Support Vector Machine for Classification of Households' Heating Type using Load Curves

SUPPORT VECTOR MACHINE FOR CLASSIFICATION OF HOUSEHOLDS' HEATING TYPE USING LOAD CURVES

Kristoffer FÜRST
Chalmers University of Technology
Sweden
kristoffer.furst@chalmers.se

Peiyuan CHEN
Chalmers University of Technology
Sweden
peiyuan@chalmers.se

Irene Yu-Hua GU
Chalmers University of Technology
Sweden
irenegu@chalmers.se

ABSTRACT

The distribution system operator lacks the knowledge of the heating system used by their customers to make sound grid planning decisions. Energy declaration from buildings and the large-scale rollout of smart meters provides an excellent opportunity to classify the heating system used. This paper proposes a machine-learning-based approach using a support vector machine (SVM) with daily load curves (mean and standard deviation of consumption) extracted from smart meter measurements. Three heating types are analysed: district heating, exhaust air heat pump, and direct electric heating. The performance was compared among the classifiers using daily load curves extracted over one year, for each month, each week, and each day of the year. The highest average accuracy of 92.6% was obtained for the SVM classifier using daily load curves extracted for each week of a year as features. Furthermore, the classifier showed a higher performance than using an ensemble of SVM or random forest classifiers (90.6%/90.5%) proposed in the literature. Lastly, an error analysis of the misclassification was carried out, including building characteristics and geographical analysis.

INTRODUCTION

Distribution system operator (DSO) needs to know their customers' heating types when making grid planning decisions. In 2018, about 50% of the heating in detached houses in Sweden came from electricity [1]. However, there is no obligation for electricity consumers to notify the grid operator of energy efficiency measures or the type of heating system used. It is difficult to make sound grid planning if consumers change their heating types without notifying their DSO. On the other hand, the large-scale rollout of smart meters for consumers, with publicly available data, provides a great opportunity to characterise the heating types of end-users. This information can be used by the DSO to make better grid planning decisions.

Different approaches have been used to classify electrical consumers including unsupervised [2] - [4] and supervised machine learning approaches [5] - [8]. The unsupervised methods cluster different load patterns. However, to classify new consumers' heating types, a post-analysis and interpretation are required to identify the characteristic of the different clusters. Classification by supervised machine learning on the other hand does not require post-analysis and interpretation. Authors [5] and [6] used a support vector machine (SVM) classifier to classify different properties of electricity consumers using a corresponding set of features extracted from smart meter measurements. The feature set in [6] was further extended in [7] and [8], evaluating multiple classifiers including random forest (RF) and SVM. SVM was found to be one of the best-performing classifiers among supervised classifiers for heat pump detection [8]. Though the authors in [7] and [8] used an extensive set of features extracted from smart meter measurements (91 features per week), the relationship between different months was not considered in a single classifier. Instead, the authors in [7] proposed an ensemble classifier, combining the result from a classifier of each week. Furthermore, [6] - [8] collected consumer class from survey data, which is usually time-consuming.

This paper aims to develop a new SVM classifier with daily load curves as features for classifying household heating types by using smart meter measurements and publicly available building data. The main contribution of this paper includes:

- identifying residential heating types using SVM together with daily load curves extracted from smart meter measurements,
- comparing classification performance using daily load curves for a year, each month, each week, and each day.

The performance of the proposed SVM classifier is evaluated through testing using smart meter measurement data collected from households in a city in Sweden. Analyses of misclassification were also performed.

CLASSIFICATION MODEL FRAMEWORK

The proposed framework seen in Fig. 1 uses a supervised machine learning classifier for the heating type of electricity consumers. The proposed method uses a soft-margin SVM with load curves extracted from smart meter measurements as feature input. Further details are described in the following subsections.

