Analysis of the Consumption of Household Appliances for the Detection of Anomalies in the Behaviour of Older People

Miguel A. Patricio[1(✉)], Daniel González[2], José M. Molina[1], and Antonio Berlanga[1]

[1] Applied Artificial Intelligence Group, Universidad Carlos III de Madrid, Madrid, Spain
mpatrici@inf.uc3m.es
{molina,aberlang}@ia.uc3m.es
[2] BQ Engineering, Madrid, Spain
daniel.gonzalez@bq.com,
http://www.giaa.inf.uc3m.es,
http://www.bq.com

Abstract. Nowadays, modern societies are facing the important problem of ageing of their population. On many occasions, older people must leave their homes to be cared for by their relatives or to enter specialised centres for the elderly. On the other hand, something similar happens with disabled people who need the support of other people for their daily activity. This phenomenon brings with it important social and economic consequences. In the activities of the daily life of the elderly it is necessary to have the monitoring of different aspects of their physical activity, such as the detection of critical situations (such as falls) or dangerous situations (such as flooding or gas leaks). The aim of this paper is to analyse the consumption of the different household appliances in order to model a normal behaviour within the daily activities of a house. By means of the consumption of the electrical appliances the aim is to detect anomalous behaviours that induce the appearance of possible problems due to the change in the consumption pattern.

Keywords: Anomaly detection · Support for daily activities · Elderly · Behaviour analysis

1 Introduction

The world population is ageing rapidly. Projections are that people 60 years old or older will outnumber children by 2030 and adolescents and youth by 2050 [10]. Therefore, concepts such as "independent living", "active ageing", "ageing at home" form the nucleus of proposals for integrated care services for the elderly.

J. M. Ferrández Vicente et al. (Eds.): IWINAC 2019, LNCS 11486, pp. 60–68, 2019.
https://doi.org/10.1007/978-3-030-19591-5_7

In recent years, research is being conducted to monitor the electricity consumption of homes. In this sense, there are works related to non-intrusive appliance load monitoring (NIALM) [2]. NIALM is the process of dis-aggregating a household's total electricity consumption into its contributing appliances.

Neural Networks have been used for the detection of anomalies in Electric Power Systems with auto-associative neural networks [8] or autoencoders [11]. Autoencoders are a kind of neural network that allows to represent a compressed version of the input. The aim of the paper is to analyse the typical or normal behaviour modelling from the consumption of different appliances, so that they can be used in the prediction of anomalous behaviours within the daily activities of the elderly. The best appliance will be evaluated to be used in a detection system of anomalous behaviours within the daily activity of older people.

The paper is organised with a first part where the problem to be solved is presented based on the information of the UK-DALE dataset. Next, we describe the complete process carried out for the modeling of behaviors based on the analysis of the electrical consumption of household appliances. In the last part of the paper, the results obtained and a discussion of the results are described.

