ISSN : 1978-8282, Online ISSN: 2655-4275

Electrical Data Classification with LVQ to Optimize Power Efficiency and Cost

Indrianto*[1], Sarifuddin Madenda[2], Prihandoko[3], Nur Azizah[4]
[1] Fakultas Telematika Energi, PLN Technology Institute, Indonesia
[2,3] Faculty of Computer Science and Information Technology, Gunadarma University, Indonesia
[4] Faculty of Science and Technology, University of Raharja, Indonesia
E-mail: *[1]**indrianto@itpln.ac.id**, [2]sarif@staff.gunadarma.ac.id, [3]prihandoko@gmail.com, [4]nur.azizah@raharja.info

Abstract

Rapid population development causes such a high energy demand among the community in Indonesia, especially those which concern with electrical energy. The use of electrical energy must be supported by good technology in the provision to distribution of electricity. With energy arrangements, energy can be saved and financing will be lighter. Energy saving can be done by making an energy management application based on incoming data from customers. The LVQ method in this study is used to classify customer data that enters the server. The data comes in low, medium, and high usage categories. The findings of this study can be further processed to enhance the classification results obtained from the LVQ method, allowing for a more optimal classification that maximizes energy efficiency. The results indicate that the LVQ algorithm achieved an accuracy rate of 91.66%. The highest electricity consumption occurred between 5:30 PM and 10:00 PM, while the lowest consumption was recorded between 12:30 AM and 8:00 AM, as well as from 11:00 PM to 12:00 AM. Meanwhile, medium electricity loads were observed between 8:30 AM and 5:00 PM. This is crucial for managing energy resources to ensure efficiency and optimization.

Keywords *— LVQ, Data Electricity, Management Energy*

1. INTRODUCTION

The introduction outlines the background of the problem being solved, the issues related to the problem being solved, the research reviews that have been done before by other researchers that are relevant to the research conducted. Electricity is an essential necessity for society that must be provided by the government [1]. People need electricity for cooking, household needs, and so on [2,3]. Because it is one of the important needs, many sources of electricity are needed, be it from fossil energy such as hydroelectric, wind or other power plants, or from other sources with the theme of renewable energy such as solar, wind, garbage, and others. Currently, the population growth in the world is so large it is estimated that in 2050, the population growth will increase dramatically. For example, Africa 40%, Asia 47.5%, and Europe 73.4%. Large population growth requires a large amount of food, clothing, and energy. Thus, population growth affects energy use, especially electrical energy.

Fossil-based electrical energy usually requires fuel to supply machines that will produce energy [4]. The more electrical energy released, the more fuel needed to drive the engine [5]. For this reason, a digitization of these plants is needed [6,7]. With this digitalization, it can be monitored and the production costs in each day can be calculated [8,9]. Thus, digitalization is very beneficial for life, especially for generators in power plants. For this reason, it is the focus of Indonesian National Electrical Company (PJL) in modernizing the energy sector in Indonesia [10]. The digitalization of building energy will incorporate smart technologies, including Machine Learning (ML) and the Internet of Things (IoT) [11,12]. With this technology, using sensor-based remote controls, monitoring and controlling process can be carried out more efficiently and provides better predictions in plant maintenance and performance. This can be taken as an effort to handle the operational efficiency of power plants in Indonesia [13]. Thus, digitalization can enable better integration between different power plants and can improve overall efficiency [14].

The use of digital technology in power plant generation in Indonesia is currently only used in mobile applications such as mobile applications for customers, smart electricity, and others [15]. Actually, the use of algorithms such as machine learning or deep learning can be applied to power plants, although for now it is still under development and research [16,17]. Most power plants in Indonesia still use conventional power such as water and fossil (fuel oil). Power plants in Indonesia are mostly fossil fuel plants [18]. So that it becomes an obstacle in carrying out the digitization process and the cost of electricity production becomes high. To be able to save electricity production costs, it can be done by digitalization. Machine learning and deep learning can be used in digitizing these plants. The processed data can be in the form of data derived from the history of electricity use or data derived from existing sensors [19].