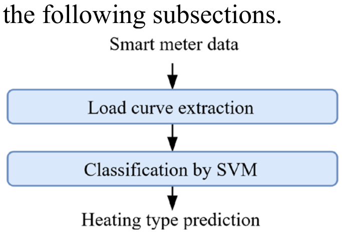

Fig. 1: Framework of consumer classification

Daily load curve extraction

Typical daily load curves are often used to describe and cluster the characteristics of electricity consumers, as in [2] - [4]. In [9], a daily load curve is defined by the average and standard deviation of electricity. These curves can be obtained for different seasons, temperature intervals, day of the week, etc. Fig. 2 shows an example of daily load curves extracted for each month of a year, including both the average value (top) and standard deviation (bottom) of electricity consumption. Consumers with three different heating types are plotted: district heating (DH), exhaust air heat pump (EAHP) and direct electric heating (DEH). It can be seen that the averaged daily load curve differs between the DH and the electric-based heating types (EAHP and DEH) in both the magnitude and shape within the day, as well as the magnitude within the year. EAHP and DEH also shows similarities when the load (and its variation) elevates during the winter period.

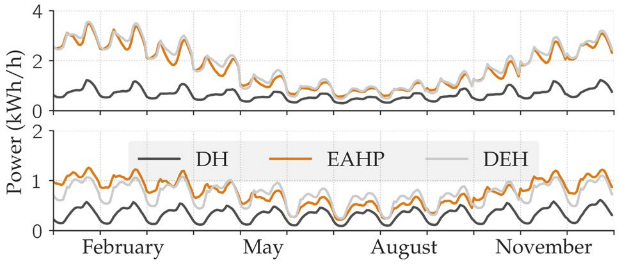

Fig. 2: Average daily load curves extracted for each month of the year. Top: average consumption; Bottom– standard deviation

This paper will evaluate the classification performance based on daily load curves. Note that the load curves are treated as static features by the classifier. That is, the time sequence of the time-dependent features is not considered.

Overview of support vector machine

SVM is a binary classifier (two classes) that often shows good performance in various classification tasks [10]. The aim is to find a hyperplane that separates the two classes with the highest margin. A soft-margin SVM allows *training* samples to violate this margin. These are penalised by $C \cdot \xi_i$, where C is a constant (specified before training) and ξ_i a slack variable for the sample $\boldsymbol{x}_i$. The value of ξ_i increases as the sample are further away from the “right side” of the hyperplane. Samples that do not violate the margin are not penalised, in other words, $\xi_i = 0$. Allowing errors in the training set makes it more robust against individual training samples. The objective function of the SVM can be defined as [11]:

$$\min_{w,b} \frac{1}{2} \|\boldsymbol{w}\|^2 + C \sum_{i=1}^{N} \xi_i \tag{1}$$

subject to:

$$y_i \cdot (\boldsymbol{w} \cdot \boldsymbol{x}_i + b) \geq 1 - \xi_i \tag{2}$$

for $i = 1, \ldots, N$ and $\xi_i \geq 0$. N is the number of training samples and $\boldsymbol{w}$ is the normal vector to the hyperplane. $\boldsymbol{w} \cdot \boldsymbol{x}_i + b = 0$ is the hyperplane separating the two classes, where b is a constant and:

$$y_i = \begin{cases} +1 \textbf{ if } \text{class } j \\ -1 \textbf{ otherwise} \end{cases} \tag{3}$$

By extension, a radial basis function (RBF) kernel is used to deal with non-linear separable classes, with the feature vector $\boldsymbol{x}_i$ mapped from the feature space to a higher-dimensional space. The SVM seeks a hyperplane that best separates the classes in that higher-dimensional space. The kernel is a Gaussian function and is given as [11], [12]:

$$K(\boldsymbol{x}, \boldsymbol{x}') = e^{-\gamma \|x - x'\|^2} \tag{4}$$

where $\|\boldsymbol{x} - \boldsymbol{x}'\|$ is the Euclidean distance between two points, and γ the hyperparameter to tune. The greater the value of γ, the closer the points need to be deemed similar. This results in a more complex decision boundary, which could lead to overfitting. Conversely, a low value of γ results in a less complex decision boundary, which could lead to underfitting of the model.

Training and classification

SVM with multi-class classification

To solve a multi-class classification problem, we use a one-vs-rest classifier, whereby each class in M is compared with the remaining $M - 1$ classes, resulting in M binary classifier if more than two classes. An unseen sample is assigned to the class with the maximum number of votes from the M binary classifiers [12].

