J.*, vol. 11, no. 3, pp. 324–345, 1991.

[55] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 3104–3112.

[56] T. Salimans and D. P. Kingma, "Weight normalization: A simple reparameterization to accelerate training of deep neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 901–909.

[57] Y. Wu, S. Zhang, Y. Zhang, Y. Bengio, and R. R. Salakhutdinov, "On multiplicative integration with recurrent neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 2856–2864.

[58] G. De Pietro, L. Gallo, R. J. Howlett, and L. C. Jain, *Intelligent Interactive Multimedia Systems and Services 2017*. Cham, Switzerland: Springer, 2018.

[59] J. Collins, J. Sohl-Dickstein, and D. Sussillo, "Capacity and trainability in recurrent neural networks," *Stat*, vol. 1050, p. 28, 2017.

[60] P. Refaeilzadeh, L. Tang, and H. Liu, "Cross-validation," in *Encyclopedia of Database Systems*. Boston, MA, USA: Springer, 2009, pp. 532–538.

[61] S. Ruder. (2016). "An overview of gradient descent optimization algorithms." [Online]. Available: https://arxiv.org/abs/1609.04747

[62] N. M. Nasrabadi, "Pattern recognition and machine learning," *J. Electron. Imag.*, vol. 16, no. 4, p. 049901, 2007.

[63] P.-T. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," *Ann. Oper. Res.*, vol. 134, no. 1, pp. 19–67, 2005.

[64] J. De Miguel Dýez, J. I. Alonso, J. M. Parýs, J. R. González-Moro, P. de Lucas Ramos, and G. G. Alonso-Vega, "Fiabilidad del diagnóstico de la epoc en atención primaria y neumologýa en españa. Factores predictivos," *Arch. de Bronconeumol.*, vol. 39, no. 5, pp. 203–208, 2003.

[65] P. Mallia and S. L. Johnston, "Influenza infection and COPD," *Int. J. Chronic Obstructive Pulmonary Disease*, vol. 2, no. 1, pp. 55–64, 2007.

[66] C. Chiatti, P. Barbadoro, A. Marigliano, A. Ricciardi, F. Di Stanislao, and E. Prospero, "Determinants of influenza vaccination among the adult and older Italian population with chronic obstructive pulmonary disease: A secondary analysis of the multipurpose ISTAT Survey on Health and Health Care use," *Human Vaccines*, vol. 7, no. 10, pp. 1021–1025, 2011.

[67] X. Querol *et al.*, "PM10 and PM2.5 source apportionment in the Barcelona Metropolitan area, Catalonia, Spain," *Atmos. Environ.*, vol. 35, no. 36, pp. 6407–6419, 2001.

[68] P. Kumar *et al.*, "The rise of low-cost sensing for managing air pollution in cities," *Environ. Int.*, vol. 75, pp. 199–205, Feb. 2015.

JAVIER MEDINA QUERO received the M.Sc. and Ph.D. degrees in computer science from the University of Granada, Spain, in 2007 and 2010, respectively. He is currently a Post-Doctoral Researcher with the University of Jaén, Spain, and the research group Intelligent Systems Based on Fuzzy Decision Analysis (Sinbad2). His research interests include fuzzy logic, e-health, intelligent systems, ubiquitous computing, and ambient intelligence.

MIGUEL ÁNGEL LÓPEZ MEDINA received the B.Sc. degree in computer science from the University of Jaén, Jaén, Spain, in 2006, and the M.Sc. degree in computer science from the University of Córdoba, Córdoba, Spain, in 2016. He has been with the Andalusia Health Service as a Health Data Analyst Engineer since 2011. His research interests include intelligence systems, fuzzy logic, and big data.

ALBERTO SALGUERO HIDALGO received the B.Sc. and M.Sc. degrees from the University of Granada, Granada, Spain, in 2001 and 2004, respectively, and the Ph.D. degree from the University of Cádiz, Spain, in 2013, all in computer science. He has been an Interim Professor of computer sciences with the University of Cádiz since 2010. His research interests include ontologies and how they can be used to improve information systems. More specifically, the last years he has been focused on applying ontologies to the recognition of activities of daily living.

MACARENA ESPINILLA received the M.Sc. and Ph.D. degrees in computer science from the University of Jaén, Jaén, Spain, in 2006 and 2009, respectively. She is currently an Associate Professor with the Department of Computer Science, University of Jaén. Her research interests include ambient assisted living, activity recognition, behavior modeling, mobile and ubiquitous health, ambient intelligence, group decision-making, decision support systems, intelligence systems, fuzzy logic, and linguistic modeling.