National Journal of Multidisciplinary Research and Development
www.nationaljournals.com
ISSN: 2455-9040
Received: 03-05-2022, Accepted: 18-05-2022, Published: 04-06-2022
Volume 7, Issue 1, 2022, Page No. 33-36

Application of machine learning models in solar energy prediction

Anuradha, Taruna Jain
Scholar, Department of Electronics and Communication Engineering, SPGOI Rohtak, Haryana, India
HOD (ECE), Department of Electronics and Communication Engineering, SPGOI Rohtak, Haryana, India

Abstract
Solar power is the conversion of energy from sunlight into electricity, either directly using photovoltaics (PV), indirectly using concentrated solar power, or a combination. Concentrated solar power systems use lenses or mirrors and solar tracking systems to focus a large area of sunlight into a small beam. Photovoltaic cells convert light into an electric current using the photovoltaic effect. Pheotovoltaics were initially solely used as a source of electricity for small and medium-sized applications, from the calculator powered by a single solar cell to remote homes powered by an off-grid rooftop PV system. Generally, two alternatives exist when it comes to the energy from the sun: concentrated solar power (CSP) and photovoltaic (PV) power. In the former case, often referred to as solar thermal power generation, standard heat-based systems are in place to change heat in the form of steam to power. In this research work we used nonlinear regression analysis techniques. This paper therefore discusses about the different regression techniques used in my research. In my research work data processing will be done by using the weather parameters such as solar irradiation, module temperature, ambient temperature etc. and the performance of the model will be evaluated using suitable and widely used performance indicators.

Keywords: photovoltaic, regression, support vector machines

Introduction
Solar influence is the adaptation of energy from daylight into electricity, whichever directly using photovoltaic (PV), ultimately using intense solar power, or a mixture. Focussed solar power schemes use lenses or glasses and tracking arrangements to focus a huge area of sunshine into a small sunbeam. Photovoltaic lockups convert light into an electrical current by means of the photovoltaic result [1]. Photovoltaics were firstly solely used as a cause of electricity for insignificant and medium-sized solicitations, from the calculator power-driven by a solitary solar cell to isolated homes power-driven by an off-grid roof PV regression. As the expense of sun oriented power has fallen, the quantity of grid associated or oriented PV frameworks has developed into the millions in real-world and utility-based photovoltaic power base stations with several megawatts are being assembled. Sun based cells are quickly turning into a reasonable, low-carbon innovation to tackle sustainable power source from the Sun.

Photovoltaic Power System
A sun based cell, or popularly known as photovoltaic cell, is a gadget that varies over light into electric flow exploiting the photovoltaic impact. The main sun oriented cell was built in the 1881 [2]. In 1957 scientist built up the practice of silicon surface passivation by warm oxidation at Bell Labs. The surface passivation procedure has since been basic to sun oriented cell efficiency. The variety of a photovoltaic power agenda, delivers direct current (DC) control which fluctuations with the daylight's force. For down to earth utilize this generally expects change to certain ideal voltages or rotating current and using inverters. Multiple sun oriented cells are associated inside modules. Modules are wired together to frame clusters, at that point attached to an inverter, which yieldscontroller at the ideal voltage, and for AC, the model recurrence/phase [3]. Numerous private PV frameworks are associated with the network any place accessible, particularly in created nations with enormous markets. In these matrix associated frameworks related to the PV, utilization of vitality hoarding is discretionary. In specific submissions, for example, satellites, encouragements, or in creating nations, batteries or extra influence generators are regularly encompassed as back-ups. Such remain solitary control frameworks grant activities nearby evening time and at different times of constrained daylight [4].

Solar Energy
Sun oriented vitality is brilliant light and warmth from the Sun that is tackled and utilizing a scope of consistently advances, for example, sun oriented warm, photovoltaic, sun based warm vitality, sunlight based engineering, liquid salt power plants, and counterfeit photosynthesis [5]. It is a significant sustainable power source and its advances are comprehensively portrayed as either latent sun powered or dynamic sun-powered, relying upon how they catch and circulate sun based vitality or convert it into sun-powered power. Dynamic sunlight based procedures incorporate the utilization of photovoltaic frameworks, concentrated sun-powered

control and sun based water warming to saddle the vitality. Latent sunlight based methods incorporate arranging a structure to the Sun, choosing materials with positive warm mass or light-scattering properties, and planning spaces that normally flow air.

