Research Article

Estimation Of Power Consumption For Household Electric Appliances

Pacha Shobha Rani[1], K. Balasaranya[2], Manipriya.G[3],Niveditha.V[4],Swathi.B[5]

[1]Associate Professor,Department of CSE, R.M.D. Engineering College
[2] Assistant Professor,Department of CSE, R.M.D. Engineering College
[3]Programmer Analyst Trainee, Cognizant Technology Solutions India Pvt.Ltd.
[4]Assistant System Engineer Trainee, Tata Consultancy Service
[5]Trainee, Johnson Control Limited
[1]psr.cse@rmd.ac.in, [2]balasaranya1701@gmail.com, [3]manipriyagopal1999@gmail.com,
[4]nivedithavijayakumar0@gmail.com,[5]swathibabu1110@gmail.com.

Article History Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: A non-nosy checking framework assesses the conduct of individual electricapparatuses from the estimation of the absolute family unit load request bend. The all-out burden request bend is estimated at the passageway of the electrical cable into the house. The force utilization of individual apparatuses can be assessed utilizing a few AI procedures by investigating the trademark recurrence substance from the heap bend of the family unit. We have just built up the observing arrangement of ON/OFF states. This framework could build up adequate precision. In the following stage, the observing framework ought to have the option to appraise the force utilization for a climate control system with an inverter circuit. In this paper, we present aftereffects of applying a few relapse strategies, for example, multilayered perceptron's (MLP), spiral premise work systems (RBFN) and bolster vector regressors (SVR) to gauge the force utilization of a climate control system. Our trials show that RBFN can accomplish the best exactness for the non-meddling checking framework.

KEY WORDS –Anomaly detection, power monitoring, support vector machine, semi-supervised learning.

1. Introduction

The improvement of a checking framework, which can screen every family unit electric machine, is critical to obtain essential data for building the strategy of the vitality preservation, the figure of electric vitality request to design constructing new force plants, furthermore,making better new client administrations. It is normal that the observing framework for family electric apparatuses is economical and non-meddlesome, on the grounds that the ordinary checking strategy which isset up estimating types of gear for every machine is costly and powersthe clients burden. The nonintrusive observing methods the estimationof intensity load is done in the outside of a house.We have built up a non-meddlesome observing framework, which canassess the ON/OFF states of family unit electric apparatuses incorporateinverter.For the framework clients, the framework can accomplish an adequate exactness utilizing enormous edge classifiers. Consequently,the framework can gauge the force utilization of traditional electricmachines, which show a consistent power utilization, in the event thatwe know the force utilization of every family unit electric machine. Bethat as it may, an electric apparatus with an inverter circuit (inverter typemachine, for example, a forced air system doesn't show a consistentforce utilization. In this way, we need to build up the non-meddlesomeobserving framework, which can gauge the force utilization of invertertype machines.

2. Literature Survey

A non-intrusive monitoring system estimates the behavior of individual electric appliances from the measurement of the total household load demand curve. The total load demand curve is measured at the entrance of the power line into the house. The power consumption of individual appliances can be estimated using several machine learning techniques by analyzing the characteristic frequency contents from the load curve of the household. We have already developed the monitoring system of ON/OFF states. This system could establish sufficient accuracy. In the next phase, the monitoring system should be able to estimate the power consumption for an air conditioner with an inverter circuit. In this paper, we present results of applying several regression methods such as multilayered perceptrons (MLP), radial basis function networks (RBFN) and support vector regressors (SVR) to estimate the power consumption of an air conditioner. Our experiments show that RBFN can achieve the best accuracy for the non-intrusive monitoring system.

3. Existing System

we compare with several regression algorithms to estimate the power consumptionof an air conditioner as follows: multi-layered perceptrons (MLP), radial basisfunction networks (RBFN) and support vector regression (SVR). MLP is aconventional nonlinear regression method. So it is a good basis to compare withthe other methods. In RBFN, each basis function, which is able to have differentparameters, acts like a hidden unit of MLP. SVR has an advantage in highdimensional space using kernel functions. Therefore, we expect that RBFN and SVR have better results than MLP.

