Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 165 (2019) 90–96

Procedia Computer Science

www.elsevier.com/locate/procedia

INTERNATIONAL CONFERENCE ON RECENT TRENDS IN ADVANCED COMPUTING 2019, ICRTAC 2019

Household Energy Management Model to Maximize Solar Power Utilization Using Machine Learning

Gautam M , Siram Raviteja, R. Mahalakshmi*

Amrita School of Engineering,
Amrita Vishwa Vidyapeetham, Bengaluru, India.

Abstract

Electricity from solar energy is the new trend in electricity production companies. Many plants have been set up to harness energy from the sun. Since solar energy can be harnessed from open-air, starting the same from home is the best way. Most commonly solar energy is harnessed from a rooftop-mounted solar panel but is connected only for a few selective loads. All the other loads when in need of electricity cannot be utilized from solar energy. Since this is the format for electricity usage, the maximum usage of this is not done. This framework is proposed to maximize the usage of solar power by connecting both solar and grid to the same node and employing a unique switching strategy using decision tree machine learning algorithm in python environment. The hardware setup has been done for the same by using data acquisition techniques, collecting real-time data consisting of demand and solar power which is updated into the database in the local setup server. This real-time data is exported and prediction for the switching configuration is done and sent to an Arduino board to predict the source that has to be given to the particular demand in real-time.

Keywords: Solar power management; Energy management system; Machine learning; Decision tree; Data mining;

* Corresponding author. Tel.: +91 9035406063;
E-mail address: deivamaha@gmail.com

10.1016/j.procs.2020.01.075

1. Introduction

Out of all the forms of energy, the solar energy is most abundant and to an extent this energy can be easily harness able for electricity purposes. Reduction in fuel for non-renewable sources, the movement towards solar energy [1] has led to a great change in electricity production. Even though many plants are setup to harness this solar energy, home is the best place to harness this as solar energy is present everywhere and less losses for transportation of this power. The existing technology present is that solar panel setup at home is connected to a different set of loads and grid is connected to a different set of loads. Both the sources can be connected to the same set of loads using a grid tie inverter which enables the solution for effective switching between the sources, which is not cost effective. Since the solar is connected only to a set of loads the usage of this power is limited and not used to its maximum extent.

The implementation of real-time data acquisition for the load switching mechanism using the prediction analysis based on machine learning algorithm for the power system management is proposed in this framework [2-3]. Real-time data acquisition hypothesis is chosen for the utility in order to switch between the sources based on the availability of the power either at the solar side or the grid side. The most important of any processing application is to obtain the required dataset for its implementation. In the framework, the process of implementing data mining mechanism is done for acquiring real time data or the real-world data. This process enables an enhanced methodology for parallel programming, where the huge amount of data is acquired and so the algorithms can easily process for the requirement and yield best outcomes [4]. This large amount of data has to be stored in a database for the further usage. The server database provides the data fragmentation and provides a proper integration with the data mining techniques [5]. The process of creating the database with respect to the server through the Structured Query Language (SQL) process is to provide a secure feature of coding and so the database works in real time without any creation of problem and so the whole process works in a safer manner. This particular process of SQL programming creates a secured gateway for the access of the server's database and the acquired data can be stored in it for the utilization of further processes [6]. The process of implementation of decision tree algorithms in the power system management is carried out in this particular framework and is not reported in any literatures. This particular methodology is brought to a small-scale analysis of the load variations and its management where a room of a particular consumer's utility is considered and so the Energy Management System (EMS) process is undertaken for the variations and proper analysis is done [7-9]. All these approaches are used for an effective management of energy by switching between the solar and grid systems employing data mining and machine learning techniques. Chapter II discusses the machine learning technique for the framework and its respective observations. Chapter III explains the hardware prototype model.

2. Machine Learning

The data that has been mined from the sensors with the help of various data mining techniques are further put into many classification and regression tasks with the help of machine learning algorithms. This approach has been employed in this framework. Decision tree algorithm has been employed as this is one of the algorithms which mimics the exact decision-making ability of that of humans. This algorithm works on bases of 2 methodologies, one is least entropy for decision making and other is the most Gini index for decision making. Entropy in machine learning is defined as how randomly the output can be predicted with the given patterns or inputs. The least entropy means that least randomness which in turn means the output is easily predicted. Decision tree algorithm first splits the columns given as inputs into separate ones. The most entropy column i.e., the least value in the column is chosen as the top most node as shown in Figure <1> and then following which the next highest entropy and so on. This process continues till the entropy of that node becomes zero. When entropy is zero gives us the meaning that the output prediction has been reached.

