Clustering Model For Solar Irradiation Prediction Using Machine Learning Algorithm

Prachi Sharma, Dr.Megha Kamble

Abstract: With every year the sources of energy are getting reduced and requirement of energy in increasing every year worldwide. Solar irradiance is renewable source of energy provided by sun naturally. So the research community is also proposing methods to convert solar radiance into energy and use it in place of depleting non-renewable source of energy. In need to understand from where this sun energy can be obtained in huge amount, it is necessary to predict the collection of energy for particular place. Solar irradiance varies with meteorological data longitude, latitude, wind speed and change of weather at different areas. Machine learning is a popular set of techniques to predict solar irradiance using weather forecasting data such as wind speed, wind direction and many such parameters in the day span. This is a survey paper which contains the review of previous work done such as ANN, SVM, NN in the field of predicting solar irradiance using weather data parameters in the different parts of world. This paper will help in understanding the use of different machine learning techniques for predicting solar irradiance and also provide future direction to retain renewable energy source like solar energy. A comparative analysis of different techniques is done using output parameters RSME ,MSE,MAE,MAPE and stated the future direction in this research article.

Keywords: Solar irradiance, machine learning, neural networkprediction,estimation

1 INTRODUCTION

The Solar irradiation is the amount of power per unit area received from the sun. Predicting solar irradiance is necessary to minimize the cost of energy and provide high-quality power in electrical power grids with distributed solar photovoltaic generations. In the case of residential and small commercial users who deploy on-site photovoltaic(PV) generations, the irradiance data cannot be obtained easily due to expensive solar irradiance meters. However, the local meteorological departments provide the availability of weather forecasting data like wind speed, visibility, humidity, dew point, and temperature on the internet with which the forecasting of irradiance data can be done. In PV system design, it is essential to know the amount of sunlight available at a particular location at a given time. The two common methods which characterise solar radiation are the solar radiance (or radiation) and solar insolation. The solar radiance is an instantaneous power density in units of kW/m2. The solar radiance varies throughout the day from 0 kW/m2 at night to a maximum of about 1 kW/m2. The solar radiance is strongly dependant on location and local weather. Solar radiance measurements consist of global and/or direct radiation measurements taken periodically throughout the day. The measurements are taken using either a pyranometer (measuring global radiation) and/or a pyrheliometer (measuring direct radiation). In well established locations, this data has been collected for more than twenty years. An alternative method of measuring solar radiation, which is less accurate but also less expensive, is using a sunshine recorder. These sunshine recorders (also known as Campbell-Stokes recorders), measure the number of hours in the day during which the sunshine is above a certain level (typically 200 mW/cm2). Data collected in this way can be used to determine the solar insolation by comparing the measured number of sunshine hours to those based on calculations and including several correction factors. The solar insolation is the total amount of solar energy received at a particular location during a specified time period, often in units of kWh/(m2 day) While the units of solar insolation and solar irradiance are both a power density (for solar insolation the "hours" in the numerator are a time measurement as is the "day" in the denominator), solar insolation is quite different than the solar irradiance as the solar insolation is the instantaneous solar irradiance averaged over a given time period. Solar insolation data is commonly used for simple PV system design while solar radiance is used in more complicated PV system performance which calculates the system performance at each point in the day. Solar insolation can also be expressed in units of MJ/m2 per year. The physical methods are expensive for estimation of solar energy and electricity generation at one particular location. Data analytics and machine learning are powerful alternate mechanisms which are presented in this paper for estimation. Various methods are studied for estimating solar radiation based on Neural network, machine learning, AI, real time series forecasting. ANN and ARIMA methods are found to provide quality prediction in some variability conditions. Three promising methods in the literature include the SVM, regression trees and random forests [2]. Literature provides worldwide instances from various regions such as Uganda, Saudi Arabia where estimation of solar radiance is done. An artificial neural networks model is proposed to predict the monthly average regular total solar irradiation depending on latitude, longitude, altitude sunshine duration, relative humidity, and maximum temperature in [3]. Similarly, comparisons are made between Bayesian Neural Network and empirical models to estimate the regular solar irradiation using the meteorological database from Saudi Arabia[4]. Moreover, an automatic relevance determination (ARD) is proposed to select the optimum input parameters of the NN[4]. A multi-point prediction system has been introduced for reducing the amount of calculation by establishing clusters through dynamic time warping algorithm depending on the likelihood