The disadvantage of existing system includes:
We only did with non-intrusive monitoring system but not for monitoringsystem.

1. PROPOSED SYSTEM

Inthis we try to improve the accuracy while compared to the previous one and here we try to find the sufficientaccuracy for monitoring System which can estimate the power consumption of electric appliances, to actual fields.

The advantage of this proposed system is in this we find the accuracy for monitoring System.

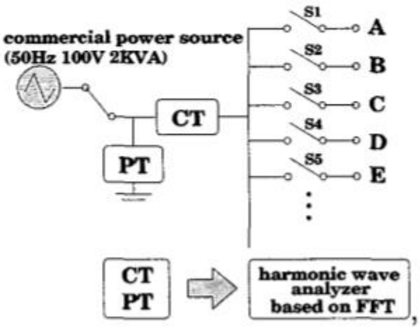


Figure:Block Diagram

Algorithm:Random Forest Algorithm:

Random forest is a type of supervised machine learning algorithm based on ensemble learning. Ensemble learning is a type of learning where you join different types of algorithms or same algorithm multiple times to form a more powerful prediction model. The random forest algorithm combines multiple algorithm of the same type i.e. multiple decision *trees*, resulting in a *forest of trees*, hence the name "Random Forest". The random forest algorithm can be used for both regression and classification tasks.

HOW RANDOM FOREST WORKS

The following are the basic steps involved in performing the random forest algorithm

1. Pick N random records from the dataset.
2. Build a decision tree based on these N records.
3. Choose the number of trees you want in your algorithm and repeat steps 1 and 2.
4. For classification problem, each tree in the forest predicts the category to which the new record belongs Finally, the new record is assigned to the category that wins the majority vote.

4. Domain specification

MACHINE LEARNING

Machine Learning is a system that can learn from example through self-improvement and without being explicitly coded by programmer. The breakthrough comes with the idea that a machine can singularly learn from the data (i.e., example) to produce accurate results.Machine learning combines data with statistical tools to predict an output. This output is then used by corporate to makes actionable insights. Machine learning is closely related to data mining and Bayesian predictive modeling. The machine receives data as input, use an algorithm to formulate answers.A typical machine learning tasks are to provide a recommendation. For those who have a Netflix account, all recommendations of movies or series are based on the user's historical data. Tech companies are using unsupervised learning to improve the user experience with personalizing recommendation.

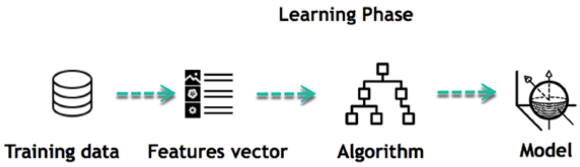
Machine learning is also used for a variety of task like fraud detection, predictive maintenance, portfolio optimization, automatize task and so on.

MACHINE LEARNING VERSUS TRADITIONAL LEARNING

Traditional programming differs significantly from machine learning. In traditional programming, a programmer code all the rules in consultation with an expert in the industry for which software is being developed. Each rule is based on a logical foundation; the machine will execute an output following the logical statement. When the system grows complex, more rules need to be written. It can quickly become unsustainable to maintain