Imbalanced classes

To balance the importance between the multiple classes, a weighting factor inversely proportional to the number of samples N_k of a given class k was used,

$$w_k = \frac{N}{M N_k}, \tag{5}$$

where N the total number of samples. The minimisation in (1) is instead [12]:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \cdot \sum_{y_i=1} w_k \xi_i \,, \tag{6}$$

Classification process

Once the decision boundary is found using (6) and all training samples, the trained classifier can be used to test-run. That is, classifying consumers’ heating type based on data that have not been used to develop the classifier.

RESULTS

Setup of experiments

Data description

The dataset consists of hourly smart meter measurements collected from households in a Swedish city over one year (2017). The labels (heating type) and heated area (for error analysis) were collected from the building’s energy declaration [13]. One and two-family households with only one smart meter and with one of the three common heating types are considered: DH, EAHP and DEH.

Pre-processing

Missing values were linearly interpolated if only two successive values were missing. Consumers with missing data after the interpolation were filtered out as it is out of the scope to approximate the smart meter measurements. Furthermore, due to a change of sampling frequency (faulty communication), consumers with smart meter measurements with less than 0.1 kWh/h difference between two adjacent hours for a rolling window of 20 hours were also removed. Lastly, both the training and test set were normalised using the mean m^l and standard deviation s^l of each feature l in the *training set*. The normalised sample $\tilde{x}_i^l$ were obtained by:

$$\tilde{x}_i^l = \frac{x_i^l - m^l}{s^l} \quad (7)$$

Dataset partition and cross-validation

A fivefold cross-validation was used to evaluate the performance. The data were split randomly into five equal-sized and non-overlapping folds according to the number of consumers/households. That is, data sequences from individual consumers were used only in training or testing, with 80% used for training and 20% used for testing. Each fold was used as the test set only once, with the remaining four folds used as the training set. The classifier was re-trained using the corresponding training set and optimised hyperparameters. Table 1 summarises the size of the training and test set for the different heating types studied.

Table 1: Number of individual customers exclusively used in each training/test set fold partition.

Heating type	Training	Testing	Total
DH	1286 (41%)	322 (10%)	1608 (51%)
EAHP	426 (14%)	107 (3%)	533 (17%)
DEH	793 (26%)	198 (6%)	991 (32%)
Total	2505 (80%)	627 (20%)	3132 (100%)

Hyperparameter tuning

A grid search of the hyperparameters was performed, with $C \in \{10^0, 10^1, 10^2, 10^3, 10^4, 10^5\}$ and $\gamma \in \{10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}\}$. Using a two-fold cross-validation on the training set, the parameter combination with the highest accuracy was selected for the final classifier.

Criteria for performance evaluation

The total accuracy, and the precision, recall and F1-score of each heating type c were used to evaluate the performance of the model, defined as:

$$\text{Accuracy} = \frac{\sum_{c=1}^{C} TP_c}{\sum_{c=1}^{C} TP_c + FP_c} \quad (8)$$

$$\text{Precision}_c = \frac{TP_c}{TP_c + FP_c} \quad (9)$$

$$\text{Recall}_c = \frac{TP_c}{TP_c + FN_c} \quad (10)$$

$$\text{F}_{1,c} = 2\frac{\text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c} \quad (11)$$

where C is the total number of classes, TP_c the true positive, FP_c the false positive, and FN_c the false negative for class c.

Results and discussion

Overall performance

The effectiveness of the proposed model was evaluated by performing fivefold cross-validation. Table 2 compares the performance of SVM with daily load curves extracted over a year, for each month, each week and each day of the year, respectively, resulting in 48 features (24 hours mean and standard deviation values), 576 features (48×12 months), 2496 features (48×52 weeks), and 8760 features (24×365 days). Note that daily load curves extracted over each day are the same as the original hourly smart meter measurement data but normalized according to (7). Comparing the results, Table 2 shows that increasing the number of features increases the performance of the classifier, with daily load curves for each week showing the best performance (92.6% accuracy). Though using the normalised hourly smart meter measurement directly as features showed a lower accuracy (91.8%), the SVM still showed a high performance, even though the number of features surpasses the number of training samples (risk for overfitting), and the data was non-stationary. It was noted that tuning the hyperparameters was essential, whereas poorly tuned hyperparameters showed a substantially reduced performance as the number of features increased.

