Solar Power Predictor using Ensemble Learning

[1]Abhishek Innani, [2]Shweta Dhatwalia, [3]Yash Tripathi, [4]Vikky Jambhulkar

[1-3]Student, Department of Computer Engineering A.R.M.I.E.T Thane, India
[4]Professor, Department of Computer Engineering A.R.M.I.E.T Thane, India

---

Abstract - *This research helps in getting the most effective prediction from ensemble learning, machine learning and deep learning techniques in solar irradiance prediction. It was seen from the literature survey, an ensemble extra tree model, random forests, gradient booster , ada booster is studied more frequently as ensemble models. However, machine learning and deep learning algorithms such as k-nearest neighbours, decision tree, support vector regression SVR, Elastic net, linear regression, lasso regression and artificial neural networks (ANN) is also possible. So, this study is the first detailed evaluation of ensemble models, machine learning and deep learning models in solar irradiance estimation domain. First frequently used base models are created and tested with the use of 2 years data. The base models are compared with their different counterparts in terms of average coefficient of determination (R2) and root mean squared error RMSE. The comparative results show that ensemble, machine and deep learning models which has the best prediction in terms of RMSE. The part of these study even is to analyse intermediate variables and do the weather forecasting by using autoregressive Integrated Moving Average ARIMA and Long Short Term Memory LSTM Model. This paper proposes ARIMA model and LSTM model for time series weather forecasting using one variable weather data. The models were tested using weather time series data from Mumbai.*

Key Words*: Solar irradiance, Machine learning, Ensemble learning, Deep learning, Weather Forecasting, Auto-regressive Integrated Moving Average, Long Short-Term Memory.*

1. INTRODUCTION

In recent years, the rapid boost of variable renewable resource energy generations particularly from wind and solar energy resources in the power generation grid has led to these generations becoming a noteworthy source of uncertainty with load behaviour still being the main source of variability. Generation and load balance are required in the economic scheduling of the generating units and in electricity market trades. Energy prediction/forecasting can be used to mitigate some of the challenges that arise from the uncertainty in the resource. Solar power forecasting/prediction is witnessing a growing attention from the research community for higher efficiency. Increasing demand for energy is leading toward integration of renewable solar energy with the non-renewable energy resources. Unlike other non-renewable energy resources solar energy is intermittent. To make efficient use of freely available energy accurate solar power forecasting is important. In this paper, we applied deep neural network and different machine learning algorithms like linear regression, SVR, random forest on solar dataset to predict solar power. Paper presents comparative analysis of Deep Neural Network against machine learning algorithms. Hence it helps in making decision to select appropriate forecasting technique. Due to a variability and uncertainty of photovoltaic (PV) output power, photovoltaic operators may be subject to significant penalties for forthcoming energy markets. Thus, the accurate prediction of PV output power plays a very important role in energy market. This project proposes a novel solar prediction scheme for one-hour ahead prediction of solar irradiance based on various meteorological factors and support vector machine (SVM).

We aim on developing an artificial neural network model to produce solar power forecasts. Sensitivity analysis of several input variables for best selection, and comparison of the model performance with multiple linear regression and persistence models can be shown.

2. LITERATURE SURVEY

[1]Solar forecasting could be a stepping stone to those challenges. alternative energy forecasting depends on the factors like knowledge of the path of the sun, the condition of the atmosphere, the scattering process and also the characteristics of a solar power plant which utilizes the sun‘s energy to make alternative energy. Solar photovoltaic systems transform solar power into wattage. The output power depends on the incoming radiation and on the solar battery characteristics. Photovoltaic power production is increasing nowadays. Forecast information is crucial for an efficient use, the management of the electricity grid and for solar power trading. Various solar forecasting research activities get motivated thanks to the factors that accurate solar forecasting techniques improves the standard of the energy delivered to the grid and minimize the extra cost related to weather dependency.

[2]As regenerative power generation mainly depends on the weather situation; most forecasting algorithms also take the weather forecast into consideration. The standard process thereby is:

• A numerical weather prediction (NWP) is made for a particular time horizon and a particular location using physical weather forecasting models.