WORKING OF MACHINE LEARNING

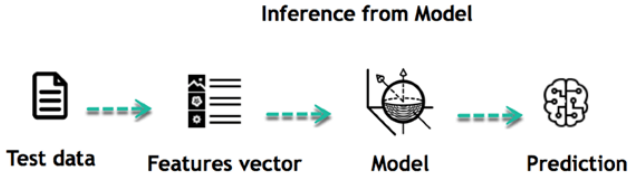
Machine learning is the brain where all the learning takes place. The way the machine learns is similar to the human being. Humans learn from experience. The more we know, the more easily we can predict. By analogy, when we face an unknown situation, the likelihood of success is lower than the known situation. Machines are trained the same. To make an accurate prediction, the machine sees an example. When we give the machine a similar example, it can figure out the outcome. However, like a human, if its feed a previously unseen example, the machine has difficulties to predict.The core objective of machine learning is the **learning** and **inference**. First of all, the machine learns through the discovery of patterns. This discovery is made thanks to the **data**. One crucial part of the data scientist is to choose carefully which data to provide to the machine. The list of attributes used to solve a problem is called a **feature vector.** You can think of a feature vector as a subset of data that is used to tackle a problem.The machine uses some fancy algorithms to simplify the reality and transform this discovery into a **model**. Therefore, the learning stage is used to describe the data and summarize it into a model.



For instance, the machine is trying to understand the relationship between the wage of an individual and the likelihood to go to a fancy restaurant. It turns out the machine finds a positive relationship between wage and going to a high-end restaurant: This is the model.

Inferring

When the model is built, it is possible to test how powerful it is on never-seen-before data. The new data are transformed into a features vector, go through the model and give a prediction. This is all the beautiful part of machine learning. There is no need to update the rules or train again the model. You can use the model previously trained to make inference on new data.

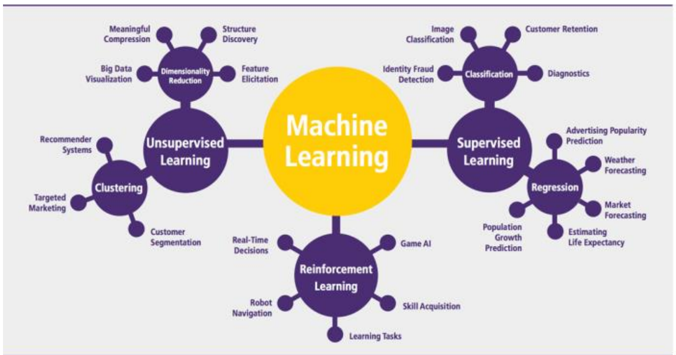


The life of Machine Learning programs is straightforward and can be summarized in the following points:

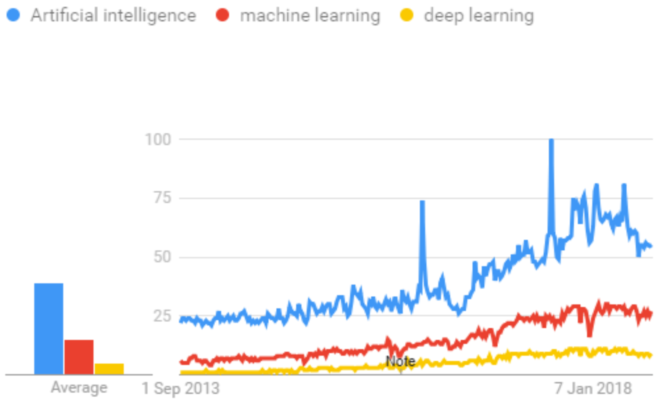
1. Define a question
2. Collect data
3. Visualize data
4. Train algorithm
5. Test the Algorithm
6. Collect feedback
7. Refine the algorithm
8. Loop 4-7 until the results are satisfying

Use the model to make a prediction

Once the algorithm gets good at drawing the right conclusions, it applies that knowledge to new sets of data.



With machine learning, you need fewer data to train the algorithm than deep learning. Deep learning requires an extensive and diverse set of data to identify the underlying structure. Besides, machine learning provides a faster-trained model. Most advanced deep learning architecture can take days to a week to train. The advantage of deep learning over machine learning is it is highly accurate. You do not need to understand what features are the best representation of the data; the neural network learned how to select critical features. In machine learning, you need to choose for yourself what features to include in the model.



Supervised learning

An algorithm uses training data and feedback from humans to learn the relationship of given inputs to a given output. For instance, a practitioner can use marketing expense and weather forecast as input data to predict the sales of cans.You can use supervised learning when the output data is known. The algorithm will predict new data.

