# Long Short Term Memory Based Model for Abnormal Behavior Prediction in Elderly Persons

Meriem Zerkouk[1] and Belkacem Chikhaoui[2]

[1] Department of Computer Science, USTO-MB University, Oran, Algeria.
zerkouk.meriem@gmail.com

[2] LICEF research center, Department of Science and Technology, TELUQ University,
Montreal, Canada
belkacem.chikhaoui@teluq.ca

**Abstract.** Smart home refers to the independency and comfort that are ensured by remote monitoring and assistive services. Assisting an elderly person requires identifying and accurately predicting his/her normal and abnormal behaviors. Abnormal behaviors observed during the completion of activities of daily living are a good indicator that the person is more likely to have health and behavioral problems that need intervention and assistance. In this paper, we propose a method, based on long short-term memory recurrent neural networks (LSTM), to automatically predicting an elderly person's abnormal behaviors. Our method allows to model the temporal information expressed in the long sequences collected over time. Our study aims to evaluate the performance of LSTM on identifying and predicting elderly persons abnormal behaviors in smart homes. We experimentally demonstrated, through extensive experiments using a dataset, the suitability and performance of the proposed method in predicting abnormal behaviors with high accuracy. We also demonstrated the superiority of the proposed method compared to the existing state-of-the-art methods.

**Keywords:** Smart home, ADL, LSTM, Anomaly detection.

## 1 Introduction

Internet of Things (IoT), which is an emerging domain, promises to create a world where all the objects around us are connected to the Internet and communicate with each other with minimum human intervention. The crucial goal is to create a better world for human beings, where objects around us are context awareness, that can respond to questions: what we want, what we need and where we are. Smart homes represent one of the main application domains of IoT that received particular attention from researchers [1].

life, and 2) to track residents' behaviors and to monitor their health by using sensors embedded in their living spaces [2]. The data collected from smart homes needs to be deeply analyzed and investigated, in order to extract useful information about residents' daily routines, and more specifically (specific) activities of daily living.

Activity recognition [3], as a core feature of smart home, consists of classifying data recorded by the different integrated environmental and/or wearable sensors, into well-defined and known movements. However, dependent persons are usually exposed to different types of problems causing them mostly to perform activities of daily living in a wrong way. Therefore, detecting abnormal behaviors is of great importance for dependent people in order to ensure that activities are performed correctly without errors [4]. This will also ensure their safety and well-being.

Detecting an anomaly in the activities of daily living (ADL) of a person is usually performed by detecting nonconformities from their usual ADL patterns. This has been conducted in various works using classical machine learning algorithms [5,6]. Tele-health care requires systems with high accuracy, less computational time and less user intervention because data are becoming larger and more complex [7]. However, deep learning architectures provide a way to automatically extract useful and meaningful spatial and temporal features of a raw data without the need for data labeling which is time consuming, complex and error prone. This makes deep learning models easily generalizable to different contexts. LSTM is a powerful deep learning model for sequence prediction and anomaly detection in sequential data [8]. LSTM models are able to extract temporal features with long time relationships. This property is of great importance in smart homes in order to understand person's behaviors, they change over time, and particularly any deviations from normal execution of activities of daily living.

In this paper, we propose an LSTM model to identify and predict elderly people's abnormal behaviors. The rationale of using LSTM model in our work is threefold: 1) it is capable of handling multivariate sequential time-series data, 2) it can identify and accurately predict abnormal behavior in time-series data [9, 10], and 3) it can automatically extract features from massive time-series data, which makes it possible to be easily generalizable to other types of data. Therefore, the contributions of our paper can be summarized as follows:

1. Proposing an LSTM model for automatic prediction of abnormal behaviors in smart homes.
2. Managing the problem of imbalanced data by oversampling minority classes.
3. Conducting extensive experiments to validate the proposed LSTM model.

The paper is organized as follows: Section 2 presents an overview of anomaly detection models and related work in machine learning algorithms. Section 3 presents material and methods to conduct our work and the obtained results. Section 4 gives a comparison of different machine learning algorithms. Finally, Section 5 discusses the outcomes of the experiments and perspectives for future work.

## 2 Related work

Tracking user behavior for abnormality detection has gain a large attention and becomes one of the main goals for certain researchers [11]. Abnormal behavior detection approaches are based mainly on machine learning algorithms, and more specifically supervised learning techniques [12]. Supervised classification techniques need labelled data points (samples) for the models to learn. This kind of classification requires to train a classifier on the labelled data points and then evaluate the model on new data points. Therefore, in case of normal and abnormal classes, the model learns the characteristics of these data points and classify them as normal or abnormal. Any data point that does not conform to this normal class, will be classified as an anomaly by the model. Various classification techniques have been applied for abnormal behavior detection.