Machine learning usage is expected to help digitize power plants. LVQ (Learning Vector Quantization) is a supervised learning algorithm used for classification [20,21]. The data inputted on the LVQ model will be trained so that the algorithm will determine the class on the nearest prototype. The distance matrix will be used to measure the similarity between the sample data and the prototype. Thus, LVQ can be used to classify the power produced by plants so that the fossil fuels used by these plants can be saved. It can also be cost-effective and cut the operational cost to produce for 30 minutes everytime the plant operates [22]. So, the existing data is processed by making a data management that can produce data classifications in the form of High, Medium, and Low usage. With a classification like this, it can be saved the use of fuel to produce electrical energy. This is because if the use of high engine use can be maximized (3 plants are turned on), at medium times the use of the machine can be used half of the existing engine (2 plants are turned on) and at low times it can be much smaller (only 1 plant is turned on). Thus, the use of fuel can be saved, this is because the energy released is adjusted to the usage of plants available. This method can also enable us to calculate the cost of operation per every 30 minutes.

Based on the matter above, the researchers tried to conduct research on data processing from the electrical usage data at the PJL substation. The data that the researchers get is derived from power usages in Java and Bali. The data comes from PJL statistical data. The data is processed using LVQ to produce a classification that will be used to recommend for plant arrangements (plants 1, 2, and 3) that will supply electricity as needed. LVQ is an algorithm that is widely capable and can be used to classify data that has the potential to be high-dimensional [23]. LVQ will produce a classification based on the High, Medium, and Low electricity usage. The results of LVQ can be used to calculate existing production costs.

Research conducted in the last 5 years based on the use of LVQ in power plants can be seen at [24-30]. The utilization of LVQ for other fields are from [37-39], energy management using LVQ is taken from [35], and the use of LVQ for classification above 80% is from [40]. From the last 5 years of research, it can be seen that the use of LVQ to classify electricity data based on the history of electricity usages is still rare, as well as the estimated price of electricity production. From the reference above, the difference between the research carried out is that researchers make detections based on data stored in the database. Then the database is processed using the LVQ algorithm, by the author classifying based on the electrical power contained in the database with the categories of Low Power, Medium Power and High Power.

2. RESEARCH METHOD

Research Methods (can include analysis, architecture, methods used to solve problems, implementation), in this discussion the author can describe how the research will be carried out.

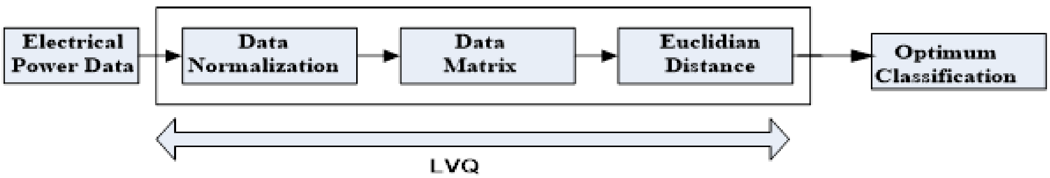

Figure 1. Research Methods

The structural Learning Vector Quantization (LVQ) is one of the classification methods of Artificial Neural Networks. LVQ works with each output unit presenting a class [41]. In simple terms, LVQ is a classification method where the target or number of groups is predetermined. The main objective of this algorithm is to approximate the distribution of vector classes, thereby reducing classification errors. Learning Vector Quantization (LVQ) operates as a supervised learning technique within a competitive layer. This competitive layer automatically learns to categorize input vectors. The classification results from this layer are determined solely by the proximity between input vectors. If two input vectors are close to each other, they will be assigned to the same class by the competitive layer.

The steps involved in the LVQ training algorithm are as follows:

1. Initialize the weight values Wj and the initial learning rate $\alpha(1)$
2. Repeat the following steps while the stopping condition has not been met
3. For each training input vector x, proceed with steps 4-5
4. Determine j such that $|x-wj|$ is Minimized
5. Adjust the weight Wj using the following update rule:
   If T = Cj then :

$$Wj(t+1) = Wj(t) + \alpha(t)[x(t) - Wj(t)] \qquad (1)$$

If T≠Cj then :

$$Wj(t+1) = Wj(t) - \alpha(t)[x(t) - Wj(t)] \qquad (2)$$

Reduce the average learning α

6. Stopping Condition Test with

X	: The set of training vectors (X1,... Xi... Xn)
T	: The correct category or class assigned to the training vectors
Wj	: The weight vector at the j-th output unit (W1j,... Wij,...,Wnj)
Cj	: The category or class associated with the j-th output unit
‖x-wj‖	: The Euclidean distance between the input vector and the corresponding weight vector for the jth output unit.

