

Energy-based decision engine for household human activity recognition

Anastasios Vafeiadis*, Thanasis Vafeiadis*,
Stelios Zikos, Stelios Krinidis,
Konstantinos Votis, Dimitrios Giakoumis,
Dimosthenis Ioannidis, Dimitrios Tzovaras
Information Technologies Institute
Center for Research & Technology Hellas
Thessaloniki, Greece Email:
{anasvaf, thanvaf,

szikos, krinidis, kvotis, dgiakoum, djoannid, tzovaras}@iti.gr

Abstract—We propose a framework for energy-based human activity recognition in a household environment. We apply machine learning techniques to infer the state of household appliances from their energy consumption data and use rulebased scenarios that exploit these states to detect human activity. Our decision engine achieved a 99.1% accuracy for real-world data collected in the kitchens of two smart homes.

I. INTRODUCTION

Knowing the true activity of occupants in a building at any given time is fundamental for the effective management of various building operation functions ranging from energy savings to security targets, especially in complex buildings with different internal kind of use. As the activities of occupants within the building vary throughout the day, it is difficult to characterize the different activities in different time periods. In general, activity monitoring in buildings is of high interest, since it significantly contributes to the improvement of a building's energy efficiency [1] and increases the quality of life of people in Ambient Assisted Living (AAL) environments [2]. Therefore, there is a need for detailed activity knowledge.

Human activity can be estimated using various sources, such as movement sensors [3], occupancy sensors [4], cameras [5], audio [6] as well as appliance current consumption [7].

In this paper, we propose a decision engine that is able to identify the activities based on the energy consumption rate of household appliances using smart plugs to support non-intrusive load monitoring (NILM) and machine learning. The human activity is recognized using only the energy consumption rate information from several appliances in a domestic environment by using only smart plugs. The decision engine for human activity recognition applies popular machine learning classifiers (supervised) on household appliances aiming to infer the appliance status (ON/OFF) along with a real time appliance activity proportion measurement for each appliance to determine daily household activities related to these appliances. To determine the most effective classifier for

Liming Chen, Raouf Hamzaoui

Faculty of Technology
De Montfort University
Leicester, UK

Email: {liming.chen, rhamzaoui}@dmu.ac.uk

each appliance we run a series of Monte Carlo simulations testing different settings for each classification method.

When it comes to human activity recognition, users' privacy is an important issue. Most of the proposed methods address the problem of human activity recognition using intrusive techniques for data collection (e.g. cameras, wearables), energy consumption rate thresholding on workload of appliances, time series analysis, probabilistic techniques and machine learning. Our work proposes the use of unobtrusive and easy-install tools (smart plugs) for data collection and a decision engine that combines energy signal classification using dominant classifiers (compared in advanced with grid search) and a probabilistic measure for appliance usage. It helps preserving the privacy of the resident, since all the activities are stored in a local database.

The remainder of the paper is organized as follows. In Section II, we review the literature. In Section III, we describe the energy consumption rate dataset and how we pre-process it. In Section IV, we formulate the decision engine for the activity recognition. In Section V, we describe our simulation setup and give our results. In Section VI, we draw our conclusions.

II. RELATED WORK

Many approaches have been proposed to address the problem of activity recognition in domestic environments. Most methods operate on the basis of multi-parametric data, taken from multiple modalities; i.e., various kind of sensors installed in the house environment.

Kim et al. [8] compared the performance of Hidden Markov Models (HMMs), Conditional Random Fields (CRFs) and the Skip-Chain CRF, of eating activities in a home environment. Nazefrad et al. [9] compared the performance of HMMs and CRFs for activity recognition in a smart home environment, using real time data from motion and temperature sensors. Giakoumis et al. [10] proposed an activity recognition scheme

for daily activities such as cooking, eating, dishwashing and watching TV, based on depth video processing and Hidden Conditional Random Fields (HCRFs), achieving an overall accuracy of 90.5% in a natural home environment.

