MACHINE LEARNING-BASED VIRTUAL SENSORS FOR GUIDING USER BEHAVIOUR: A CASE STUDY ON HOUSEHOLD APPLIANCES

Ilare, Dennis (1,2);
Cascini, Gaetano (1);
Manzoni, Stefano (1);
Mansutti, Alessandro (2)

1: Politecnico di Milano;
2: ROLD

ABSTRACT

The Agenda 2030 calls for collective awareness, starting with individuals. The interaction between users and household appliances produces a relevant amount of data that can be elaborated through Machine Learning algorithms to guide users towards sustainable behaviours. In particular, the data already available on household appliances can be conveniently used to create Virtual Sensors, increasing the overall information about the system. This paper focuses on the description of the pipeline for the creation of Virtual Sensors and applies it to a no-frost refrigerator. The Data Acquisition phase is described and feeds the Model Creation phase. For the case study, the data have been discretized and labelled to train a Random Forest algorithm. The validation of the model has been done on an independent dataset. An analysis of the minimum prediction accuracy required for the model is reported. Furthermore, experimental data shows the effect of hot load positioning on the compressor's working time rate.

Keywords: Virtual Sensors, Sustainability, Machine learning, Artificial intelligence, Case study

Contact:
Ilare, Dennis
Politecnico di Milano
Italy
dennis.ilare@polimi.it

Cite this article: Ilare, D., Cascini, G., Manzoni, S., Mansutti, A. (2023) 'Machine Learning-Based Virtual Sensors for Guiding User Behaviour: A Case Study on Household Appliances', in *Proceedings of the International Conference on Engineering Design (ICED23),* Bordeaux, France, 24-28 July 2023. DOI:10.1017/pds.2023.250

1 INTRODUCTION

1.1 Context

Household appliances are among the most diffused products in houses around the world; hence, they can significantly affect users and their approach towards sustainability. Indeed, Arçelik[1] reports that household appliances are responsible for 40% of home energy consumption. On the one hand, household appliances have been objects of energy and water consumption in nominal conditions for many years. On the other hand, the efficiency of systems depends not only on the design of the system itself but also on the usage and lifestyle behaviours of users. As an example, the user interaction with the refrigerator plays an essential role in its performance and the quality of food preservation, as reported by (Hoffmann, et al., 2021) . A further example, provided by (Maggipinto, et al., 2019), deals with the identification of laundry fabric to select the most suitable washing program: in actuality, the user commonly selects the generic washing program even if this behaviour affects the performance in terms of water and energy consumption. While everyone can easily perceive the energy consumption aspect, the users are not always aware of the correct behaviours to guarantee an optimal delivery of the household appliance function. For example, (James, et al., 2017) in their review, show a significant number of users who are unaware of the correct positioning of food inside their refrigerator. The user effect remains valid also at a higher level: (Iweka, et al., 2019) explain that identical residences can show up to 37% disparity in electricity because of different energy-related behaviours of people (i.e. how often and how much time they spend cooking, how they manage space and water heating, air conditioning, entertainment and lighting etc.).

Household appliances have always been equipped with sensors and Artificial Intelligence algorithms have the potential to elaborate all the data coming from them to guide users towards sustainable behaviours. The low computing capacity available on household appliances makes the application of AI on household appliances limited to computationally compliant algorithms.

The digitalization process in household appliances has been focused on the possibility of remote monitoring and control, while "smarter" appliances are currently limited to prototypes or high-class products. These applications have limited positive impact on a large scale because of their insufficient diffusion.

1.2 General aim and specific goal

Data from dedicated sensors represent a straightforward solution to recognize user behaviours in situ (i.e. in the same place where the user is interacting). Indeed, the correlation between a defined user behaviour and a specific sensor designed to intercept it is pretty natural. At the same time, embedding new multiple sensors in the appliance sector is rarely accepted for both costs and system complexity reasons. This consideration is particularly true considering medium-low-cost appliances. The latter is, indeed, the most diffused category across users, making it the most promising class to be considered. The possibility to integrate user behaviour recognition without the necessity of adding hardware relies upon a software solution. This paper proposes to use Virtual Sensor technology, i.e. *a type of software that, given the available information, processes what a physical sensor otherwise would* [2], on a medium-cost, frost-free refrigerator. In particular, multiple Virtual Sensors are built to predict the distribution temperature inside the refrigerator. Since a lot of sensors for control purposes are available on refrigerators (and more in general on household appliances), it is possible to use them as input data for creating Virtual Sensors through Machine Learning (hereafter "ML"), aiming to enrich the amount of information inside it. Such additional information can be conveniently combined with the data coming from physical sensors, formally creating a Digital Twin (i.e. a digital representation of a physical system). Finally, the Digital Twin allows to profile the user interaction with the household appliance and, through the application of ML, to recognize behaviours that can reduce the lifetime of the appliance or can be energetically inconvenient. In a more schematic way: starting from physical sensors acquisitions (step 1), it is possible to combine them using ML to create Virtual Sensors (step 2). Using the first two steps, it is possible to create a Digital Twin (step 3) and use it to intercept some specific user behaviour (step 4).