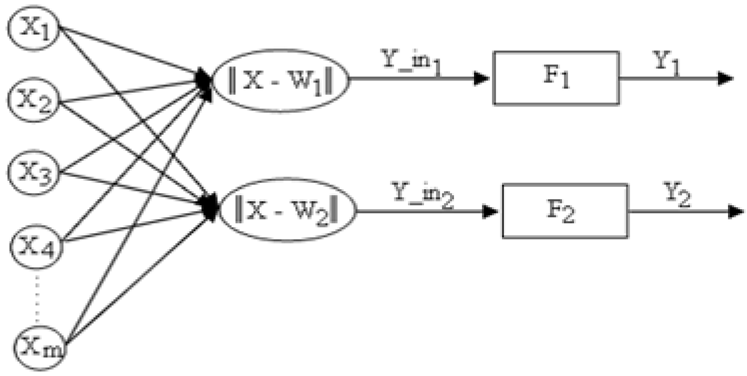

Figure 2. LVQ

The description of Figure. 2 is as follows X represents the input, while W1 is the weight vector that links each neuron in the input layer to the first neuron in the output layer. Similarly, W2 is the weight vector that connects each neuron in the input layer to the second neuron in the output layer. The LVQ algorithm operates by calculating the Euclidean distance between input vectors. If the computed distance is the shortest or smallest among all comparisons, the algorithm assigns the data to the class or target associated with that minimum distance.

3. RESEARCH RESULTS AND DISCUSSION

3.1. RESULT

3.1.1. Design

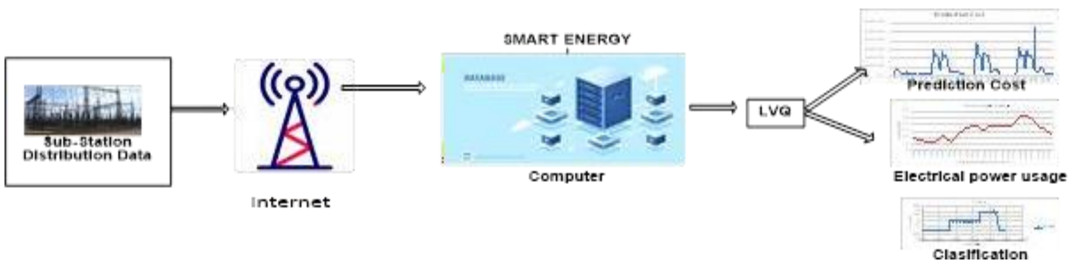

Figure 3. Design

Data coming from the substation is taken and sent to the server via the internet. Then, as seen in the Fig.3. image. Data is stored on a server which is then processed, containing power consumption data every 30 minutes, then with Learning Vector Quantization (LVQ) the data is classified according to High, Medium, and Low usages. The results of LVQ can be used to calculate the amount of power and classified as well as the cost of production every 30 minutes.

3.1.2. Algoritma LVQ

The processing steps in learning/training on the LVQ method can be illustrated through the Algoritma below:

Algoritma : LVQ Algorithm

Input : Vector
Output : Weight

Random weight initialization
Set the α (learning rate), maximum number of epohs, and decrease α For each
epoh until it reaches the maximum number of epohs:
Random training data sequence For
setiap data training (x, y):
Find the best neurons:
Calculate the distance between the input vector x and each weight vector in
the layer
Select the neuron with the closest weight to the input (minimum distance) Calculate
y_pred prediction value using the best neurons
Hitung error: error = y - y_pred For
each wi weight on the best neurons:
Update weight: wi = wi + α * error * xi

Update neighbor weight: wi_neighbor = wi_neighbor - α
* error * xi
Update α: α = α – decrease α

Check the stop criteria (e.g., achieve a certain error tolerance or reach the maximum number
of epohs)
If the stop criteria are met, exit the loop
If not, move on to the next epoh

In Algorthm. We can see. If the epoch value is less than the maximum epoch value, the distance calculation process will be carried out with input vectors and weights using the Euclidean formula and the results of these calculations will be seen the smallest value or distance from the input vector with each weight, this smallest distance will be processed to obtain the final weight value.