There are two categories of supervised learning:

- Classification task
- Regression task

Classification

Imagine you want to predict the gender of a customer for a commercial. You will start gathering data on the height, weight, job, salary, purchasing basket, etc. from your customer database. You know the gender of each of your customer, it can only be male or female. The objective of the classifier will be to assign a probability of being a male or a female (i.e., the label) based on the information (i.e., features you have collected). When the model learned how to recognize male or female, you can use new data to make a prediction. For instance, you just got new information from an unknown customer, and you want to know if it is a male or female. If the classifier predicts male = 70%, it means the algorithm is sure at 70% that this customer is a male, and 30% it is a female.The label can be of two or more classes. The above example has only two classes, but if a classifier needs to predict object, it has dozens of classes (e.g., glass, table, shoes, etc. each object represents a class)

Regression

When the output is a continuous value, the task is a regression. For instance, a financial analyst may need to forecast the value of a stock based on a range of feature like equity, previous stock performances, macroeconomics index. The system will be trained to estimate the price of the stocks with the lowest possible error.

Algorithm Name	**Description**	**Type**
Linear regression	Finds a way to correlate each feature to the output to help predict future values.	Regression
Logistic regression	Extension of linear regression that's used for classification tasks. The output variable 3is binary (e.g., only black or white) rather than continuous (e.g., an infinite list of potential colors)	Classification
Decision tree	Highly interpretable classification or regression model that splits data-feature values into branches at decision nodes (e.g., if a feature is a color, each possible color becomes a new branch) until a final decision output is made	Regression Classification
Naive Bayes	The Bayesian method is a classification method that makes use of the Bayesian theorem. The theorem updates the prior knowledge of an event with the independent probability of each feature that can affect the event.	Regression Classification
Support vector machine	Support Vector Machine, or SVM, is typically used for the classification task. SVM algorithm finds a hyperplane that optimally divided the classes. It is best used with a non-linear solver.	Regression (not very common) Classification

Random forest	The algorithm is built upon a decision tree to improve the accuracy drastically. Random forest generates many times simple decision trees and uses the 'majority vote' method to decide on which label to return. For the classification task, the final prediction will be the one with the most vote; while for the regression task, the average prediction of all the trees is the final prediction.	Regression Classification
AdaBoost	Classification or regression technique that uses a multitude of models to come up with a decision but weighs them based on their accuracy in predicting the outcome	Regression Classification
Gradient-boosting trees	Gradient-boosting trees is a state-of-the-art classification/regression technique. It is focusing on the error committed by the previous trees and tries to correct it.	Regression Classification

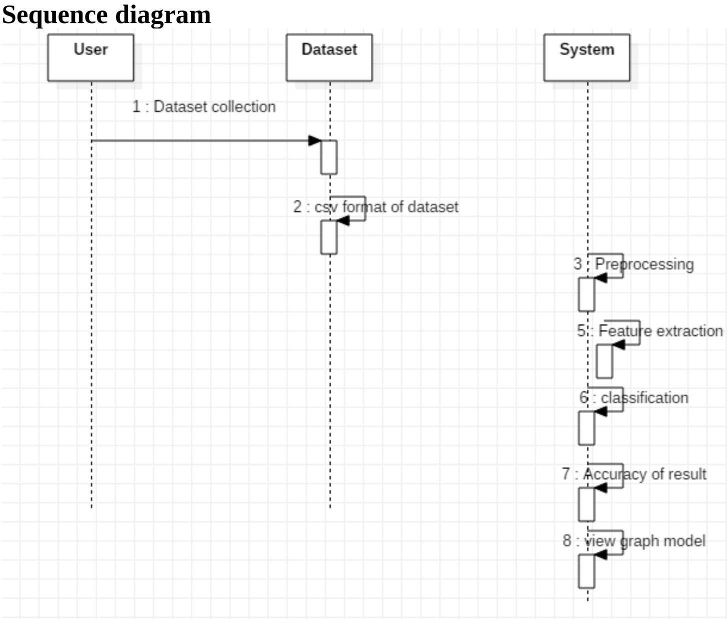
Unsupervised learning

In unsupervised learning, an algorithm explores input data without being given an explicit output variable (e.g., explores customer demographic data to identify patterns)You can use it when you do not know how to classify the data, and you want the algorithm to find patterns and classify the data for you.