For all models, it is seen that the model had more difficulties distinguishing the electricity-based heating types than DH. The higher performance of the DH class was expected as the average consumption (see Fig. 2) was in general lower than the electricity-based heating sources. Furthermore, the high F1-score of DH (high precision *and* recall) shows that the models had higher difficulties distinguishing EAHP from DEH. As also seen in Fig. 2, the average profile of the electricity-based heating sources showed similar characteristics.

Comparing to existing classifiers

Table 2: Performance of the fivefold cross-validation using a daily load curve extracted over a year, for each month, each week, and each day of the year, without weekday differentiation. All performance values are averaged ± standard deviation.

Daily load curve	Precision (%)			Recall (%)			F1-score			Accuracy
	DH	EAHP	DEH	DH	EAHP	DEH	DH	EAHP	DEH	Total
Over a year	97.8 ± 0.6	55.1 ± 3.2	79.5 ± 1.0	92.5 ± 1.3	70.7 ± 1.5	74.3 ± 3.7	95.1 ± 0.6	61.9 ± 2.1	76.8 ± 2.1	83.0 ± 0.9
Over each month	98.3 ± 0.6	79.2 ± 2.1	90.3 ± 2.0	94.3 ± 0.2	86.1 ± 3.4	91.9 ± 2.2	96.3 ± 0.3	82.5 ± 0.8	91.1 ± 1.1	92.1 ± 0.4
Over each week	98.5 ± 0.7	83.0 ± 1.4	88.7 ± 1.6	94.9 ± 0.7	83.3 ± 2.0	93.7 ± 1.4	96.7 ± 0.6	83.1 ± 0.6	91.1 ± 0.8	92.6 ± 0.3
Over each day	98.5 ± 1.0	85.5 ± 3.0	85.3 ± 1.3	95.1 ± 0.9	77.3 ± 4.2	94.2 ± 1.5	96.8 ± 0.5	81.1 ± 1.4	89.6 ± 0.6	91.8 ± 0.3

The proposed classifier with daily load curves was compared to the classifier proposed in [7] and [8], using *only* smart meter measurements. The work in [7] further improved household classification using smart meter measurements as proposed in [5] and [6]. Note that the authors evaluated multiple machine learning algorithms, however, for this comparison, we only consider SVM and RF which showed the highest performance in [8]. Furthermore, a grid search using a two-fold cross-validation was used where the best-performing hyperparameters were selected. Searching grid was $C \in \{10^0, 10^1, 10^2, 10^3, 10^4, 10^5\}$ and $\gamma \in \{10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}\}$ for SVM and for RF the minimum number of samples required to be a leaf node $\{2,4,8,16,32\}$. Table 3 summarises the performance of the test set. The results show that the proposed classifier using a daily load curve for each week of the year results in a better average test accuracy (92.6%) than the SVM and RF developed in [7] (90.6%/90.5%). Even though the precision *or* recall was lower for some heating types of the proposed classifier, the F1-score was the highest for all three heating types.

Model and error analysis

An important step in model development is analysing classification errors. This is done to better understand the misclassifications and see if there are any errors/biases.

Building characteristics

The classification error was evaluated by assessing the heated area and age of the building and whether there are systematic patterns between the correctly and incorrectly classified consumers, see Fig. 3. It is seen that for DH, the average heated area (left figure) was greater for the wrongly classified consumers than for the correctly classified ones. This could indicate that using the heated area in the modelling could further increase the performance of the model. However, the difference between the correctly and incorrectly classified consumers was small for exhaust air heat pumps and direct electric heating. Further studies are needed to analyse the background of this pattern.