[1] Arçelik A.Ş., 2021. 2021 Annual Report, s.l.: Arçelik A.Ş..

[2] Jackson, C., 2019. What is a virtual sensor? [Online] Available at: https://www.lifecycleinsights.com/what-is-a-virtual-sensor/ [Accessed 15 11 2022]

This paper focuses on the application of the first two steps of explained method on a low-medium market frost-free refrigerator. It is divided into four further sections. **Section 2** examines the **State of the Art** about Virtual Sensors for refrigerators and user recognition models and defines the research questions of this work. In **Section 3,** the path for the **Model Creation** of the Virtual Sensors in the refrigerator is presented, with an explanation of the data acquisition pipeline and the definition of the mathematical model. The **results** and evaluation criteria for the Virtual Sensors are outlined in **Section 4**. Eventually, the **conclusions** are reported in **Section 5**.

2 STATE OF THE ART

2.1 Virtual sensors for refrigerators

It is not common to find the formal application of Virtual Sensors at the market level; however, it is more and more common to find *smart refrigerators* on the global market.

The shifting from a typical refrigerator to a *smart refrigerator* can be done with different modalities and characteristics, depending on producers, geographical position and integration level. In particular, we can recognize two main approaches: (i) an integrated solution that acts at the platform level, and (ii) a retrofit solution that adds new functions to a refrigerator, independently from its nominal characteristics.

Regarding the first approach, we can see two main solutions proposed by global producers, i.e. the producer of the product as a whole, made up of many components integrated into a single solution. Indeed, Whirlpool[3] has developed a refrigerator with Adaptive Intelligence technology, claiming continuous sensing and adaptation of the refrigerator based on room conditions and usage patterns. This feature has been conceived in particular for the Indian market. On the other hand, Panasonic[4] claims the use of AI in refrigerators to adapt and optimize the inverter functioning (for better compressor control). A different approach is applied for the European market by different producers - e.g. Bosch[5] and Liebherr[6] - where the effort is more focused on the remote control and integration of the refrigerator in the smart home ecosystem.

Differently from integrated solutions, retrofit ones allow for adding smart functionalities to a common refrigerator. An example is Fridge Eye[7], a smart camera which enables users to remotely control the food inside the refrigerator by using a dedicated app to integrate image recognition of groceries. To have satisfactory results, more cameras are needed. A similar solution has been developed through a collaboration between Microsoft and Liebherr[8], where deep learning algorithms have been used for image recognition inside the refrigerator. SYNAP IoT[9] is a plug & play solution for industrial refrigeration (B2B market), claiming intelligent predictive maintenance as an additional feature.

In more narrow terms, there are few examples of Virtual Sensors in refrigerators in literature. In (Vitor, et al., 2020) the authors introduce the concept of controlling the defrost timing in a single-compartment refrigerator without adding sensors to the product. Different variables such as temperatures and opening times have been used to predict the right defrost timing just using a software solution based on regression methods. The method is interesting since it uses the data already available

[3] Whirlpool, 2022. Intellifresh 340L 3 Star Convertible Frost Free Double-Door Refrigerator. [Online] Available at: https://www.whirlpoolindia.com/intellifresh-340-l-3-star-frost-free-double-door-refrigerator--convertible-freezer---inverter-compressor-/p [Accessed 15 11 2022].

[4] Croma, 2022. Panasonic Refrigerators. [Online] Available at: https://www.croma.com/lp-panasonic-refrigerator [Accessed 15 11 2022]

[5] BSH Elettrodomestici S.p.A., 2022. Frigoriferi French Door. [Online] Available at: https://www.bosch-home.com/it/prodotti/frigoriferi-congelatori/frigoriferi-french-door [Accessed 15 11 2022].

[6] Liebherr, 2022. SmartDevice: il tuo ingresso nella Smart Home. [Online] Available at: https://home.liebherr.com/it/ita/alla-scoperta-di-liebherr/smartdevice/smartdevice.html [Accessed 15 11 2022].

