Classifying Future Scope in Energy Resources and Predicting Power Demand using Multilayer Perceptron

M.Anita Priscilla Mary [1], Dr.M.S.Josephine [2], Dr.V.Jeyabalaraja [3], Dr.S.Nirmala Sugirtha Rajini [4]

[1] Research Scholar, Department of Computer Applications, Dr.M.G.R Educational and Research Institute, Maduravoyal, Chennai-600095, Tamil Nadu, India.
[2]Professor, Department of Computer Applications, Dr.M.G.R Educational and Research Institute, Maduravoyal, Chennai-600095, Tamil Nadu, India.
[3] Professor, Department of Computer Science & Engineering, Vellammal Engineering College, Surapet, Chennai-600066, Tamil Nadu, India.
[4]Professor, Department of Computer Applications, Dr.M.G.R Educational and Research Institute, Maduravoyal, Chennai-600095, Tamil Nadu, India.

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

Abstract: To identify the future scope in Energy resources using various Energy resources like Hydro, Wind, Thermal, Biomass and others using Classification methods like Bayesian Classification and EM algorithm. Forecasting of Energy demand is identified by using the previous year's statistical data. Applying Time series forecasting methods which is implemented using WEKA. The best algorithm is identified by comparing the Evaluation parameters of various Time Series methods and the best method is used for predicting the future demand.

Keywords: Classification, Forecasting, Methods, Evaluation Parameters

1. Introduction

Energy resources are of different categories which are used for generation of power. In order to identify which Energy resource will have future scope in power generation to meet over the future demand. In data mining classification methods are applied to the various energy resources. Machine learning actualizes multifarious prediction mechanism to inspect data in order to attain predictions. WEKA is used for agglomeration as it has congeries of machine learning algorithms for data mining functions [1]. In order to find the future scope in energy resources two algorithms are enforced, first one is Bayesian Classification and the next one is EM algorithm. In order to find the effective utilization of Electrical Energy resources, power demand is analyzed in various Energy sectors like Domestic, Commercial, Agriculture, Industries and others using K-Means Clustering Algorithm. In order to find the energy sectors were it needs effective utilization of Energy resource, analysis made on existing K-means algorithm the disadvantage is identified thus Proposed Altered K-means clustering algorithm is defined for the analysis of various Energy sectors. Forecasting of Energy demand is identified by using the previous years of statistical data using Artificial intelligent methods. Election of the prime algorithm is invoked by discovering the evaluation parameters Mean Absolute Error, Root Relative Squared error, Direction Accuracy Relative Absolute Error, Mean Absolute Percentage Error, etc are contemplated to find the average of error values. The best algorithm for forecasting is identified and prediction of Energy demand for future is identified by applying Times Forecasting method in WEKA. Energy demand and Energy availability for the future is predicted in order to find the Solar Energy Generation to meet over the future demand [2].

2. Related work

Monire Nouranji, Allireza Souri, Majid Samed Zamini (2016) in their study they have proposed the numerous application in malware detection. Dynamic analysis approach has been explored for classifying the malware lineament. The data mining accession is more adequate for observe malware and behavioural collocation of malware can be useful to detect malware in a behavioural antivirus.[3]

Joao Virotea, Rui Neves-Silva (2012) The objective is to ensure the return on investment from these projects. The procure results manifest that the recommended energy consumption model lucubrate incumbent behavioural patterns from the building. It reliably mimeographs them, predicts the framework of energy consumption and constitutes probable areas of energy waste [4].

Joao Virotea, Kwok-Wing Chau, Chun-Tian Cheng, Lin Qiu(2009) in their paper they have proposed about performance evaluation measures evaluation measures, the coefficient of correlation (R), Nash-Sutcliffe efficiency coefficient (E), Root Mean Squared Error, Mean value of Absolute Percentage Error which need to be estimated in the enforcement of models. The results evince that the best acquirement

can be accomplished by ANFIS, GP and SVM, in terms of particular evaluation criteria during the training and validation phases [5].

Vijayamohanan Pillai N (2008) this paper proposes the electricity demand foretell acess and proposes a scholastic time series structure. It establishes the foretell, compel analysis and the methodologies to overcome compel are analyzed [6]

.