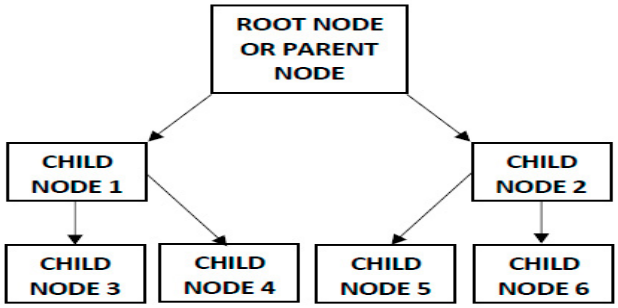


Fig. 1. Decision tree node splitting algorithm representation.

2.1. Data set

The data from a particular consumer's house load system has been collected from the sensors. These Sensors have been installed for 180 days at the consumer's house. Data collected data from these are the grid voltage, grid current, solar voltage and solar current. Totally these four parameters that have been collected for 180 days are then converted to their respective power equivalents for one hour which resulted in hourly data for 180 days i.e. three months. This data is then split into training and testing data where training data consists of a small percentage of total data and the rest as testing data in order to compare with the training data which is kept as a reference data.

This particular data included the four conditions of the project so as to carry out the machine learning process in detailed and robust manner for the prediction of the source that can be connected to the consumer's house load system. By this particular data-set, the consumer can now be able to predict the usage and so that particular individual can now utilize more amount of the renewable energy source that is connected to his house. Now that particular individual can also control his usage and save his electricity bill. By this observations and inferences that are been carried out, the user can now send the data to the electricity board and so the board can now decide on expanding the system to other areas for providing efficient power to the rural areas and also the area with deficit amount of power. The dataset and the plot of obtained data in excel are shown in Figure <2> and Figure<3>.

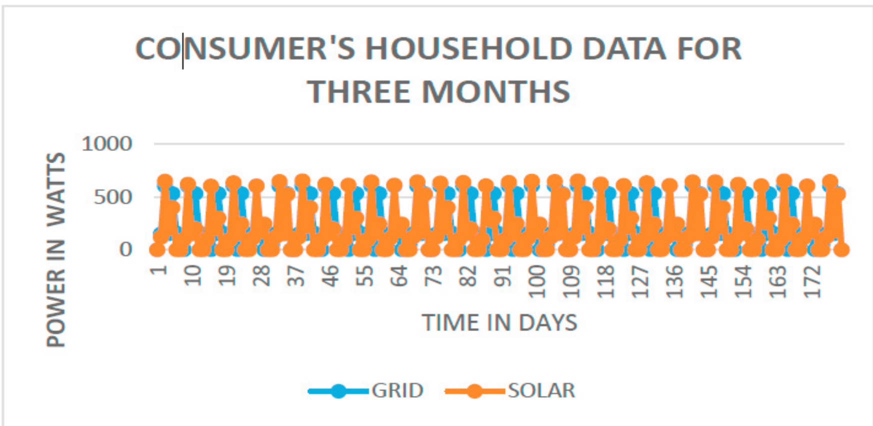


Fig. 2. Data mining of consumer load system for three months.

2.2. Training data

The data-set that is present with the user is termed as a data model and the part of that particular data model is called a training data which means the training data is the subset of data model. The training data set in algorithm programming process is the definite model set for the learning mechanisms of the machine learning algorithms. This is the definite model of the acquired data for the process of implementing and making the machine get experienced alongside with various Application Program Interface (API) and programming to train the machine to work automatically. In practice, the definite training data-set or the model often consist of relation between the input data to the respective output data which is commonly shown as the target value. The training set contains the known target and the model gets experienced on this data model in order to be generalized to other data-set later for further processing. A training data-set is set of examples used for learning and experiencing the machine for proper functioning, which is to fit the parameters. The current model is run with the training data model and produces an outcome, which is then compared with the target, for each input data in the training data-set. Based on the outcome of the comparison and the specific learning algorithm being used, the parameters of the model are adjusted. The model sees and learns from this particular data-set that is been provided by the user in Table 1.