Furthermore, Fig. 3 shows a major difference in building age (right figure) *between* the different *heating types*; EAHP were normally found in newer buildings; DEH as the only heating source was found mainly in houses from the 70s; DH was seen in both newer and older houses. The age of the building could be included in the model. However, there is a risk that the classifier will show a bias, such as classifying a consumer with an older house with an EAHP installed as DEH. Instead, the age of a building may be seen as an indication of whether there is an increased probability of changing the heating system. Nonetheless, it is seen that the building of correct classified samples was on average slightly newer. Further analysis should be devoted to how to incorporate the age of the building into the model without inducing bias.

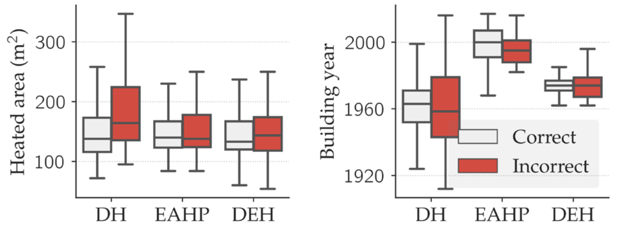

Fig. 3: Classification performance as a function of left) heated area and right) age of the building. The boxplot was based on the test classification from all folds.

Geographical analysis

The geographical analysis of the misclassifications in Fig. 4. shows that some areas were over-represented with misclassifications. This could indicate a geographical bias, which could be further investigated. For instance, it was found out that in one area with a high misclassification rate, the consumers had collectively changed their heating system from DH to EAHP. On one hand, the example shows that the energy declaration may be outdated since it was issued. This can cause errors in the training process and especially in the evaluation of the model. On the other hand, it also shows that the classification model can indicate whether a consumer's heating system has changed. This information is particularly useful to DSOs in their grid-planning process. Further discussions can be found in the report in [14].

CONCLUSION

The proposed SVM framework has successfully classified the heating type of electricity consumers using smart meter measurements only. Households with one out of three common heating types were evaluated, including DH, EAHP and DEH. The proposed SVM using daily load curve for each week of the year as features showed the best

Table 3: Comparing to existing classifier by fivefold cross-validation. All performance values are averaged ± standard deviation.

Method	Precision (%)			Recall (%)			F1-score			Accuracy
	DH	EAHP	DEH	DH	EAHP	DEH	DH	EAHP	DEH	Total
Proposed SVM with daily load curve over each week	98.5 ± 0.7	83.0 ± 1.4	88.7 ± 1.6	94.9 ± 0.7	83.3 ± 2.0	93.7 ± 1.4	96.7 ± 0.6	83.1 ± 0.6	91.1 ± 0.8	92.6 ± 0.3
SVM with ensemble classifier [7]*	95.1 ± 1.2	93.8 ± 3.2	83.3 ± 1.6	96.4 ± 0.7	64.2 ± 1.5	95.6 ± 0.4	95.7 ± 0.9	76.2 ± 1.9	89.0 ± 1.0	90.6 ± 0.5
RF with ensemble classifier [7]*	96.3 ± 0.4	91.0 ± 3.8	82.5 ± 1.3	94.8 ± 0.8	67.7 ± 4.7	95.9 ± 0.7	95.5 ± 0.5	77.6 ± 3.9	88.7 ± 0.6	90.5 ± 0.5

* Only using feature extracted from smart meter measurements, not including weather-based features

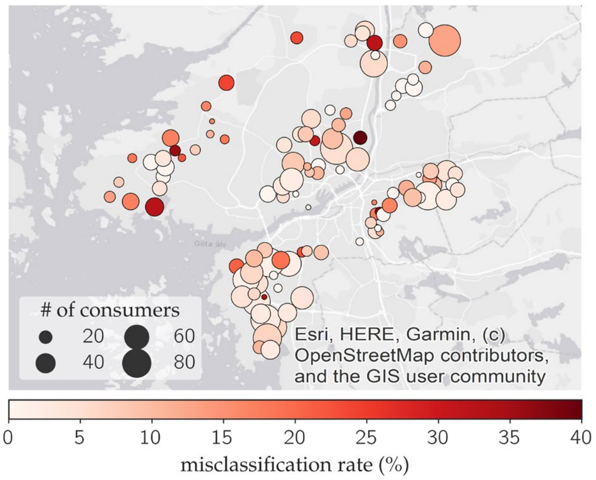

Fig. 4: Misclassification rate per zip code, based on the test classification from all folds. (This map was created using ArcGIS® software by Esri. ArcGIS® and ArcMap™ are the intellectual property of Esri and are used herein under license. Copyright © Esri. All rights reserved. For more information about Esri® software, please visit www.esri.com)