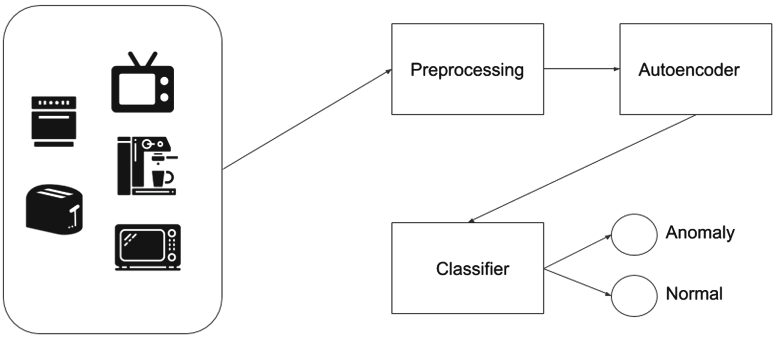
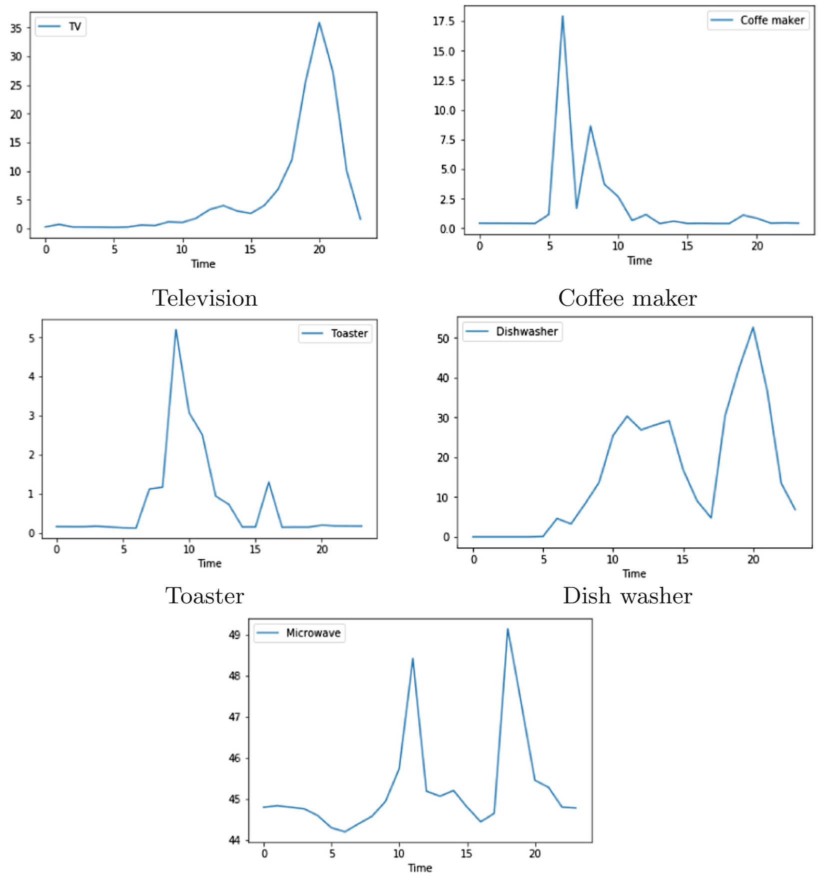


Fig. 1. Phases in the process of detecting anomalous behaviours from the electrical consumption of household appliances.

2 Problem Statement

The problem that arises in this work is divided into different phases (see Fig. 1). Starting from the real electrical consumption of different domestic appliances, a pre-processing of the signals will be carried out. This pre-processing is related to the elimination of outliers (days considered as vacations, for example), as well as the organisation of the information in a set of variables that indicate the summary of the daily behaviour of a house in terms of the electrical consumption of household appliances. The pre-processed signals will be used to model a normal behaviour in the electrical consumption of the different appliances. For each of the appliances, a typical or normal behaviour model will be obtained. In order to validate all the models associated with each appliance, a set of anomalous values will be obtained by shifting the original consumption and varying the amplitude values of the signals in a random way. Each of the models will be evaluated through the use of a classifier. This classifier will be designed using the value of the internal neurons of the autoencoder when at the entrance of the network both typical and abnormal behavioural samples are presented (Fig. 2).



Television Coffee maker

Toaster Dish washer

Microwave

Fig. 2. Histogram of mean power demand on each appliance measure in watts.

2.1 Dataset

The dataset used in this work is the UK-DALE dataset [7], where the power demand of five houses has been collected every 6 s for 4 years. The electrical consumption is measured in watts. Only one house has been considered since the habits of use of each appliance can vary significantly between houses.

The appliance used in this paper are the television, the coffee maker, the toaster, the microwave and the dishwasher since this are the most common appliances in all homes and their consumption habits are relevant for the detection of anomalies.

2.2 Data Preprocessing

This section describes the actions carried out for the preprocessing of each one of the electrical consumption distributions of the different domestic appliances. It has been decided to group the power demand in hours, since the average use of the appliances usually lasts several minutes or even hours. Below it is shown the average electricity demand per hour in the selected home for the five selected appliances. It can be seen that the appliance usage pattern of each appliance varies significantly.