[7] Indiegogo, 2022. Fridge Eye - Turn Your Fridge Into a Smart Fridge.. [Online] Available at: https://www.indiegogo.com/projects/fridge-eye-turn-your-fridge-into-a-smart-fridge#/updates/all [Accessed 15 11 2022].

[8] Hughes, O., 2016. Microsoft's smart fridge project might soon be able to tell you when you're out of milk. [Online] Available at: https://www.ibtimes.co.uk/microsofts-smart-fridge-project-might-soon-be-able-tell-you-when-youre-out-milk-1579783 [Accessed 15 11 2022].

[9] SYNAP IoT, 2022. NO.1 BOX. [Online] Available at: https://www.synapiot.com/thebox [Accessed 15 11 2022].

on the appliance to optimize its functioning. The model is developed on a single-compartment refrigerator, where the variability of the temperatures usually is uniform compared to other kinds of domestic refrigerators. In (Andrade-Ambriz, et al., 2022) the authors use Neural Networks with acoustic signals to predict the amount of frost on a domestic refrigerator. They create four different classes of frost, starting from "No frost" up to "High frost" and obtaining an overall prediction accuracy of 94% on their dataset. The main weakness in this study is the necessity of adding a new kind of sensor to the appliance: household appliance producers are quite reluctant to the introduction of additional sensors, in particular, if they are different also in terms of technology. This gap is usually overcome only if there are particularly advantageous economic aspects or in the presence of a brand-new function. There are different authors (Zhang, et al., 2018), (Anon., 2021), (Jain, et al., 2021) that are focused on the application of image recognition for food inside refrigerators. Despite this not being the exact application of the Virtual Sensor on a refrigerator, it is a topic of particular relevance because of its affinity in terms of applied algorithms and possible development. The main drawback of this kind of approach is the necessity of additional cameras inside the refrigerator. Even if some models are already equipped with cameras, this feature is generally dedicated to top-level market refrigerators. These solutions are limited in terms of diffusion and, consequently in terms of environmental impact.

2.2 Virtual sensors for usage mode and user behaviour

The increasing awareness of the users regarding environmental sustainability is pushing the producers to develop solutions aiming at reducing the energy consumption of the house. In this case, a Virtual Sensor has to be meant as the ability of sensing behaviours and usage modes without sensors specifically designed for this purpose. To do that, a form of intelligence is needed.

Different solutions are nowadays present on the market. A first example is Sense[10], a Machine Learning based plug & play solution that helps real-time monitoring of the electrical loads in the house, claiming a reduced energy consumption at the house level. This solution uses the concept of electrical signature (i.e. the ability to recognize an electric load and assign it to an object or a specific action of the object), assigning different loads to different objects connected to the electrical grid and allowing monitoring also in terms of correct functioning of the object itself. This kind of solution is widely adopted by others at the market level, increasing user visibility on the most consuming elements inside the house. The main limitation of this kind of solution is the missing information about the user's specific actions, which may cause increased energy consumption. In this way, the user will recognize which of the objects electrically connected is causing a waste or an overload, but it will be less intuitive for the user to understand the specific behaviour that causes the anomaly. A second example is hOn[11], a mobile application that allows the monitoring of different household appliances, giving the possibility of having some tips for their correct usage (e.g. how to load a dishwasher efficiently). To have feedback from the application, the user has to interact with the mobile phone while using the household appliance (e.g. take a photo to have feedback about dishwasher loading). This solution guarantees that the monitoring of the appliances is completely automatized. One disadvantage of this approach is the necessity of a considerable effort from the user to receive specific tips about correct actions without the possibility of real monitoring of the behaviour.

Different studies about frequent communication of electricity consumption to raise user awareness and foster virtuous user behaviour, such as (Riche, et al., 2010) and (Grønhøj & Thøgersen, 2011), have been conducted. The focus was on providing more frequent communication than the one usually based on the consumption bill, to provide the users with a tool to check their habits and to act to improve them. To the best of our knowledge, there is still a lack of research about giving suggestions to the user at the time and at the place where the interaction with the household appliance is happening. However, this kind of information could be beneficial to increase user awareness towards more sustainable behaviours.

In general, to overcome the limitations of current solutions, two aspects should be taken into account together: (i) the precision in terms of time and place of the suggestions, (ii) a well-integrated communication between the appliance and the user to reduce as much as possible the user effort.

---

[10] sense, 2016. How Does Sense Detect My Devices? [Online] Available at: https://blog.sense.com/articles/how-does-sense-detect-my-devices/ [Accessed 17 11 2022].

[11] hOn, 2022. Home - hOn | Your smart life companion. [Online] Available at: www.hon-smarthome.com [Accessed 17 11 2022].