3. Proposed System

Classification is applied to various energy resources in order to identify which energy resource is going to meet over future demand. Fourteen years of statistical data from the year 2005 to 2019 in various Energy resources of Tamil Nadu is used for the analysis. Classification is applied to classify various energy resources using classes attributes Hydro, Thermal, Gas, CGS, Wind, Solar, Biomass and others. It is converted into a tabular data where each column of the table represents a different value. It is categorized into columns such as year and class in order to apply classification. In order to identify the effective utilization of Electrical Energy resources, the fourteen years of statistical data of various Energy sectors are used for analysis. It is categorized into year and sectors to find which sector needs effective utilization of Energy resources. In order to find future demand 9 years of statistical data is taken to predict the future demand. It is represented into two columns year and demand in order to apply Time series forecasting methods.

4. Implementation and Methodology

4.1 Implementation

WEKA is a free source machine learning groupware that can be pervade through a menu-driven interface, standard terminal function, or a Java API. It is extensively used for teaching, research, and industrial function. It contains a deluge of constructive tools for authoritative machine training tasks, and additionally gives diaphanous access to well-known toolboxes such as scikit-learn R, and Deeplearning4. Python is a persuasive programming language which has high aligned data architecture and has effectual accession to object oriented programming is used for the implementation of clustering.

Methodology

Classification is applied to various energy resources and identified which energy resource is going to meet over the future demand. Naive Bayesian Classification is applied for various energy resources and EM algorithm is finally applied to the consolidated data to find which resource will helpful to meet over the future demand [7] which is implemented using WEKA. Effective utilization of Energy resource is identified by applying proposed Altered K-Means algorithm to various Energy sectors. Forecasting of Energy demand is identified using the Artificial intelligent methods. The Selection of best algorithm is implemented by finding the evaluation parameters such as Mean Absolute Error, Root Relative Squared error, Relative Absolute Error etc., comparing the averages of the error values. The best algorithm for forecasting is identified and prediction of Energy Forecasting method.

4.3 Classification

4.3.1 Naive Bayesian Classification

Applying Naive Bayesian Classification to various Energy resources such as Hydro, Thermal, Gas, Solar, and others various evaluation parameters are analyzed using WEKA[8]. In ROC (Receiving Operating Character) Area is calculated as a metric to evaluate classifier output quality. This gives an idea of how classifiers are performing. ROC is calculated as a comparison of TPR (True Positive rate) and FPR (False Positive Rate).

Table 1: Classification by Receiving Operating Character(ROC) Area

CLASS	ROC AREA
HYDRO	0.036
THERMAL	0.036
GAS	0.036
CGS	0.012
WIND	0.024
SOLAR	0.000
BIOMASS	0.000
COGEN	0.060
OTHERS	0.083

Analyzing on the values of ROC Area it is identified that Energy resources Solar and Biomass has the value Zero. It specifies that these two areas are under the development for Energy generation. Hence we need to development in Solar and Biomass in order to over the future demand.

4.3.2 Comparative Study on Solar and Biomass

EM algorithm empowers a probability distribution to each adduce which indicates the probability of it relationship to each of the clusters. EM algorithm is correlated to previous year generation data of Solar and Biomass.

