Iraqi Journal of Science, 2023, Vol. 64, No. 12, pp: 6509- 6521
DOI: 10.24996/ijs.2023.64.12.33

ISSN: 0067-2904

Energy Consumption Prediction of Smart Buildings by Using Machine Learning Techniques

Sofiene Haboubi*, Oussama Ben Salem
Signals Images and Information Technologies Lab. University of Tunis El Manar, National Engineering School of Tunis, Tunisia

Received: 22/5/2021 Accepted: 4/12/2022 Published: 30/12/2023

Abstract

This paper presents an IoT smart building platform with fog and cloud computing capable of performing near real-time predictive analytics in fog nodes. The researchers explained thoroughly the internet of things in smart buildings, the big data analytics, and the fog and cloud computing technologies. They then presented the smart platform, its requirements, and its components. The datasets on which the analytics will be run will be displayed. The linear regression and the support vector regression data mining techniques are presented. Those two machine learning models are implemented with the appropriate techniques, starting by cleaning and preparing the data visualization and uncovering hidden information about the behavior of the smart building appliances using the energy consumption feature. Afterwards, the implementation of two regression models to predict total energy consumption began. On a hospital database, the two techniques' performances are compared and validated. The results achieved are promising and prove the reliability of the IoT smart building platform.

Keywords: Smart building, Energy consumption, Internet Of Thing, Cloud Computing, Support Vector Regression, Support Vector Machine.

التنبؤ باستهلاك الطاقة في المباني الذكية باستعمال تقنيات التعلم الآلي

سفيان الحبوبي *, أسامة بن سالم
مخبر البحث في الإشارات والصور وتقنيات المعلومات، جامعة تونس المنار ، المدرسة الوطنية للمهندسين بتونس، تونس

الخلاصة

في هذه الورقة العلمية، تم تقديم منصة إنترنت الأشياء للمباني الذكية مع الحوسبة الضبابية والحوسبة السحابية القادرة على إجراء تحليلات تنبؤية في الوقت الفعلي تقريبًا على مستوى عُقد الضباب. شرح الباحثون بدقة إنترنت الأشياء في المباني الذكية، وتحليلات البيانات الضخمة وتقنيات الحوسبة الضبابية والحوسبة السحابية. ثم عرضوا منصة الطاقة الذكية ومتطلباتها ومكوناتها. ثم تم تقديم مجموعات البيانات التي سيتم إجراء التحليلات عليها. إثر ذلك عرض الباحثون تقنيات التنقيب عن بيانات بالانحدار الخطي وبدعم الانحدار المتجه. يتم تنفيذ هذين النموذجين من التعلم الآلي بالتقنيات المناسبة، بدءًا من التنظيف وإعداد تصور البيانات والكشف عن المعلومات المخفية حول سلوك أجهزة المباني الذكية باستعمال المؤشرات المميزة لاستهلاك الطاقة. بعد ذلك يتم تنفيذ نموذجي الانحدار للتنبؤ بكامل استهلاك الطاقة. تتم مقارنة أداء التقنيتين والتحقق من صحتهما. النتائج المتحصل عليها في هذا البحث تعتبر واعدة وتثبت موثوقية منصة أنترنت الأشياء في المباني الذكية.

*Email: sofiene.haboubi@enit.utm.tn

1. Introduction

Advances in wireless sensor network technologies have contributed a lot to realizing the modern smart home. With noticeable computing power and adapted communication protocols for these small devices, data is being generated at an alarming rate.

The need to capture and analyze this data has become vital to understanding and uncovering hidden valuable information that can have a major impact on our economy, our lifestyle, and even our safety. Utility companies, for example, use IoT big data analytics to recommend electricity bill reduction plans based on a personalized profile for each home generated by smart cities. Thus, it leads to reduced costs for both utility companies and homeowners, and it also saves energy [1].

IoT big data analytics has many advantages because it helps understand the behavior of smart homes and their occupants, but it is quite difficult because it deals with massive amounts of data that are generated every second. Hence, a system able to manage the growing volume of data and analyze it in near real-time to provide actionable insights has to be set up. Yassine et al. [2] proposed a platform for IoT smart home big data analytics using fog and cloud computing.

The purpose of our study is to take a different approach from other proposals:

- Including historical data, weather, and whether the day is a holiday to predict electricity consumption.
- Test energy consumption prediction in a hospital environment with these valuable features.

2. Literature Review

2.1 IoT Frameworks

The Internet of Things is deeply rooted in our modern lifestyle, and it is affecting our daily routine from how we perceive reality to how we interact with our environment. Connected things have increased in the last few years, from smart watches that track our daily activity to smart cars that provide us with the shortest routes.

The term IoT was first introduced to the world by the British technology pioneer Kevin Ashton in 1999. Even though it has been around for more than two decades, there's a controversy among researchers, innovators, and academicians over how to define the IoT.
A complete definition of the Internet of Things is proposed by [3].

It is basically a giant network of connected devices that gather data and communicate through well-adapted protocols such as MQTT [4] and Zigbee [5]. It describes the next generation of the internet, in which physical objects are identified using wireless sensor networks (WSN) [6] and can exchange and process data in response to predefined situations.

WSN technology became popular over the last few years due to its unique characteristics: low cost, energy efficiency, small physical size, computational power, communication capabilities, distributed sensing, etc. Thus, it is the perfect fit to build a smart home where all appliances are connected.