Pirzada and al [13] explored KNN as a classifier which works well to classify data in categories. They performed a binary classification where they classify activity as good or bad to distinguish the anomaly in the user behavior. Eventually, the proposed KNN applied to predicts whether the class belongs to regular (good) or irregular (bad) class. The performed work monitor health conditions of elderly person living alone using sensors in unobtrusive manner.

Aran et al. [14] proposed a method to automatically observe and model the daily behavior of the elderly and detect anomalies that could occur in the sensor data. In their proposed method, the anomaly relies on signal health related problems. For this purpose, they have created a probabilistic spatio-temporal model to summarize daily behavior. They define anomalies as significant changes from the learned behavioral model and detected, the performance is evaluated by cross-entropy measure. Once the anomaly is detected, the caregivers are informed accordingly.

Ordonez et al [15] presented an anomaly detection method based on Bayesian statistics that identify anomalous human behavioral patterns. Their proposed method assists automatically the elderly person's with disabilities who live alone, by learning and predicting standard behaviors to improve the efficiency of their healthcare system. The Bayesian statistics are choosen to analyze the collected data, the estimation of the static behavior is based on three probabilistic features that introduce, namely sensor activation likelihood, sensor sequence likelihood and sensor event duration likelihood.

Yahaya et al. [16] proposed novelty detection algorithm known as One-Class Support Vector Machine (SVM) which is applied for detection of anomaly in activities of daily living. The anomaly is situated in sleeping patterns, which could be a sign of Mild Cognitive Impairment (MCI) in older adults or other health-related issues.

Palaniappan et al. [17] interested in detecting abnormal activities of the individuals by ruling out all possible normal activities. Authors define abnormal activities as unexpected events that occur in random manner. Multi-class SVM method is used as classifier to identify the activities in form of a state transition table. The transition table helps the classifier in avoiding the states which are unreachable from the current state.

Hung et al. [18] proposed a novel approach that mix SVM and HMM to a homecare sensory system. RFID sensor networks are used to collect elder's daily

activities, Hidden Markov Model (HMM) used to learn the data, and SVMs used to estimate whether the elder's behavior is abnormal or not.

Bouchachia et al. [19] proposed an RNN model to deal the problem of activity recognition and abnormal behavior detection for elderly people with dementia. The proposed method suffered from the lack of data in the context of dementia.

The aforementioned methods suffer from one or more of the following limitations:

1. The presented methods focus on the spatial and the temporal anomalies in user assistance. However, we noted that the abnormal behavior does not treated in the case of the smart home
2. These methods require feature engineering, which is difficult specifically when data becomes larger.
3. The abnormality identification and prediction lack of good accuracy.

These points motivate us to propose a method, which tries to overcome these limitations and to be useful in smart homes for assistance.

## 3 Proposed method

In this section, we present a description of our problem related to the identification and prediction of elderly person's abnormal behaviors.

### 3.1 Problem description

Abnormality detection is an important task in health care monitoring, particularly for monitoring elderly in smart homes. Abnormality consists in finding unexpected activities, variations in normal patterns of activities, finding the patterns in data that do not conform to the expected behavior [20] because humans usually perform their ADLs in a sequential manner. According to [21], the abnormality can be categorized into temporal, spatial, and behavioral abnormality. Our work focuses on the behavioral anomaly because this kind of abnormality depends on the same on time (when performing the activity) and location (where performing the activity). Each activity is defined by a sequence of sub-activities and if the person violates the expected sequence then it is an abnormality.

### 3.2 LSTM for anomaly description

LSTMs [22] are a recurrent neural network architecture, the principal characteristic is the memory extension that can be seen as a gated cell, where gated means that the cell decides whether or not to store or delete information, based on the importance it assigns to the information. The assignment of importance happens through weights, which are also learned by the algorithm. This simply means that it learns over time, which information is important.

LSTM architecture consists of three layers: Input layer, hidden layer and output layer. The hidden layers are fully connected to the input and output layers. A layer in

LSTM is composed from blocs, and each block has three gates: input, output and forget gates. Each gate is connected to each other. These gates decide whether or not to let new input in (input gate), delete the information because it isn't important (forget gate) or to let it impact the output at the current time step (output gate).