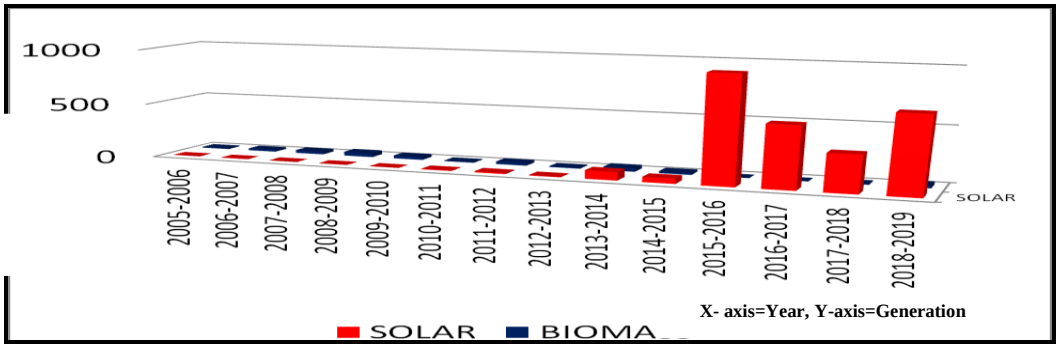

Figure 1: Comparative study on Solar and Biomass

Table 2. Applying EM algorithm to Solar and Biomass generation

CLASS	MEAN	STD.DEV
SOLAR	191.5399	17.4639
BIOMASS	17.4639	11.8754

Applying EM algorithm for the generation data of Solar and Biomass, it is identified that the mean and standard deviation value of Solar generation is higher than Biomass generation[9]. Hence it is identified that the Solar Energy is one of the resource to meet over the future electricity demand. Since growth of Solar energy generation is higher compared to Biomass. Solar energy plant can be easily implemented in Energy sectors such as Domestic, Agriculture, Commercial and Industry is has faster growth. Solar power plant does not produce harmful gases and the cost of implementing solar power plant is less compared to Biomass. Hence Solar power generation has the best scope to meet over the future demand[10].

4.4 Clustering

The advantages and disadvantages of K-means clustering algorithm is identified, the disadvantages of K-means clustering algorithm is rectified by Proposed Altered K-means clustering algorithm. Accuracy Score, Iterations, Inertia and Elapsed time are the various measures used for performance analysis.

Table 3: Comparison of Existing K-means Algorithm and Proposed Altered K-means Algorithm

Measures	Existing K-Means	Proposed Altered K-Means
Accuracy Score	0.833	0.98
Iterations	3	4
Inertia	143.3345	89.6384
Elapsed Time	-0.00010	-0.000333

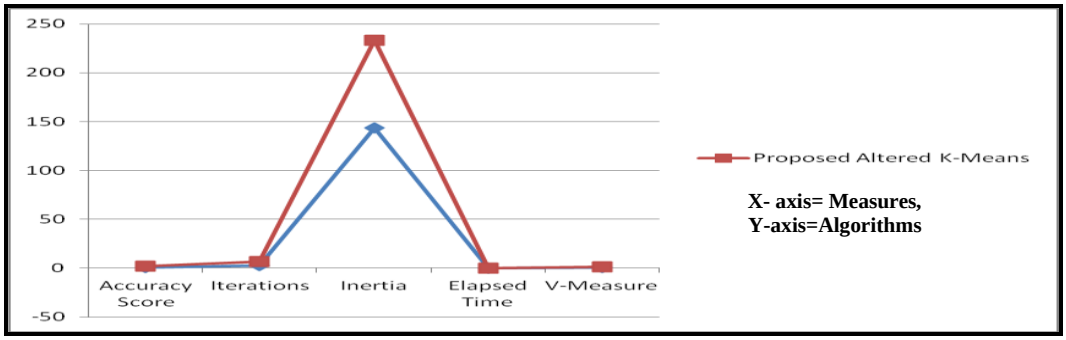

Figure 2: Comparitive study on Existing K-means Algorithm and Proposed Altered K-means Algorithm

Applying Proposed Altered K-means Algorithm using Python to various Electrical Energy sectors it is identified that Domestic, Agriculture and Industries needs effective utilization of energy resource to reduce power consumption.

4.5 Time series forecasting

Time series forecasting refines of using a model to accomplish foresight the future phase based on the known past phase. The algorithms of Time Series forecasting Linear Regression, Gaussian Process, Multilayer Perceptron, Support Vector Machine[11].