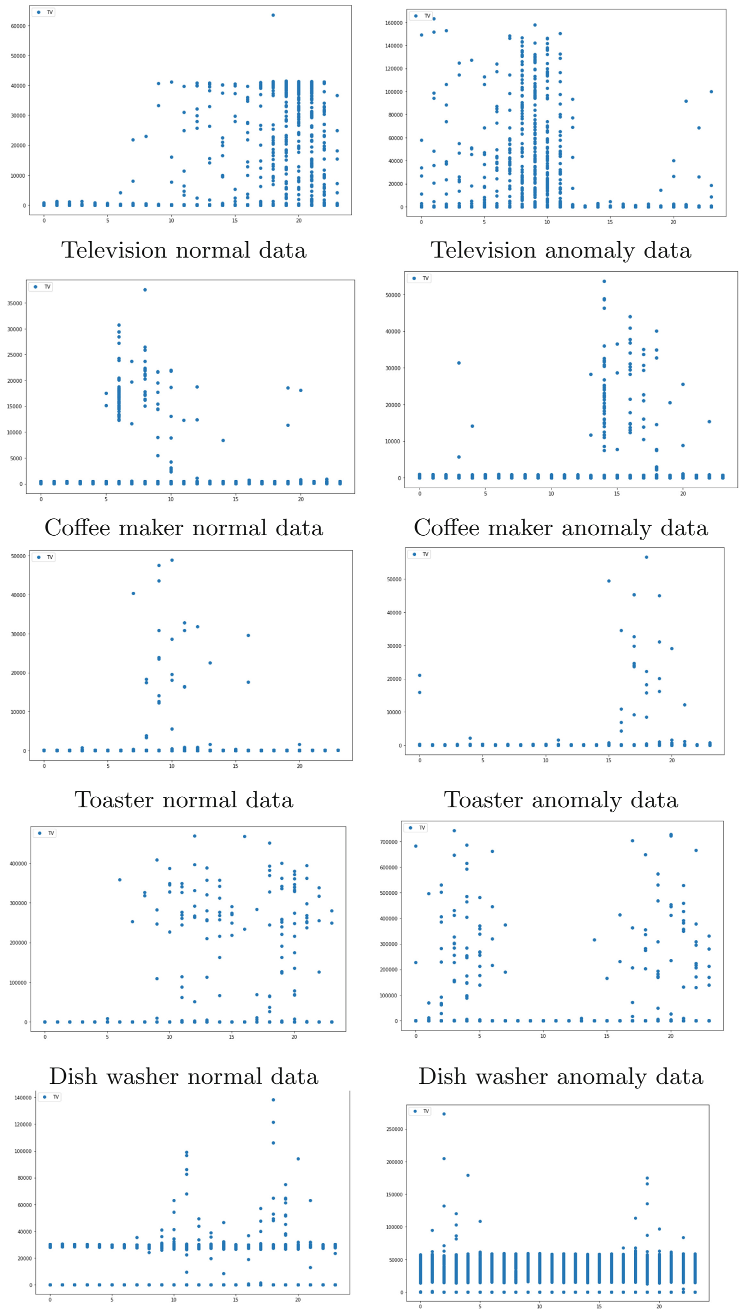
To work effectively with the data, a Data Engineering work has been done. The proposed model has 24 variables, each variable reflects the total watt consumption of the appliance at each hour of a day. The average power consumption at each hour is not used because some relevant information can be lost with this metric. After a work of Data Science, there has been detected some outliers. These outliers are the days in which no use of the appliance has been registered and it might be vacation days or a failure in the sensor or in the appliance, so it has been decided to eliminate this days from the dataset because this entries may affect the performance of the model.

Once the data processing has been done, it is necessary to generate the anomalies. As stated in [5], a nocturnal activity may be associated with Dementia or Alzheimer, so the data has been displaced 8 hours forward. In addition, these values have been multiplied by a random value between $[0.25, 2]$, since a variation in the use of appliance can be associated with anomalous behaviour. The data has been separated by 80% for training and 20% for validation. Below at Fig. 3 it is shown a comparison between the distribution of typical data and anomalies for each appliance.

2.3 Autoencoders for Anomaly Detection

An autoencoder is a neural network where the number of input nodes is equal to the number of output nodes. The autoencoders has an intermediate layer with a lower number of neurons. The autoencoders are part of unsupervised learning, since they do not require labels for their training. This architecture tries to find an arbitrary function $f(W, b)x = x$ that makes the input equal to the output, being W and b the weights and biases of the neural network [1].

The autoencoder of this paper has 24 input nodes as shown in Fig. 4, where each input node is one hour of the day. In the intermediate layer it has two nodes. The popularity of this architecture has increased since at the intermediate layer, the data is represent in a smaller dimension, so the autoencoders are able to simplify the representation of the data and it makes the classification task



Television normal data

Television anomaly data

Coffee maker normal data

Coffee maker anomaly data

Toaster normal data

Toaster anomaly data

Dish washer normal data

Dish washer anomaly data

Microwave normal data

Microwave anomaly data

Fig. 3. Normal and anomaly data for each appliance.

easier. At the following figures, it is shown how the autoencoder represent the data of each appliances. As shown in the graphs, the autoencoder represents the data in a simplified way, the classification of said data is now easier (Fig. 5).

