Detecting Anomalies in Electricity Consumption of Buildings

Davide Azzalini, Enrico Ragaini, Francesco Amigoni
Politecnico di Milano
{davide.azzalini, enrico.ragaini, francesco.amigoni}@polimi.it

Abstract

Anomaly detection systems have great potential for preventing energy waste in office and residential buildings. Avoiding such waste translates not only in savings and higher levels of comfort for inhabitants, but also plays a big role in the protection of the environment. In this report, we present some preliminary results of the application of several AI-based methods to the detection of anomalies in data on electricity consumption collected in two office buildings.

1 Introduction

Massive amounts of operational data relative to buildings are being collected and available for data analysis. It is therefore very promising to develop data-driven approaches to obtain insights and actionable knowledge to better manage buildings. Forecasting future consumptions and detecting anomalies on energy usage are two of the most common examples of analyses that can be performed [Bolchini *et al.*, 2017]. In this report we focus on *anomaly detection* in electrical energy consumption of buildings.

Several challenges must be faced when developing data-driven models of the behavior of buildings from an electrical energy consumption standpoint. For instance, energy consumption depends on many factors whose dynamics are often interrelated in non-trivial ways. An example is the temperature regulation through internal heating, ventilation, and air conditioning systems (HVACs), to maintain high comfort levels for those living or working inside the building.

By exploiting anomaly detection systems for energy consumption, building managers could easily be notified about internal problems and promptly act to solve them. The potential impact of adopting such solutions is not just in the decrease of the costs and the increase of inhabitants' comfort, but it is also on the environment. The prevention of unnecessary energy waste in buildings could play a big role in the urgently needed reduction of world-wide power consumption to alleviate the burden on both power grids and emission of greenhouse gasses[1]. In this report, we present some preliminary results on the detection of anomalies in electrical energy consumption we obtained in applying AI methods to data coming from two office buildings.

[1]https://ec.europa.eu/energy/eu-buildings-factsheets_en

2 Related Work

Anomaly detection [Chandola *et al.*, 2009] is the task of finding patterns in data that do not conform to the expected behavior. Anomaly detection has been a field of interest both within academia and industry for a very long time, suffice to say that the first works in this direction date back to as early as the late XIX Century [Edgeworth, 1887]. The importance of anomaly detection is rooted in the fact that detected anomalies can lead to significant and often critical insights that can be exploited in a wide variety of application areas. The application domains to which anomaly detection has been applied over the years are very varied and include credit card or insurance fraud [Raj e Portia, 2011], healthcare [Esteva *et al.*, 2017], cyber-physical systems [Goh *et al.*, 2017], surveillance [Nawaratne *et al.*, 2019], intrusion detection in computer networks [Mukherjee *et al.*, 1994] and many others [Chandola *et al.*, 2009; Ruff *et al.*, 2021; Pang *et al.*, 2021; Chalapathy e Chawla, 2019; Blázquez-García *et al.*, 2021].

Time series anomaly detection is the highly related to the problem considered here. A *time series* consists in a sequence of observations that have been recorded in an orderly fashion and that are correlated in time [Hamilton, 1994]. Approaches to detect anomalies in time series can be broadly categorized according to type of time series they are able to deal with [Blázquez-García *et al.*, 2021]. *Univariate* time series are ordered sequences of real-valued observations, while *multivariate* time series can be thought as ordered sets of k-dimensional vectors, where k is the number of observations available at each timestamp. Multivariate approaches represent a more powerful tool as they can model also the fact that each variable could depend not only on its past values but also on the other variables (both at the current time as well as in the past).

