Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 221 (2023) 1276–1283

Procedia Computer Science

www.elsevier.com/locate/procedia

10th International Conference on Information Technology and Quantitative Management

Study on behavioral risk for aging based on smart mattress monitoring data

Hongrui Wang[a,b], Aihua Li[a]*, Ting Chen[c], Jianbing Liu[c]

[a] School of Management Science and Engineering, Central University of Finance and Economics, Beijing 102206, China
[b] School of Sociology and Population Studies, Renmin University of China, Beijing 100872, China
[c] Research Institute for Smart Aging, Beijing Academy of Science and Technology, Beijing 100035, China

Abstract

With an aging population and a gap in demand for professional caregivers, China's elderly care system is facing severe challenges. Using data mining to expand the smart senior care scenario can effectively improve elderly care services. Based on smart mattress datasets, this paper uses machine learning classification and unsupervised anomaly detection models to analyze the possible risks in the daily behavior of older people from the perspectives of both sleep apnea problems and abnormal physiological information. The model results show that the Stacking algorithm based on data fusion can effectively identify the risk of sleep apnea. In contrast, the Prophet and DBSCAN models can carry out anomaly mining of physiological information of single and combined variables, respectively. Ultimately, based on the research, this paper provides targeted recommendations regarding data collection and integration applications.

Keywords: smart senior care; anomaly detection; data fusion

1. Introduction

As the trend of population aging continues to deepen, elderly services have become a hot topic in society. On the one hand, official data from China's Health and Wellness Commission shows that by the end of 2021, there were 267.36 million people aged 60 and above in China, accounting for 18.9% of the total population. It shows that aging has become an essential trend in China's social development. On the other hand, as the life expectancy of the aging

* Corresponding author. Tel.: 86-10-62288622
E-mail address: aihuali@cufe.edu.cn

10.1016/j.procs.2023.08.116

population increases, the proportion of people suffering from diseases and physiological problems gradually increases. According to the fourth sample survey on the living conditions of the elderly in China's urban and rural areas, there are roughly 40.63 million disabled and semi-disabled older adults in China, accounting for 18.3% of the total elderly population. Official figures from China's Health and Welfare Commission show that by the end of 2021, there were 358,000 elderly service institutions and facilities of all kinds in China, with 8.159 million beds for elderly services; based on the national standard ratio of 1:4 of elderly care workers to older adults served, at least 2.03 million elderly care workers will be needed. At the same time, at the end of 2021, there were only 415,800 professional and technical staff in China's elderly care institutions, leaving a shortfall of nearly 1.6 million professionals.

In summary, with the reduced family care staffing and the shortage of professional elderly care workers in demand, China's elderly care system is facing severe challenges. Innovative elderly care services can effectively alleviate this situation. For example, smart mattresses can monitor vital signs and sleep problems for the elderly, which has application value. Therefore, it is essential to vigorously promote intelligent elderly care services to solve the shortage of elderly care service resources. This paper aims to use machine learning methods to analyze and process multi-source heterogeneous data from smart mattresses, detect abnormal behaviors of the elderly that may be at risk, build corresponding analysis and warning models, and improve the level of elderly care services.

2. Related Research

2.1. Behavioral Risks in older people

Current exploration of behavioral risk in older people can be divided into two categories, depending on the abnormality: explicit abnormalities and implicit abnormalities. Explicit anomalies are mainly related to vital signs and can be identified by drastic changes in sensor data, such as abnormal physiological parameters [1,2] or fall detection [3-6], while implicit anomaly detection tends to be more challenging; these types of anomalies generally do not cause significant changes in sensor data and usually need to be mined from daily life patterns such as sleep, diet and toileting [7]. In contrast, from a data mining perspective, academics often use machine learning methods to study these two types of anomalous behavior by training the behavioral data collected by the sensor units to form detection and risk models. Further, we can classify detection methods as supervised and unsupervised learning.

2.2. Smart Mattress

Considering the availability of data, analyzing data for smart mattresses is the primary objective of this research paper. Smart mattresses are intelligent products based on the advanced Internet of Things and various sensor technologies, which are scientifically combined and designed to integrate multiple monitoring functions or auxiliary functions into a single mattress. Due to age and other factors, the elderly are prone to develop their conditions during the bed-ridden period, and sleep problems are frequent and not easy to detect; in response to this feature, the clinical application of smart mattresses in hospitals or nursing homes and other institutions have gradually been promoted, advancing the development process of intelligent elderly care. From the technical point of view, scholars have mainly researched smart mattresses from three perspectives: physiological information monitoring [8], sleep problem detection [9,10] and long-term care functions [11].