As we mentioned previously, our motivations in using LSTM rely on the fact that it enables to remember their inputs over a long period of time which allows to remember the data sequences. Abnormality detection aims to identify a small group of samples, which deviate remarkably from the existing data. That why, we choose LSTM to identify and accurately predict the abnormality behavior from a long sequential data given that persons perform their ADL in a sequential manner, less human intervention in the identification and prediction process.

The LSTM input layer development requires reshaping the data. It needs the input data to be 3-dimensions as training sample, time step, and features. We add for this layer an activation function (ReLu). To avoid overfitting problem in LSTM architectures, we used the dropout method [26] and improve model performance. In our proposed model, the dropout is applied between the two hidden layers and between the last hidden layer and the output layer. We setup the dropout at 20% as recommended in literature [26].

The last layer (dense layer) defines the number of outputs which represents the different activities and anomaly (classes). The output is considered as vector of integer which is converted into binary matrix. The anomaly prediction is formulated as multi classification problem which requires to create 7 output values, one for each class, Softmax as activation function and categorical_crossentropy is used as the loss function. The figure 1 indicate the LSTM architecture development.

**Figure 1:** LSTM development.
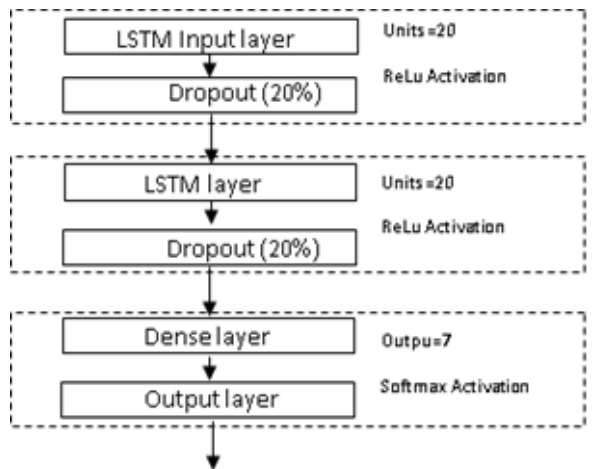
# 4 Experiment study

In this section, we present our dataset that we want to analyze, overcome to the problem of imbalanced data by oversampling our data with SMOTE. After that, we identify and predict the abnormal behavior based on LSTM model.

## 4.1 Dataset

This research uses SIMADL [23] dataset generated by OpenSHS [24] which is an open source simulation tool that offered the flexibility needed to generate the inhabitant's data for classification of ADLs. OpenSHS was used to generate several synthetic datasets that includes 29 columns of binary data representing the sensor values, each binary sensor has two states, on (1) and off (0). The sensors can be divided into two groups, passive and active. The passive sensors react without explicitly the participants interact with them. Instead, they react to the participant movements and positions.

The sampling was done every second. Seven participants were asked to perform their simulations using OpenSHS. Each participant generated six datasets resulting in forty-two datasets in total. The participants self-labelled their activities during the simulation. The labels used by the participants were: Personal, Sleep, Eat, Leisure, Work, Other and Anomaly. The simulated anomalies are behavioral and are described in table 1. Note that each user has his/her own behavioral abnormality to simulate.

**Table 1:** Anomalies description

| Participants | Description |
|---|---|
| Participant 1 | Leaving the fridge door open. |
| Participant 2 | Leaving the oven on for long time. |
| Participant 3 | Leaving the main door open. |
| Participant 4 | Leaving the fridge door open. |
| Participant 5 | Leaving the bathroom light on. |
| Participant 6 | Leaving tv on. |
| Participant 7 | Leaving light bedroom and wardrobe open. |

## 4.2 Imbalanced data

The distribution of the classes (that represent the different ADL) is not uniform, leads to imbalanced classes. This situation appears because of rare abnormal behavior which is clear in the figure 2. As shown in figure 2, the class anomaly represent a minority. We tackle this problem in order to improve our classification performance. Dealing with imbalanced datasets requires strategies such as oversampling techniques before providing the data as input to the LSTM model. Oversampling strategy consists in augmenting the minority class samples to reach a balanced level with the majority class.

### 4.2.1 Oversampling

We deal with abnormality, called also anomaly, detection problem as a supervised learning that refers to correctly classifying rare class samples as compared to majority samples.

Therefore, anomalies are a minority in the whole behavior, which create an imbalanced data problem. Therefore, we have to oversample our data and after that, we can classify correctly.