Table 1. Data-set of the power in hours of a particular consumer's house.

Hours	Grid power (W)	Solar power (W)
1	0	0
2	150	120
3	603	650
4	150	200
5	533	400
6	160	0
7	50	0
8	0	100
9	603	620

2.3. Testing data

This particular data-set is the subset of the training data-set or it can also be the subset of the whole data model. This corresponds to the final evaluation that the model goes through after the training phase or the process has been completed. A test data-set is a model that is a part of the training data-set, but that follows the same probability distribution as the training data model. If a model fit to the training data-set also suits the test data-set as well and the minimal overfitting has taken place. This is the data typically used to provide a proper evaluation of the final outcome that are completed and fit on the training data-set. It is only used once a model is completely trained. By using this set, we can get the working accuracy of our model. Example of testing data-set is shown in Figure<3>.

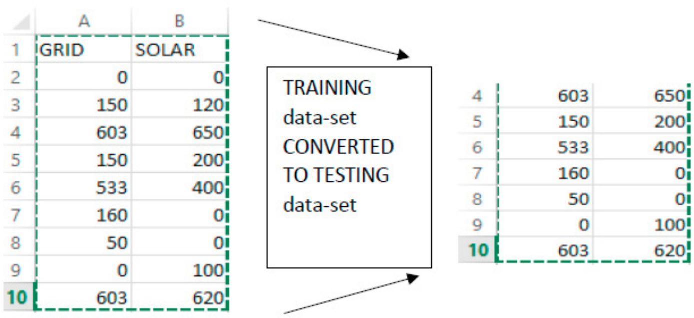


Fig. 3. Snapshot of the testing data-set of consumer's house from excel.

The overall processing of the data model to training data-set and testing dataset can be shown as in Figure<4> and the evaluation values are shown in the Figure <5>.

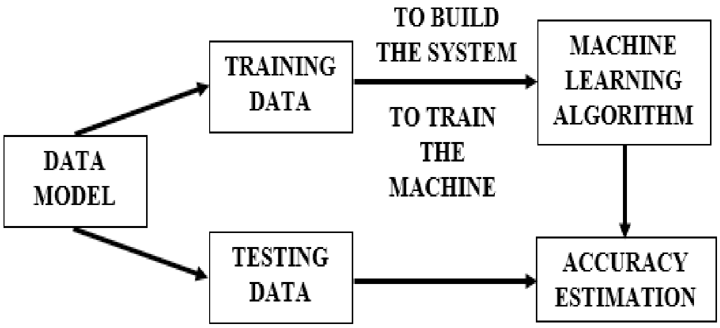
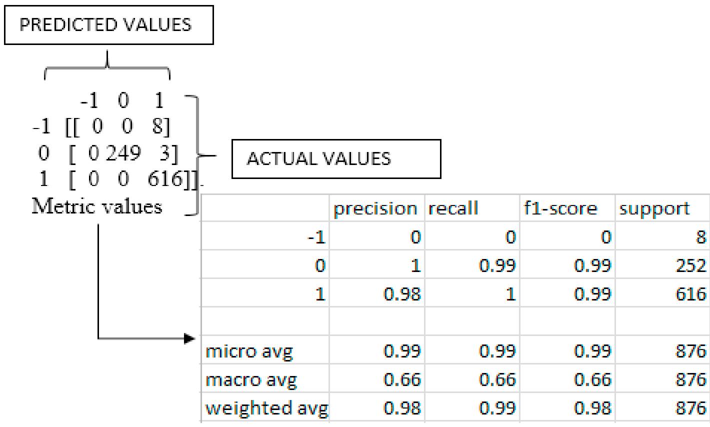


Fig. 4. Model of training data-set and testing data-set.



	precision	recall	f1-score	support
-1	0	0	0	8
0	1	0.99	0.99	252
1	0.98	1	0.99	616
micro avg	0.99	0.99	0.99	876
macro avg	0.66	0.66	0.66	876
weighted avg	0.98	0.99	0.98	876

Fig. 5. Snapshot of metric values along with confusion matrix from python.