2.3. Research Review

As the trend of population aging continues to deepen, research on how to identify the behavioral risks of the elderly has received widespread attention from academia. Identifying behavioral risks in the elderly is mainly based on wearable devices and various types of sensor data in distributed environments, using machine learning methods and mainly focusing on unexpected events such as fall risk. As a typical class of distributed environmental sensors, scholars generally recognize the positive role played by smart mattresses in extracting the physical state of the elderly and detecting sleep problems. In addition, with the development of IoT technology, many scholars have started focusing on fusing multiple sources of heterogeneous sensor data to form an integrated system that can analyze risk [7,12-13]. However, existing studies are mainly from a technical perspective and use laboratory simulation data, and research on deeper mining of actual elderly care data is still relatively limited. Therefore, this paper is innovative and valuable in studying actual elderly care data from the perspective of elderly care, extracting the daily activity patterns of the elderly from smart mattress data, and then detecting the risk of abnormal behavior.

3. Model and Data Description

3.1. Modeling

In this paper, we divided our study into two parts. On the one hand, we classified and predicted the in-bed state based on mattress data, and data fusion methods were used to identify the risk of apnea during sleep. On the other hand, we used the Prophet model and the DBSCAN model to detect abnormal physiological information during mattress use. The critical models involved in the empirical study are as follows.

(1) Stacking

The Stacking algorithm, as a typical representative of integrated learning, is divided into two levels: first, the base learner is trained with the initial data set to obtain new feature data, and then the secondary learners are trained based on the newly generated feature data to output the final model results.

(2) Prophet

The Prophet algorithm is a time series prediction algorithm open-sourced by Facebook in 2017 that can effectively analyze large-scale data. The algorithm is based on a time series decomposition model, which decomposes the time series into three parts: growth trend $g(t)$, cyclical trend $s(t)$ and holiday effect $h(t)$ [14]. The model's specific formula is as follows: $y(t)$ is the target forecast data, and $\varepsilon(t)$ is the error term.

$$y(t) = g(t) + s(t) + h(t) + \varepsilon(t) \tag{1}$$

(3) DBSCAN

The DBSCAN (Density-Based Spatial Clustering of Application with Noise) algorithm is a typical density-based clustering method. The method requires low data set shape and can effectively handle anomalous data, but the clustering quality could be better when the sample density distribution is uneven.

3.2. Data Description

This paper focused on analyzing the smart mattress data provided by a research institution in Beijing, covering the daily mattress use of five elderly users with different physical conditions and mobility. The data type was second-level data transmitted in 3 seconds, including variables such as heart rate, breathing rate, in-bed status and in-bed position, as shown in Table 1.

Table 1. Variable profile of the data set.

Variable	Type	Meaning
state	nominal	mov=large body movement; cou=medium body movement; sle=slight body movement; apn=no breathing; off=away from bed; on=in bed
hr	numerical	Heart rate per minute
resp	numerical	Respirations per minute
weight	numerical	In-bed weight to assist in determining if in-bed
pos_x; pos_y	numerical	Indicates the bed position, with the x-axis oriented from left to right and the y-axis oriented from bottom to top; takes values in the range [1,9]
time	numerical	In principle, the time interval between data transfers is 3 seconds

For the mattress data, most of the device data time interval was from 0:00 am on 3 November to 10:00 am on 14 November 2021. However, there needed to be more data for device number Z59653, with only partial data from 3 to 6 and 11 to 14 November 2021 during the day. Table 2 shows the attributes and known labels for each mattress.

Table 2. Dataset overview.

User number	Off status percentage	Apn status percentage	Missing Data percentage	Data volume	User activity capacity
Z59435	0%	1.7%	0.3%	330,037	fully bedridden
Z59479	48.4%	0.1%	0.6%	326,144	non-fully bedridden
Z59593	0.1%	1.5%	0.6%	327,070	fully bedridden
Z59653	51.3%	1.6%	77.6%	74,634	non-fully bedridden
Z59719	69.9%	2.4%	0.4%	326,696	non-fully bedridden

3.3. Data pre-processing

Observing the raw data set, we found that the data interval was not exactly 3 seconds, and there were also some missing periods. Considering the need for subsequent analysis, we needed to smooth the data and process the missing values. Firstly, each numerical variable of the second-level data was averaged by minute and then smoothed to obtain the minute-level data. Then small missing segments in the minute data, ranging from 1 to 10 minutes, were smoothed using the cubic spline interpolation method.