2.3 Research question

Given the context described in Section 1 and the solutions currently existing on the market depicted in Section 2, this research aims to verify the possibility of using Virtual Sensors based on an equivalent set of physical sensors currently present on a medium-class appliance to recognize off-design conditions that could be interesting for the user. To achieve the goal of this research, it is necessary to study the minimum needed accuracy of the Virtual Sensors to guarantee that a in situ recognition of user behaviour is technically feasible.
This paper will attempt to answer the following research questions: *(i) using as inputs the data coming from the same sensors currently present on a refrigerator: what is the achievable accuracy in temperature distribution prediction? (ii) Does this accuracy allow the recognition of off-design behaviours?*

3 MODEL CREATION

The Virtual Sensors solution can be conveniently applied on a medium-low market refrigerator, where the cost represents a key factor and therefore it is not sustainable to add new sensors. The refrigerator results are interesting due to its continuous functioning and because of the relevance of user interaction with its performance. Moreover, medium-class refrigerators are largely diffused and with a possible relevant impact.
This section starts from the acquisition pipeline for the physical sensors already present on the refrigerator (e.g. temperature sensors in different compartments, door switch etc.) and describes the process up to the creation of a Virtual Sensor for the temperature of every shelf of the refrigerator compartment.

3.1 Data acquisition pipeline

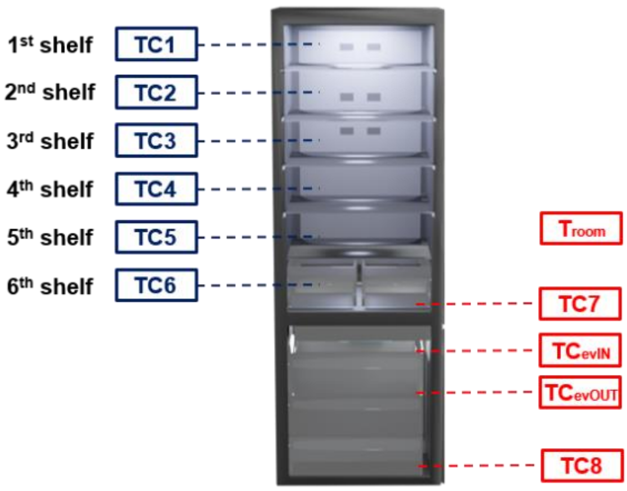

Figure 1. Position of the temperature sensors for the acquisition phase. The sensors already installed on the refrigerator are indicated with red labels, while the ones needed for the creation of the Virtual Sensors on each shelf are indicated in blue.

The case study described in this paper is focused on a bottom-up no-frost refrigerator. This technology is based on a freezer compartment placed on the bottom and equipped with an evaporator. In the upper part of the appliance, there is the refrigerator compartment. These two compartments can communicate thanks to a damper: a mechanical valve that opens only when the upper compartment needs to be cooled down. Once the damper is open, a fan allows the circulation of cool air from the freezer compartment to the refrigerator compartment. The peculiarity of no-frost refrigerators is the presence of a heating element on the evaporator. This element melts the frost that accumulates on the evaporator during the refrigeration cycle. Such a function reduces the human effort for refrigerator maintenance. This type of refrigerator is largely diffused, particularly in Europe, because of a fair trade-off between costs and performance. Its diffusion on the market and the potential environmental impact make it interesting as a starting point for our study.

https://doi.org/10.1017/pds.2023.250 Published online by Cambridge University Press

The process described in Section 1 includes an initial step that needs data from the physical sensors in the household appliances. This data can be obtained by reading the motherboard of the refrigerator or by placing sensors in substitution of the ones already present in it. In our situation, the second solution is more convenient not only for the sake of simplicity but also to guarantee a uniform data type. In fact, for the creation of Virtual Sensors, it is needed to acquire data in the positions where we want to predict a physical measure, but where we do not have currently available sensors. In the refrigerator case, we would like to predict the temperatures on each shelf, but the products, which are currently available, do not have sensors in these positions and it is necessary to place additional ones. Referring to Figure 1, the system under analysis is equipped (in red in Figure 1) with a temperature sensor for the refrigerator compartment (TC6), the freezer compartment (TC7), the room temperature (T_{room}), the evaporator inlet (T_{evIN}), the evaporator outlet (T_{evOUT}). Combining these signals with others (i.e. the state of the door and the state of the damper), we want to predict all the temperatures on the other shelves (in blue in Figure 1). The choice to place substituting sensors on the present ones guarantees that all the collected data will have the same characteristics in terms of the type of sensors and acquisition frequency. All the temperature sensors are thermocouples of type T, while the Boolean signals are acquired using an electromechanical switch - i.e. an electrical component consisting of movable electrical contacts, actuated by mechanical parts.