3. Hardware Implementation

The framework here is used to predict the switching configuration of control switch 1 and 2 as represented in the block diagram in Figure< 6> of the relays in order to predict the source that has to be given to the respective load. The overall block diagram of the framework is shown is Figure <7>. There are various stages in which this total process takes place. The chronological order of the process is in the below sub chapters.

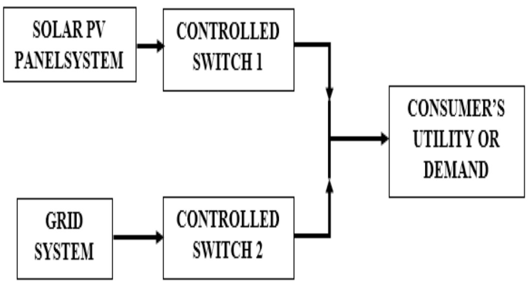


Fig. 6. Block diagram of switching position.

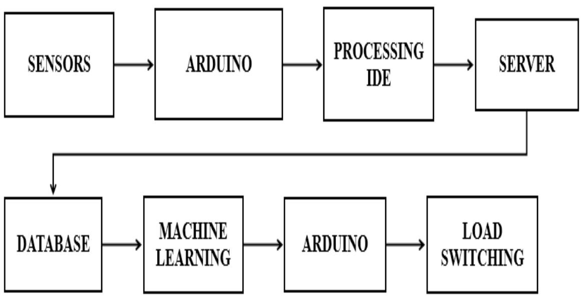


Fig. 7. Block diagram of the overall system.

3.1. Sensor data and Arduino

There are DC voltage, DC current, AC current and temperature sensors used. The DC sensors are placed near the solar panel's output, temperature sensor near the solar panel and AC sensors near the load as in Figure <8>. DC sensors determine the DC voltage using a voltage sensor SKU VOLT, which works based on voltage divider circuit and DC current using current sensor ACS712, which works on the principal of Hall Effect sensing. The AC sensors determine the AC current using a current sensor KG190, which works on the principal of induction in the circuit. Temperature sensors are placed near the solar panel in order to determine the temperature near the solar panel. These data are taken in through the input pins in Arduino Uno board and pre-processed in the Arduino IDE programming to find the respective DC power produced by the solar panel, real time demand of the home and temperature near the solar panel.

3.2. Server and Database

These data collected from the sensors are to be stored for further use. A local server has been created using Wamp server software package. Server provides with a lot of services of which one of them is database. A MySQL (Structured Query Language) database has been created in this server. The connection between the Arduino data collected and database is done using a platform called processing IDE. This platform is used to bridge the gap between the Arduino and the database by collecting all the data that is been printed in the Arduino IDE and storing it into an array. This array is later appended into the database using queries. The appended data viewable in the form of table consisting of the powers and temperature values with the current time stamp in the database visualizer as shown in Figure <8>. The most recent one row data is then exported as a (comma separated values) CSV file for every five minutes using an event scheduler query written for an indefinite loop. This data is also deleted after every five minutes as this exporting of data cannot be rewritten using a batch file created in the local computer where this whole process takes place.

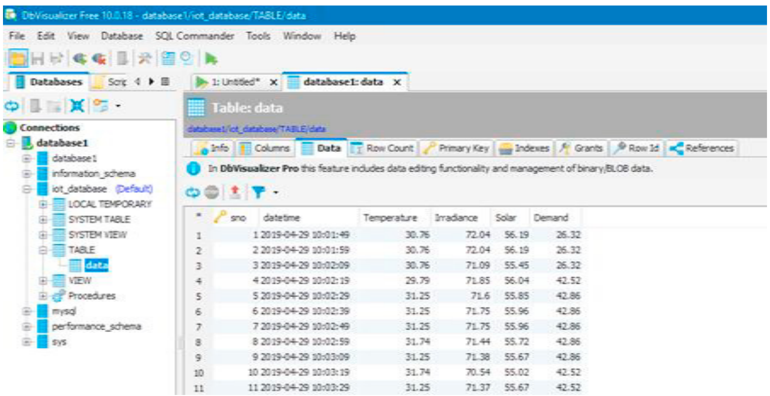


Fig. 8. Database visualization.