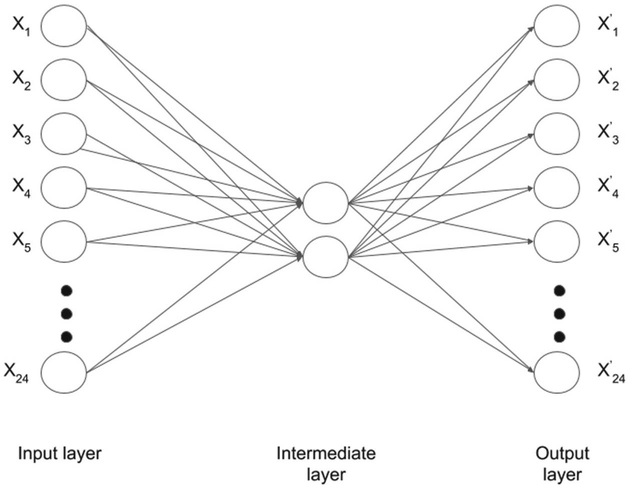
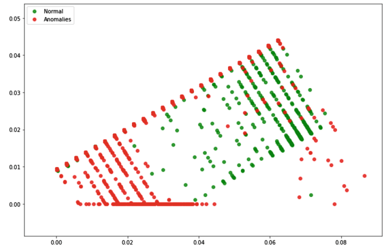


Fig. 4. Architecture of the autoencoder used.

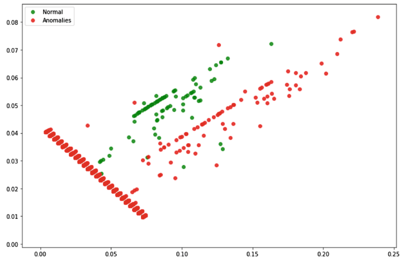
2.4 Anomalies Classification

For the classification task, there has been selected a Random Forest Classifier. The Random Forest is a Bagging method [3], where some Decision Trees are put in parallel. This classifier is part of Ensemble Learning. To find the better parameters for the model, the Grid Search [9] has been used, where the data is split and this framework evaluate the performance of some possible values defined for the parameters. The metrics used for evaluate the performance of the model are:

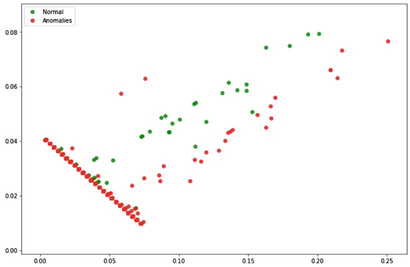
- AUC (ROC): Measures the capacity of the model to differentiate the classes [4].
- Precision: Success on all the data.
- Recall: True positives on all the positive data.
- F1 Score: Average of the precision and recall [6].



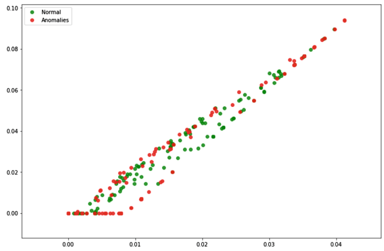
Television



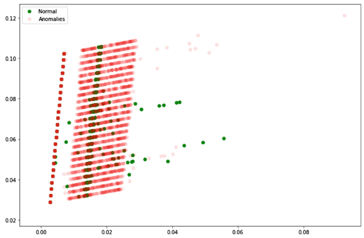
Coffee maker



Toaster



Dish washer



Microwave

Fig. 5. Representation of the data of each appliance at the intermediate layer.

3 Experimentation

With all the treatment of the dataset and with the autoencoder trained, the results of the described model are shown. In the following table, the results of each appliance with the optimal parameters found with the Grid Search are showed (Table 1).

The appliance that has given better results has been the microwave with an 88% of accuracy, being the one that has obtained the best result in the other metrics too. The confusion matrix this appliance is showed below (Fig. 6).