• *Prachi Sharma Research Scholar, CSE dept. L.N.University Bhopal, India prachisharma3110@gmail.com*
• *Dr.Megha Kamble CSE dept. L.N.University Bhopal, India megha_vi@yahoo.com*

of time series data[5]. In Durban, clustering of solar irradiance pattern was performed through two forecasting methods namely k-means clustering followed by a rule. k-means clustering regularly cover profiles. Then the second part acts as a rule for predicting cloud cover files, categorized as per the average in the morning, afternoon above or below 50%. Four classes were found in the literature in both the methods in association with the irradiance classes for forecasting irradiance class for the upcoming day. An unsupervised clustering-based (UC-based)[1] solar forecasting method is designed for a short duration global horizontal irradiance(GHI) forecasting. Many solar power technologies through text mining are worked upon. There are 2280 international patents and 5610 literature reviews formed where initially this review resulted in the formation of a solar power knowledge oncology schema, Followed by non-supervised machine learning techniques finally applying word-embedding algorithm and cross-validation of results are provided in the literature. A convolutional neural network is presented to model the time consuming broadband radioactive transfer modeling.

.

2 Ease of Use

2.1 Problem and Challenges in estimation

To benchmark different forecasting techniques of solar PV panel energy output are available. After rigorous experiments done in the literature, machine learning and time series techniques can be used to dynamically learn the relationship between different weather conditions and the energy output of PV systems. The models that will be implemented and compared have been selected based on their tendency to perform well in previous research of energy forecasting. The focus will be placed in benchmarking ML and time series techniques. Many of the models presented in the introduction section are generic and therefore do most of them have a wide range of different model set-ups. The aim of this article is to give a general overview of the relative performance of the methods rather than investigating a specific model in depth.
The facts with existing data availability of meteorological data are provided here. One is that the forecasts are produced every 6 hours and span six hours forward in time, which means that it is not possible to produce more extended predictions than 6 hours. Another consequence is that the training set is reduced for longer time horizons as one cannot use unavailable weather forecasts. It is likely that some models can be improved on estimation. The issue with the weather data is that the produced forecast is not for the exact coordinates of the PV installation. This is not ideal, however, it is likely that the impact is limited as the weather variables were considered as more important in some cases, as the lagged variable may reflect the actual weather condition at the location of the PV panel. Another issue is that the forecasts are delivered four times per day and only have a horizon of six hours, which limits the number of observations as the forecast horizon grows.

2.2 Objective

The major objectives of this research article are as follows:
To understand conceptual and theoretical concepts of data mining in solar irradiation prediction.
Review of methods to design optimal algorithm for solar irradiance prediction for number of places and then clustering the regions according to level of solar radiance
To develop an effective model to predict the solar irradiation using an enhanced machine learning technique.
To develop an efficient approach of optimal multiplicative algorithm for clustering model for regionwise solar irradiation