• Supported this NWP, a forecasting algorithm predicts the longer term power generation for a particular regenerative energy generator.

The quality of the ability forecasting algorithm heavily depends on the NWP, which implies it decreases with increasing time horizon. Forecasting algorithm horizons range from short-term forecasts within the range of hours, mid-term forecasts up to some days, to long-term forecasts within the range of some weeks.

Of course, different time horizons are of interest for various market participants. Specifically, the day-ahead forecast (24h–48h), a typical mid-term forecast, is of interest to several actors within the field, and that we will, therefore, put a stress on this kind of forecast. Forecasting algorithms typically are either supported physical models, which use knowledge about the weather situation and the power station to get a forecast. These models typically calculate turbine / alternative energy curves. A bonus of these models is simple comprehensibility and high accuracy if all parameters are often measured precisely. On the opposite hand, models from machine learning or statistics have proven to achieve success within the domain of power forecasting. In contrast to physical models, these sorts of models don't explicitly model the physical process including the ability generation, but the input-output relationship of NWP parameters to the ability output. Thereby, they will abstract from inherent errors during measurement and may be applied to a good form of tasks in a very flexible manner. Typical approaches within the area of machine learning are (multi-)linear regression techniques or artificial neural networks (ANN)

3. EXISTING SYSTEM

There are two ways for solar energy forecasting which are available for Industries. Short-term solar energy forecasting is provides predictions up to 7 days ahead. This type of forecast/prediction is additionally valuable for grid operators so as to create decisions of grid operation, as well as, for electric market operators. Under this attitude, the meteorological resources are estimated at a different temporal and spatial resolution. This suggests that meteorological variables and phenomena are looked from a more general perspective, not as local as now-casting services do. During this sense, most of the approaches make use of various numerical weather prediction models (NWP) that provide an initial estimation of weather variables. Long-term solar energy forecasting usually refers to forecasting of the annual or monthly available resource. This is often useful for renewable energy producers and to barter contracts with financial entities or utilities that distribute the generated energy. In general, these long-term forecasting is sometimes done at a lower scale than any of the 2 previous approaches. Hence, most of those models are run with mesoscale models fed with reanalysis data as input and whose output is post-processed with statistical approaches supported measured data

4. PROPOSED SYSTEM

In this project, we generate a comparative study showing analysis of various algorithms. This study shows analysis and provides the algorithm giving the simplest result together with comparison. The most objective of our project is that a lot of power generating companies can use our alternative energy predictor on an outsized scale to achieve profit for the corporate by predicting the facility generation of a particular day or a specific week. Generating renewable energy from solar panels emits little pollution into the air, and thus alternative energy is a much cleaner source of energy than the burning of fossil fuels, gasoline and sources which causes pollution. Urban or Rural areas that commit to use alternative energy to power the localities and society would thus enjoy a cleaner quality of air within the region. Because solar panels emit very low amount of pollution into the environment, alternative energy doesn't damage the atmosphere or cause heating. Thus, if areas commit to use alternative energy to generate electricity, the shift will help diminish the results of world warming

In this project, we've got used four approaches for getting the simplest prediction out of any learning:

1. Using Deep Learning
2. Using Machine Learning
3. Using Ensemble Learning
4. Using Traditional Method (ARIMA, LSTM)

1. Deep Learning

Deep learning has been an element of a broader family of machine learning method supported artificial neural network with representation learning. Learning will be supervised, semi-supervised or unsupervised.

Deep learning architectures is as follows deep neural networks, deep belief networks, recurrent neural network and convolutional neural network is applied to fields including computer vision, speech recognition, language processing, social network filtering, artificial intelligence, drug design, image analysis, material inspection and board games programs, where they need produced results appreciate and in some cases surpassing human expert performance.

Artificial neural network (ANNs) were inspired by scientific discipline and distributed communication nodes in biological systems. ANNs have various

differences from biological brains. Specifically, artificial neural networks tend to be static and symbolic, while the biological brains of most living organisms is dynamic and analog.