A smart platform is a homogeneous setup of connected appliances that can be controlled remotely by the homeowner in order to provide comfort and a good experience of assisted living. A smart platform can be basically divided into three layers [7]:

- The sensing layer: It consists of WSNs and smart meters that collect data and measure activity from home devices and appliances and then send it to the network layer.
- Network Layer: It is adapted to receive data from sensors in accordance with IoT communication protocols like MQTT, then it sends the collected data to the next layer.
- Application Layer: This is where the processing and analysis of the collected data happens, and then the insights can be clearly visualized for the user so that he can easily interact with these devices through an interface. This is the layer that we will be working on in this paper.

2.2 Big Data analytics

In today's information age, where information is critical to standing out from the crowd in every field, 2.5 quintillion bytes of data are created every day, and this number is growing exponentially with the growth of the Internet of Things (IoT) [8]. Most of this generated data is in its raw format and cannot be exploited unless it is processed and analyzed; the ability to extract insights from it will lead to innovation and productivity [9].

IBM data scientists break big data into four dimensions: volume, variety, velocity, and veracity [10]. These four dimensions make it easier to define the nature and the challenges of big data:

- Volume: is the most important aspect of big data; from historical, unprocessed data to newly generated data, these massive amounts of data necessitate scalable storage capacities and adaptable DBMS.
- Variety: Data exists in different forms, such as photos, videos, texts, financial transactions, etc. and most of this data is unprocessed.
- Velocity: Data is being generated at an incredible speed due to mobile technologies. Enterprises and organizations are struggling to keep up with this rate.
- Veracity: Data isn't always true; thus, it can be misleading. This can be due to poor quality of data or false information shared on social media platforms.

Big Data analytics is basically analyzing huge amounts of data to uncover hidden patterns and correlations, which can be done by applying data mining algorithms to large datasets. Hence, this greatly contributes to faster and more efficient decision-making [11-14]. Big Data analytics is made up of five steps:

- Identifying the problem: What is the problem that needs to be solved.
- Collecting: identify what kind of data is required to solve our problem and collect it.
- Preprocessing: Since data is usually unprocessed, it needs to be cleaned and prepared, like by replacing missing values or converting a timestamp to a human-readable date format.
- Analyzing: Once the data is prepared, apply various methods to analyze it, like data mining algorithms, to reveal valuable information.
- Visualizing: To highlight the revealed information, you need to visualize it in graphs or tables.

Big Data analytics can be classified into four main types:

- Descriptive analytics: It takes into consideration the historical and generated values to describe and summarize the data so that it can be understood by the user.
- Predictive Analytics: It uses machine learning models to forecast what will happen in the future. This is the type of analysis that will be performed in this paper.
- Prescriptive analytics: It is a step further than predictive analytics since it uses simulation algorithms to provide a recommendation for a possible action to be taken by the user.
- Diagnostic Analytics: It takes a deeper look at the data through data mining algorithms to determine what the causes are behind a certain result.

2.3 Cloud and Fog computing

Cloud computing is a model where a user has access to physical computing resources remotely. It has numerous advantages, including high computing power and scalable storage capacities. Cloud computing can be classified into three main types [15, 16]:

- Infrastructure as a Service (IaaS): It is the infrastructure part of the cloud that provides accrediting access to servers, networks, and storage space.
- Platform as a Service (PaaS): It provides, in addition to the infrastructure part, integrated software like middleware and frameworks.
- Software as a Service (SaaS): It facilitates software access by providing ready-to-use applications accessible via the user's web browser.

In addition to the services provided, cloud computing has become popular due to its scalability, where we can easily add or remove servers based on our needs. Another interesting feature is that you can access these services instantly, and last but not least, cloud computing reduces costs significantly since it is cheaper than buying and installing hardware and servers.

Unlike cloud computing, which is quite popular, fog computing is a new term for a lot of people. Fog computing is related to the IoT because it extends cloud services closer to edge devices [17, 18]. It is not another version of cloud computing; it is just a bridge between the edge and the cloud.

In the IoT, fog computing has many advantages since it is close to the end user; it offers great mobility, real-time interactions, and low latency [19, 20].

3. Materials and Methods

The experiments are performed on:

- two years of data generated by appliances in a smart home in Vancouver, British Columbia, Canada.
- and healthcare facility in San Diego, CA, USA.

This analysis has the goal of predicting energy consumption in a defined period of time based on historical data for energy consumption, the day, the hour of the day, and the weather.

This analysis is performed in the fog node as long as it is dealing with time series and continuous streams of data that need to be accomplished in near real-time in order to meet the prerequisites of applications like electricity bill reduction plans.

3.1 The Almanac of Minutely Power dataset (AMPds2)

The data from the AMPDs2 was collected from a smart home in Vancouver, British Columbia, Canada. It has 730 days of captured data per meter [21]. It is publicly available for download from Harvard Dataverse in different formats; this dataset has 1,051,200 readings per meter and contains twenty-one power meters, two water meters, and two natural gas meter datasets [22]. It also contains hourly weather data for the same period. For this purpose, the researchers are only interested in the electricity datasets and the weather dataset.