More recently, Stankovic et al. [11] proposed an innovative methodology that characterizes the energy consumption rate of domestic life by making the linkages between appliance enduse and activities through an ontology built from qualitative data about the household and NILM data. Lavin and Klabjan [12] proposed a clustering technique for time series, on energy usage data provided by several U.S. power utilities, aiming to compare and contrast those with similar energy usage tendencies and to identify potentials for energy efficiency along with the open and close hours for business.

Cottone et al. [13] trained an HMM as an automated activity recognizer using sensors network readings initially converted into meaningful events, by applying a lossy compression algorithm based on minimum description length. The aim of their work was to level out peaks of energy consumption rate by identifying the appliances whose service is effectively needed by users, and postponing the use of the others until the combined demand for energy falls below some predefined threshold. In the work of Xu et al. [14] an alternative scalable two-stage methodology for household consumption segmentation is proposed that considers both the shape of a load profile (the time and magnitude of its peaks in household appliances consumption) along with its overall consumption to determine different typical consumer behavior patterns. Rao et al. [15] proposed an approach combining machine learning (Support Vector Machines) edge analysis and time series models (autoregressive moving average) on the identification of active appliances and on the prediction of future power usage utilizing demographic data in addition to aggregate power usage over time. Deshmukh and Lohan [16] proposed a framework for creating the appropriate features and labels from the training data and used these features to predict the appliance status (ON/OFF) and appliance energy consumption rate using a variety of classifiers. Finally, Belley et al. [17] proposed an algorithm for human activity recognition extracting features from the active and reactive power of each device using a Gavazzi meter. This method does not consider the case where a smart home would be equipped with several devices of a specific model. Furthermore, it would require the purchase of materials to measure the harmonics in power system in some cases, such as the television.

Contemporary activity recognition methods in smart homes rely mostly on sensors, which are further separated into wearable [18] and environment-related ones [19]. Recent work [20] shows that ontologies and semantic technologies have been used for activity modeling and representation, as well as knowledge-driven approaches [21]. Wearable-based techniques depend on user interaction with the sensor and, in most cases, on user motion measured with accelerometers.

III. DATA COLLECTION AND ANALYSIS

In order to infer the daily activities of a resident from electricity meters, one has to know the operating state of an electrical appliance. Estimating the operating state of an

electrical appliance within a household, based on its power consumption, requires an extensive data collection procedure.



(a) Kitchen environment setup CETH KRPIS smart home with selected devices of interest



(b) Kitchen environment setup CETH/ITI ground floor kitchen with selected devices of interest

Fig. 1: Data collection environments

As an initial step of this research, we focused on the power consumption of electrical appliances in a kitchen environment. From the first house (House A) (Fig. 1a), we collected data from the oven, the cooker hood, the dishwasher, the fridge and the main consumption (which includes the Heating, Ventilation, and Air Conditioning (HVAC), lights and other appliances) of the entire apartment. Regarding the second house (House B) (Fig. 1b), we collected data from one fridge, since our goal in that specific setup, was to check if it is possible to detect when a resident opens and closes the fridge door. In what follows, the infrastructure for data collection and the approach that was followed for pre-processing of the dataset are described.

A. Data Collection Infrastructure

Fig. 2 shows our data collection infrastructure. We installed a Gavazzi smart electricity meter in the oven and the main consumption panel of House A. The Gavazzi meter communicates with a Raspberry Pi via BACnet and then the Raspberry Pi sends the raw data to the InfluxDB database via a RESTful web service. We also measured the electricity consumption of selected devices (fridge, cooker hood, dishwasher and oven) via a wireless network of smart plugs that use the ZigBee protocol (<https://www.plugwise.com>). The installed smart plug modules communicate with each other forming a network of mesh topology. Furthermore, a special built-in module

was used in order to monitor the power consumption of the electrical kitchen appliance.

An aggregator application was developed and installed on a PC (MQTT Broker). It requested the current power consumption from each module for given time steps, received the corresponding messages, which include the measured energy consumption rate of the connected appliance in Watts, the time stamp (in UTC; later converted to local time), the ID of the device, and then stored the data directly into the database (InfluxDB).

