Indonesian Journal of Data and Science

Volume 5 Issue 2
ISSN 2715-9936
https://doi.org/10.56705/ijodas.v5i2.130

Research Article

A Survey on Machine Learning Techniques for The Prediction of Solar Power Production

L. O. Lamidi [1], A. M. Oyelakin [2,*], M. B. Akanbi [3]

[1] Part Time Lecturer in Physics, Centre for Part Time and Professional Studies. Al-Hikmah University, Ilorin, Nigeria, -
[2] Department of Computer Science, Crescent University, Abeokuta, Nigeria, moruff.oyelakin@cuab.edu.ng
[3] Department of Computer Science, Kwara State Polytechnic, Ilorin, Nigeria, -

Correspondence should be addressed to A. M. Oyelakin; moruff.oyelakin@cuab.edu.ng

Received 12 June 2024; Accepted 20 July 2024; Published 31 July 2024

Abstract:

Renewable energy sources are needed globally to support the available non-renewable energy sources our day-to-day living. There is high demand for renewable energy sources in both the developed and developing economies. Solar power is a good example of renewable energy source and people are currently embracing it globally for both domestic and industrial uses. Generally, these energy sources are meant to support the hydro, thermal and other energy sources that are available in different countries of the world. With the popularity of solar energy for both domestic and industrial usage, it can be argued that the estimation of the production level of such energy source is necessary so as to achieve proper planning and management. Due to the fact that the availability of the solar energy power depends largely on a number of environmental and weather conditions, predicting its production or generation can be very important. This study surveyed different works in the area of using machine learning techniques for solar power production prediction. The articles sourced were from notable research repositories. This study focuses on articles that were published between 2013 and 2023 on the subject matter. Different types of machine learning (ML) algorithms that have been used to build models from solar energy datasets are reported in this study. It is believed that the work can provide better insights for the researchers working in the problem area. Thus, the insights in this study can lead to building of improved machine learning-based models for solar power forecasting.

Keywords: Energy Forecasting, Solar Energy Generation, PV panel, Machine learning, .
Dataset link: -

1. Introduction

Renewable energy sources are needed worldwide so as to support the available non-renewable energy ones in this present time we find ourselves. With the popularity of solar energy for both domestic and industrial usage, it can be argued that the estimation of the production level of such energy source is necessary so as to achieve proper planning and management. It is essential to know the availability of the energy sources in order to achieve proper planning and administration [1].Solar power is an alternative energy that is produced through the use of photovoltaic (PV) systems. As argued by [2], the generated power of these PV systems is dynamic and depends on environmental situations. Several ML approaches have been used for the planning of the solar-based energy prediction using promising features. [2] identified irradiance, humidity, PV surface temperature, and wind speed as some of the attributes for the energy prediction task. The argument was based on the importance of the identified features.

Aside this, [3] pointed out that solar power suffers from uncontrollable fluctuation since it is highly depending on other weather variables. And thus, there is need to predict its production accurately. ML algorithms are classified as supervised, unsupervised or semi-supervised [4]. The choice of which type of algorithm to make use of depends on the problem and the innovative approach of the researcher. [5] pointed out the importance of solar energy and then

argued the promising of building ML-based models based on solar power data. Learning models were built for the prediction of power production and the results were found promising. ML involves training an algorithm to perform tasks by learning from data patterns rather than explicitly programming systems to do so [4], [6].

[7] also mentioned that solar power forecasts are applied so as to have adequate management of the electricity grid and for trading of power. It has, however been identified that there are drawbacks of the solar photovoltaic (PV) energy source which can as a result of continuous changes of the climatic conditions. Apart from this, it has been found out that solar energy output power prediction is a hot research area in the contemporary society [8]. This study seeks to identify and survey the various ML-based methods that have been used for the prediction of solar energy production or generation. [9] have emphasised the trends and gaps that exist in ML-based solar energy forecasting. This study is targeted at having a review of literature that can provide further directions on this research area. Thus, the study carried out a survey of various machine learning approaches that have been used for the prediction of solar energy production.