3.2 EnergyPlus dataset

The data were retrieved from the energy consumption simulation of a healthcare facility in San Diego, CA, USA, with EnergyPlus Version 9.0.1 software [23]. The building is a reference building developed by the US Department of Energy. The simulated data concerns energy consumption from January 2014 to January 2020. The values in this dataset are calculated per hour, which gives a total of 52,584 hourly timesteps.

3.3 Regression analysis techniques

Smart homes generate a continuous stream of data, and since they want to predict energy consumption based on historical data [24], the hour of the day, the month, and the weather, there is nothing better than a regression model to get the job done. A regression analysis is used to predict a variable by investigating the relationship between the dependent and independent variables. It is most often used for predicting prices based on given features or energy forecasting.

Regression models are a supervised machine learning technique [25, 26]. In this work, we will be interested in only two of those techniques: linear regression and support vector regression.

Linear regression is the most common and popular regression technique. In this technique, an input variable X is used to predict an output variable Y. The relationship between those two variables is linear and can be written like: $Y=a+bX$.

The researchers want to predict energy consumption of HVAC system [27] based on the temperature values. They can see that there's a correlation between these two variables. An increase in temperature means an increase in HVAC system energy consumption.

Y represents the HVAC system energy consumption, and X represents the temperature. If they want to predict Y for a value of X, they need to calculate the coefficients a and b. In most cases, measuring the error by a cost function is mean squared error (MSE), which is the average of the squared difference between that blue line and the data points. So, for a linear regression model, try to find the coefficients when the cost function is at its minimum.
Once the regression line has been drawn with minimal error, one must be cautious in selecting a value X in order to avoid predicting a Y value for a value of X that is outside the data range [28].

Support Vector Machine (SVM) [11] is a supervised learning model that is quite popular for its use in classification problems.

Support Vector Regression (SVR) is almost similar to SVM, but it is used for regression problems. Even though it is not as popular as SVM, SVR is an effective algorithm for real-value dataset predictions.

3.4 Data Cleaning and Preparation

Data cleaning and preparation is a vital step in any type of analytics since it ensures a good quality of data that is ready to be mined by a specific algorithm [29]. In this study, we are going to predict energy consumption, and to accomplish that, we only need the apparent power column, so only keep the date/time column and the apparent power column of each dataset. Merging the datasets of each appliance into one final dataset where each column represents the apparent power of a certain appliance. Identifying the standby power threshold of each appliance by plotting the graph of its energy consumption, then setting the standby power threshold to zero in order to be able to identify when an appliance is switched on.

Once the standby power thresholds are set to zero, the hourly weather dataset is imported. From the weather dataset, grab only the temperature column and then replace the missing value with the median of the measured temperatures.

After that, merge the weather dataset with the appliance apparent power measurements dataset. By doing that, a lot of missing values are generated since the two datasets have different lengths (minutely time series versus hourly time series), so replace each missing temperature value with the preceding value since the temperature is the same for the same hour. Adding a month and an hour of the day column is necessary because appliance and user energy consumption behavior varies by hour and month.

Another significant variable often neglected is whether or not a day is a day off (weekends and holidays), because electricity usage during a day off is quite different from a typical day. Then, create a categorical variable column with a value of one if the day is a weekend or a holiday, otherwise a value of zero.

3.5 Data Visualization

Data visualization is not less significant than other steps since it helps us uncover valuable information about the behavior of the smart building appliances and the occupants and visualize temperature variations and their correlations with energy consumption.

Another important feature that affects energy consumption is the temperature. Figure 1 shows the temperature variations in Vancouver, British Columbia, Canada, and San Diego, USA, over the span of a year.

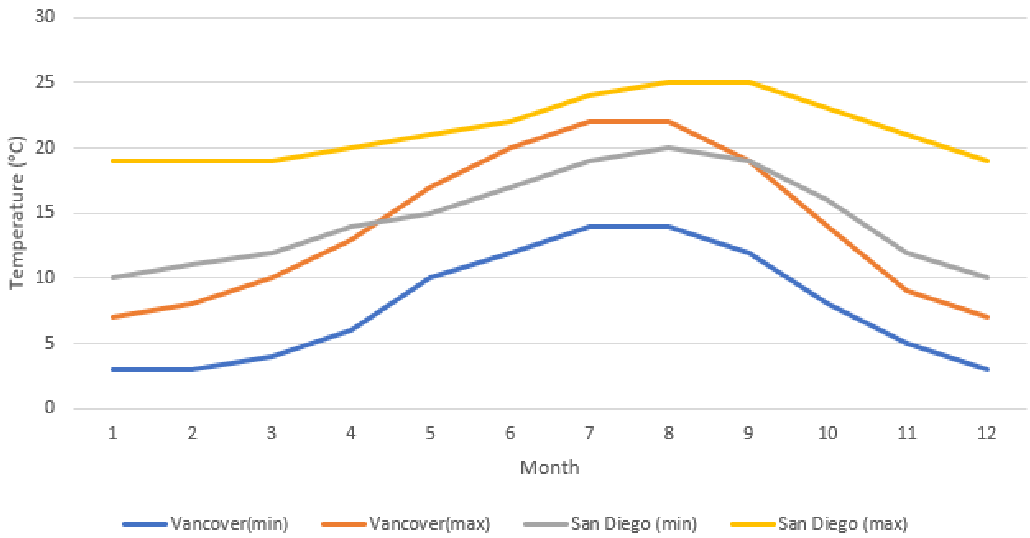

Figure 1: Caption of temperature variations in Vancouver, CA and San Diego, USA.