Algorithm	**Description**	**Type**
K-means clustering	Puts data into some groups (k) that each contains data with similar characteristics (as determined by the model, not in advance by humans)	Clustering

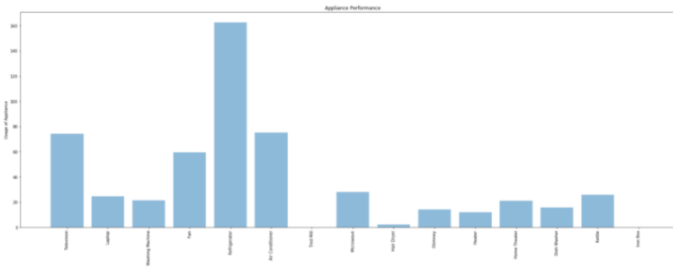
Gaussian mixture model	A generalization of k-means clustering that provides more flexibility in the size and shape of groups (clusters	Clustering
Hierarchical clustering	Splits clusters along a hierarchical tree to form a classification system. Can be used for Cluster loyalty-card customer	Clustering
Recommender system	Help to define the relevant data for making a recommendation.	Clustering
PCA/T-SNE	Mostly used to decrease the dimensionality of the data. The algorithms reduce the number of features to 3 or 4 vectors with the highest variances.	Dimension Reduction

Sequence diagram

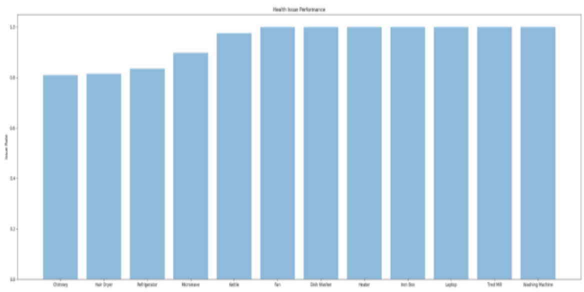


5. Result

In [8]:
```
objects_col = ('Television','Laptop','Washing Machine','Fan','Refrigerator','Air Conditioner','Tred Mill','Microwave',
            'Hair Dryer','Chimney','Heater','Home Theater','Dish Washer','Kettle','Iron Box')
y_pos = np.arange(len(objects_col))
performance_col = coltotvalue

plt.figure(figsize=(30,10))
plt.bar(y_pos, performance_col, align='center', alpha=0.5)
plt.xticks(y_pos, objects_col,rotation='vertical')
plt.ylabel('Usage of Appliance')
plt.title('Appliance Performance')

plt.show()
```

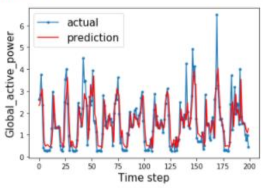


In [5]:
```
objects = pred_applicances1
y_pos = np.arange(len(objects))
performance = pred_result1

fig = plt.figure(figsize=(30,10))
plt.bar(y_pos, performance, align='center', alpha=0.5)
plt.xticks(y_pos, objects)
plt.ylabel('Issue Rate')
plt.title('Health Issue Performance')

plt.show()
```



In [4]:
```
#visualizing the clusters

plt.scatter(x[y_kmeans==0, 0], x[y_kmeans==0,1], s=50, c='red', label='Different Work') # s is size here
plt.scatter(x[y_kmeans==1, 0], x[y_kmeans==1,1], s=50, c='green', label='')
plt.scatter(x[y_kmeans==2, 0], x[y_kmeans==2,1], s=50, c='black', label='')
plt.scatter(kmeans.cluster_centers_[:,0], kmeans.cluster_centers_[:,1], s=50, c='yellow', label='centroids')

plt.title('Clusters Human Activities')
plt.xlabel('Clustered Data')
plt.ylabel('Performance Rate')
plt.legend()
plt.show()
```