For the second house (House B), we followed a similar procedure using the plugwise smart plugs that collected the energy consumption rate data for the specific device that we monitored (fridge).

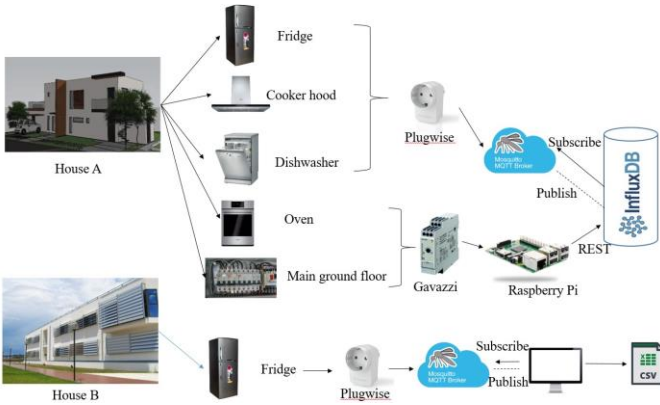


Fig. 2: Data collection infrastructure

B. Data Pre-Processing

After retrieving the raw data for House A, over a period of one month, a pre-processing step was performed in order to create the final aggregated dataset, which includes events per one-minute intervals of all the measured features. It is worth mentioning that due to technical issues with the smart plugs or the InfluxDB database, we had to overcome the sparsity of the raw data matrix. In order to solve this problem, we filled the missing values with the mode of the values of the last 15 minutes, until a new value was sent to the database. Regarding House B, we collected the electricity consumption of a fridge over a period of 10 days. The plugwise smart plug was sending data every 5 seconds, a time interval that was sufficient to detect whether someone opens and closes the door of the fridge.

The next step was to aggregate the features, consisting of energy consumption rate in Watts for each of the four appliances of interest (oven, fridge, dishwasher, cooker hood). Firstly, we had to round the time (index), since there was a delay of a few ms between the "subscription" and the "publish" of the event to the MQTT broker. Secondly, we manually labeled the dataset regarding the target feature or the state of operation (ON/OFF). The fridge was considered to

be always ON, even when the compressor was not operating. The rest of the devices were labeled as OFF (0) when the reading of the sensor was between 0 and 2.1348 W (a value around 2 W was considered as a 0 for all the appliances from the manufacturer) and ON (1) when the reading of the sensor was greater than 3 W. Hence, the dimensions of the overall dataset is 1440(minutes) \times 4(number of appliances) (for each day, without taking into account the target feature).

Fig. 3 shows indicative instances of the power consumption for the four appliances from House A. We noticed that we could detect a difference in power consumption regarding the LED state of the cooker hood (measured 4 W). In addition, after measuring the power consumption of the oven (Gavazzi meter sent data every 15 minutes), we could check if there are any "matching" times between the two devices, in order to infer the activity of cooking. Furthermore, the operation of the dishwasher was periodic and therefore quite trivial to infer the activity of washing the dishes. The most challenging appliance was the fridge, since our goal was to detect the appliance usage, in terms of door opening and closing events (based on the fridge light consumption). The fridge located in House A was a state of the art machine, in terms of energy efficiency and consequently it was not possible to detect when the resident opened and closed the door, even when we increased the data collection time to 20 s.

On the other hand, the fridge that we monitored located in House B was an older model than the one in House A. After sampling at 5 s, we noticed that we could detect when a resident opens and closes the door of the fridge (light turned on) only when the compressor was not operating (Fig. 4), since the consumption increased from 2 W to 6 W. Otherwise, it was not possible to detect any activity, since there was no difference in the power consumption.

C. Appliance State Proportion Feature

A key feature to the proposed activity decision engine is the appliance state proportion, which defines the probability of an appliance to be at ON stage on a pre-decided overlapping sliding window [22]. Assume again a set of M activities. In our approach, for each activity $i \in \{1, 2, \dots, M\}$, and each decision time t , a feature $F_t^{(i,j)}$ is calculated for each sensor j , as the proportion of time, or probability, that sensor j is activated at time t , that is:

$$F_t^{(i,j)} = \frac{T_{ON}^{(i,j)}}{T} \quad (1)$$

where $0 \leq T_{ON}^{(i,j)} \leq T$ is the total amount of time that sensor j is activated at time t , the latter having a total duration of T . In our work, T was equal to 2 minutes, as it was found to be sufficient for activity detection.