To guarantee uniformity of data type and acquisition synchronization, a LabVIEW application for Real-Time acquisition hardware has been used. In particular, the Real Time acquisition hardware ensures the acquisition frequency. For this application, the sampling rate has been fixed at 1 Hz while the door is closed and 10 Hz while the door is open. The dynamics of the phenomena inside the refrigerator are generally relatively slow, but the opening of the door is an important event and, at least in the first analysis, it is crucial to have a complete view of the system. The modules used for the acquisition of all the signals allow for a maximum sampling frequency of 75 Hz with high accuracy in time guaranteed by the Real Time module. All the acquisitions are collected as time series.

The data collection has been structured aiming at the creation of a dataset comprehending different load conditions. Even if the refrigerator is quite a simple technology, its interaction with the environment and its dependency on the user interaction (e.g. frequency and time of door opening) makes it very complex to create a complete dataset. For this reason, we defined different conditions that represent common situations as the starting point. Our goal is to create a model that can predict temperatures in these normal conditions, while anomalies will be highlighted as outliers. The tested conditions are the linear combination of the following variables:

- Initial Condition of the Refrigerator Compartment: ICRC = {Empty; Full}
- Initial Condition of the Freezer Compartment: ICFC = {Empty; Full}
- Shelf Position of the new load inside the Refrigerator Compartment: SP = {None; 1st; 2nd; 3rd; 4th; 5th; 6th; Full}

For all the tests, we decided to use bottles of water, because of their availability, their well-known thermal capacity and the possibility of reusing them. The RC is considered full with the following litres of water for each shelf: 6 L for the 1st shelf, 13L for the 2nd, 3rd and 4th shelves, and 8 L for the 5th and 6th shelves. Each of these values is obtained as the maximum amount of 1.5 L bottles and 5 L tanks that can be physically placed on the shelf itself. The FC is considered full with 20L of iced water. The load is always represented by water at ambient temperature (between 20 and 25°C) in compliance with the amount that can be loaded on different shelves (same distribution showed for the full RC). The tests are always conducted starting from a steady-state condition, i.e. after at least 4 hours since the last opening of the refrigerator. As it can be observed in Table 1, 30 different conditions have been tested for more than 20 days of acquisition and a total amount of 18000 samples (considering the first 10 minutes of each test with a sampling frequency equal to 1 Hz). Considering only the first 10 minutes after the door opening (combined with the slow dynamic of the system) allows approximating as time-independent every sample and to use algorithms that are compliant with the available computational resource.

3.2 Modelling

Once the acquisition phase is completed, the dataset with all the physical quantities of interest is ready for analysis and modelling. Before arriving at the creation of a mathematical model, every data-driven approach needs some preliminary steps. One of the typical steps is the cleaning of the data, e.g. the handling of missing or noisy data. This step is necessary with data coming from different sources or coming from acquisition systems that work in an operating environment so they can have communication issues. Since we collected data in a controlled environment, under test conditions, we do not have particular issues with missing data and so we do not need any data-cleaning phase.

Table 1. List of variables used to model the problem

Variable names	**Variable Type (Feature/Target, Direct/Indirect, Numeric/Boolean)**			**Notes**
TC1, TC2, TC3, TC4, TC5, TC6	Target	Direct	Numeric	Temperatures of the different shelves of RC
TC_{evIN}, TC_{evOUT}	Feature	Direct	Numeric	Temperatures of the inlet and outlet of the evaporator
TC7	Feature	Direct	Numeric	The temperature of the RC control
TC8	Feature	Direct	Numeric	The temperature of the FC control
TC_{room}	Feature	Direct	Numeric	The temperature of the room
DAMPER	Feature	Direct	Boolean	State (open or closed) of the damper
DOOR	Feature	Direct	Boolean	State (open or closed) of the door
$Dt_{lastOpen}$	Feature	Indirect	Numeric	Time lasted from the last door opening
$TC7_{lastOpen}$	Feature	Indirect	Numeric	The temperature of the TC7 the last time the door was opened
$TC7_{xSecAgo}$	Feature	Indirect	Numeric	The temperature of TC7 x seconds before the prediction

Another typical step is feature engineering, where indirect variables are created starting from direct variables (as indicated in Table 2). The term *feature* indicates those variables that the model needs to generate a prediction. The term *target* represents the variable generated by the model given some inputs. In Table 1, the Variable Type column resumes what is the role of a certain variable in the model (Target vs Feature), how the variable has been obtained (Direct vs Indirect) and the data type (Numeric vs Boolean).