In [22]:
```
aa=[x for x in range(200)]
plt.plot(aa, inv_y[:200], marker='.', label="actual")
plt.plot(aa, inv_yhat[:200], 'r', label="prediction")
plt.ylabel('Global_active_power', size=15)
plt.xlabel('Time step', size=15)
plt.legend(fontsize=15)
plt.show()
```



6. Conclusion

In this we compare with several regression algorithms to estimate the power consumption of an airconditioner as follows: multi-layered perceptron's (MLP), radial basis function networks (RBFN)and support vector regression (SVR). MLP is a conventional nonlinear regression method. we findthe sufficient accuracy for monitoring System which can estimate the power consumption ofelectric appliances, to actual fields.

Reference

1. C. M. Bishop.: Newral Networks for Pattern Recognition. Oxford University Press, 1995.
2. W. Hart.: Non-intrusive appliance load monitoring. Proceedings of the IEEE, vol. 80, no. 12, 1992.
3. J.Moody et al.: Fast learning in networks of locallytuned processing units. Neural Computation 1 (2), , 692498, 2001. 281-294, 1989.
4. H.Murata et al.: Applying Kernel Based Subspace classification to a Non-Intrusive Monitoring System for Household Electric Appliances. ICANN2001
5. T.Onoda et al.: Applying Support Vector Machines and Boosting to a Non-Intrusive Monitoring System for Household Electric Appliances with Inverters. NC'2000, 2000.
6. D. E. Rumelhart et al.: Learning Internal Rep resent ations by Error Propagation. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. l: Foundations, 318-362, Bradford Books/MIT Press, Cambridge, MA., 1986.
7. V. N. Vapnik.: The Nature of Statistical Learning TheonJ. Springer, 1995.S. Zhai, Y. Cheng, W. Lu, and Z. Zhang, ‘‘Deep structured energy basedmodels for anomaly detection,’’ in Proc. Int. Conf. Mach. Learn., 2016,pp. 1–10.
8. C. Fan, F. Xiao, Y. Zhao, and J. Wang, ‘‘Analytical investigation ofautoencoder-based methods for unsupervised anomaly detection in build-ing energy data,’’ Appl. Energy, vol. 211, pp. 1123–1135, Feb. 2018.
9. E. Mocanu, P. H. Nguyen, M. Gibescu, and L. Wil Kling, ‘‘Deep learn-ing for estimating building energy consumption,’’ Sustain. Energy, GridsNetw., vol. 6, pp. 91–99, Jun. 2016.
10. R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, ‘‘Deep learning and its applications to machine health monitoring,’’ Mech. Syst. Signal Process., vol. 115, pp. 213–237, Jan. 2019.
11. J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and M. Mohamad, ‘‘Nontech- nical loss detection for metered customers in power utility using support vector machines,’’ IEEE Trans. Power Del., vol. 25, no. 2, pp. 1162–1171, Apr. 2010.
12. S. S. S. R. Depuru, L. Wang, V. Devabhaktuni, and R. C. Green, ‘‘Highperformance computing for detection of electricity theft,’’ Int. J. Electr.Power Energy Syst., vol. 47, pp. 21–30, May 2013.
13. S. Makonin, F. Popowich, I. V. Bajić, B. Gill, and L. Bartram, ‘‘ExploitingHMM sparsity to perform online real-time nonintrusive load monitoring,’’IEEE Trans. Smart Grid, vol. 7, no. 6, pp. 2575–2585, Nov. 2016.
14. P. Jokar, N. Arianpoo, and V. C. Leung, ‘‘Electricity theft detection in AMIusing customers’ consumption patterns,’’ IEEE Trans. Smart Grid, vol. 7,no. 1, pp. 216–226, 2016.
15. A. Iwayemi and C. Zhou, ‘‘SARAA: Semi-supervised learning for auto-mated residential appliance annotation,’’ IEEE Trans. Smart Grid, vol. 8,no. 2, pp. 779–786, Mar. 2017.
16. E. Bair, ‘‘Semi-supervised clustering methods,’’ Wiley Interdiscipl. Rev.,Comput. Statist., vol. 5, no. 5, pp. 349–361, 2013.
17. M. Wytock and J. Z. Kolter, ‘‘Contextually supervised source separationwith application to energy disaggregation,’’ in Proc. AAAI Conf. Artif.Intell., 2014, pp. 1–7.
18. E. Elhamifar and S. Sastry, ‘‘Energy disaggregation via learning powerletsand sparse coding,’’ in Proc. AAAI Conf. Artif. Intell., 2015, pp. 1–7.
19. K. Yan, C. Zhong, Z. Ji, and J. Huang, ‘‘Semi-supervised learning for early detection and diagnosis of various air handling unit faults,’’ Energy Buildings, vol. 181, pp. 75–83, Dec. 2018.
20. Y. Zhang, W. Chen, and J. Black, ‘‘Anomaly detection in premise energy consumption data,’’ in Proc. IEEE Power Energy Soc. Gen. Meeting,Jul. 2011, pp. 1–8.
21. J. Chou and A. S. Telaga, ‘‘Real-time detection of anomalous power con- sumption,’’ Renew. Sustain. Energy Rev., vol. 33, pp. 400–411, May 2014.
22. D. F. M. Cabrera and H. Zareipour, ‘‘Data association mining for iden- tifying lighting energy waste patterns in educational institutes,’’ EnergBuildings, vol. 62, pp. 210–216, Jul. 2013.
23. M. Hu, Z. Ji, K. Yan, Y. Guo, X. Feng, J. Gong, X. Zhao, andL. Dong, ‘‘Detecting anomalies in time series data via a meta-feature basedapproach,’’ IEEE Access, vol. 6, pp. 27760–27776, 2018.