The next step that will be done after one epoch or one iteration is complete is to reduce the learning rate (α) by 0.1 in the initial learning rate value. The process continues until the nth data (xn) until it finds the final weight that will be used as a weight for further data testing. The data used in this study is data on the load power consumption. The data variable is power load consumption. The data is grouped into variables with each calculated starting at 00.30 a.m. to 12.00 p.m with a time gap of 30 minutes. In addition to input data, the LVQ method target/desired class has been determined in advance, where the class in this power setting there are 3 target classes, namely class 1. Low Power, 2. Medium Power, and 3. High Power. The target class conditions obtained will later be used to determine the class to be used. Power that must be expended for consumption needs. The following is the data that will be processed in the learning vector quantization method with a time range of 30 minutes.

3.2. DISCUSSION

In the data obtained above, the next stage is to normalize the data with the aim of obtaining data with smaller values with a limit of 0-1. The data normalization stage is carried out before entering the data training process stage. Data normalization is also performed on training data and test data. Based on the experimental results, Minimum value x = 11711, Maximum value x = 17503. By using normalization data, weights will be obtained with graph values as follows:

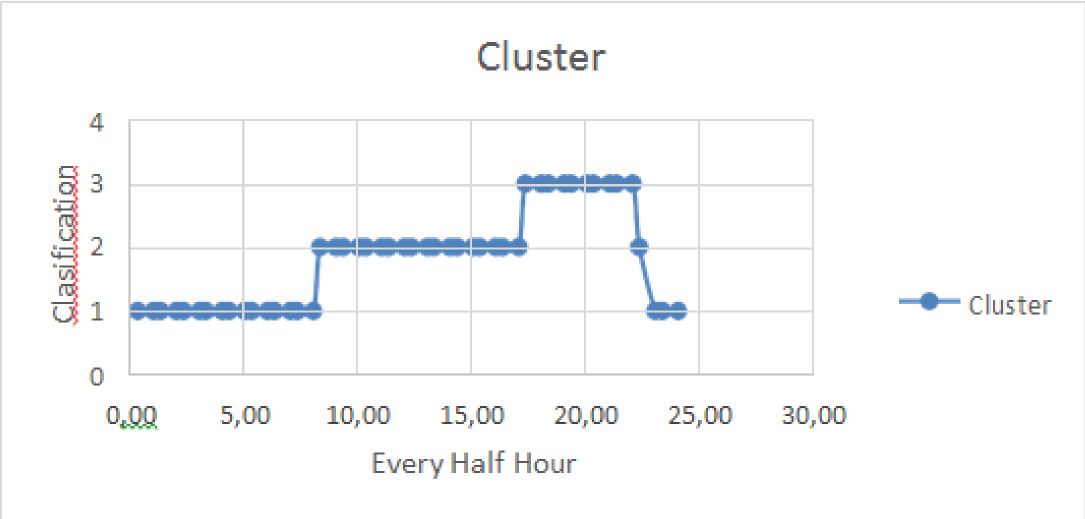

Figure 4. Cluster

With cluster division in Fig.4 it will be obtained. High Cluster On high clusters, will activate 3 generator, Intermediate Cluster On an intermediate cluster, will activate 2 generators., and Low Cluster On low clusters, will activate 1 generator. Above it can be seen that the highest load is at 5.30 p.m. to 10.00 p.m. and the lowest is at 12.30. a.m. to 08.00 a.m. and 11.00 a.m. to 12.00 p.m., and the medium load is at 8.30 a.m. to 5.00 p.m.