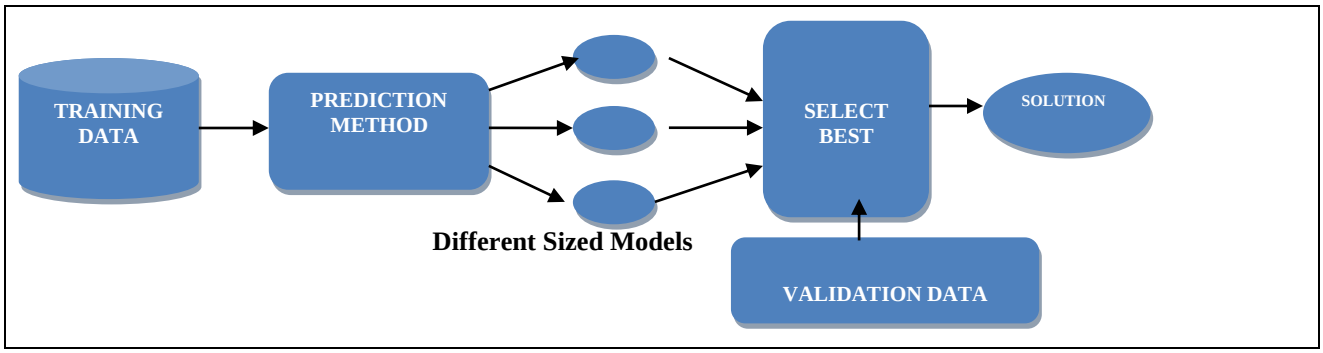

Figure 3: Finding Best Model among Predictive Model using WEKA

4.5.1 Evaluation Parameters of Time series Algorithms

Evaluation parameters such as Mean squared error, Root mean squared error, Mean absolute percentage error, Relative absolute error, Root relative squared error, Direction accuracy and Mean absolute error are used for the evaluation.

Table 4 : Comparison of Evaluation Parameters

Evaluation Parameter	Gaussian Process	Linear Regression	Multilayer Perceptron	Support Vector Machine
Mean Squared Error	118698314.2	2091080.1	29557.9	6108522.1
Root Mean Squared Error	10206.4	1434.8	156.1	2412.48
Mean Absolute Percentage Error	88.34	16.86	1.46	21.75
Relative Absolute Error	88.3	16.8	1.5	21.7

Root Relative Squared Error	98.8	16.8	1.5	25.6
Direction Accuracy	26.66666	80	80	80
Mean Absolute Error	9130.1	1350.6	145.4	1986.2

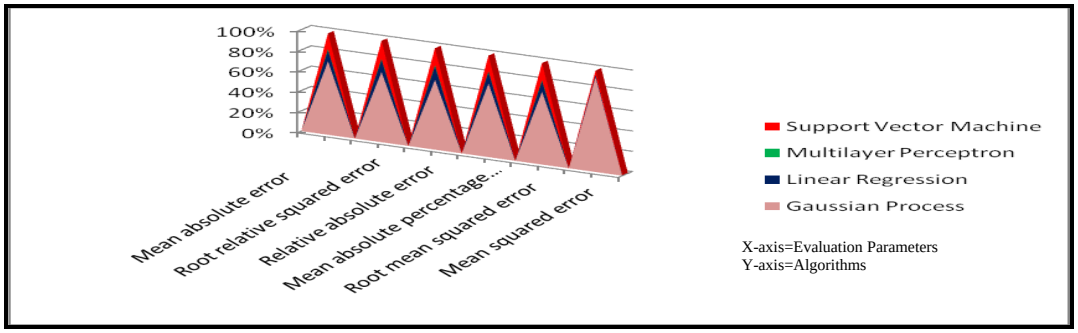

Figure 4: Visualization of various algorithms

Analyzing on the evaluation parameters of various algorithms it is identified that error value is minimal in Multilayer perceptron and the error values are higher in Gaussian process. When the parameters like Mean Squared error are taken for comparison in Multilayer perceptron has the value less compared to other algorithms. Similarly when Root mean squared error, Mean Absolute Error, Root Relative Squared Error, Relative Absolute Error and Mean Absolute Percentage Error are less in Multilayer perceptron compared to other algorithms.

Observing the Evaluation parameters from Table-4 Multilayer Perceptron has the lowest error values compared to other algorithms. Hence Multilayer Perceptron is the best algorithm for applying Time Series forecasting[12].

4.5.2 Multilayer Perceptron

A Multilayer Perceptron is a source of feed forward Artificial Neural Network. It subsists minimal of three stratums of nodes an input layer, a hidden layer and an output layer. Precluding the input nodes, each node is a neuron that uses a non aligned stimulating objective. Multilayer Perceptron exploits a supervised training approach called back propagation[13][14].