4. Empirical Results Analysis

4.1. Risk of apnea during sleep

4.1.1. Preparation of empirical study

The empirical preparation was divided into the following three parts.

(1) Sample. Considering that different physical conditions may impact the identification of apnea status, we divided the sample into two categories (fully bedridden and non-fully bedridden) and conducted testing separately.

(2) Data. In order to accurately monitor apnea in the elderly, data related to five older people in the out-of-bed state were first excluded. In addition, max-min standardization and smote sample balancing were performed.

(3) Model. The base classifiers of the Stacking algorithm were set to linear discriminant analysis, decision tree and neural network model. In contrast, the secondary classifier was a logistic regression model. In order to accurately test the algorithm performance, we used a five-fold cross-validation method to divide the training set and test set and toke the mean of five cross-validation results as the final results. This paper used five indicators to evaluate the model performance: accuracy, precision, recall, F1 score and AUC value.

4.1.2. Empirical results

The model results of apnea detection for two fully bedridden elderly and three non-fully bedridden elderly are shown in Tables 3 and 4, respectively.

Table 3. Model classification results (fully bedridden).

Model	Accuracy	Precision	Recall	F1 Score	AUC value
Logistic Regression	0.5506	0.5605	0.6221	0.5846	0.5633
Linear Discriminant Analysis	0.5510	0.5614	0.6221	0.5850	0.5639
Decision Tree	0.8072	0.7509	0.9196	0.8264	0.8238
Multi-Layer Perceptron	0.5120	0.5095	0.6097	0.5676	0.5350
Random Forest	0.8139	0.7578	0.9264	0.8324	0.8825
AdaBoost	0.5776	0.5676	0.6723	0.6145	0.6210
Stacking	**0.8806**	**0.9188**	**0.8543**	**0.8854**	**0.9441**

Table 4. Model classification results (non-fully bedridden).

Model	Accuracy	Precision	Recall	F1 Score	AUC value
Logistic Regression	0.8364	0.8388	0.9218	0.8671	0.8544
Linear Discriminant Analysis	0.8336	0.8331	0.9223	0.8645	0.8542
Decision Tree	0.8795	0.8663	0.9657	0.9028	0.8825
Multi-Layer Perceptron	0.8357	0.8476	0.9092	0.8667	0.8784
Random Forest	0.8838	0.8702	0.9700	0.9071	0.9167
AdaBoost	0.8441	0.8565	0.9127	0.8717	0.8694
Stacking	**0.9205**	**0.9622**	**0.8883**	**0.9238**	**0.9563**

For two elderly groups, the apnea detection results were analyzed in terms of both commonalities and differences:

In common, the data fusion approach significantly improved the classification performance of the apnea detection model. Compared to the traditional base classifier model, the integrated classification model based on data fusion generally achieved better recognition results, regardless of the model integration approach.

In different, the performance of single and integrated models for elderly groups had different mobility abilities. Overall, decision tree, random forest and Stacking models performed excellently across elderly groups. In contrast, logistic regression, linear discriminant analysis and neural network models showed significantly better performance in identifying apnea in fully bedridden elderly than in non-fully bedridden elderly, with some variability.

4.2. Risk of abnormal physiological information

4.2.1. A case of the non-fully bedridden elderly

Due to different modeling principles, we used the Prophet model to examine a single variable of physiological information and the DBSCAN model to examine integrated variables. Considering the data quality, the non-fully bedridden elderly (Z59479) was first selected to detect anomalies using the minute-level data.

Based on the model results, we selected the user's respiratory rate and heart rate data on 13 November for plotting. As shown in Figure 1, the anomalies identified by the models are highlighted in red, while the grey markers are the time series data without anomalies. It is easy to see that the model can effectively detect extreme values, focusing on moments when the respiratory frequency is 30 breaths/minute and above. Moreover, the DBSCAN model is sensitive to higher levels of heart rate, focusing on moments when the heart rate is 70 beats per minute and above, although in the range of average heart rate levels of the elderly, but is a relatively rare occurrence for this older person and warrants ongoing attention by caregivers.