Machine Learning

ML algorithms "learn" and are used to build models from data by focusing on a specific problem and identifying patterns from the data. Using large amounts of data, learning algorithms are used to be models that are able to accurately describe the data it's ingesting. In this work, emphasis is on regression problems that are purely in the area of solar power production prediction. The problem that can be solved in using ML techniques can be classification, Regression or Clustering based. [10] and [11] identified different machine learning algorithms based on the categories of supervised, unsupervised and semi-supervised. The authors equally emphasised different problems: classification, regression and clustering problems that are abound in ML field. The solar power prediction problem in this paper is regression-based. In this study, it has to be pointed out that the use of ML techniques for solar power prediction is a problem that is based on regression analysis. The period that the analysis covers depend largely on the kind of dataset that is being used for the analysis.

2. Method:

This study sourced for literature from different research repositories using some keywords relevant to the problem under consideration. Thereafter, the relevant studies were surveyed. About seventy-two (72) literature items on the subject-matter were downloaded from the different research sources. Then, the works that were deemed very close and relevant in this study was sorted out and were used in both the conceptual review and empirical review of this paper. That amounts to about thirty studies. Then, the literature sourced was organised under the survey section by arranging them from most recent to the fairly old ones. The year covered in this study is from 2013 till 2023.

Research Sources/Repositories Used

The research repositories or sites used for sourcing for the relevant studies include: Science Direct, Research gate, IEEE Explore, ACM Conferences, Google search engine, and Google scholar.

Keywords used for the searches

The searches were conducted using a wide range of search strings. The search phrase used was: ("solar power" OR "solar energy"), ("machine learning techniques for forecasting solar power production" OR "machine learning models for predicting solar energy production") and ("solar power production" OR "solar power generation"). This search phrase was considered so as to obtain a good number of relevant studies in research repositories for the survey being carried out. According to [3], renewable energy in different nations is regarded as a good source of electricity production that is more reliable. This study focused on studies or literature that were published in the English language and contained in journals (both printed and electronic), white papers, conference proceedings, and books.

Surveyed Works

[9] classified literature that focuses on ML-based solar energy forecasting. The approach used is to categorise the materials based on different aspects and attributes relevant to PV systems. [12] reported recent studies that used shallow and deep learning machine learning approaches to forecasting Renewable Energy Generation. The work equally provided the trends and future prospects that can be explored. [13] built eight different machine learning models for the prediction of solar power generation. The study achieved good performances both in each of the models and on the average. [14] surveyed works that reported current developments in the field. Specifically, the researchers examined data cleaning methods, learning algorithms, selection of parameter and energy prediction model performance assessment.

[15] proposed ML-based models for solar energy prediction. Four various regression models were built for solar power prediction by the authors. They claimed that the performances of the models are promising. However, it was argued that the Elastic regression model outperforms others. Also, PCA was applied to the dataset within the scope of the study, and better results were obtained in the elastic regression model. [16] predicted power generation in a solar plant with the use machine learning algorithms. The accuracy of forecasting was checked directly with the practical data which is generated and simulated data using MATLAB/Simulink by applying various machine learning algorithms. The study argued that ML approach used was promising for the problem at hand.

[17] built solar power forecasting models using two selected learning algorithms. The models were evaluated and compared. The author mentioned that the results were promising. Similarly, [18] proposed a deep learning (DP) technique for solar power prediction. The DL method is based on the Long Short-Term Memory (LSTM) algorithm. The prediction result shows that the LSTM network produced the best results for each category of days.

[1] proposed four solar power prediction models using four different ML algorithms for the solar energy hourly prediction in Nigeria. The analysis was carried out in different climatic zones of Nigeria. It was claimed that Support Vector Machine has the best performance at twenty (24) hours ahead while Decision Tree showed the worst performance. Authors pointed out that the best results were arrived at using Lagos data while poor results were recorded with Abuja data.

[19] used some ML and DL algorithms for building solar power prediction models. The dataset that was used for the study was collected for a period of one month. The performances of the various algorithms were compared using the Root Mean Square Error (RMSE) values as metric. Furthermore, [14] employed ensemble of machine learning approach for building a model that can predict solar energy production. The authors argued that the ensemble model performed better than the conventional individual models. They further argued that the ML-based model outperformed the other category.