Among the different possibilities, we decided to model the problem as a classification problem. In this approach, the classes are temperature classes, where each of them represents an interval of temperatures. This choice derives from the need to take into account the computational power and the type of logic available on medium-cost refrigerators motherboards. Moreover, the adopted approach allows fixing the tolerance of the Virtual Sensor through the discretization of the variables. The possibility of fixing the upper bound and lower bound of each class allows us to manage and optimize both the computational power and the prediction error of the final application. The steps from this point up to the creation of the model are:

1. Discretization of the target variables for the creation of the prediction classes: in our case, we applied an equal-width uniform discretization, where every class contains all the values within an interval of ±2 °C from the mean value of the class.
2. Separation of the dataset in a train set and a test set.
3. Creation of the data-driven model using the *Random Forest* algorithm (Breiman, 2001). This algorithm was chosen because its output consists of a series of Boolean choices, as *Decision Trees* are general. This means it can be represented in a matrix form. This representation fits very well with the most diffused motherboard of household appliances. Moreover, in the interest of limited available computational and memory resources, the hyper-parameters of the algorithm have been fitted with an optimization process allowing a maximum of 25 estimators and a maximum of 6 nodes for each of them. This means that the Forest is composed of a maximum of 25 *Decision Trees* with a maximum of 6 nodes each, and the prediction class is assigned by majority voting criteria. Moreover, this type of algorithm integrates the possibility of having the weights of the feature in the prediction process, integrating the feature selection process. The hyper-parameter selection, the feature selection and the estimation of the performance on the training dataset are performed using cross-validation on the train set.

https://doi.org/10.1017/pds.2023.250 Published online by Cambridge University Press

4. Evaluation of the performance on the test set. This step allows the evaluation of the performance on a completely independent dataset.
5. The whole process is repeated for each shelf.

Table 2. Prediction accuracy of the model on an independent test set

	TC2	TC3	TC4	TC5	TC6
Prediction Accuracy	0.85	0.87	1.00	0.76	0.89

Some preliminary analysis and the initial trials of model creation suggested that to achieve reasonable performance an additional temperature sensor, acting as input, is needed. For this reason, we decided to transform the TC1 target (the temperature sensor in the top part of the refrigerator) into an input for our model. From a practical point of view, this means that our system needs an additional sensor in this position to work properly. This slight modification is considered acceptable, given the kind of additional sensor needed - already existing inside the refrigerator - and the considerable advantages that brings with it.

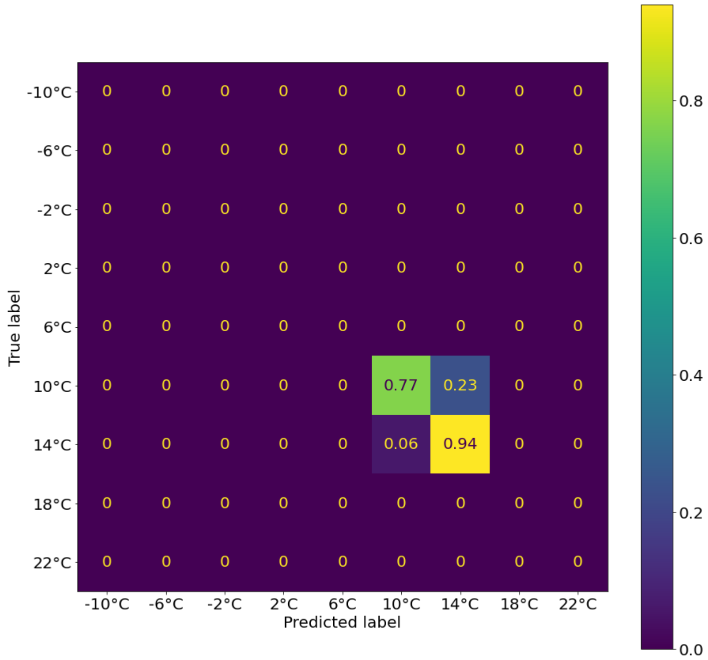

Figure 2. Example of confusion matrix of TC2 on an independent test set

Table 2 reports the prediction accuracy for all the shelves on an independent test set. The accuracy for this kind of classification is calculated as the ratio between the correctly classified samples for a certain class over the total number of samples of the test.

These results come from a single independent test set. This means that a brand-new dataset has been acquired and used for a preliminary performance evaluation. Further tests with different conditions must be performed to evaluate better the overall performance and the ability of the model to infer different ambient conditions and different user interactions.