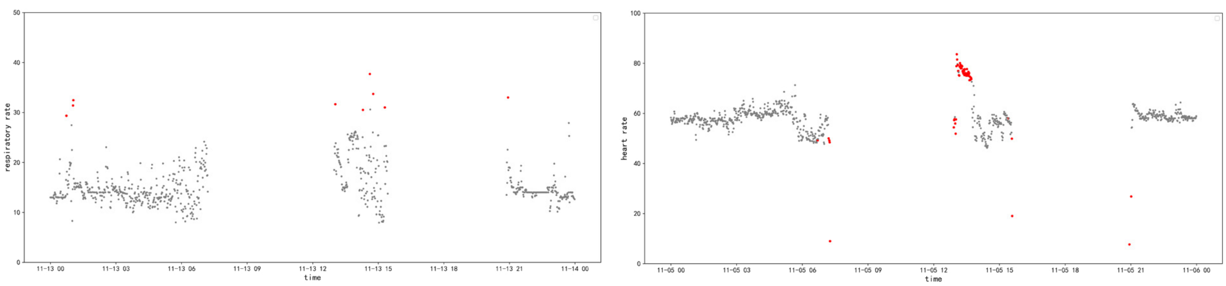

Fig. 1. (a) Prophet model anomaly detection example; (b) DBSCAN model anomaly detection example(Z59479).

In addition to combining medical knowledge and mining physiological information for moments of abnormality, the detection of integrated information can also detect some inter-sensor data mismatches, which are shown in Table 5. For example, the first three data bars at noon on 12 November showed that the smart mattress continued to transmit information about the user's heart rate detection when the respiratory rate and weight indicated that the user was out of bed. The DBSCAN model effectively detected these anomalies. The reason for this could be either an inherent bias in the data collection performed by the sensor or an unexpected situation that was not anticipated by the experiment.

Table 5. Data mismatch between sensors.

Time	Heart rate	Respiratory rate	Weight	Position x	Position y
2021-11-12 12:36	19.7	-1	1	3.05	3.8
2021-11-12 12:37	21.9	-1	1	2	2.4
2021-11-12 12:38	37.15	-1	1	2.95	3.5
2021-11-12 12:44	-1	-1	8.1	1.35	1.6

4.2.2. A case of fully bedridden elderly

For the non-fully bedridden elderly, the detection models were built in two dimensions, single and combined variables, and the Prophet and DBSCAN models performed well. Based on this, we selected the fully ambulatory user (Z59435) to investigate the robustness of the models. The detection and comparison results are shown in Table 6.

Table 6. Data mismatch between sensors.

Model	Detection dimension	User Z59479 (non-fully bedridden)		User Z59435 (fully bedridden)	
		Number of outliers	Percentage of outliers	Number of outliers	Percentage of outliers
Prophet	Heart rate	0	0.00‰	222	13.49‰
	Respiratory rate	51	3.10‰	571	34.70‰
	Weight	0	0.00‰	154	9.36‰
	Position x	0	0.00‰	256	15.56‰
	Position y	1	0.06‰	416	25.28‰
DBSCAN	——	234	14.23‰	193	11.73‰

The results of the Prophet model showed a significant increase in abnormal physiological information for fully bedridden users compared to non-fully users. At the same time, the respiratory rate had the highest frequency of abnormalities, which was consistent with the results for non-fully users. Meanwhile, the DBSCAN model detection results showed little difference in the amount of abnormal information between the two types of users, pending subsequent analysis in conjunction with the visualization results.

Based on the model results, we selected the user's respiratory rate data on 8 and 11 November for plotting. As shown in Figure 2, the red markers are the abnormal points determined by the models. The results show that the Prophet model focuses on the very high respiratory rate values at 30 breaths/minute and above and the shallow values at 10 breaths/minute and below, which can effectively detect abnormal values in physiological information. Moreover, the DBSCAN model is more sensitive to extreme respiratory rate values at 30 breaths/min and above and can also detect some moments of low respiratory rate, which has some effect on detecting abnormal physiological information.