As shown in the graph (Figure 2), hourly resampled data shows the same behavior of the appliance as minutely measured. The resampled smart home dataset in a monthly mean was used to verify once again the veracity of the data and estimate the fridge's energy consumption compared to the whole house consumption. Figure 3 shows the fridge's share of the whole house's energy consumption over a one-year span.

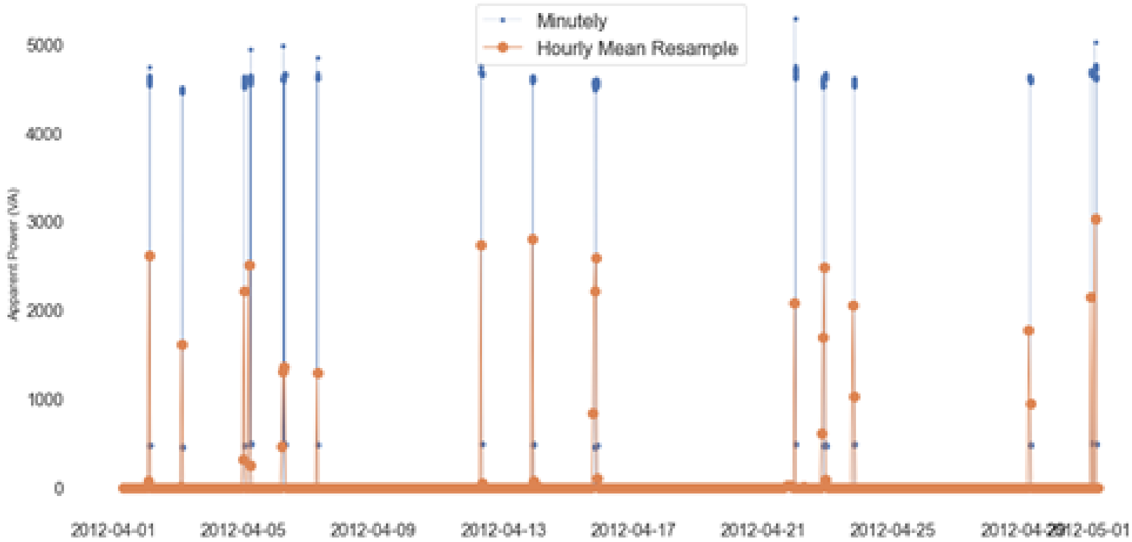

Figure 2: The minutely energy consumption of the kitchen oven versus its hourly.

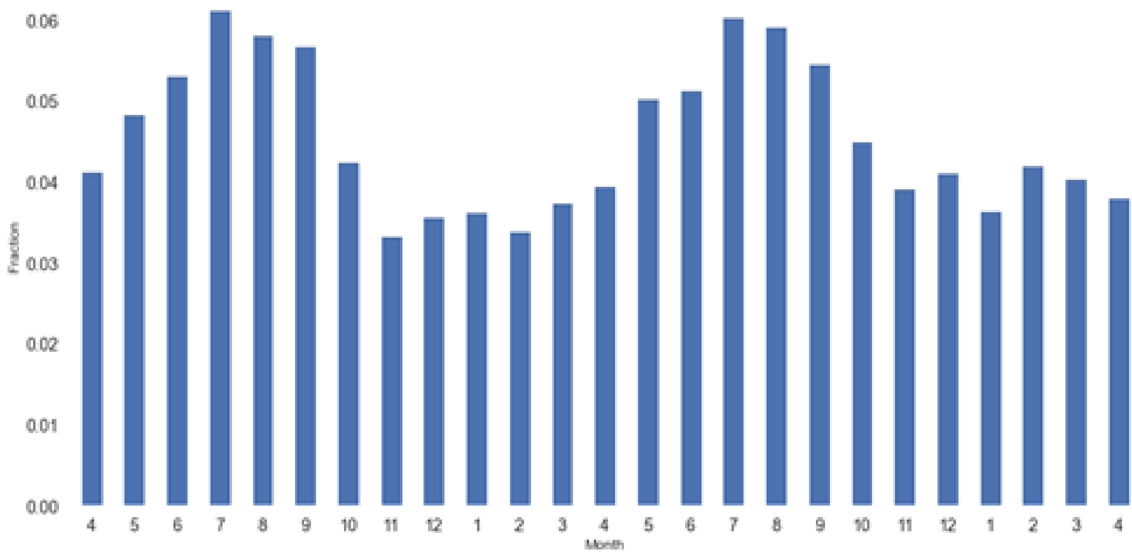

Figure 3: The Fridge energy consumption.

During the summer, the fridge consumes a greater portion of the building's energy. Hence, the results confirm that the temperature has a significant influence on energy consumption.

To clearly see this correlation, visualize the temperature variations versus the energy consumption in the same graph. Figure 4 shows the variations in the temperature values versus the variations in the heat pump energy consumption. As it is shown in the graph, when the temperature increases, the heat pump’s energy consumption decreases, which makes total sense.