Examples of recent anomaly detectors for electricity consumption data are [Fan *et al.*, 2018; Pereira e Silveira, 2018; Chou e Telaga, 2014; Araya *et al.*, 2017]

3 Method

Recently, deep learning models have been employed to re-address several data-related tasks, including those relative to

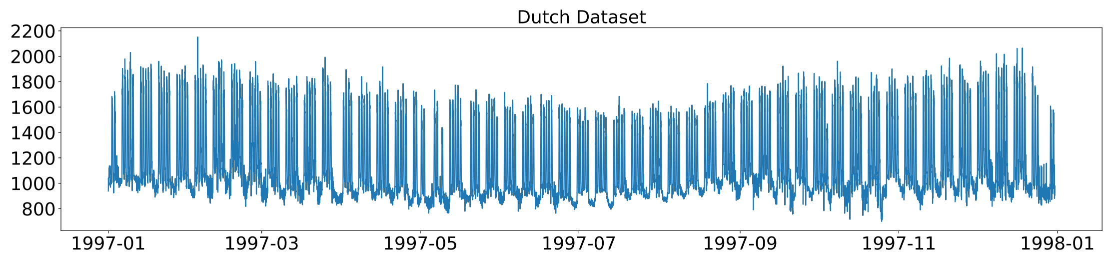

Figura 1: Dutch dataset.

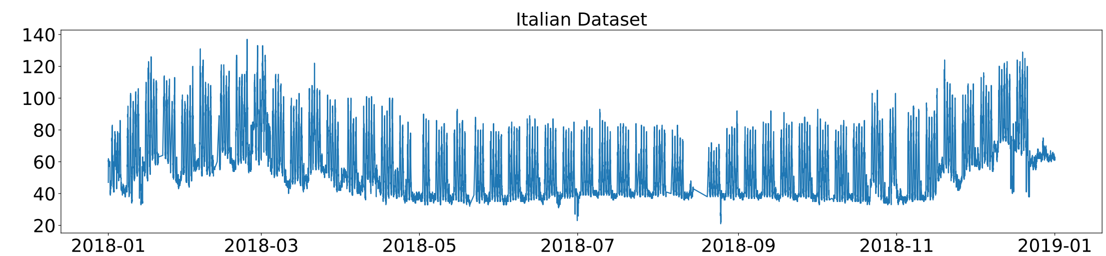

Figura 2: Italian dataset (just lighting consumption for the year 2018 is shown).

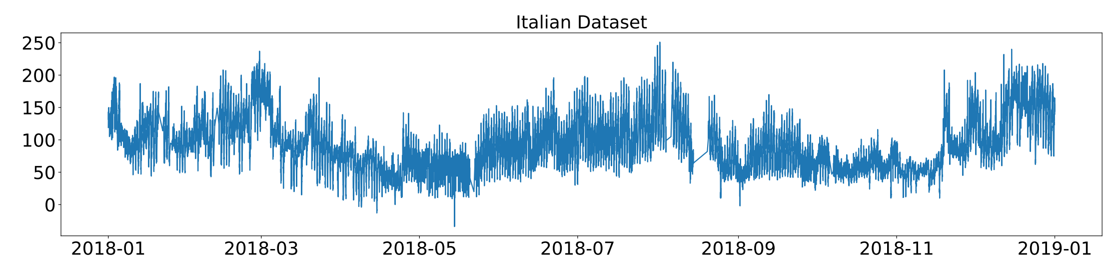

Figura 3: Italian dataset (just HVAC consumption for the year 2018 is shown).

electricity consumption anomaly detection [Fan *et al.*, 2018; Pereira e Silveira, 2018], providing significant improvements over classical state-of-the-art methods [Ruff *et al.*, 2021; Pang *et al.*, 2021; Chalapathy e Chawla, 2019]. In this report we focus on autoencoders.

Autoencoders (AEs) [Hinton e Salakhutdinov, 2006] are particular kinds of artificial neural networks which are trained to reconstruct their input, in a self-supervised manner. An AE is composed of an *encoder* network and a *decoder* network. The encoder takes as input the training data $x \in \mathbb{R}^d$, where d is the dimension of the data, and compresses these data into a *latent space* $z \in \mathbb{R}^h$, where h is the dimension of the encoding, usually $h < d$. Then, the decoder tries to map back the latent internal representation z to the original input space $\hat{x} \in \mathbb{R}^d$, through reconstruction. The encoder structure can be considered as a bottleneck, in which data pass and are compressed to extract a meaningful encoded representation. The decoder does the opposite. The two networks are characterized by f_ϕ, the encoding function, and f_θ, the decoding function, where $f_\phi : \mathbb{R}^d \to \mathbb{R}^h$ and $f_\theta : \mathbb{R}^h \to \mathbb{R}^d$. Finding weights (parameters) ϕ and θ for the two functions can be done by backpropagation, minimizing the loss function $\mathcal{L}_{AE}(x, \hat{x}) = \parallel x - \hat{x} \parallel^2$, called *reconstruction error*, given input x and model output $\hat{x}$.