Fig. 5 depicts our proposed methodology for an energy sensor-based (smart plugs) decision engine for human activity recognition. Assume a set of N sensors (smart plugs) that provides the input data (energy consumption rate of N household appliances) for the N classifiers, where each classifier

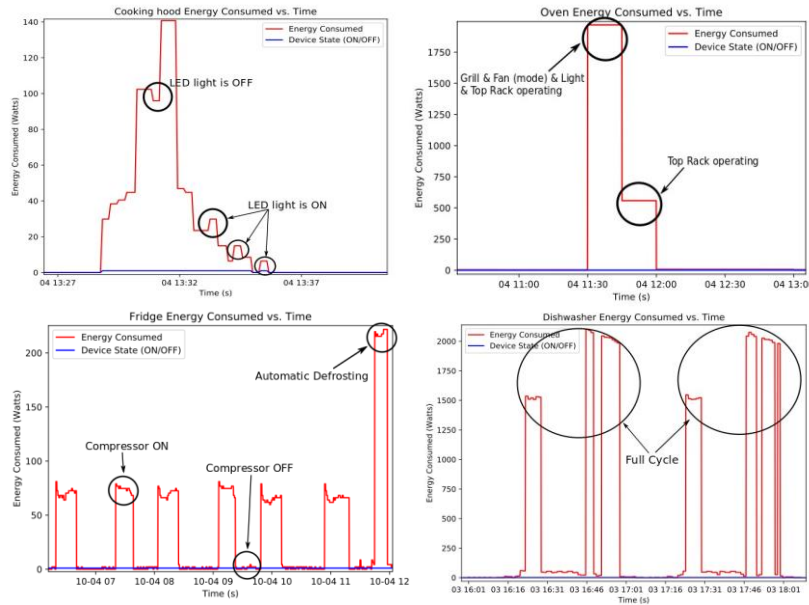
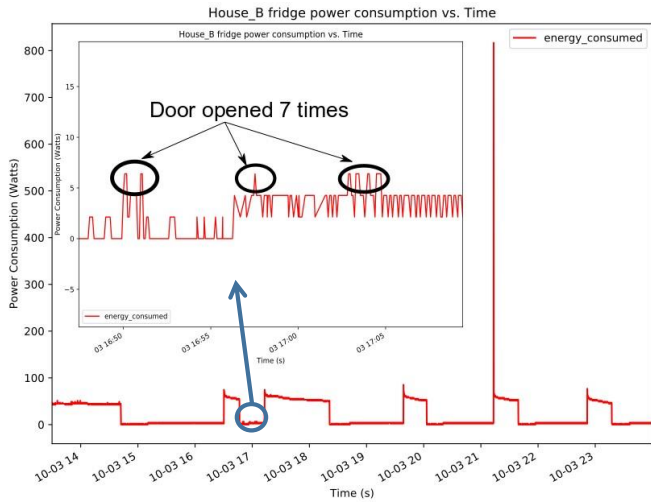


Fig. 3: Power consumption plots of the selected appliances

Fig. 4: Fridge power consumption from House B

is dedicated to a specified appliance and $M \geq N$ activities (an activity may relate with more than one appliance). The input data for each classifier is the energy consumption rate measurements of the whole set of appliances while the supervisory signal (used during training) is a vector of the specified appliance states (0 and 1, where 0 denotes the OFF state and 1 denotes the ON state of an appliance). The decision engine calculates the probability of presence of a specific activity using Rule-Based Scenarios (RBS) that get as inputs the appliance operating state from the classifier and the appliance state proportion, at the decision time t . An appliance was

considered to be active if the state proportion feature was above a threshold of 0.5. In total we collected 21,600 measurements over the 30 days (720 2-minute measurements per day).