Figure 2 shows an example of a confusion matrix for predicting TC2 on the same independent test set used for the prediction accuracy reported in Table 2. This kind of representation is helpful in understanding the model's strengths and weaknesses better. It represents the percentages of class prediction. It allows visualizing in which class assignation the model is robust and in which it is not. Furthermore, it shows how far from the correct prediction our model is. The example of Figure 2 shows that for what concerns TC2 prediction, our model has some issues in the classification of temperatures above 14°C.

https://doi.org/10.1017/pds.2023.250 Published online by Cambridge University Press

4 RESULTS

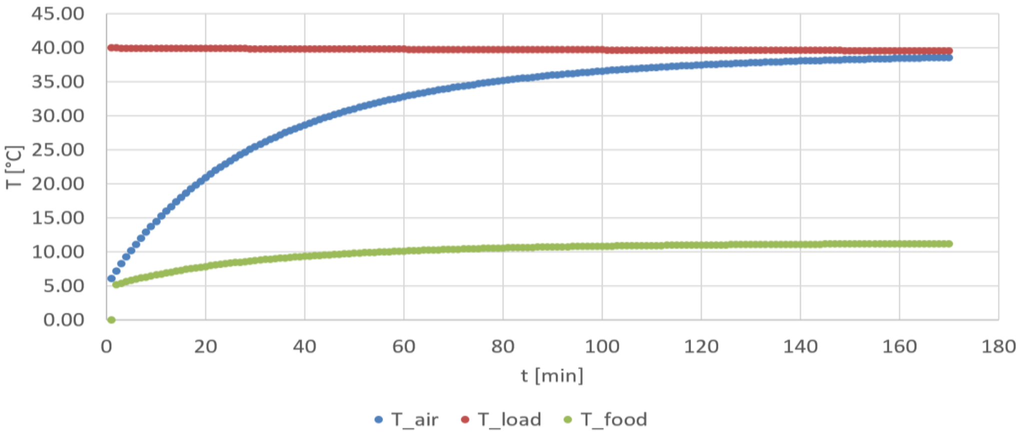

Figure 3. Theoretical effect of 1kg of water at 40°C on the air inside the refrigerator and already cold food

To evaluate the results and the performance of the model created, it is important to define a target in terms of the minimal prediction accuracy needed and the expected impacts. The following section is needed to understand whether the definition of the temperature classes (performed in Section 3.2) is tight enough to be able to intercept a specific user behaviour and to preliminary evaluate the impact of this ability.
To define the minimal prediction accuracy needed by the model, some theoretical calculations have been done. Assuming a completely isolated refrigerator compartment, with 250 L (the entire compartment) of cold air at 5°C, with null airflow (i.e. with the damper closed, so without new cold air coming from the freezer) and considering just natural convection as heat transfer, it is possible to evaluate how a hot load acts on the temperature distribution of the compartment. The hot load is considered equivalent to the wrong behaviour of the user in the use of the appliance. Figure 3 shows that with these conditions, the hot water causes a fast increase in the air temperature (10°C in less than 10 minutes). This theoretical consideration leads us to define that a minimum prediction accuracy in the order of magnitude of 1°C is enough to intercept such kind of possible behaviour from a user.
This requirement is compliant with the model created in Section 4. The developed model has a prediction tolerance of 2°C - imposed by the discretization process - and the accuracy shows that the model can predict most of the situations with the available data.
The ability to intercept a hot load positioned in the refrigerator or other inconvenient behaviours is interesting because these actions can be the root of increased energy consumption or incorrect preservation of the food. For a precise evaluation of these effects, a considerable number of tests and data are necessary. For a preliminary evaluation of the consequences of a hot load inside the refrigerator, some tests have been conducted. A hot load of 1kg of water at 40°C was placed in a no-frost refrigerator in different positions, namely on the top shelf (1st shelf in Figure 1) and on the bottom shelf (6th shelf in Figure 1). These tests have been compared with the functioning of the empty refrigerator in terms of compressor functioning time percentage.

Table 3. Compressor functioning time percentage on real tests depends on the position of the hot load in the refrigerator compartment.