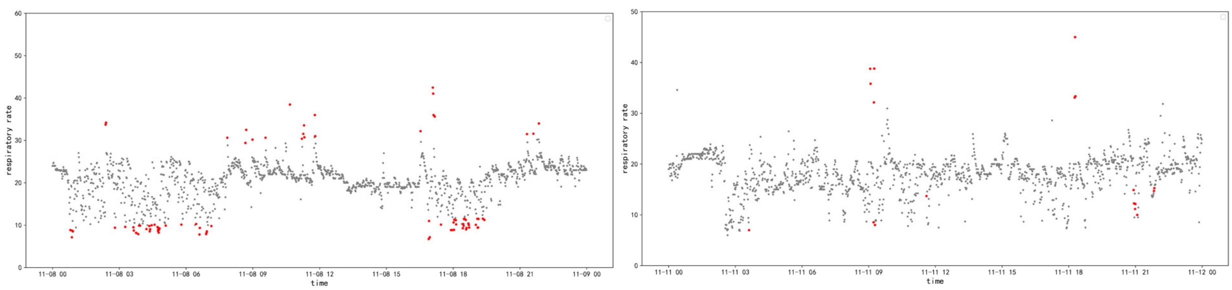

Fig. 2. (a) Prophet model anomaly detection example; (b) DBSCAN model anomaly detection example(Z59435).

4.2.3. A case of visualization improvements

We implemented an interactive visualization of the Prophet model using Altair in Python, as shown in Figure 3. In the corresponding window, red marker dots indicate anomalous moments, and the confidence interval is filled in green. Clicking on the corresponding position shows the specific anomalous value and the corresponding moment, which provides a preliminary exploration for future integration applications in the smart senior care system.

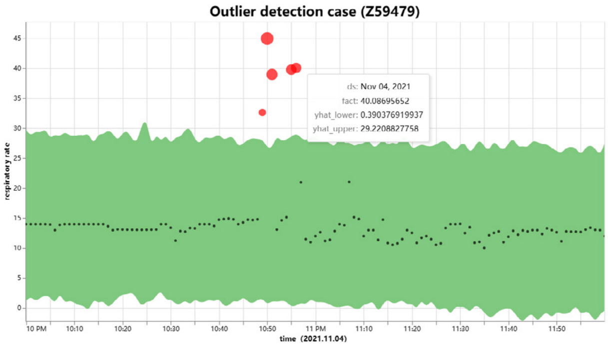

Fig. 3. Interactive visualization of anomaly detection results for Prophet model.

4.2.4. Summary

In summary, the Prophet and DBSCAN models are robust and can effectively detect anomalies in elderly groups with different mobility. Compared to non-fully ambulatory users, fully ambulatory older users have significantly more abnormal physiological information regarding univariate detection. This discrepancy may be due to health factors, with the fully ambulatory elderly caught in some obvious medical conditions, or due to the caregiver's all-day care behavior interfering with the mattress data, and is subject to subsequent in-depth analysis in conjunction with additional information. In addition, given that the DBSCAN model clusters multidimensional variables and traverses data points corresponding to multidimensional variables, there is a possibility that a single dimension of physiological information may not be adequately reflected. Therefore, the visual analysis of the combined variable detection has certain limitations compared to the univariate detection and requires a specific analysis of the risk factors for the corresponding abnormal moments in conjunction with professional nursing experience.

5. Conclusion

Based on the smart mattress data, this paper addresses the behavioral risk issues of the aging-in-place group and forms the following two conclusions.

(1) Risk identification of sleep apnea. The performance of single and fusion models differed for elderly groups with different mobility, with logistic regression, linear discriminant analysis and neural network models performing better in identifying apnea in fully bedridden elderly. The fusion classification models generally achieved better performance than the single classification models. Among them, the Stacking model can accurately detect the presence of apnea emergencies in elderly users and has specific practical application value.

(2) Risk identification of abnormal physiological information. We selected the Prophet and the DBSCAN model to conduct abnormal mining of physiological information for single and combined variables, respectively, which is feasible. We also found that single variable detection can correspond to more precise application scenarios, thus helping caregivers improve their work efficiency. In contrast, integrated variable detection has certain advantages in mining the sensor data matching problem.

In summary, data fusion methods enhance apnea detection results, while abnormality detection from a single or combined variable perspective is effective. The following suggestions are obtained based on the conclusions.