We measured the energy consumption of the cooker, cooker hood, oven, washing machine, and dish washer to monitor the three following activities: cooking, washing the dishes, and washing clothes. Therefore for our particular example N was 4 (oven, cooker hood, dishwasher and washing machine) and M was 4 (cooking, washing clothes, washing dishes and doing nothing; when no activity of the aforementioned was performed). Furthermore, we applied a probability boost of 0.3 for the activity of cooking when the cooker hood is ON and a 0.7 for the activity of cooking when the oven is ON, since the state of cooker hood is not related directly with cooking.

A key-feature of the proposed decision engine is the use of an efficient and effective classification technique. To identify a suitable classification technique, we tested the following machine learning methods: Support Vector Machines (SVMs) with their basic kernels (Linear, Polynomial and Radial Basis Function (RBF)) [23], Decision Trees (DT) [24], Naive Bayes (NB) [25], Logistic Regression (LR) [26], Artificial

Neural Networks (ANNs) and specifically a Back-Propagation Network (BPN) [27]. Along with these standalone machine learning algorithms, the ensemble learning methods of Random Forest (RF) [28] and Gradient Boosting (GB) [29] were also tested for their predictive performance. We considered the aforementioned classifiers, since a simple thresholding would not be robust against noisy signals and we would lose significant information from the 2 minute-windows.

For the two-class classification scenario of a single household appliance (appliance status OFF/ON), in order to assess our models, the measures of precision, recall, accuracy and Matthews Correlation Coefficient (MCC) were used, which are computed from the contents of the confusion matrix of the

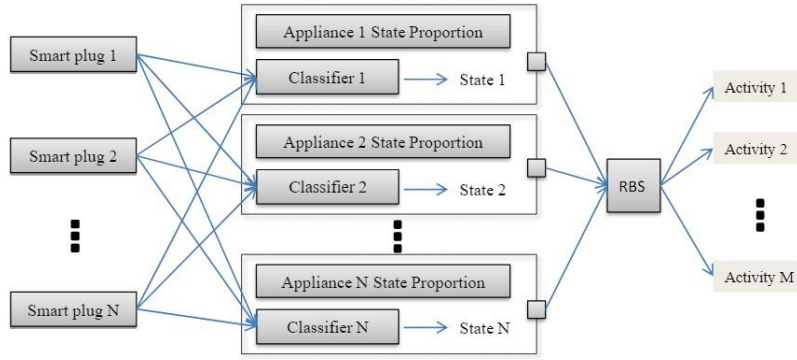


Fig. 5: Overview of energy sensor-based decision engine human activity detection

TABLE I: Accuracy of proposed decision engine

Activities	Number of correct predictions	Ground-truth in 2-minute intervals for 30 days	Accuracy (%)
Cooking	632	685	92.3
Washing Clothes	240	250	96.1
Washing the Dishes	920	965	95.4
Doing Nothing	19,621	19,700	99.6
<i>Average</i>	-	-	99.1

classification predictions. Precision is the ratio of predicted true positive cases to the sum of true positives and false positives, recall is the proportion of the true positive cases to the sum of true positives and false negatives and accuracy is the proportion of the total number of predictions that were correct.

Precision or recall alone cannot describe a classifier's efficiency, especially in cases where the labels in training target feature are not balanced. Therefore, MCC is used as balanced evaluation measure and specifically a correlation coefficient among the actual classification and predicted output of the classifier. It returns a value between -1 and +1, where +1 represents a perfect prediction, 0 random prediction and -1 indicates total disagreement between prediction and observation. MCC is calculated directly from the confusion matrix and is given by the equation:

$$MCC = \frac{TP*TN - FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (2)$$

where TP (True Positives), FP (False Positives), TN (True Negatives) and FN (False Negatives).