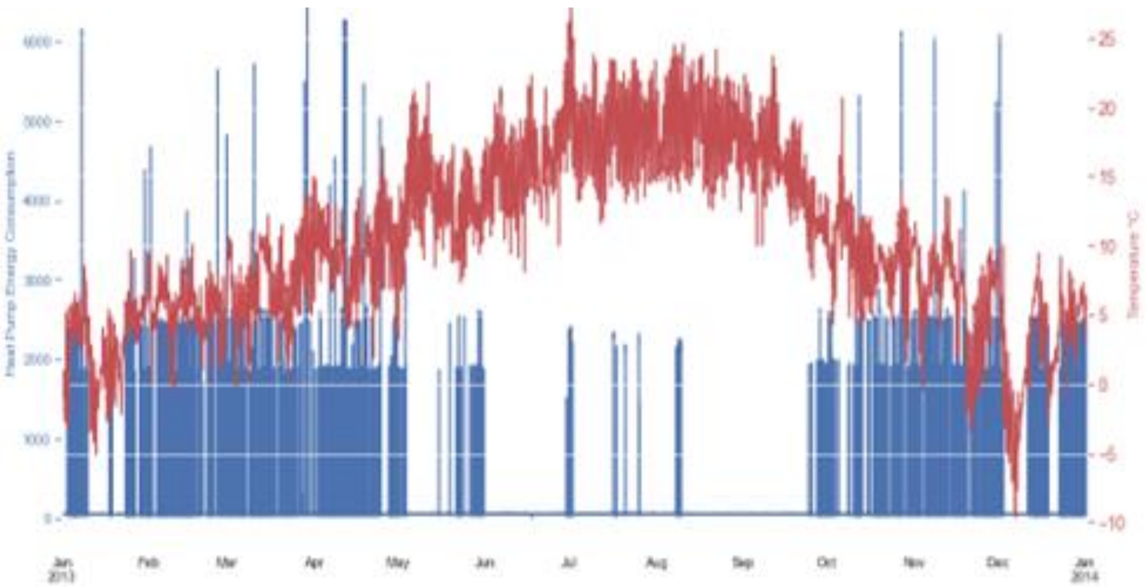

Figure 4: Heat Pump energy consumption and Temperature in 2013.

3.6 Regression Models Implementation

The regression models will do the predictive analytics in order to predict the whole energy consumption based on the weather, the hour of the day, whether or not the day is a day off, the month, and the historical data of the energy consumption. The probability distribution for total energy consumption is checked.

The correlation matrix between all the columns is plotted. Figure 5 represents the probability distribution of the total energy consumption.

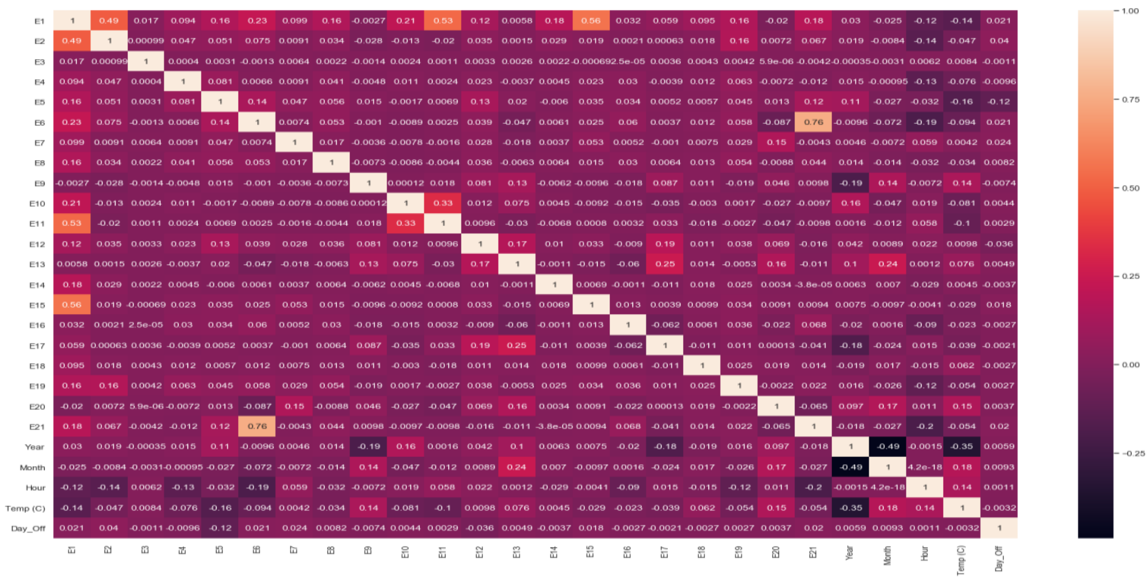

Figure 5: Correlation matrix between the dataset features.

starting with linear regression and then performing support vector regression. First, we split our dataset into features that the model will train on and a target variable that the model will predict. Then, split it into a training set and a test set.

An additional step is to scale the variables. Feature scaling is used to normalize the data, which speeds up the calculations in certain algorithms like support vector machines. Even though it is not necessary for linear regression since the outcome is the same whether the data is scaled or not, using it for both regression models allows us to objectively compare the results.

- Linear Regression: Calculate the R squared, which is the proportion of the variance in the dependent variable that is predictable from the independent variable. According to Cohen [30], when the R squared is equal to or greater than 0.29, it is considered substantial. R squared for the linear regression model on this dataset equals 0.296.
- Support Vector Regression: Tuning the model parameters (Gamma and C) of a Gaussian function to optimize the result. The Gamma parameter is the inverse of the radius of influence of samples selected by the model as support vectors. The C parameter gives the model freedom to select more samples as support vectors. R squared for the Support Vector Regression model on our dataset equals 0.308.