With the data from the microwave, not only the best precision is achieved, but also the lowest percentage of false negatives is achieved. It is very important that the number of false positives be low since this type of failure occurs when the model is not able to detect an anomaly. The false positives in this model are not so important, since they do not produce dangerous situations for the user.

Table 1. Results of each appliance on autoencoder.

	AUC (ROC)	Precision	Recall	F1 Score
Television	0.73	0.71	0.73	0.72
Coffee maker	0.73	0.73	0.71	0.72
Toaster	0.70	0.72	0.65	0.69
Dish washer	0.62	0.65	0.60	0.63
Microwave	**0.86**	**0.88**	**0.84**	**0.86**

	p	n
p′	256	90
n′	39	249

Television

	p	n
p′	118	43
n′	48	124

Coffee maker

	p	n
p′	107	41
n′	57	130

Toaster

	p	n
p′	178	178
n′	150	153

Dish washer

	p	n
p′	291	39
n′	54	274

Microwave

Fig. 6. Confusion matrix of the model with each appliance.

4 Conclusions

The autoencoders are a neural network that simplify the classification work. Finding an autoencoder that suits in a data allows to build a strong classifier that achieves a good performance. The model use combine an Autoencoder to represent the data in a simpler way a Random Forest Classifier, whose parameters has been tuned by Grid Search. The proposed model is capable of detect anomalies in the behaviour of the user with only the power demand of the TV. The model only requires to be trained with previous data of the power demand of this appliance. The appliance that appears to be more relevant in anomalies detection is the toaster and this may be due to the fact that the daily use of the toaster follows a very stable pattern, and the use of other appliances is more random.

This method is a non-intrusive way of caring for people with Dementia or Alzheimer, as it will be able to detect in a precise way situations where the user requires attention and avoid possible dangerous situations derived from these diseases. For the future work, a better performance of the model can be achieved with other anomalies data. It would be interesting to test the model with real anomaly data or use a more precise method to generate anomalies. On the other hand, a combination of some appliances might achieve a better performance and it should be tested.

Acknowledgements. This work was funded by private research project of Company BQ and public research projects of Spanish Ministry of Economy and Competitivity (MINECO), references TEC2017-88048-C2-2-R, RTC-2016-5595-2, RTC-2016-5191-8 and RTC-2016-5059-8.

References

1. Baldi, P.: Autoencoders, unsupervised learning, and deep architectures. J. Mach. Learn. Res. (2012). https://doi.org/10.1561/2200000006
2. Belley, C., Gaboury, S., Bouchard, B., Bouzouane, A.: An efficient and inexpensive method for activity recognition within a smart home based on load signatures of appliances. Pervasive Mob. Comput. (2014). https://doi.org/10.1016/j.pmcj.2013.02.002
3. Breiman, L.: Random forests LEO. Mach. Learn. (2001). https://doi.org/10.1023/A:1010933404324
4. Davis, J., Goadrich, M.: The relationship between Precision-Recall and ROC curves. In: Proceedings of the 23rd International Conference on Machine Learning - ICML 2006 (2006). https://doi.org/10.1145/1143844.1143874
5. Evans, L.K.: Sundown syndrome in institutionalized elderly. J. Am. Geriatr. Soc. (1987). https://doi.org/10.1111/j.1532-5415.1987.tb01337.x
6. Goutte, C., Gaussier, E.: A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In: Losada, D.E., Fernández-Luna, J.M. (eds.) ECIR 2005. LNCS, vol. 3408, pp. 345–359. Springer, Heidelberg (2005). https://doi.org/10.1007/978-3-540-31865-1_25
7. Kelly, J., Knottenbelt, W.: The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. Sci. Data (2015). https://doi.org/10.1038/sdata.2015.7
8. Martinelli, M., Tronci, E., Dipoppa, G., Balducelli, C.: Electric power system anomaly detection using neural networks. In: Negoita, M.G., Howlett, R.J., Jain, L.C. (eds.) KES 2004. LNCS (LNAI), vol. 3213, pp. 1242–1248. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30132-5_168
9. Pedregosa, F., et al.: Scikitlearn: machine learning in Python. J. Mach. Learn. Res. (2011). https://doi.org/10.1007/s13398-014-0173-7.2
10. United Nations: World Population Ageing 2015. Technical report (2015). ST/ESA/SER.A/390
11. Yuan, Y., Jia, K.: A distributed anomaly detection method of operation energy consumption using smart meter data. In: 2015 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), pp. 310–313 (2015)