V. RESULTS

We generated 100 Monte Carlo iterations for different parameter scenarios in each classifier to eliminate the bias. For every iteration, we used a random sampling crossvalidation where the percentage of samples in the training and the testing datasets was 70% and 30%, respectively. For the SVM with polynomial kernel, the parameter, θ , which is a free parameter taking integer values, is assigned as: $\theta=(start=30,end=60,step=6)$ and the polynomial degree takes

the values $p=(start=2,end=7,step=1)$. We found that after the 4th degree, we overfitted the dataset. For the SVM with linear kernel, we used the default configuration of scikit-learn [30]. For the SVM with radial basis function kernel, σ varied same as θ and the constant C as $C=(start=100,end=1000,step=100)$. The parameter σ of the RBF kernel handles the non-linear classification. For the DT we used the default optimized version of the Classification and Regression Trees (CART) algorithm. For NB we used the Gaussian algorithm for classification and for LR we used the default configuration of scikit-learn. The BPN had a single hidden layer and the number of neurons varies as $n=(start=100,end=200,step=20)$. The RF and GB have an ensemble of $estimators=(start=20,end=100,step=20)$ DTs. The combination of all values of parameters and the size of 100 Monte Carlo iterations for each case, results in an overall of 11000 tested cases grid search.

Since more than one classifier achieved the best performance, we selected the SVM with polynomial kernel classifier and performed activity inference for random times. Table I summarizes the accuracies for the monitored activities (cooking, washing clothes, washing dishes) over 30 days. When an appliance was switched on, our decision engine was not able to instantly detect that it was operating and relate it to an activity. However, after 4 minutes of operation of that particular device it was able to predict the activity related to the operating appliance correctly and achieved an average accuracy of 99.1%. Furthermore, since the fridge required a

high sampling rate in order to determine when the door is open or closed, it is not a device that can be directly related to an activity, such as cooking. However, it can be used as a "supplementary" appliance to increase the confidence of the predicted activity.

VI. CONCLUSION

We presented a framework for human activity context inference, based on energy consumption rate from selected appliances. The results are very promising towards unobtrusive activity detection for ambient assisted living. While our experiments were done in a kitchen environment, our approach is flexible enough to be applied to other smart home environments. Since most activities within a house are related with the use of an electrical appliance, this unimodal approach has a significant advantage using inexpensive smart plugs and smart meters for each appliance.

Nevertheless, the main limitation of our approach is that it will not work with activities that are not related to electrical appliances, such as sleeping and taking a shower. Therefore there is a strong need for combining the smart meter sensors with acoustic, CO₂, occupancy sensors to increase the activity recognition performance. As future work, we plan to increase our dataset, increase the time for appliance monitoring, add more devices from other smart home environments and test the algorithm in new "untrained" home environments.

ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 676157, project ACROSSING and innovation actions grant agreement No. 723059, project enCOMPASS.

REFERENCES

- [1] L. Perez-Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," *Energy and buildings*, vol. 40, no. 3, pp. 394–398, 2008.
- [2] P. Rashidi and A. Mihailidis, "A survey on ambient-assisted living tools for older adults," *IEEE journal of biomedical and health informatics*, vol. 17, no. 3, pp. 579–590, 2013.
- [3] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SigKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.
- [4] C. Duarte, K. Van Den Wymelenberg, and C. Rieger, "Revealing occupancy patterns in an office building through the use of occupancy sensor data," *Energy and Buildings*, vol. 67, pp. 587–595, 2013.
- [5] A. Jalal, Y.-H. Kim, Y.-J. Kim, S. Kamal, and D. Kim, "Robust human activity recognition from depth video using spatiotemporal multi-fused features," *Pattern recognition*, vol. 61, pp. 295–308, 2017.
- [6] A. Vafeiadis, K. Votis, D. Giakoumis, D. Tzovaras, L. Chen, and R. Hamzaoui, "Audio-based event recognition system for smart homes," in *Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCOM/IoP/SmartWorld), 2017 Intl IEEE Conferences*. IEEE, 2017.