24. J. Alcalá, O. Parson, and A. Rogers, ‘‘Detecting anomalies in activitiesof daily living of elderly residents via energy disaggregation and Coxprocesses,’’ in Proc. ACM Int. Conf. Embedded Syst. Energy-Efficient BuiltEnviron., 2015, pp. 225–234.
25. S. Rahimi, A. D. C. Chan, and R. A. Goubran, ‘‘Nonintrusive load mon-itoring of electrical devices in health smart homes,’’ in Proc. IEEE Int.Instrum. Meas. Technol. Conf., May 2012, pp. 2313–2316.
26. M. Hori, T. Harada, and R.-I. Taniguchi, ‘‘Anomaly detection for an elderlypersonwatching system using multiple power consumption models,’’ inProc. Int. Conf. Pattern Recognit. Appl. Methods, 2017, pp. 669–675.
27. S. R. Gaddam, V. V. Phoha, and K. S. Balagani, ‘‘K-Means+ID3: A novelmethod for supervised anomaly detection by cascading K-means clusteringand ID3 decision tree learning methods,’’ IEEE Trans. Knowl. Data Eng.,vol. 19, no. 3, pp. 345–354, Mar. 2007.
28. X.-D. Wang, R.-C. Chen, F. Yan, Z.-Q. Zeng, and C.-Q. Hong, ‘‘Fastadaptive K-means subspace clustering for high-dimensional data,’’ IEEEAccess, vol. 7, pp. 42639–42651, 2019.139724 VOLUME 7, 2019
29. X. Wang et al.: SEPAD in Real Time Using Semi-Supervised Learning
30. S. Welikala, C. Dinesh, M. P. B. Ekanayake, R. I. Godaliyadda, andJ. Ekanayake, ‘‘Incorporatiappliance usage patterns for non-intrusiveload monitoring and load forecasting,’’ IEEE Trans. Smart Grid, vol. 10,
31. no. 1, pp. 448–461, Jan. 2019.