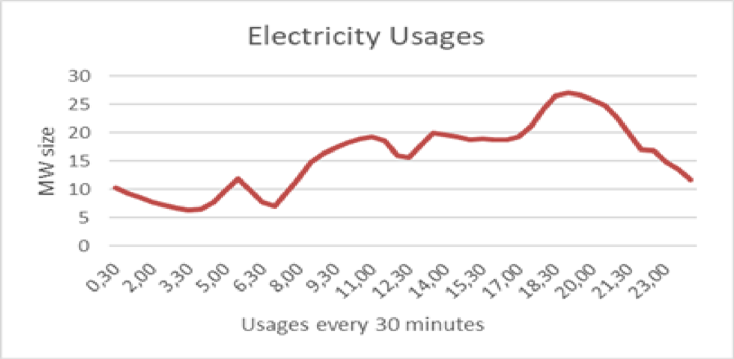

Figure . 5 Electricity Consumption

With power consumption as seen in Fig.5. It can be seen that the highest energy use is at 05.00pm to 10.00pm with energy use reaching 25 MW.

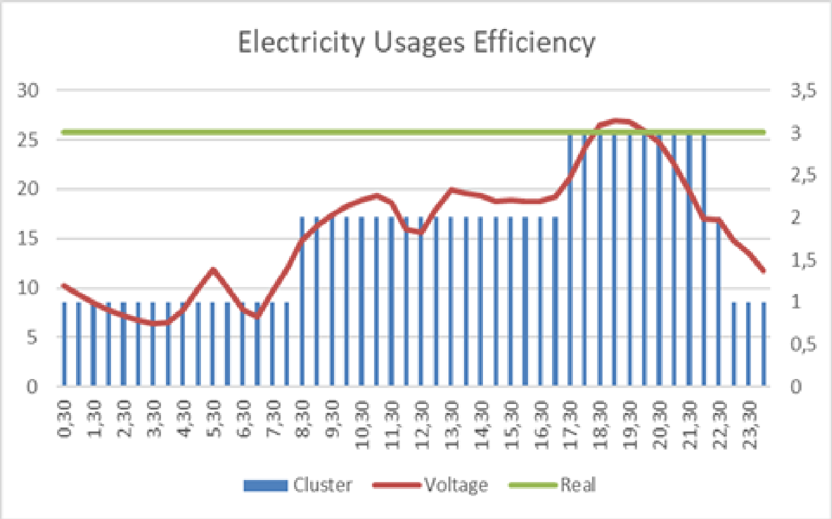

Figure. 6. Efficiency of Electricity Use

Based on Fig.6. The highest efficiency occurs at 10:30 p.m. to 4:30 p.m. and the lowest occurs at 5:30 p.m. to 9:30 p.m. So that the savings from the power used are available at 11.30 p.m. to 4.30 p.m.

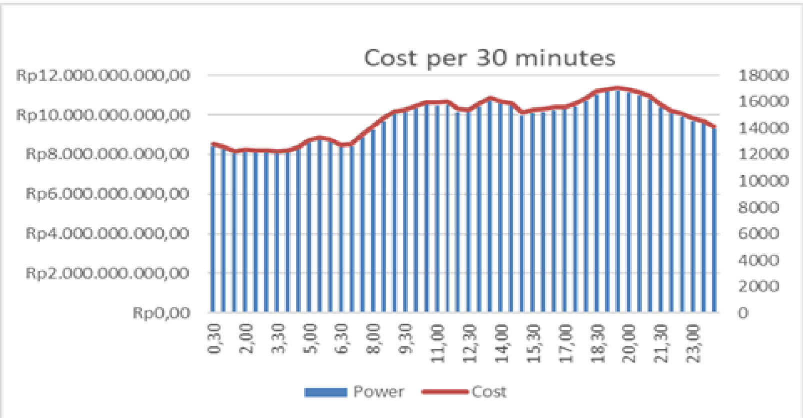

Figure. 7. Cost Per 30 minutes

On Fig. 9. It can be seen that the lowest cost usage is at 11:30 p.m. to 4:30 p.m. and the average savings per day for half an hour is IDR 9,840,532,000, assuming the cost per KWh is IDR 600.