The main idea behind the current use of AEs for anomaly detection is to train them only on nominal data (following a semi-supervised approach) so that they will not be able to accurately reconstruct anomalous behaviors (that the AEs have never seen), which will thus produce high reconstruction errors.

We test different AE architectures, including Variational AutoEncoders (VAEs) [Kingma e Welling, 2014], and different combinations of convolutional and recurrent layers [Goodfellow *et al.*, 2016].

4 Experimental Results

We perform experiments on two datasets collected from office buildings (that typically are inhabited only during working hours).

The first dataset contains electrical power consumption data of a research facility in the Netherlands over the year 1997, with a granularity of 15 minutes.

The second dataset contains electrical power consumption data of an office building in Italy over a period of three years, with a granularity of 15 minutes. Being a rather big building, it is equipped with several smart circuit breakers that collect information regarding parking lighting, electric vehicle charging stations, offices' lighting and power outlets, HVAC systems, elevators, kitchens, and cafeteria's refrigerators. For this building, also the outside temperature and humidity are collected, which represent valuable information as their correlation to electricity consumption has been studied extensively [Hor *et al.*, 2005]. Hence, this Italian dataset, unlike the Dutch one, is multivariate.

Figure 1 depicts the Dutch dataset, while Figures 2 and 3 display consumptions due to lighting and HAVC of the Italian dataset for the year 2018, respectively. A clear weekly seasonality can be observed in both datasets in the form of high consumptions during working hours on weekdays and low consumptions otherwise (easily observable in the plots where five peaks are followed by two days of low consumptions, i.e., weekends). In the Italian dataset, additionally, a yearly seasonality can be observed, which is different for the two series depicted. In Figure 2, it can be seen as consumption due to lighting increases during Winter as a consequence of the reduction of daylight hours, while, in Figure 3, which depicts HVAC consumption, the yearly seasonality is high both during Winter (when offices need to be heated) and Summer (when offices need to be cooled).

An electricity demand anomaly detector is hence a systems (algorithm) that need to be able to model multiple seasonalities and correlations among different dimensions.

Examples of anomalies we detect are the breaking of the weekly seasonality (e.g., during holidays) or abrupt spikes. Figure 4 contains examples of such anomalies.

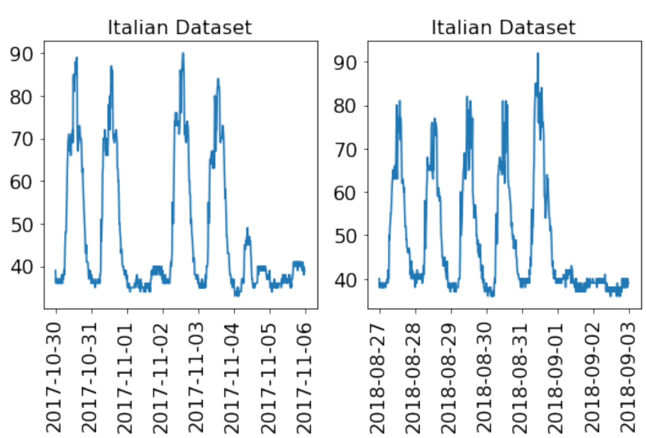

Figura 4: Anomalies on the Italian dataset. The All Saints' holiday (November 1), on the left. Abrupt spike (on Friday), on the right.

Among the various AEs architectures tested, we have noticed that more complex ones (especially those comprising recurrent layers) do not necessarily imply better detection performances. We suspect bigger datasets are needed to exploit the full potential of these models.

5 Conclusion

We have presented some preliminary results regarding the detection of anomalies in electrical energy consumption of two office buildings. Our results suggest that AEs represent a powerful tool for detecting anomalies in electricity demand time series.

Riferimenti bibliografici

[Araya *et al.*, 2017] Daniel B Araya, Katarina Grolinger, Hany F ElYamany, Miriam AM Capretz, e Girma Bitsuamlak. An ensemble learning framework for anomaly detection in building energy consumption. *Energy Build*, 144:191–206, 2017.