A subset of data is taken from the minority samples as an example and then new synthetic similar data points are created. These synthetic data points are then added to the original dataset. The new dataset is used to train the classification models. The main objective of balancing classes is to either increasing the samples of the minority class or decreasing the samples of the majority class. In oversampling, we increase the minority class samples. This is done in order to obtain approximately the same number of instances for both the classes as demonstrated in figure 2. Our motivation in the use of this strategy is to avoid overfitting. We use SMOTE statistical method [25] to oversample our classes as indicated in figure 2. We note that the x-axes indicate the number of classes and y-axes indicate the number of input data.

**Figure 2:** Imbalanced classes Vs Balanced classes
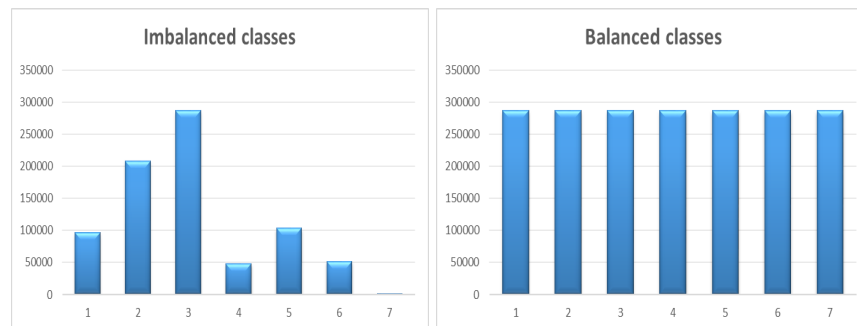
### 4.3 Network architecture and hyper-parameters tuning

The crucial task is to find a suitable network structure for training the data, specifically to choose the right amount of nodes and layers. Many experiments were run by varying LSTM networks architecture as shown in table 2 to find the suitable units number. We varied this number from 20, 30, 50, 60, 100 to 200. The number of layers was experimentally fixed to four layers.

To compile and fit the model, we experimentally fixed mini-batch size to 128 samples, with ADAM [27] as optimizer, which is an algorithm that can used instead of the classical stochastic gradient descent procedure to update the network weights iterative based on training data.

Our experiment with LSTM were implemented in Python language using Keras library [28] with Tensorflow [29]. The figure 1 shows the development of our LSTM network and the adjusted parameters to obtain the appropriate results.

### 4.4 Performance metrics analysis

High performance system should have less false positive and false negative rates. The performance of our proposed method is evaluated in terms of precision, recall and f-score [30]. Table 2 presents the obtained results.

**Table 2:** LSTM experiment.

| Units | Precision | Recall | F-score |
|-------|-----------|--------|---------|
| 20 | 0.91 | 0.91 | 0.91 |
| 30 | 0.87 | 0.85 | 0.86 |
| 50 | 0.90 | 0.90 | 0.90 |
| 60 | 0.86 | 0.86 | 0.86 |
| 100 | 0.90 | 0.90 | 0.90 |
| 200 | 0.90 | 0.90 | 0.90 |

Table 2 refers to the results obtained on SIMADL dataset for abnormal behavior detection and shows that LSTM with 20 units gives the best precision of 0.91%, recall of 0.91 % and f-score of 0.91%. The rest of the results is expressed in the table 2 for which we have varied the unit number to 30, 50, 60, 100 and 200.

To demonstrate the superiority of the proposed method, we conducted comparison with existing state-of-the-art methods. The results are summarized in table 3.

**Table 3:** LSTM comparison with the state of the art.

| Models | Precision | Recall | F-score |
|--------|-----------|--------|---------|
| LSTM | 0.91 | 0.91 | 0.91 |
| SVM | 0.90 | 0.90 | 0.90 |
| NB | 0.79 | 0.79 | 0.79 |
| KNN | 0.89 | 0.89 | 0.89 |
| NN | 0.77 | 0.77 | 0.77 |

According to table 3, we note that LSTM gives a good result by comparing to machine learning methods such as: SVM, NB, KNN and NN.

## 5    Conclusion

We proposed an LSTM based abnormal behavior prediction method. Our method identifies and predicts abnormal behaviors with a high degree of accuracy and with less user intervention in order to automate the identification and prediction process. We note that before to classify the activities, we have checked the distribution of classes, we detected an imbalanced classes, to deal with this problem we have applied SMOTE  for oversampling classes.

The future work can focus on using real dataset from environmental and physiological sensors to understand the health condition of the elderly person's for better wellbeing.