As In this study, the level of accuracy was measured using testing techniques from the results of power classification with the confussion matrix method. Accuracy in the confussion matrix serves to predict the accuracy of system modeling that has been made by measuring the level of proximity of the actual value to the predicted value. The accuracy calculation used in this study is by means of multiclass calculations or accuracy calculations involving more than 2 classes in the confussion matrix. In this study, the training data used was power consumption data in the month with a total of 31 days and 48 hours. In the training process, training data is divided into 3 classes where testing is carried out by calculating the level of accuracy of the available data. The following is an analysis table of the confussion matrix accuracy process. The calculation of the accuracy of the above data using the results of the classification detected is correctly in the same class (TP) added with the classification detected not being in the same class (TN). The sum result is divided by the values of TP, TN, FP, and FN. The result of classification testing accuracy with March data is 91.66%.

4. CONCLUSION

From the results of the simulations that have been carried out, it can be seen that all existing data has been structured in their respective classes, so that they can be easily understood. Each class indicates the level of electricity consumed by the community. Thus, the Learning Vector Quantization (LVQ) method can help facilitate the processing of large amounts of data just by making a classification of each data according to needs. In this study, the accuracy rate is 91.66% and the highest usage is at 5.30 p.m. to 10.00 p.m. and the lowest are at 12.30 a.m. to 08.00 a.m. and 11.00 p.m. to 12.00 p.m., and medium load is at 8.30 a.m. to 5.00 p.m.

5. SUGGESTED

This research certainly has many shortcomings. For this reason, this research can be developed further, especially in the use of data. The authors wish to thank Insitut Teknologi PLN and Universitas Gunadarma.

6. REFERENCES

[1] Y. Liu, L. Wu, and J. Li, “Peer-to-peer (P2P) electricity trading in distribution systems of the future,” Electr. J., vol. 32, no. 4, pp. 2–6, 2019.

[2] J. S. Chou and N. S. Truong, “Cloud forecasting system for monitoring and alerting of energy use by home appliances,” Appl. Energy, vol. 249, no. February, pp. 166–177, 2019.

[3] X. Hou, J. Wang, T. Huang, T. Wang, and P. Wang, “Smart Home Energy Management Optimization Method Considering Energy Storage and Electric Vehicle,” IEEE Access, vol. 7, pp. 144010–144020, 2019

[4] V. Foster, J.-P. Tre, Q. Wodon, and W. Bank, “Energy prices, energy efficiency, and fuel poverty,” (Unpublished Pap. Work. Bank, no. September, 2000.

[5] Piotr F. Borowski, “Digitization, Digital Twins, Blockchain, and Industry 4.0 as Elements of Management Process in Enterprises in the Energy Sector,” energies, pp. 109–133, 2021.

[6] A. J. Isaksson, I. Harjunkoski, and G. Sand, “The impact of digitalization on the future of control and operations,” Comput. Chem. Eng., vol. 114, pp. 122–129, 2018.

[7] M. Porthin, M. Liinasuo, and T. Kling, “Effects of digitalization of nuclear power plant control rooms on human reliability analysis – A review,” Reliab. Eng. Syst. Saf., vol. 194, 2020.

[8] P. E. D. Love and J. Matthews, “The ‘how’ of benefits management for digital technology: From engineering to asset management,” Autom. Constr., vol. 107, no. July, 2019.

[9] P. Augustine, The industry use cases for the Digital Twin idea, 1st ed., vol. 117, no. 1. Elsevier Inc., 2020.

[10] M. Sharifzadeh, A. Sikinioti-Lock, and N. Shah, “Machine-learning methods for integrated renewable power generation: A comparative study of artificial neural networks, support vector regression, and Gaussian Process Regression,” Renew. Sustain. Energy Rev., vol. 108, no. July 2018, pp. 513–538, 2019.

[11] M. Sharifzadeh, A. Sikinioti-Lock, and N. Shah, “Machine-learning methods for integrated renewable power generation: A comparative study of artificial neural networks, support vector regression, and Gaussian …,” Renew. Sustain. …, 2019.

[12] S. Tian, W. Yang, J. M. Le Grange, P. Wang, W. Huang, and Z. Ye, “Smart healthcare: making medical care more intelligent,” J. Glob. Health, vol. 3, no. 3, pp. 62–65, 2019.