[Blázquez-García *et al.*, 2021] Ane Blázquez-García, Angel Conde, Usue Mori, e Jose A Lozano. A review on outlier/anomaly detection in time series data. *ACM Comput Surv*, 54(3):1–33, 2021.

[Bolchini *et al.*, 2017] Cristiana Bolchini, Angela Geronazzo, e Elisa Quintarelli. Smart buildings: A monitoring and data analysis methodological framework. *Build Environ*, 121:93–105, 2017.

[Chalapathy e Chawla, 2019] Raghavendra Chalapathy e Sanjay Chawla. Deep learning for anomaly detection: A survey. *arXiv preprint arXiv:1901.03407*, 2019.

[Chandola *et al.*, 2009] Varun Chandola, Arindam Banerjee, e Vipin Kumar. Anomaly detection: A survey. *ACM Comput Surv*, 41(3):1–58, 2009.

[Chou e Telaga, 2014] Jui-Sheng Chou e Abdi Suryadinata Telaga. Real-time detection of anomalous power consumption. *Renew. Sust. Energ. Rev.*, 33:400–411, 2014.

[Edgeworth, 1887] Francis Ysidro Edgeworth. Xli. on discordant observations. *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, 23(143):364–375, 1887.

[Esteva *et al.*, 2017] Andre Esteva, Brett Kuprel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, e Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, 2017.

[Fan *et al.*, 2018] Cheng Fan, Fu Xiao, Yang Zhao, e Jiayuan Wang. Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data. *Appl Energy*, 211:1123–1135, 2018.

[Goh *et al.*, 2017] Jonathan Goh, Sridhar Adepu, Marcus Tan, e Zi Shan Lee. Anomaly detection in cyber physical systems using recurrent neural networks. In *Proc. HASE*, pages 140–145, 2017.

[Goodfellow *et al.*, 2016] Ian Goodfellow, Yoshua Bengio, e Aaron Courville. *Deep learning*. MIT Press, 2016.

[Hamilton, 1994] James Douglas Hamilton. *Time series analysis*. Princeton University Press, 1994.

[Hinton e Salakhutdinov, 2006] Geoffrey Hinton e Ruslan Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006.

[Hor *et al.*, 2005] Ching-Lai Hor, Simon J Watson, e Shanti Majithia. Analyzing the impact of weather variables on monthly electricity demand. *IEEE Trans Power Syst*, 20(4):2078–2085, 2005.

[Kingma e Welling, 2014] Diederik Kingma e Max Welling. Auto-encoding variational Bayes. In *Proc. ICLR*, 2014.

[Mukherjee *et al.*, 1994] Biswanath Mukherjee, L Todd Heberlein, e Karl N Levitt. Network intrusion detection. *IEEE Netw*, 8(3):26–41, 1994.

[Nawaratne *et al.*, 2019] Rashmika Nawaratne, Damminda Alahakoon, Daswin De Silva, e Xinghuo Yu. Spatiotemporal anomaly detection using deep learning for real-time video surveillance. *IEEE Trans Industr Inform*, 16(1):393–402, 2019.

[Pang *et al.*, 2021] Guansong Pang, Chunhua Shen, Longbing Cao, e Anton Van Den Hengel. Deep learning for anomaly detection: A review. *ACM Comput Surv*, 54(2):1–38, 2021.

[Pereira e Silveira, 2018] Joao Pereira e Margarida Silveira. Unsupervised anomaly detection in energy time series data using variational recurrent autoencoders with attention. In *Proc. ICMLA*, pages 1275–1282, 2018.

[Raj e Portia, 2011] S Benson Edwin Raj e A Annie Portia. Analysis on credit card fraud detection methods. In *Proc. ICCCET*, pages 152–156, 2011.

[Ruff *et al.*, 2021] Lukas Ruff, Jacob R Kauffmann, Robert A Vandermeulen, Grégoire Montavon, Wojciech Samek, Marius Kloft, Thomas G Dietterich, e Klaus-Robert Müller. A unifying review of deep and shallow anomaly detection. *P IEEE*, 2021.