## References

1. Perera, C., Zaslavsky, A., Christen, P. & Georgakopoulos, D. Context aware computing for the internet of things: A survey. *IEEE Commun. Surv. Tutorials* 16, 414–454 (2014).
2. Lê, Q., Nguyen, H. B. & Barnett, T. Smart Homes for Older People: Positive Aging in a Digital World. 607–617 (2012). doi:10.3390/fi4020607
3. Mehr, H. D., Polat, H. & Cetin, A. Resident Activity Recognition in Smart Homes by Using Artificial Neural Networks. (2016).
4. Sanchez-cortes, D. & Do, M. Anomaly Detection in Elderly Daily Behavior. 1, 51–67 (2016).
5. Omar, S., Ngadi, A. & H. Jebur, H. Machine Learning Techniques for Anomaly Detection: An Overview. *Int. J. Comput. Appl.* 79, 33–41 (2013).
6. Grover, A. Anomaly Detection for Application Log Data. (2018).
7. Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L. & Muller, P. A. Deep learning for time series classification: a review. *Data Min. Knowl. Discov.* (2019).
8. Wang, J., Chen, Y., Hao, S., Peng, X. & Hu, L. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognit. Lett.* 119, 3–11 (2019).
9. Chandola, V. Anomaly Detection : A Survey. 1–72 (2009).

10. Paudel, R., Eberle, W. & Holder, L. B. Anomaly Detection of Elderly Patient Activities in Smart Homes using a Graph-Based Approach. 163–169 (2016).

11. Langensiepen, C. & Lotfi, A. Advances in Computational Intelligence Systems. 840, 362–371 (2019).

12. Mukhopadhyay, S. C. Next Generation Sensors and Systems. *Next Gener. Sensors Syst.* 16, 1–330 (2015).

13. Pirzada, P., White, N. & Wilde, A. Sensors in smart homes for independent living of the elderly. *5th Int. Multi-Topic ICT Conf. Technol. Futur. Gener. IMTIC 2018 - Proc.* (2018).

14. Sanchez-cortes, D. & Do, M. Anomaly Detection in Elderly Daily Behavior. **1**, 51–67 (2016).

15. Ordóñez, F. J., de Toledo, P. & Sanchis, A. Sensor-based Bayesian detection of anomalous living patterns in a home setting. *Pers. Ubiquitous Comput.* 19, 259–270 (2015).

16. B, S. W. Y., Langensiepen, C. & Lotfi, A. Advances in Computational Intelligence Systems. in 840, 362–371 (Springer International Publishing, 2019).

17. Palaniappan, A., Bhargavi, R. & Vaidehi, V. Abnormal human activity recognition using SVM based approach. *Int. Conf. Recent Trends Inf. Technol. ICRTIT 2012* 97–102 (2012).

18. Hung, Y., Chiang, C., Hsu, S. J. & Chan, C. Abnormality Detection for Improving Elder ' s Daily. *Science (80-. ).* 186–194 (2010).

19. Arifoglu, D. & Bouchachia, A. Activity Recognition and Abnormal Behaviour Detection with Recurrent Neural Networks. *Procedia Comput. Sci.* 110, 86–93 (2017).

20. Novák, M., Jakab, F. & Lain, L. anomalyDetection-2013. **3**, (2013).

21. Zhu, C., Sheng, W., Member, S., Liu, M. & Member, S. Wearable Sensor-Based Behavioral Anomaly Detection in Smart Assisted Living Systems. 1–10 (2015).

22. Sepp, H. & J?urgen, S. 1997Hochreiter_LSTM. *Neural Comput.* 9, 1–32 (1997).

23. Alshammari, T., Alshammari, N., Sedky, M. & Howard, C. SIMADL: Simulated Activities of Daily Living Dataset. *Data* 3, 11 (2018).

24. OpenSHS by openshs: https://github.com/openshs/

25. Chawla, N.V., Bowyer, K.W., Hall, L.O., K. W. P. SMOTE: Synthetic Minority Over-Sampling Technique. Journal of Artificial Intelligence Research. *J. Artif. Intell. Res.* 16, 321–357 (2002).

26. Hinton, G. Dropout : A Simple Way to Prevent Neural Networks from Overfitting. 15, 1929–1958 (2014).

27. Kingma, D. P. & Ba, J. L. A : a m s o. 1–15 (2015)

28. FAQ - Keras Documentation: https://keras.io/

29. TensorFlow: https://www.tensorflow.org/

30. Technology, I. & Technology, I. a R Eview on E Valuation M Etrics F or D Ata C Lassification E Valuations. *Int. J. Data Min. Knowl. Manag. Process* 5, 1–11 (2015).