[20] presented learning methods for the prediction of the power output of solar panels using weather and air pollution attributes. The authors used different features, such as weather, air pollution, and weather and air pollution for the predictive analysis and concluded that the approach is good. [21] performed a comparative study has been made between the different machine learning techniques. The authors argued that the experimental and simulation results show that the proposed methods can be utilised to successfully predict the solar power output for a short duration of time. The study also reveals that ANN and GBRT performed better than KNN.

Also, [2] investigated the influence of some climatic situations on the output of a PV system. The study used selected learning algorithms for building the solar power forecasting models. The authors claimed that the approaches used in the study revealed great results. However, the performance of Random Forest Regressor was claimed to be better than SVM and Linear regression models.

[3] proposed approaches for the forecasting of solar power that was generated 36 hours in advance from Yeongam solar power plant in South Korea. Experimental results showed that the approaches were promising as good results were obtained. [22] equally predicted global solar radiation by using K-means and Support Vector regression algorithms. The study argued that the machine learning approaches used were very promising for solar power prediction.

[23] used selected ML algorithms for the prediction of Power Generation. The authors argued that the study focused on comprehensive comparative analysis and evaluation of ten recent neural networks and intelligent algorithms in short-term PV forecasting. For all the experimentations, the results obtained were said to be good. [24] explained the effects of climatic parameters on the PV system generation. Artificial Neural Networks (ANN) and regression models were built from the chosen dataset. The study reported that the ANN model outperforms all other techniques that were discussed.

Also, [25] proposed hybrid long short-term memory recurrent neural network and Recurrent Neural Network (LSTM-RNN) PV systems output forecasting. Authors argued that the approach is very effective owing of its classification results.

[26] carried out a study for the comparison of models that were built from ML-based techniques and time series approaches. The study used five different sites in Sweden for the sources of data. Authors argued that the ML-based methods are more promising compared to the time series ones using simplicity as a metric. [27] performed a comparison of support vector machine and extreme gradient boosting performances in solar radiation production. The climate data of China was employed in the study. It was argued that the results were excellent.

[28] equally carried out a study that used Neural Network (NN) algorithm for the prediction and validation of solar output power within a period of 24 hours. Authors argued that the result obtained from the NN-based solar power prediction model was promising. Similarly, [29] built a model for the forecasting of solar power in Multi-site. The chosen ML algorithm used in the study is gradient boosted regression trees. Authors argued that the results were promising based on the chosen metrics. Yu and Goh (2017) proposed a solar forecasting approach using a combination of ANN and fuzzy logic pre-processing. Authors claimed that the approach is better in solar energy prediction.

[5] built models for solar energy prediction using Support Vector Machines and Gradient Boosted Regression algorithms. The authors used three different feature selection methods. They argued that good classification results were obtained in each scenario with the new technique more promising. Apart from this, [30] proposed a method for building intelligent solar power prediction model. The model focused on predicting the availability of solar sources using historical data from Rockhampton, Australia. Real life solar radiation data is collected across six years with hourly resolution from 2005 to 2010. Authors pointed out that the experimental results showed a very high prediction accuracy

3. Results and Discussion

This work first of all introduced solar power as a good example of renewable energy sources. Then, focus was shifted to the discussions on how machine learning (ML) techniques have been found promising for forecasting solar power generation. Relevant articles were sourced from research databases. Some of the notable works on the use of different ML approaches for effective forecasting of the solar energy in different countries of the world were surveyed. It is believed that the approach used will provide further insights for researchers on how to build improved ML-based models for the solar energy production prediction. This study identified different machine learning techniques that have been used to build solar power prediction models. The study pointed out that different data sources of various sizes and patterns were used in the experimentations. The work also pointed out that different data sources of various sizes and patterns were used in the experimentations. Some datasets were large some are moderately small. The algorithms used for building the models are of various ML categories and types. The algorithms range from linear to non-linear types. The results of all the empirical papers used in the survey were reported in order to evaluate how promising they are in solar power prediction activity.