	Hot load - 1st shelf (top)	**Hot load - 6th shelf (bottom)**	**Baseline (empty)**
Compressor functioning time %	15-20%	25-35%	45-55%

Table 3 reports the different percentages of compressors functioning dependently on the load conditions. The percentage is expressed as ranges because of the variability of the refrigerator's behaviour due to different ambient conditions. From this table is quite evident that the introduction of a hot load in the refrigerator compartment causes a relevant increase in the compressor working time and, consequently a higher electrical consumption. This is especially true for the loading in the bottom part where there is the temperature sensor for the refrigerator compartment control: the hot load has a more direct effect on the air close to this sensor, causing a more frequent request for cold air.

https://doi.org/10.1017/pds.2023.250 Published online by Cambridge University Press

5 CONCLUSIONS

This paper presents an approach to increase the available data coming from a no-frost refrigerator without the necessity to add new sensors. The original contribution of this work is given by the use of Virtual Sensors on household appliances to intercept user behaviours with the final goal of environmental sustainability. This kind of application could be relevant in terms of diffusion thanks to the exploitation of Machine Learning that avoids the need for additional hardware.

State of the art of solutions existing on the market and studied in the literature about the application of Virtual Sensors inside a refrigerator have been presented. The same has been applied to the recognition of user behaviour through the exploitation of Machine Learning. The necessity of additional sensors, the application on top-level refrigerators and the necessity of a relevant user effort represent the main limitations of the current solutions. It has presented a Virtual Sensors pipeline that starts with the acquisition of data coming from the appliance and continues with ML techniques to create a data-driven model.

This ML-based model demonstrates that *(i) it is possible to predict temperatures on different shelves of the refrigerator with ± 2 °C accuracies simply using the currently available sensors and* **adding** *a temperature sensor on the top shelf*. The ±2 °C represents an upper bound since the authors have imposed it using a discretization process. The prediction accuracies on the independent test set are generally higher than 76%. It is important to remind that all the predictions have been performed on the refrigerator after at least 4 hours since the last door opening. Given that our findings are based on a limited number of tests, the result from such analyses should be treated with the utmost caution. The goodness of this performance must be confirmed by additional tests on the same refrigerator in a different ambient, to test the inference of the model and eventually to create a more general one. Moreover, the performance is obtained thanks to the addition of a temperature sensor (TC1) to the inputs of our model. Although it is not an ideal situation, we think that this slight modification is acceptable in favour of the additional function that this method is adding.

It has been illustrated that from a theoretical point of view *(ii) the accuracy of ±2 C is enough to intercept some harmful user behaviour, such as the positioning of a hot load inside the refrigerator*. We are aware that this statement will need support also from real data for stronger validation, and so the future development of the research will consist of new acquisitions of the same refrigerator in a different environment to verify the validity of the model in a different environment and eventually correct it. Furthermore, an experimental campaign to verify the possibility of intercepting specific user behaviours will be conducted.

REFERENCES

Andrade-Ambriz, Y. A. et al., 2022. Frost thickness estimation in a domestic refrigerator using acoustic signals and artificial intelligence. *Expert Systems with Applications*, 1 09.Volume 201.

Anon., 2021. Normalization of data for training and analysis by the MaskRCNN model using the k-means method for a smart refrigerator's computer vision. *Journal of Physics: Conference Series*, Volume 1889.

Breiman, L., 2001. Random Forests. *Machine Learning*, Issue 45, pp. 5-32.

Grønhøj, A. & Thøgersen, J., 2011. Feedback on household electricity consumption: learning and social influence processes. *International Journal of Consumer Studies*, pp. 138-145.

Hoffmann, T. G. et al., 2021. Impact of household refrigeration parameters on postharvest quality of fresh food produce. *Journal of Food Engineering*, Issue 306.

Iweka, O., Liu, S., Shukla, A. & Yan, D., 2019. Energy and behaviour at home: A review of intervention methods and practices. *Energy Research & Social Science*.

Jain, P. et al., 2021. Automated Identification Algorithm Using CNN for Computer Vision in Smart Refrigerators. *Computers, Materials & Continua*, Volume 71.

James, C., Onarinde, B. & James, S., 2017. The Use and Performance of Household Refrigerators. *Comprehensive Reviews in Food Science and Food Safety*, Issue 16, pp. 160-179.

Maggipinto, M. et al., 2019. Laundry Fabric Classification in Vertical Axis Washing Machines Using Data-Driven Soft Sensors. *Energies*, Issue 12.

Riche, Y., Dodge, J. & Metoyer, R. A., 2010. *Studying Always-On Electricity Feedback in the Home*. s.l., s.n.

Vitor, M. F., dos Santos Silveira, A. & Flesch, R. C., 2020. Ambient virtual sensor based defrost control for. *Applied Thermal Engineering*, Issue 166.

Zhang, W. et al., 2018. Multi-source data fusion using deep learning for smart refrigerators. *Computers in Industry*, 02, Volume 95, pp. 15-21.