4.5.3 Forecasting of Energy Demand and Energy Availability in Million Units using WEKA

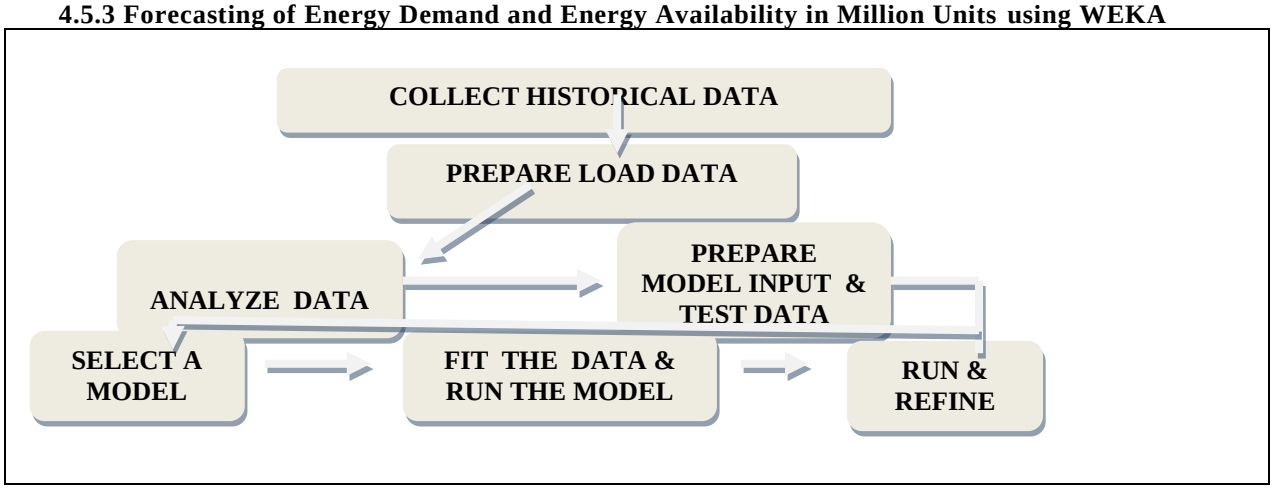

Figure 5: Energy Load Forecasting Model

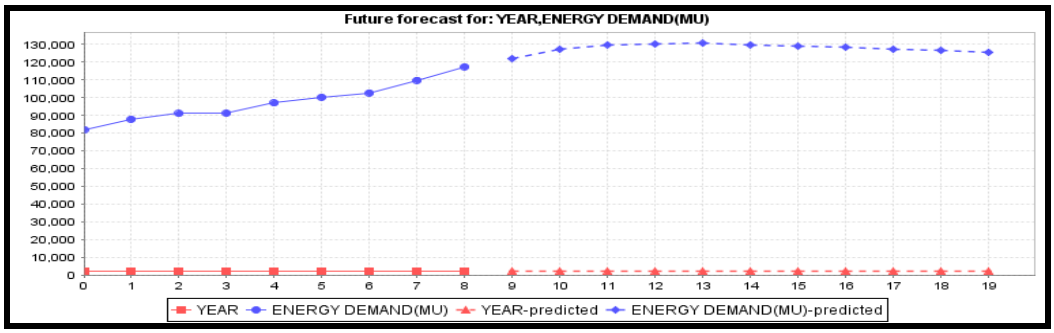

Figure 6 : Predicted Demand (MU)

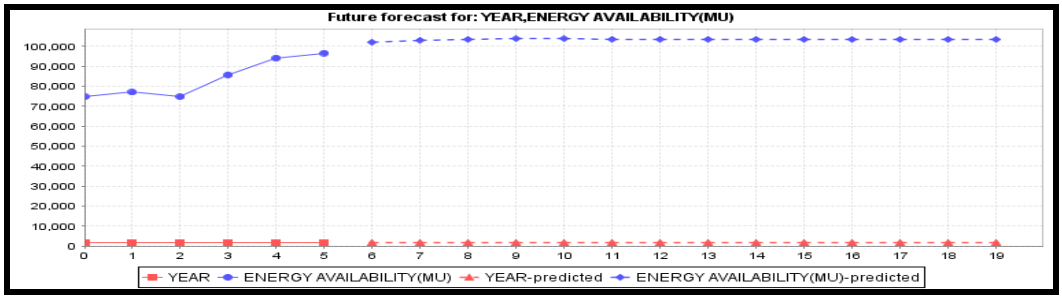

Figure 7: Predicted Availability (MU)