performance, with a classification accuracy of 92.6%. When the normalised hourly smart meter measurements were used directly as features, the performance was still good with an accuracy of 91.8%. Furthermore, compared to existing classifiers using an ensemble classifier with SVM and RF, the proposed method showed a higher performance (compared to 90.6% and 90.5%). This shows that including the load curve in *one* classifier can achieve higher performance. Moreover, the posterior analysis, using publicly available building data, showed some key findings that can be used to further investigate misclassified samples to improve the classification accuracy. First, misclassified consumers with DH tended to have a larger heated area than the correctly classified ones. Second, a geographical analysis of the misclassified consumers showed a good indication of whether there was a geographical bias. This is valuable information for the distribution system operator during their long-term grid planning and should be further investigated.

ACKNOWLEDGEMENT

The authors would like to acknowledge the Energiforsk reference group and the financial support from Energiforsk [EVU10429] and The Swedish Energy Agency [50237-1].

REFERENCES

[1] Swedish Energy Agency, “Energy statistics for one- and two dwelling buildings 2018,” 2019. [Online]. Available: https://www.energimyndigheten.se/statistik/. [Accessed 11 06 2021].

[2] P. Westermann, D. Chirag, A. Schlueter and R. Evins, “Unsupervised learning of energy signatures to identify the heating system and building type using smart meter data,” *Applied Energy,* vol. 264, 2020.

[3] Z. Jiang, R. Lin and F. Yang, “A Hybrid Machine Learning Model for Electricity Consumer Categorization Using Smart Meter Data,” *Energies,* vol. 11, p. 2235, 2018.

[4] C. Sandels and J. Widén, “End-User Scenarios and Their Impact on Distribution System Operators - A techno-economic analysis,” Energiforsk, Stockholm, 2018.

[5] H. Fei, Y. Kim, S. Sahu, M. Naphade, S. K. Mamidipalli and J. Hutchinson, “Heat Pump Detection from Coarse Grained Smart Meter Data with Positive and Unlabeled Learning,” in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2013.

[6] C. Beckel, L. Sadamori, T. Staake and S. Santini, “Revealing household characteristics from smart meter data,” vol. 78, pp. 397-410, 2014.

[7] K. Hopf, M. Sodenkamp and T. Staake, “Enhancing energy efficiency in the residential sector with smart meter data analytics,” *Electronic Markets,* vol. 28, no. 4, pp. 453-473, 2018.

[8] A. Weigert, K. Hopf, N. Weinig and T. Staake, “Detection of heat pumps from smart meter and open data,” *Energy Informatics,* vol. 3, pp. 1-14, 2020.

[9] Svenska Elverksföreningen, “Belastningsberäkning med typkurvor,” Svenska Elverksföreningen, Stockholm, 1991.

[10] M. Fernandez-Delgado, C. E. S. Barro and D. Amorim, “Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?,” *Journal of Machine Learning Research,* vol. 15, pp. 3133-3181, 2014.

[11] C. Corinna and V. Vapnik, “Support-vector networks,” *Machine Learning,* vol. 20, p. 273–297, 1995.

[12] C.-C. Chang and C.-J. Lin, “LIBSVM: A library for support vector machines,” 2019.

[13] Boverket, “Sök och hämta energideklaration,” 01 01 2020. [Online]. Available: https://www.boverket.se/sv/energideklaration/sok-energideklaration/. [Accessed 4 04 2020].

[14] K. Fürst, “Electricity consumer classification using supervised machine learning,” Energiforsk, Stockholm, 2021.