4. Conclusion

This work surveyed papers that have used the various ML-based methods that have been used for the prediction of solar energy production or generation. The study is targeted at having a comprehensive review of literature that can provide further directions on this research area. The work emphasised the popularity of ML techniques for solving problems across many areas or disciplines. This study carried out a survey of various ML-based approaches that researchers have used for the prediction of solar energy production. It is believed that the approach used will provide further insights for researchers on how to build improved ML-based models for the solar energy production prediction. Specifically. this study established that different traditional ML and Deep learning techniques have been used to build solar power prediction models. The work also pointed out that different data sources of various sizes and patterns were used in the experimentations of such works while some datasets were large some are moderately small. It is believed that this study can provide more insights to researchers in this area.

References:

[1] S. Ogunjo, O. Aderonke, and B. Rabiu, “Machine learning prediction of solar energy potential in Nigeria,” 2022, p. 030002, doi: 10.1063/5.0099501.

[2] K. Anuradha, D. Erlapally, G. Karuna, V. Srilakshmi, and K. Adilakshmi, “Analysis Of Solar Power Generation Forecasting Using Machine Learning Techniques,” *E3S Web Conf.*, vol. 309, p. 01163, Oct. 2021, doi: 10.1051/e3sconf/202130901163.

[3] B. Carrera and K. Kim, “Comparison Analysis of Machine Learning Techniques for Photovoltaic Prediction Using Weather Sensor Data,” *Sensors*, vol. 20, no. 11, p. 3129, Jun. 2020, doi: 10.3390/s20113129.

[4] R. H. Charlier, “Coastal Planning and Management,” *Int. J. Environ. Stud.*, vol. 66, no. 6, pp. 800–800, 2009, doi: 10.1080/00207230600836518.

[5] R. Aler, R. Martín, J. M. Valls, and I. M. Galván, "A Study of Machine Learning Techniques for Daily Solar Energy Forecasting Using Numerical Weather Models," 2015, pp. 269–278.

[6] G. Chartrand *et al.*, "Deep Learning: A Primer for Radiologists," *RadioGraphics*, vol. 37, no. 7, pp. 2113–2131, Nov. 2017, doi: 10.1148/rg.2017170077.

[7] D. P. Larson, L. Nonnenmacher, and C. F. M. Coimbra, "Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest," *Renew. Energy*, vol. 91, pp. 11–20, Jun. 2016, doi: 10.1016/j.renene.2016.01.039.

[8] A.-N. Sharkawy, M. Ali, H. Mousa, A. Ali, and G. Abdel-Jaber, "Machine Learning Method for Solar PV Output Power Prediction," *SVU-International J. Eng. Sci. Appl.*, vol. 3, no. 2, pp. 123–130, Dec. 2022, doi: 10.21608/svusrc.2022.157039.1066.

[9] A. Alcañiz, D. Grzebyk, H. Ziar, and O. Isabella, "Trends and gaps in photovoltaic power forecasting with machine learning," *Energy Reports*, vol. 9, pp. 447–471, Dec. 2023, doi: 10.1016/j.egyr.2022.11.208.

[10] A. Abraham Iorkaa, M. Barma, and H. GAYA Muazu, "Machine Learning Techniques, methods and Algorithms: Conceptual and Practical Insights," *Int. J. Eng. Res. Appl. www.ijera.com*, vol. 11, no. 8, pp. 55–64, 2021, doi: 10.9790/9622-1108025564.

[11] G. Cerulli, "The Basics of Machine Learning," pp. 1–17, 2023, doi: 10.1007/978-3-031-41337-7_1.

[12] N. E. Benti, M. D. Chaka, and A. G. Semie, "Forecasting Renewable Energy Generation with Machine Learning and Deep Learning: Current Advances and Future Prospects," *Sustainability*, vol. 15, no. 9, p. 7087, Apr. 2023, doi: 10.3390/su15097087.

[13] A. Balal, Y. Pakzad Jafarabadi, A. Demir, M. Igene, M. Giesselmann, and S. Bayne, "Forecasting Solar Power Generation Utilizing Machine Learning Models in Lubbock," *Emerg. Sci. J.*, vol. 7, no. 4, pp. 1052–1062, Jul. 2023, doi: 10.28991/ESJ-2023-07-04-02.

[14] C. Vennila *et al.*, "Forecasting Solar Energy Production Using Machine Learning," *Int. J. Photoenergy*, vol. 2022, pp. 1–7, Apr. 2022, doi: 10.1155/2022/7797488.