4. Results

After implementing the two regression models and calculating the R squared for both (0.296 for the AMPds2 dataset and 0.311 for the EnergyPlus9 dataset) for the linear regression model.

After plotting the predicted values of the smart building's energy consumption by the linear regression model and the actual values in the same graph, we can identify how accurate the results are. Figure 6 shows the predicted values by the linear regression model versus the real values of the total energy consumption.

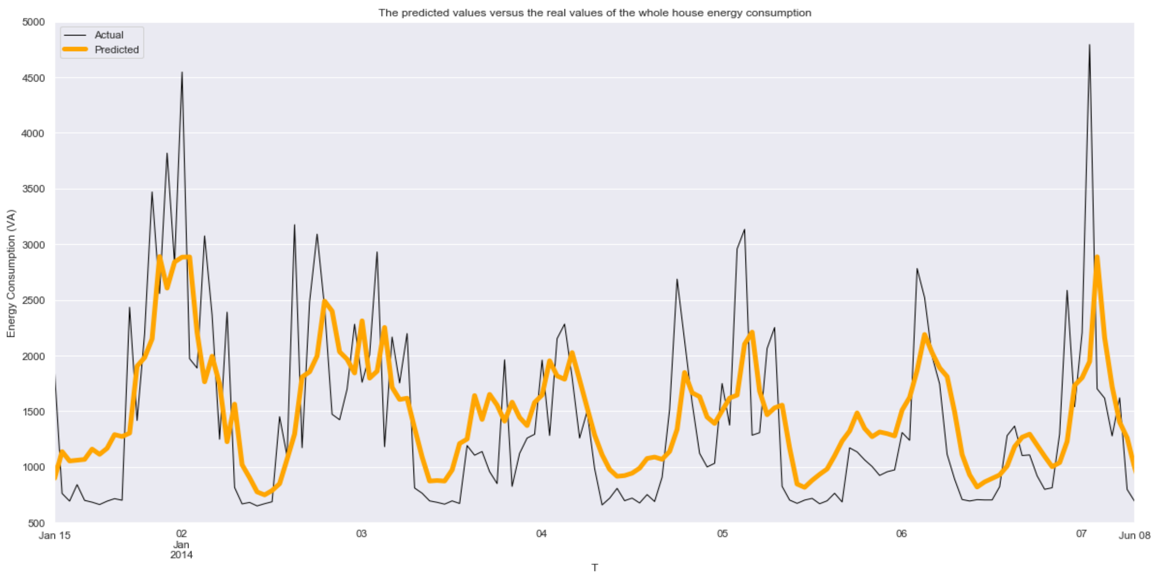

Figure 6: The whole energy consumption predicted values by the linear regression model.

As shown in the linear regression graph, the prediction was acceptable as it clearly captured the shape of the curve. It didn't perfectly predict the values, but it shows clearly the variations between the minimums and the maximums, and it almost encapsulates all the peaks.

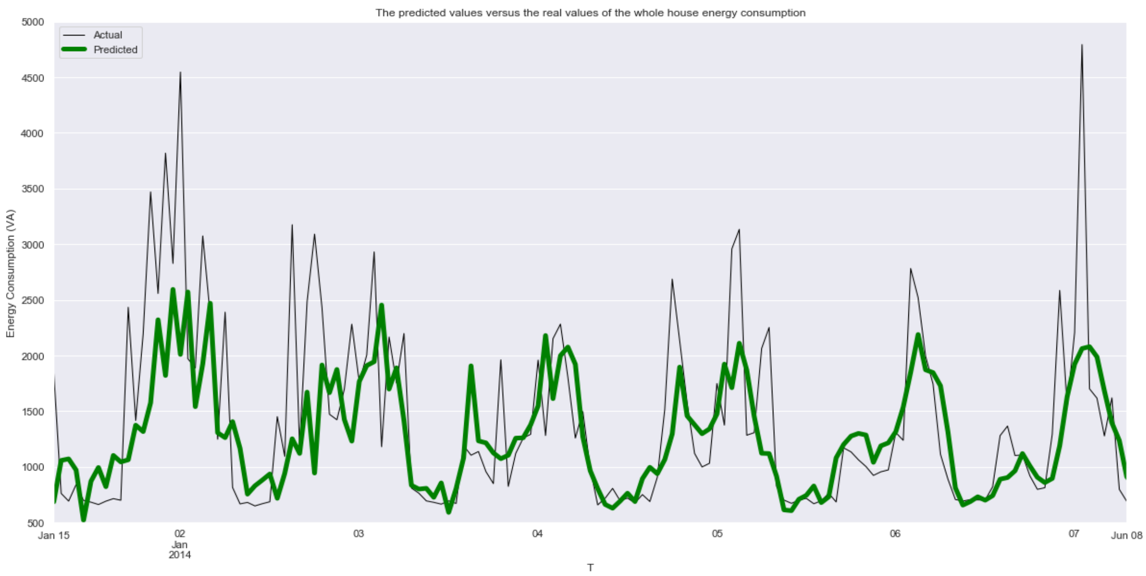

Figure 7: The whole energy consumption predicted values by the support vector regression model.