3 Literature Survey

Solar irradiance prediction is emerging and most relevant research area in the application of machine learning and also contributed by many researchers. [1] proposed a new solar prediction scheme for the hourly day ahead solar irradiance prediction using the data of weather forecasting. The prediction problem is formulated as a structured output prediction problem which combining predicts a collective set of outputs simultaneously. Long short-term memory (LSTM) networks train the proposed prediction model by considering the existing dependency between consecutive hours of the day. Various algorithms including persistence algorithm, linear least square regression, and multilayered feedforward neural networks were compared through backpropagation algorithm(BPNN) for predicting solar irradiance in [2]. The probability of developing an artificial neural network model is explored for predicting monthly average regularly total solar irradiation on a horizontal surface for regions of Uganda depending on latitude, longitude, altitude, sunshine duration, relative humidity and maximum temperature[3].A comparative study is performed between Bayesian Neural Network(BNN), classical neural Network(NN) and empirical models to estimate the regular solar irradiation(DGSR) using the experimental meteorological database in the year range between1998 to 2002 at Saudi Arabia. There were four parameters including air temperature, relative humidity, sunshine duration, and extraterrestrial irradiation being employed in the study. Automatic relevance determination(ARD) method has evaluated for selecting the optimum input parameters of the NN. The output showed BNN to perform better in comparison to other NN structures and empirical models[4]. Authors [5] discussed the multi-point predictions of solar irradiance. The proposed multi-point systems for minimizing the percentage of calculation. This system builds clusters depending on time data series through the execution of time series data using the dynamic time warping algorithm and thus a small percentage of the model is designed representing each cluster. Some more references evaluated two forecasting methods namely k-means clustering followed by a rule.k-means clustering regularly cover profiles .Then the second rule acts as a rule for predicting cloud cover files, categorized as per the average in the morning, afternoon above or below 50%.Four classes were observed in both the methods in association with the irradiance classes for forecasting irradiance class for the upcoming day. In Durban, clustering of solar irradiance pattern was performed through two forecasting methods namely k-means clustering followed by a rule.k-means clustering regularly cover profiles. Then the second part acts as a rule for predicting cloud cover files, categorized as per the average in the morning, afternoon above or below 50%. Four classes were found in both the methods in association with the irradiance classes for forecasting irradiance class for the upcoming day. Solar forecasting

accuracy is very important for improving forecasting accuracy.An unsupervised clustering-based (UC-based) solar forecasting method is designed for a short duration global horizontal irradiance(GHI) forecasting. Many solar power technologies through text mining are worked upon.Here taking 2280 international patents and 5610 literature reviews as tuyd was formed where initially this review resulted in the formation of a solar power knowledge oncology schema, Followed by non-supervised machine learning techniques finally applying word-embedding algorithm and cross-validation of results. A convolutional neural network is presented to model the time consuming broadband radioactive transfer modelling. An improvement of energy efficiency of the construction sector has become necessary to reduce gas emission through ML approaches including artificial neural network,support vector machine,Gaussian-based regressions and clustering.

4 Proposed Methodology

Most of the review papers have shown much work have been done using ML approaches such as Decision trees,Regression treses,artificial neural network,BNN,but very few rarely ensemble approaches are used in solar energy forecasting.Ensemble approaches helps to combine various advantages of diffrent ML approaches to develop one one predictive model for decreasing the variance , bias as well as improving the predictions . Initially in this study comparison of time series techniques and machine learning techniques is suggested for solar energy forecast. We find that employing time series models is a complex procedure due to the non-stationary energy time series. In contrast, machine learning techniques were more straightforward to implement. For further research, we suggest continuing comparing different machine learning techniques in depth while using feature engineering approaches of numerical weather predictions.

5 Conclusion

The performance of the proposed approach will be estimated by means of the following measure, Mean Absolute Error (MAE) Normalized Root Mean Square Error (nRMSE) Prediction Accuracy (%) Comparative performance will be carried out with existing approaches to prove the effectiveness of proposed technique for solar irradiance prediction in different regions with the help of weather data. Future scope includes implementation of the proposed work for real time dataset.

References

[1] Qing, X., & Niu, Y. (2018). Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy, 148, 461-468

[2] Voyant, C., Notton, G., Kalogirou, S., Nivet, M. L., Paoli, C., Motte, F., & Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. Renewable Energy, 105, 569-582.

[3] Mubiru, J. (2008). Predicting total solar irradiation values using artificial neural networks. Renewable Energy, 33(10), 2329-2332..

[4] Yacef, R., Benghanem, M., & Mellit, A. (2012). Prediction of daily global solar irradiation data using Bayesian neural network: A comparative study. Renewable energy, 48, 146-154.

[5] Tanaka, Y., & Takahashi, M. (2017). SVM-based Multi-point Prediction System of Solar Irradiance using DTW-based Cluster Analysis. Transactions of the Institute of Systems, Control and Information Engineers, 30(8), 305-313..