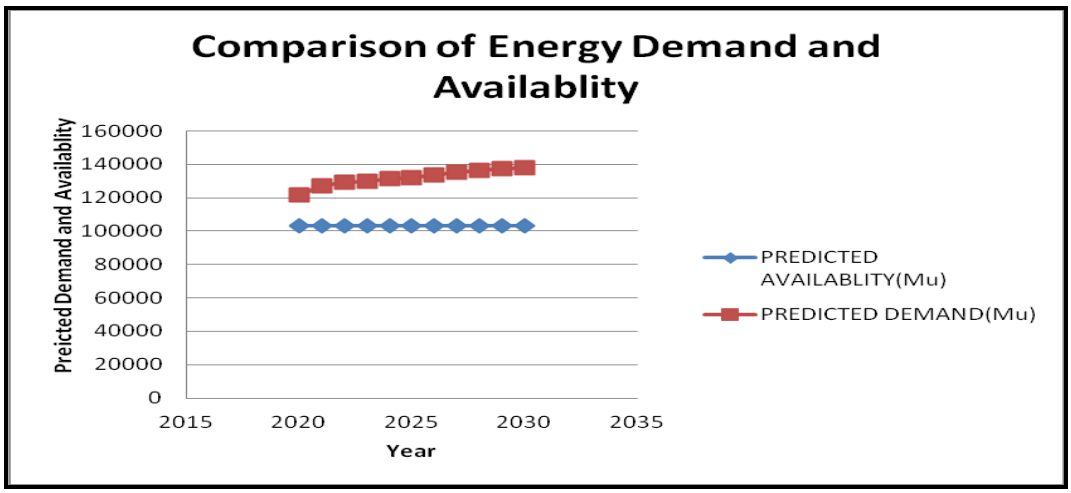

Figure 8: Comparison of Predicted Demand and Predicted Availability

Analyzing on the results of Predicted Demand and Predicted Availability it is found that Predicted Demand has the highest value compared to Predicted Availability. Hence there should be an improvement in power generation to meet over the Energy Demand. As previously analyzed from Table-2 Solar is having higher scope compared to Biomass. Development on solar power generation will help to meet over the future Energy Demand.

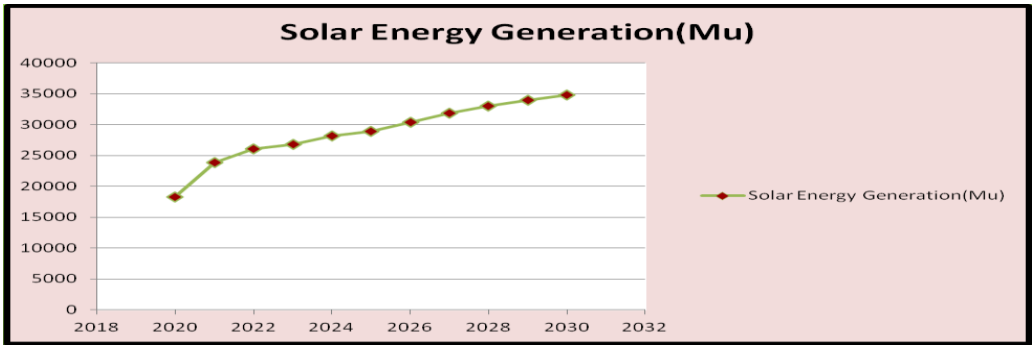

Figure 9: Solar Power generation to meet future demand

5. Results and Discussion

The analyzes was to identify which energy resource is going to have future scope in power generation. By applying Naive Bayesian classification, which energy resource has to be developed is identified. It is found that Solar and Biomass is under development stage by applying EM algorithm which resource will have faster development it is found that Solar has a higher growth compared to Biomass. A comparative analysis was applied to various Time series forecasting algorithm, in order to find which algorithm is the best algorithm to predict the future power demand and availability by using evaluation parameters. It is found that Multilayer Perceptron has the lesser error value compared to other algorithms. Hence Multilayer Perceptron is selected as best algorithm for forecasting. Applying Multilayer Perceptron algorithm Energy demand and energy availability is predicted for the future.