By calculating the R squared for the support vector regression (0.308 for the AMPDS2 dataset and 0.281 for the EnergyPlus9 dataset), then, in the same graph, plot the predicted and actual values of total energy consumption from the SVR model to see how they relate to one another. Figure 7 shows the predicted values by the support vector regression model versus the real values.

5. Discussion

The previous predictive analytics were performed in the fog node in order to predict energy consumption for a specific appliance.

The analytics performed in the fog node platform included frequent pattern mining and clustering mining, with the goal of activity recognition that can be used in e-health (for example, monitoring a patient's activity at home to identify health problems).

Our approach in this paper is to perform the regression data mining technique at the fog node in order to predict energy consumption. Utility companies can use those predictions to recommend electricity bill reduction plans. It can also be quite beneficial for companies that are in the advertisement business, where they can use an appliance energy consumption prediction to pass on targeted commercials based on a specific profile.

In this study, the researchers performed two regression data mining techniques: Linear regression, which is a popular technique, generates the best-fitting line using ordinary least squares that minimizes the sum of the squared prediction error. Support-vector regression, which is a less popular technique, uses the same logic as support-vector machines but instead of predicting classes, it predicts values.

From the obtained results of the two models' performances, it is quite clear that the SVR model generated better predicted values. Linear regression didn't capture the correct shape of the curve, unlike support vector regression. Thus, conducting predictive analytics in the fog node using the SVR model is recommended since it is helpful in dealing with limitations related to the probability distribution and the shape of the data.

According to the results of the support vector regression graph, the prediction was accurate enough in capturing all the variations.

6. Conclusion

The results obtained were promising, as the two regression models' outputs were sustainable according to [30], but the support vector regression model generated better predicted values. Hence, SVR is the appropriate predictive data mining technique to be used in the fog node for smart home electricity usage measurements.

In this study, the proposed platform confirmed her reliability to perform IoT big data analytics and introduced a new predictive analytics technique that is quite helpful for the in-time decision-making process. It helps electric utility companies recommend electricity bill reduction plans for their clients. The predictive analytics done for the entertainment equipment helps companies that are interested in targeted advertisements.

The predictive analytics performed in the fog node can manage the continuous incoming streams of data, sending the uncovered knowledge about the appliance's energy consumption to the cloud. Thus, with the distributed fog node technology, smart city energy management can be easily handled by this platform. Although the results from the predictive analytics performed in the fog node were satisfying, scalability is always a major issue since we are dealing with big data.

In the future, it will be interesting to use appropriate tools such as Apache Hadoop to perform intensive big data analytics in the cloud rather than at fog nodes. In the future, it will be interesting to use appropriate tools such as Apache Hadoop to perform intensive big data analytics in the cloud rather than at fog nodes. They can also keep working on optimizing the computing time of the fog nodes by using different data mining techniques and moving towards real-time analytics.

Apart from energy consumption management, this platform can be used for managing natural gas and water consumption, and it can even combine them with electricity to have a platform able to manage these three features. Not only predictive analytics but also prescriptive analytics can be implemented, where an interactive recommender system can be implemented for the homeowners to suggest possible actions to reduce consumption and enhance the experience of assisted living.

References

[1] S. Singh and A. Yassine, "Mining Energy Consumption Behavior Patterns for Households in Smart Grid," *IEEE Transactions on Emerging Topics in Computing*, vol. 7, no. 3, pp. 404-419, 2019, doi: 10.1109/tetc.2017.2692098.

[2] A. Yassine, S. Singh, M. S. Hossain, and G. Muhammad, "IoT big data analytics for smart homes with fog and cloud computing," *Future Generation Computer Systems*, vol. 91, pp. 563-573, 2019, doi: 10.1016/j.future.2018.08.040.

[3] S. Madakam, R. Ramaswamy, and S. Tripathi, "Internet of Things (IoT): A Literature Review," *Journal of Computer and Communications*, vol. 03, no. 05, pp. 164-173, 2015, doi: 10.4236/jcc.2015.35021.

[4] A. S. Shubhangi, A. N. Pooja, P. S. Shubhangi, D. S. Vrushali, and R. J. Yogesh, "MQTT-Message Queuing Telemetry Transport protocol," *International Journal of Research*, vol. 03, pp. 240-244, 2016.

[5] C. M. Ramya, M. Shanmugaraj, and R. Prabakaran, "Study on ZigBee technology," presented at the 2011 3rd International Conference on Electronics Computer Technology, 2011/4, 2011.