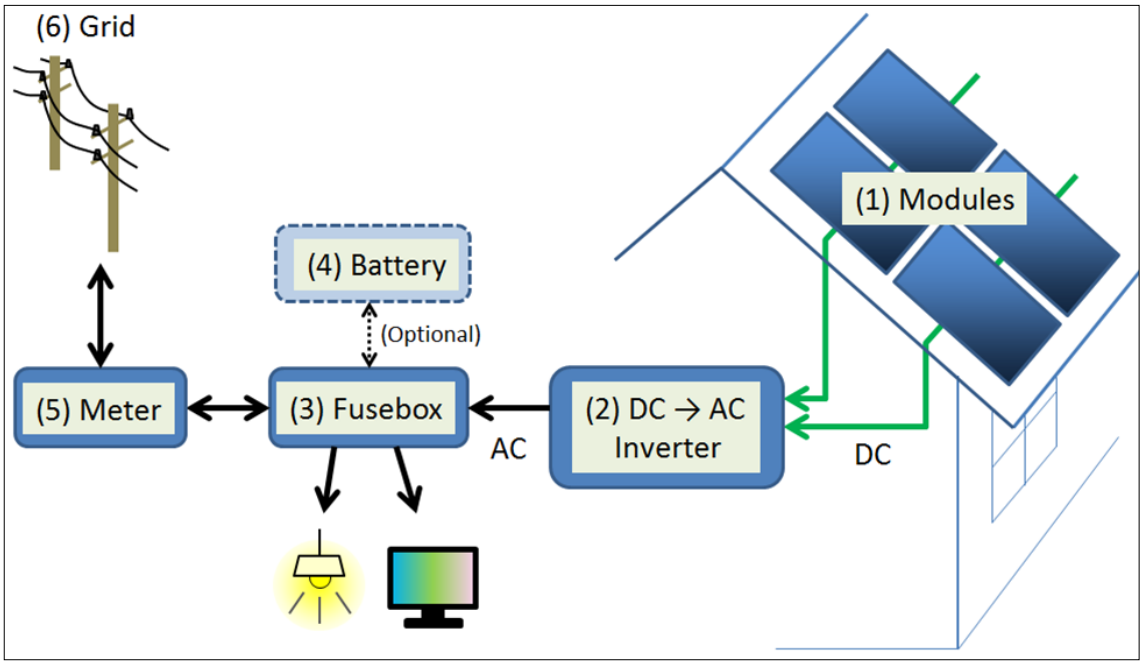

Fig 1: Photovoltaic Power Systems

Regression

Regression analysis is a predictive modelling technique that analyzes the relation between the target or dependent variable and independent variable in a dataset. The different types of regression analysis techniques get used when the target and independent variables show a linear or non-linear relationship between each other, and the target variable contains continuous values. The regression technique gets used mainly to determine the predictor strength, forecast trend, time series, and in case of cause & effect relation. A regression task can be considered a curve fitting problem, where the output variable is a non-discrete variable that takes values in an interval in the real axis or in a region in the complex numbers plane. Having a data set including training points of yi and xi, where i can represent any real number higher than 1, it is possible to estimate a function f, whose graph fits the data [14]. There are different types of regression analysis but we used nonlinear regression techniques.

- Nonlinear regression is a mathematical model that fits an equation to certain data using a generated line. As is the case with a linear regression that uses a straight-line equation (such as Ŷ= c + m x), nonlinear regression shows association using a curve, making it nonlinear in the parameter. Some nonlinear Regression techniques given below:

SVM (Support Vector Machines)

SVMs or Support Vector Machines are one of the most popular and widely used algorithm for dealing with classification problems in machine learning. However, the use of SVMs in regression is not very well documented. This algorithm acknowledges the presence of non-linearity in the data and provides a proficient prediction model [6]. The SVM regression algorithm is referred to as Support Vector Regression or SVR. Before getting started with the algorithm, it is necessary that we have an intuition of what a support vector machine actually is. In machine learning, Support Vector Machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. In Support Vector Regression, the straight line that is required to fit the data is referred to as hyper plane.

KNN (K-Nearest Neighbors Algorithm)

KNN regression is a non-parametric method that, in an intuitive manner, approximates the association between independent variables and the continuous outcome by averaging the observations in the same neighbourhood. The size of the neighbourhood needs to be set by the analyst or can be chosen using cross-validation (we will see this later) to select the size that minimises the mean-squared error. While the method is quite appealing, it quickly becomes impractical when the dimension increases, i.e., when there are many independent variables [7]. KNN algorithm can be used for regression problems. The KNN algorithm uses 'feature similarity' to predict the values of any new data points. This means that the new point is assigned a value based on how closely it resembles the points in the training set. For example, we know that ID11 has height and age similar to ID1 and ID5, so the weight would also approximately be the same [10]. Had it been a classification problem, we would

have taken the mode as the final prediction. In this case, we have two values of weight – 72 and 77. Any guesses on how the final value will be calculated? The average of the values is taken to be the final prediction.

Decision Tree

Decision Trees (DTs) are a non-parametric supervised learning method used for Regression. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. A tree can be seen as a piecewise constant approximation [8]. For instance, in the example below, decision trees learn from data to approximate a sine curve with a set of if-then-else decision rules. The deeper the tree, the more complex the decision rules and the fitter the model.

A decision tree consists of three types of nodes:

1. Decision nodes – typically represented by squares
2. Chance nodes – typically represented by circles
3. End nodes – typically represented by triangles

It requires little data preparation. Other techniques often require data normalisation, dummy variables need to be created and blank values to be removed, however that this module does not support missing values.