(1) Improve the quality of age-related data. In the process of data collection in this study, there were phenomena such as the lack of data in long segments due to the low motivation of older people to use them, and also encountered

objective problems such as sensor data mismatch and the quality of data involving the elderly needs to be improved. On the one hand, it is essential to actively bridge the digital divide, strengthen training on the use of intelligent devices for the elderly, optimize the age-appropriate design of products, and comprehensively improve the motivation of elderly groups to use innovative elderly products. On the other hand, China has not yet formed a nationwide unified data platform for basic information on the elderly, and there are problems such as data barriers. In the future, we should further develop big data technology based on data fusion and realize data sharing, common and shared through a unified information platform, to continuously improve the efficiency of using information related to the elderly.

(2) Systematize integrated applications. As a complex industry involving healthcare, living services and information software, the smart senior care industry constantly poses new problems and challenges. The existing applications usually only target specific details of the independent aspects of the elderly care sector, and the overall layout is weak. In the future, according to the actual status of the elderly market, we should fully consider the immediate needs of multiple user groups, such as the elderly, their families and healthcare professionals, and systematically design the platform to promote the integrated application and sustainable development of the smart elderly system. For example, the future can be based on distributed environmental sensors, including light sensors, smart mattresses and other sensors, to achieve a daily activity portrayal of the elderly and build a more accurate real-time risk monitoring system.

Finally, this paper has conducted some research based on actual data on behavioral risk in older people, but certain things still could be improved. In terms of data, the existing data does not involve much basic information about the elderly and activity events. With information supplementation, we can increase the scale of user detection and conduct an adequate robustness analysis of the model in the future. Regarding methodology, we can pay attention to the progress of research in anomaly detection and appropriately select algorithms for comparative research. With machine learning algorithms as the main focus, we can continuously update our risk model base to optimize the detection effect.

Acknowledgments

This research is partially supported by the National Natural Science Foundation (71932008)

References

[1] Lotfi A, Langensiepen C, Mahmoud S M, et al. (2012) "Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behavior." *Journal of ambient intelligence and humanized computing*, **3(3)**: 205–218.

[2] Lupiani E, Juarez J M, Palma J, et al. (2017) "Monitoring elderly people at home with temporal case-based reasoning." *Knowledge-Based Systems*, **134**: 116-134.

[3] Li Y, Ho K C, and Popescu M. (2012) "A microphone array system for automatic fall detection." *IEEE Transactions on Biomedical Engineering*, **59(5)**: 1291-1301.

[4] Li Z, Zang C, Yang G, and Xiang R. (2016) "Design of a fall detection system based on semi-supervised learning." *Sensors and Microsystems*, **35(10)**:67-69.

[5] Cao A, Lin C, and Wu S. (2017) "A Detection Method for the Fall Behavior of Elders Based on Hidden Markov Model." *Chinese Journal of Biomedical Engineering*, **36(2)**:165-171.

[6] Vaiyapuri T, Lydia E L, Sikkandar M Y, et al. (2021) "Internet of things and Deep learning enabled elderly fall detection model for smart homecare." *IEEE Access*, **9**: 113879-113888.

[7] Pan L, Chen Q. (2022) "Abnormal Behavior Detection Model Based on Multi-sensor Sequence for Eldercare." *Journal of Chinese Computer Systems*, **43(9)**:1984-1991.

[8] Tang K, Kumar A, Nadeem M, et al. (2021) "CNN-based smart sleep posture recognition system." *IoT*, **2(1)**: 119-139.

[9] Laurino M, Arcarisi L, Carbonaro N, et al. (2020) "A smart bed for non-obtrusive sleep analysis in real world context." *IEEE Access*, **8**: 45664-45673.

[10] Sangeetha T, Kumutha D, Bharathi M D, et al. (2022) "Smart mattress integrated with pressure sensor and IoT functions for sleep apnea detection." *Measurement: Sensors*, **24**: 100450.

[11] Hong Y S. (2018) "Smart care beds for elderly patients with impaired mobility." *Wireless Communications and Mobile Computing*, 2018.

[12] Ray P P. (2014) "Home Health Hub Internet of Things (H 3 IoT): An architectural framework for monitoring health of elderly people." *In Proceedings of IEEE ICSEMR,* Chennai.

[13] Lupiani E, Juarez J M, Palma J, et al. (2017) "Monitoring elderly people at home with temporal case-based reasoning." *Knowledge-Based Systems*, **134**: 116-134.

[14] Taylor S J, Letham B. (2018) "Forecasting at scale." *The American Statistician*, **72(1)**: 37-45.