6. Conclusion

The result of predicted Energy demand and availability proves that Energy demand is higher than the Energy availability. Hence the solution for energy resource for future demand is identified to meet over the energy demand. The analyzes proves that Solar power generation is one of the best resource to meet over the future demand. The benefits of Solar power generation does not produce solid, radioactive wastages and harmful gases. Solar Energy is a non-polluting, reliable and clean source of energy. In Economical perspective, tax incentive, an elimination of electricity bills, increased property value and high durability. Solar energy will become available, accessible and affordable to all citizens.

References

A. M Anita Priscilla Mary, MS Josephine, V Jeyabalaraja (2020), A Comparative Study and Analysis of Classification Methodologies in Data Mining for Energy Resource.In:Peng SL.,Son L., Suseendran G.,Balaganesh D., (eds) Intelligent Computing and Innovation on Data Science. Lecture Notes in Networks and Systems, Vol 118. Springer,Singapore.

B. M.Anita Priscilla Mary , Dr.M.S.Josephine , Dr.V.Jeyabalaraja, Dr.S.Nirmala Sugirtha Rajini. (2020). Identification and Performance Evaluation for Effective Utilization of Electrical Energy Resource using K-Means Clustering Algorithm. International Journal of Advanced Science and Technology, 29(9s),55-62.Retrievedfromhttp://sersc.org/journals/index.php/IJAST/article/view/13001.

C. Monire Norouzi, Alireza Souri, Majid Samad Zamini(2016) A Data Mining Classification Approach for Behavioral Malware Detection, Journal of Computer Networks and Communications, Journal of Computer Networks and Communications Volume 2016, Article ID 8069672, 9 pages http://dx.doi.org/10.1155/2016/8069672,Hindawi Publishing Corporation.

D. Neves-silva, R., & DEE Group Author (2012). Stochastic models for building energy prediction based on occupant behavior assessment. *Energy and Buildings*, *53*(NA), 183-193. https://doi.org/10.1016/j.enbuild.2012.06.001.

E. João Virotea, Kwok-Wing Chau ,Chun-Tian Cheng, Lin Qiu(2009) A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series, Journal of Hydrology, Volume 374, Issues 3–4, 15 August 2009, Pages 294-306.

F. Pillai N, Vijayamohanan. (2008). Forecasting Demand for Electricity: Some Methodological Issues and an Analysis..

G. Eduardo A. Gerlein, Martin McGinnity, Ammar Belatreche, Sonya Coleman

H. Evaluating machine learning classification for financial trading: An empirical approachExpert Systems with Applications, Volume 54, 2016, pp. 193-207

I. N. Kaur and A. Kaur, 2016 "Predictive modelling approach to data mining for forecasting electricity consumption", 6th International Conference - Cloud System and Big Data Engineering (Confluence), Noida, 2016, pp. 331-336.

J. Kishore1 GDK, Dr. Babu Reddy2 M (2017) Comparative analysis between classification algorithms and data sets (1: N & N: 1) through WEKA. ISSN (Online) 2456–3293

K. Data Mining: Concepts And Techniques 3rd Edition Authors: Jiawei Han Micheline Kamber Jian PeiHardcover ISBN: 9780123814791 eBook ISBN: 9780123814807 Imprint: Morgan Kaufmann Published Date: 22nd June 2011Page Count: 744.

L. Mustafa Göçken, Mehmet Özçalıcı, Aslı Boru, Ayşe Tuğba Dosdoğru Integrating metaheuristics and Artificial Neural Networks for improved stock price predictionExpert Systems with Applications, Volume 44, 2016, pp. 320-331

M. ian-Zhou Wang, Ju-Jie Wang, Zhe-George Zhang, Shu-Po GuoForecasting stock indices with back propagation neural networkExpert Systems with Applications, Volume 38, Issue 11, 2011, pp. 14346-14355

N. Nikolaos Kourentzes, Devon K. Barrow, Sven F. Crone Neural network ensemble operators for time series forecastingExpert Systems with Applications, Volume 41, Issue 9, 2014, pp. 4235-4244
O. Szarek D., Bielak, Wyłomańska A.Long-term prediction of the metals' prices using non-Gaussian time-inhomogeneous stochastic processPhysica A: Statistical Mechanics and its Applications, Volume 555, 2020