[13] M. S. Ibrahim, W. Dong, and Q. Yang, “Machine learning driven smart electric power systems: Current trends and new perspectives,” Appl. Energy, 2020.

[14] C. Magazzino, M. Mele, and N. Schneider, “A machine learning approach on the relationship among solar and wind energy production, coal consumption, GDP, and CO2 emissions,” Renew. Energy, vol. 167, pp. 99–115, 2021.

[15] S. Abbas et al., “Modeling, Simulation and Optimization of Power Plant Energy Sustainability for IoT Enabled Smart Cities Empowered with Deep Extreme Learning Machine,” IEEE Access, vol. 8, pp. 39982–39997, 2020.

[16] Y. Y. Hong and C. Rioflorido, “A hybrid deep learning-based neural network for 24-h ahead wind power forecasting,” Appl. Energy, 2019.

[17] L. Cheng and T. Yu, “A new generation of AI: A review and perspective on machine learning technologies applied to smart energy and electric power systems,” Int. J. Energy Res., 2019.

[18] H. Lubis, “Renewable Energy of Rice Husk for Reducing Fossil Energy in Indonesia,” J.

Adv. Res. Appl. Sci. Eng. Technol., vol. 11, no. 1, pp. 17–22, 2018.

[19] D. Hartono, S. H. Hastuti, A. Halimatussadiah, A. Saraswati, A. F. Mita, and V. Indriani, "Comparing the impacts of fossil and renewable energy investments in Indonesia: A simple general equilibrium analysis," Heliyon, vol. 6, no. 6, p. e04120, 2020.

[20] N. Qu, J. Zuo, J. Chen, and Z. Li, "Series Arc Fault Detection of Indoor Power Distribution System Based on LVQ-NN and PSO-SVM," IEEE Access, vol. 7, pp. 184020–184028, 2019.

[21] A. Dharma, N. K. Mataram, and I. N. G. Antara, "Learning Vector Quantization Neural Network (LVQNN): Can it be implemented on the Forecasting Electrical Loads?," IOP Conf. Ser. Earth Environ. Sci., vol. 248, no. 1, 2019.

[22] M. Mishra, "Power quality disturbance detection and classification using signal processing and soft computing techniques: A comprehensive review," Int. Trans. Electr. Energy Syst., vol. 29, no. 8, 2019.

[23] C. Chupong and B. Plangklang, "Incremental Learning Model for Load Forecasting without Training Sample," Comput. Mater. Contin., vol. 72, no. 3, pp. 4515–4527, 2022.

[24] G. Kumar, S. Sharma, and H. Malik, "Learning Vector Quantization Neural Network Based External Fault Diagnosis Model for Three Phase Induction Motor Using Current Signature Analysis," Procedia Comput. Sci., vol. 93, no. September, pp. 1010–1016, 2016.

[25] J. U. Bhuiyan, T. A. Begum, and M. R. Alam, "S-Transform and mahalanobis distance based approach for classifying power quality disturbances," ECCE 2017 - Int. Conf. Electr. Comput. Commun. Eng., pp. 681–685, 2017

[26] M. Sardogan, A. Tuncer, and Y. Ozen, "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm," UBMK 2018 - 3rd Int. Conf. Comput. Sci. Eng., pp. 382–385, 2018.

[27] O. V. Mandrikova, V. V. Geppener, and B. S. Mandrikova, "Method of analysis of cosmic ray data based on neural networks of LVQ," J. Phys. Conf. Ser., vol. 1368, no. 5, 2019.

[28] A. Kumar and K. Kaur, "A Novel MCDM-Based Framework to Recommend Machine Learning Techniques for Diabetes Prediction," pp. 1–15, 2023.

[29] Abdurrasyid, M. N. Indah Susanti, I. Indrianto, and K. Zhafira, "Acute Respiratory Infections Diagnosis Using Learning Vector Quantization," ICCoSITE 2023 - Int. Conf. Comput. Sci. Inf. Technol. Eng. Digit. Transform. Strateg. Facing VUCA TUNA Era, pp. 656–661, 2023.

[30] M. Straat et al., "Learning vector quantization and relevances in complex coefficient space," Neural Comput. Appl., vol. 32, no. 24, pp. 18085–18099, 2020.