[6] M. R. Ahmed, X. Huang, D. Sharma, and H. Cui, "Wireless Sensor Network: Characteristics and Architectures," *International Journal of Information and Communication Engineering*, vol. 6, no. 12, p. 4, 2012.
[7] O. Bhat, S. Bhat, and P. Gokhale, "Implementation of IoT in Smart Homes," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 6, no. 12, p. 7, 2017, doi: 10.17148/IJARCCE.2017.61229.
[8] B. Marr. "How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read." Forbes. https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/?sh=587c302c60ba (accessed 11-15-2021).
[9] A. D. Mauro, M. Greco, and M. Grimaldi, "What is big data? A consensual definition and a review of key research topics," *AIP Conference Proceedings*, vol. 1644, no. 1, pp. 97-104, 2015, doi: 10.1063/1.4907823.
[10] R. Vella, "Exploring the Potential of Digital Practices in Malta-As viewed by the National Repositories," Ph.D. Thesis, 2018. [Online]. Available: https://dx.doi.org/10.13140/RG.2.2.13708.97920
[11] H. Chen, X. Xie, W. Shu, and N. Xiong, "An Efficient Recommendation Filter Model on Smart Home Big Data Analytics for Enhanced Living Environments," *Sensors (Basel)*, vol. 16, no. 10, Oct 15 2016, doi: 10.3390/s16101706.
[12] Q. Huang, S. Jing, J. Yi, and W. Zhen, "Data Management in Smart Grid," in *Innovative Testing and Measurement Solutions for Smart Grid*, pp. 183-210, 2015.
[13] S. N. Amy and R. H. David, "DATA ANALYTICS AND PREDICTIVE ANALYTICS IN THE ERA OF BIG DATA," in *Internet of Things and Data Analytics Handbook*: Wiley, 2017, pp. 329-345, 2017.
[14] E. l. Bossé and B. Solaiman, *Information fusion and analytics for big data and IoT*. Norwood, MA.: Artech House (in eng), 2016.
[15] S. Sakr, A. Liu, D. M. Batista, and M. Alomari, "A Survey of Large Scale Data Management Approaches in Cloud Environments," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 311-336, 2011, doi: 10.1109/surv.2011.032211.00087.
[16] F. Mohammed, A. M. Ali, A. S. A.-M. Al-Ghamdi, F. Alsolami, S. M. Shamsuddin, and F. E. Eassa, "Cloud Computing Services: Taxonomy of Discovery Approaches and Extraction Solutions," *Symmetry*, vol. 12, no. 8, 2020, doi: 10.3390/sym12081354.
[17] E. Badidi, "A Framework for Software-as-a-Service Selection and Provisioning," *International Journal of Computer Networks and Communications (IJCNC) Vol 5, No 3, pp. 189-200, May 2013*, 2013/6 2013. [Online]. Available: http://arxiv.org/abs/1306.1888.
[18] A. M. Mohammed and F. A. Omara, "A Framework for Trust Management in Cloud Computing Environment," presented at the 2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE), 2020. [Online]. Available: http://dx.doi.org/10.1109/ITCE48509.2020.9047791.
[19] H. Atlam, R. Walters, and G. Wills, "Fog Computing and the Internet of Things: A Review," *Big Data and Cognitive Computing*, vol. 2, no. 2, p. 10, 2018/4 2018, doi: 10.3390/bdcc2020010.
[20] J. Almutairi and M. Aldossary, "Modeling and Analyzing Offloading Strategies of IoT Applications over Edge Computing and Joint Clouds," *Symmetry*, vol. 13, no. 3, Mar 2021, doi: 10.3390/sym13030402.
[21] S. Makonin, "AMPds2: The Almanac of Minutely Power dataset (Version 2)," 6 2019, doi: 10.7910/DVN/FIE0S4.
[22] S. Makonin, B. Ellert, I. V. Bajic, and F. Popowich, "Electricity, water, and natural gas consumption of a residential house in Canada from 2012 to 2014," *Sci Data*, vol. 3, no. 1, p. 160037, Jun 7 2016, doi: 10.1038/sdata.2016.37.
[23] M. K. Heris. "Particle Swarm Optimization in MATLAB." Yarpiz. https://yarpiz.com/50/ypea102-particle-swarm-optimization (accessed 3-3-2022).
[24] Osório, k. Shafie, Carvalho, and Catalão, "Analysis Application of Controllable Load Appliances Management in a Smart Home," *Energies*, vol. 12, no. 19, p. 3710, 2019 2019, doi: 10.3390/en12193710.

[25] S. Wang, Y. Yao, F. Zhu, W. Tang, and Y. Xiao, "A Probabilistic Prediction Approach for Memory Resource of Complex System Simulation in Cloud Computing Environment," *Symmetry,* vol. 12, no. 11, 2020, doi: 10.3390/sym12111826.

[26] R. A. Giancristofaro and L. Salmaso, "Model performance analysis and model validation in logistic regression," *Statistica,* vol. 63, no. 2, pp. 375-396, 2007, doi: 10.6092/issn.1973-2201/358.

[27] A. Ghahramani, P. Galicia, D. Lehrer, Z. Varghese, Z. Wang, and Y. Pandit, "Artificial Intelligence for Efficient Thermal Comfort Systems: Requirements, Current Applications and Future Directions," *Frontiers in Built Environment*, vol. 6, Apr 28 2020, doi: 10.3389/fbuil.2020.00049.

[28] K. H. Zou, K. Tuncali, and S. G. Silverman, "Correlation and simple linear regression," *Radiology,* vol. 227, no. 3, pp. 617-22, Jun 2003, doi: 10.1148/radiol.2273011499.

[29] G. G. de Carvalho Chrysostomo, M. V. B. de Aguiar Vallim, L. S. da Silva, L. A. Silva, and A. R. de Aguiar Vallim Filho, "A Framework for Big Data Analytical Process and Mapping—BAProM: Description of an Application in an Industrial Environment," *Energies,* vol. 13, no. 22, 2020, doi: 10.3390/en13226014.

[30] J. Cohen, *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale, N.J: L. Erlbaum Associates, 1988.