- [7] G. Bauer, K. Stockinger, and P. Lukowicz, "Recognizing the use-mode of kitchen appliances from their current consumption," *EuroSSC*, vol. 9, pp. 163–176, 2009.
- [8] E. Kim, S. Helal, and D. Cook, "Human activity recognition and pattern discovery," *IEEE Pervasive Computing*, vol. 9, no. 1, 2010.
- [9] E. Nazerfard, B. Das, L. B. Holder, and D. J. Cook, "Conditional random fields for activity recognition in smart environments," in *Proceedings of the 1st ACM International Health Informatics Symposium*. ACM, 2010, pp. 282–286.
- [10] D. Giakoumis, G. Stavropoulos, D. Kikidis, M. Vasileiadis, K. Votis, and D. Tzovaras, "Recognizing daily activities in realistic environments through depth-based user tracking and hidden conditional random fields for mci/ad support," in *ECCV Workshops (3)*, 2014, pp. 822–838.
- [11] L. Stankovic, V. Stankovic, J. Liao, and C. Wilson, "Measuring the energy intensity of domestic activities from smart meter data," *Applied Energy*, vol. 183, pp. 1565–1580, 2016.
- [12] A. Lavin and D. Klabjan, "Clustering time-series energy data from smart meters," *Energy efficiency*, vol. 8, no. 4, pp. 681–689, 2015.
- [13] P. Cottone, S. Gaglio, G. L. Re, and M. Ortolani, "User activity recognition for energy saving in smart homes," *Pervasive and Mobile Computing*, vol. 16, pp. 156–170, 2015.
- [14] S. Xu, E. Barbour, and M. C. Gonzalez, "Household segmentation by load shape and daily consumption," in *Proceedings of ACM SigKDD 2017 conference, Halifax, Nova Scotia, Canada*, 2017.
- [15] K. M. Rao, D. Ravichandran, and K. Mahesh, "Non-intrusive load monitoring and analytics for device prediction," in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, vol. 1, 2016.
- [16] A. Deshmukh and D. Lohan, "Cs446 project: Electric load identification using machine learning," Tech. Rep., 2015.
- [17] C. Belley, S. Gaboury, B. Bouchard, and A. Bouzouane, "An efficient and inexpensive method for activity recognition within a smart home based on load signatures of appliances," *Pervasive and Mobile Computing*, vol. 12, pp. 58–78, 2014.
- [18] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu, "Sensorbased activity recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 790–808, 2012.
- [19] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.
- [20] L. Chen, C. Nugent, and G. Okeyo, "An ontology-based hybrid approach to activity modeling for smart homes," *IEEE Transactions on human-machine systems*, vol. 44, no. 1, pp. 92–105, 2014.
- [21] L. Chen, C. D. Nugent, and H. Wang, "A knowledge-driven approach to activity recognition in smart homes," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 6, pp. 961–974, 2012.
- [22] G. Papamakarios, D. Giakoumis, M. Vasileiadis, A. Drosou, and D. Tzovaras, "16 human computer confluence in the smart home paradigm: Detecting human states and behaviours for 24/7 support of mild cognitive impairments," *Human Computer Confluence Transforming Human Experience Through Symbiotic Technologies*, pp. 275–293, 2016.
- [23] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*. ACM, 1992, pp. 144–152.
- [24] S. Jun Lee and K. Siau, "A review of data mining techniques," *Industrial Management & Data Systems*, vol. 101, no. 1, pp. 41–46, 2001.
- [25] I. Rish, "An empirical study of the naive bayes classifier," in *IJCAI 2001 workshop on empirical methods in artificial intelligence*, vol. 3, no. 22. IBM, 2001, pp. 41–46.
- [26] H.-F. Yu, F.-L. Huang, and C.-J. Lin, "Dual coordinate descent methods for logistic regression and maximum entropy models," *Machine Learning*, vol. 85, no. 1, pp. 41–75, 2011.
- [27] D. E. Rumelhart, G. E. Hinton, R. J. Williams et al., "Learning representations by back-propagating errors," *Cognitive modeling*, vol. 5, no. 3, p. 1, 1988.

- [28] A. Liaw, M. Wiener *et al.*, "Classification and regression by randomforest," *R news*, vol. 2, no. 3, pp. 18–22, 2002.
- [29] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.
- [30] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.