[15] M. Y. ERTEN and H. AYDİLEK, "Solar Power Prediction using Regression Models," *Uluslararası Muhendis. Arastirma ve Gelistirme Derg.*, vol. 14, no. 3, pp. 1–1, Dec. 2022, doi: 10.29137/umagd.1100957.

[16] M. Rupesh, J. Swathi Chandana, A. Aishwarya, C. Anusha, and B. Meghana, "Prediction of Solar Power Using Machine Learning Algorithm," 2022, pp. 529–539.

[17] B. Zazoum, "Solar photovoltaic power prediction using different machine learning methods," *Energy Reports*, vol. 8, pp. 19–25, Apr. 2022, doi: 10.1016/j.egyr.2021.11.183.

[18] M. Elsaraiti and A. Merabet, "Solar Power Forecasting Using Deep Learning Techniques," *IEEE Access*, vol. 10, pp. 31692–31698, 2022, doi: 10.1109/ACCESS.2022.3160484.

[19] S. P. C. Machina, S. S. Koduru, and S. Madichetty, "Solar Energy Forecasting Using Deep Learning Techniques," in *2022 2nd International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC)*, Jan. 2022, pp. 1–6, doi: 10.1109/PARC52418.2022.9726605.

[20] C. S. H. Chuluunsaikhan Tserenpurev, Nasridinov Aziz, Choi Woo Seok, Choi Da Bin and K. Y. Myoung, "Predicting the Power Output of Solar Panels based on Weather and Air Pollution Features using Machine Learning, Journal of Korea Multimedia Society," 2021.

[21] P. R. Vishnu, C. S. Roy, and A. Srihari, "Solar Power Output Prediction using Machine Learning Techniques," 2021.

[22] T. R. Ayodele, A. S. O. Ogunjuyigbe, A. Amedu, and J. L. Munda, "Prediction of global solar irradiation using hybridized k-means and support vector regression algorithms," *Renew. Energy Focus*, vol. 29, pp. 78–93, Jun. 2019, doi: 10.1016/j.ref.2019.03.003.

[23] D. Su, E. Batzelis, and B. Pal, "Machine Learning Algorithms in Forecasting of Photovoltaic Power Generation," in *2019 International Conference on Smart Energy Systems and Technologies (SEST)*, Sep. 2019, pp. 1–6, doi: 10.1109/SEST.2019.8849106.

[24] D. Van Tai, "Solar photovoltaic power output forecasting using machine learning technique," *J. Phys. Conf. Ser.*, vol. 1327, no. 1, p. 012051, Oct. 2019, doi: 10.1088/1742-6596/1327/1/012051.

[25] M. Abdel-Nasser and K. Mahmoud, "Accurate photovoltaic power forecasting models using deep LSTM-RNN," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 2727–2740, Jul. 2019, doi: 10.1007/s00521-017-3225-z.

[26] E. Isaksson and M. K. Conde, "Solar Power Forecasting with Machine Learning Techniques," *Kth R. Inst. Technol.*, p. 46, 2018.

[27] J. Fan *et al.*, "Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China," *Energy Convers. Manag.*, vol. 164, pp. 102–111, May 2018, doi: 10.1016/j.enconman.2018.02.087.

[28] S. Leva, A. Dolara, F. Grimaccia, M. Mussetta, and E. Ogliari, "Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power," *Math. Comput. Simul.*, vol. 131, pp. 88–100, Jan. 2017, doi: 10.1016/j.matcom.2015.05.010.

[29] C. Persson, P. Bacher, T. Shiga, and H. Madsen, "Multi-site solar power forecasting using gradient boosted regression trees," *Sol. Energy*, vol. 150, pp. 423–436, Jul. 2017, doi: 10.1016/j.solener.2017.04.066.

[30] M. R. Hossain, A. M. T. Oo, and A. B. M. S. Ali, "Hybrid Prediction Method for Solar Power Using Different Computational Intelligence Algorithms," *Smart Grid Renew. Energy*, vol. 04, no. 01, pp. 76–87, 2013, doi: 10.4236/sgre.2013.41011.