3.3. Machine learning

This data that has been exported in CSV format is the input file for prediction using machine learning. A machine learning approach has been employed in order to predict the switching configuration of the switches placed in this system to determine the source appropriate for the particular load.

3.4. Relay switching

This exported CSV file from the machine learning is fed as the input to another Arduino which is programmed in a way that it reads the CSV file and determines the switching configuration and inputs the same to a relay which switches between the sources. The hardware setup of the total process is shown is Figure<9>.

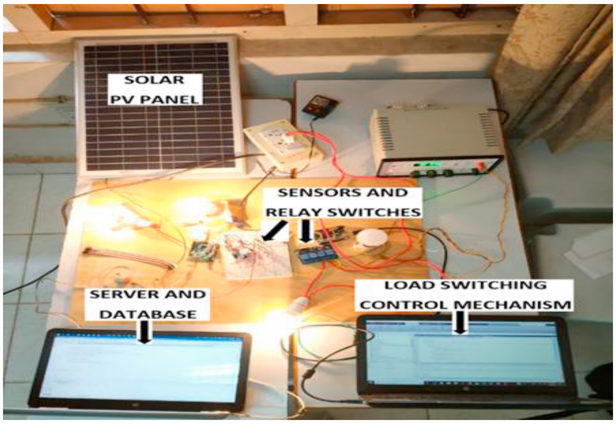


Fig. 9. Hardware prototype.

4. Conclusion

This framework provides one of the most efficient method to maximize the usage of solar power in order to meet the demand and uses the grid power only when demand is not able to be met by the solar alone. This framework made the switching possible by closing the relay switches corresponding to solar power and utilizing this form of energy until it is not able to meet the rising demand. When the demand exceeds the limit more than the production of solar power, then grid switch is turned on. The whole is made possible through decision tree machine learning algorithm employed in this framework. This strategy employed here allows the reduction is electricity cost and as well as increase in use of renewable energy. This hardware prototype is exclusively implemented for houses. This when employed on a large scale, decreases the non-renewable energy consumption, thus reducing all environmental factors.

References

[1] R. Mahalakshmi, Ashwin Kumar A. and A. Kumar. (2014) "Design of Fuzzy Logic based Maximum Power Point Tracking controller for solar array for cloudy weather conditions," *2014 Power and Energy Systems: Towards Sustainable Energy*, Bangalore, pp. 1-4.

[2] K. P. N and P. V. D. (2017) "Machine Learning Based Residential Energy Management System," *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, Coimbatore, pp. 1-4.

[3] Amarasinghe, Gihan & Abeygunawardane, s. (2018). "Application of Machine Learning Algorithms for Solar Power Forecasting in Sri Lanka", *2018 2nd International Conference On Electrical Engineering (EECON).* pp.87-92.

[4] Junliang Fan et. al (2019). "Empirical and machine learning models for predicting daily global solar radiation from sunshine duration: A review and case study in China." *Renewable Energy Sustainable Energy Reviews 100 (1),* pp. 186-212.

[5] M. Sousa, M. Mattoso and N. F. F. Ebrecken (1998) "Data mining: a tightly-coupled implementation on a parallel database server," *Proceedings Ninth International Workshop on Database and Expert Systems Applications (Cat. No.98EX130)*, Vienna, pp. 711-716.

[6] K. Kamtuo and C. Soomlek. (2016) "Machine Learning for SQL injection prevention on server-side scripting," *2016 International Computer Science and Engineering Conference (ICSEC)*, Chiang Mai, pp. 1-6.

[7] E. Borioli, E. Ciapessoni, D. Cirio and E. Gaglioti. (2009) "Applications of Neural Networks and Decision Trees to Energy Management System Functions," *2009 15th International Conference on Intelligent System Applications to Power Systems*, Curitiba, pp. 1-6.

[8] Dittawit Kornschnok, Finn Arve Aagesen. (2014) "Home energy management system for electricity cost savings and comfort preservation", *Proceedings of IEEE Fourth International Conference on Consumer Electronics-Berlin*, pp. 309-313.

[9] P. Dongbaare, S. O. Osuri, S. P. Daniel Chowdhury. (2017) "A smart energy management system for residential use", *IEEE PES Power Africa Accra*, pp. 612-616.