Table 1: Differences between SVM, KNN, Decision Tree (all techniques used/best technique used)

SVM	KNN	Decision Tree
SVM supports both linear and non-linear solutions using kernel trick	KNN is a non -parametric model	Decision trees supports non linearity
SVM algorithm used to create the best line or decision boundary that can segregate n-dimensional space	KNN is slow in real time as it have to keep track of all training data and find the neighbor nodes	Decision trees will be having better average accuracy.
It perform well when the training data is less, and there are large number of features	In KNN, there is no predefined form of the mapping function	Decision trees handles colinearity
SVM chooses the extreme points/vectors that help in creating the hyperplane.	KNN algorithm at the training phase just stores the dataset	possible to validate a model using statistical tests.

So the proposed work deals with the efficient self-learning model which will increase the regulation in the energy constraints of the photovoltaics cell and will increase the regression rate in prediction of the energy levels so that the performance will be controlled and resource will be manage accordingly.

Proposed Methodology

Below is the proposed solution for the whole research work:

1. Initially we will collect the suitable solar plant data and study the characteristics of energy generation from solar power systems.
2. Then data processing will be done to select the weather parameters such as solar irradiation, module temperature, ambient temperature etc.
3. Then an extensive data exploration will be performed to set input and output targets.
4. Then the selected input and output targets will be divided into training and test sets.
5. Then basic regression learning algorithm will be used to test their predictive power in predicting solar generation.
6. Then an efficient regression based machine learning model will be developed to predict the generation with maximum possible accuracy.
7. Then the performance of the model will be evaluated using suitable and widely used performance indicators.

Python tool used for research and development all given process.

Conclusion

A review of solar energy, regression, SVM, KNN, Decision Tree is presented briefly. We also discussed about our proposed methodology used in my research work. This brief representation is very useful for my research. Then basic regression learning algorithm will be used to test their predictive power in predicting solar generation.

References

1. Bayindir Ramazan, Mehmet Yesilbudak, Medine Colak, Naci Genc. "A novel application of Naive Bayes classifier in photovoltaic energy prediction." In 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA) IEEE, 2017, 523-527.
2. Cammarano Alessandro, Chiara Petrioli, Dora Spenza. "Pro-Energy: A novel energy prediction model for solar and wind energy-harvesting wireless sensor networks." In 2012 IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Systems (MASS 2012) IEEE, 2012, 75-83.

3. Martín Luis, Luis F, Zarzalejo Jesús Polo, Ana Navarro, Ruth Marchante, Marco Cony. "Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning." Solar Energy,2010:84(10):1772-1781.
4. Wan Can, Jian Zhao, Yonghua Song, Zhao Xu, Jin Lin, Zechun Hu. "Photovoltaic and solar power forecasting for smart grid energy management." CSEE Journal of Power and Energy Systems,2015:1(4):38-46.
5. Izgi Ercan, Ahmet Öztopal, Bihter Yerli, Mustafa Kemal Kaymak, Ahmet Duran Şahin. "Short–mid-term solar power prediction by using artificial neural networks." Solar Energy,2012:86(2):725-733.
6. Tao Cai, Duan Shanxu, Chen Changsong. "Forecasting power output for grid-connected photovoltaic power system without using solar radiation measurement." In The 2nd International Symposium on Power Electronics for Distributed Generation Systems, IEEE, 2010, 773-777.
7. Sharma Navin, Pranshu Sharma, David Irwin, Prashant Shenoy. "Predicting solar generation from weather forecasts using machine learning." In 2011 IEEE international conference on smart grid communications (SmartGridComm), IEEE, 2011, 528-533.
8. Abedinia Oveis, Nima Amjady, Noradin Ghadimi. "Solar energy forecasting based on hybrid neural network and improved metaheuristic algorithm." Computational Intelligence,2018:34(1):241-260.
9. Mirzapour Farzaneh, Mostafa Lakzaei, Gohar Varamini, Milad Teimourian, Noradin Ghadimi. "A new prediction model of battery and wind-solar output in hybrid power system." Journal of Ambient Intelligence and Humanized Computing,2019:10:(1):77-87.
10. Elsheikh Ammar H, Swellam W, Sharshir Mohamed, AbdElaziz AE. Kabeel Wang Guilan, Zhang Haiou. "Modeling of solar energy systems using artificial neural network: A comprehensive review." Solar Energy,2019:180:622-639.
11. Vandana khanna, Bijoy kishore das, Dinesh bisht,"Matlab/Simelectronics model based study of solar cells", International journal of Renewable Energy Research, 2013, 3(1).
12. TaoMa Hongxing Yang, Lin Lu." Solar photovoltaic system modeling and performance prediction", Renewable and Sustainable Energy Reviews,2014:36:304-315.
13. Rakhi Sharma GN, Tiwari. "Technical performance evaluation of stand-alone photovoltaic array for outdoor field conditions of new delhi", in Applied energy,2012:92:644-652.
14. Ayompe LM, Duffy A, McCormack SJ, Conlon M. "Measured performance of a 1.72kw rooftop grid connected photovoltaic system in Ireland", in Energy conversion and management,2011:52:816-825.
15. Brenna M, Faranda R, Leva S. "Dynamic analysis of a new network topology for high power grid connected PV systems", IEEE, 2010.