Algorithm used – Artificial neural network

2. Machine Learning

At a high level, machine learning is that the process of teaching a system the way to make more accurate predictions when fed with data. It's a field of study that overlaps with and inherits ideas from many related fields like computer science. It's programming system to optimise a performance criteria using example data or past experience. Machine learning can automatically detect patterns in data, so to use the uncovered patterns to predict future data or other outcomes of interest. Machine learning supported statistics is essentially attempting to go looking out the connection between input and output variables. Machine learning is closely associated with Statistics, AI and processing. In general, any machine learning problems are often assigned to 1 of two classifications: Supervised and Unsupervised learning. In our case we are using six supervised learning algorithms for machine learning. As ours may well be a comparison study between different algorithms, we during this scenario are trying to go looking out the foremost effective prediction for solar power prediction for the study.

Algorithms Used **-** K Nearest Neighbors (KNN), Decision Tree, Linear Regression, Lasso Regression, Support Vector Regression (SVR), Elastic Net.

3. Ensemble Learning

In statistics and machine learning, ensemble methods use multiple learning algorithms to urge better predictive performance than may well be obtained from any of the constituent learning algorithms alone. Unlike a statistical ensemble in statistical mechanics, which is usually infinite, a machine learning ensemble consists of only a concrete finite set of various models, but typically allows for a lot more flexible structure to exist among those alternatives.

An ensemble is itself a supervised learning algorithm; because it should be trained then want to create predictions. The trained ensemble, therefore, represents one hypothesis. This hypothesis, however, isn't necessarily contained within the hypothesis space of the models from which it's built. Thus, ensembles is also shown to own more flexibility within the functions they're going to represent. This flexibility can, in theory, enable them to over-fit the training data more over one model would, but in practice, some ensemble techniques (especially bagging) tend to chop back problems associated to over-fitting of the training data.

Algorithms Used - Random Forest, Gradient Booster, Extra Tree, Ada Booster.

4. Traditional Method

Traditionally, there are several techniques to effectively forecast the following lag of some time series data like uni-variate Auto-regressive (AR), uni-variate Moving Average (MA), Simple Exponential Smoothing (SES), and more notably Auto-regressive Integrated Moving Average (ARIMA) with its many variations. Particularly, ARIMA model has demonstrated its out-performance in precision and accuracy of predicting the following lags of some time series. With the recent advancement within the computational power of computers and more importantly developing more advanced machine learning algorithms and approaches like as deep learning, algorithms are developed to forecast statistic data. The research question investigated during this text is that whether and also the way newly developed deep learning-based algorithms for forecasting statistic data, like "Long Short-Term Memory (LSTM)", are superior to the traditional algorithms. The empirical studies conducted and thru this text show that deep learning-based algorithms like LSTM outperform traditional-based algorithms like ARIMA model. More specifically, the common reduction in error rates obtained by LSTM is between 84 - 87 percent when put next to ARIMA indicating the prevalence of LSTM to ARIMA. Furthermore, it absolutely was noticed that the quantity of coaching times, called "epoch" in deep learning, has no effect on the performance of the trained forecast model and it exhibits a very random behaviour.

Algorithm Used – ARIMA, LSTM

5. SYSTEM DESIGN AND ARCHITECTURE

In each of these solutions, the building of model undergoes three major stages:

- Stage 1 - Preprocessing data: The input of dataset either given by the user or used by the developer is preprocessed by adding bias to generate multiple datasets for efficient model construction after which the model is trained on this preprocessed data set.

- Stage 2 - Weather Forecasting: Weather may be forecasted on the given data set to generate more data for efficient model training. It used forecasting algorithms like LSTM, ARIMA

- Stage 3 - Energy Prediction: Finally, Solar Energy generation is predicted using Machine Learning algorithms like LR, RR, EN, Decision Tree, SVR and Ensemble Learning algorithms like AdaBoost, RF,

Gradient Boost, XT and finally Deep Learning by using ANN.

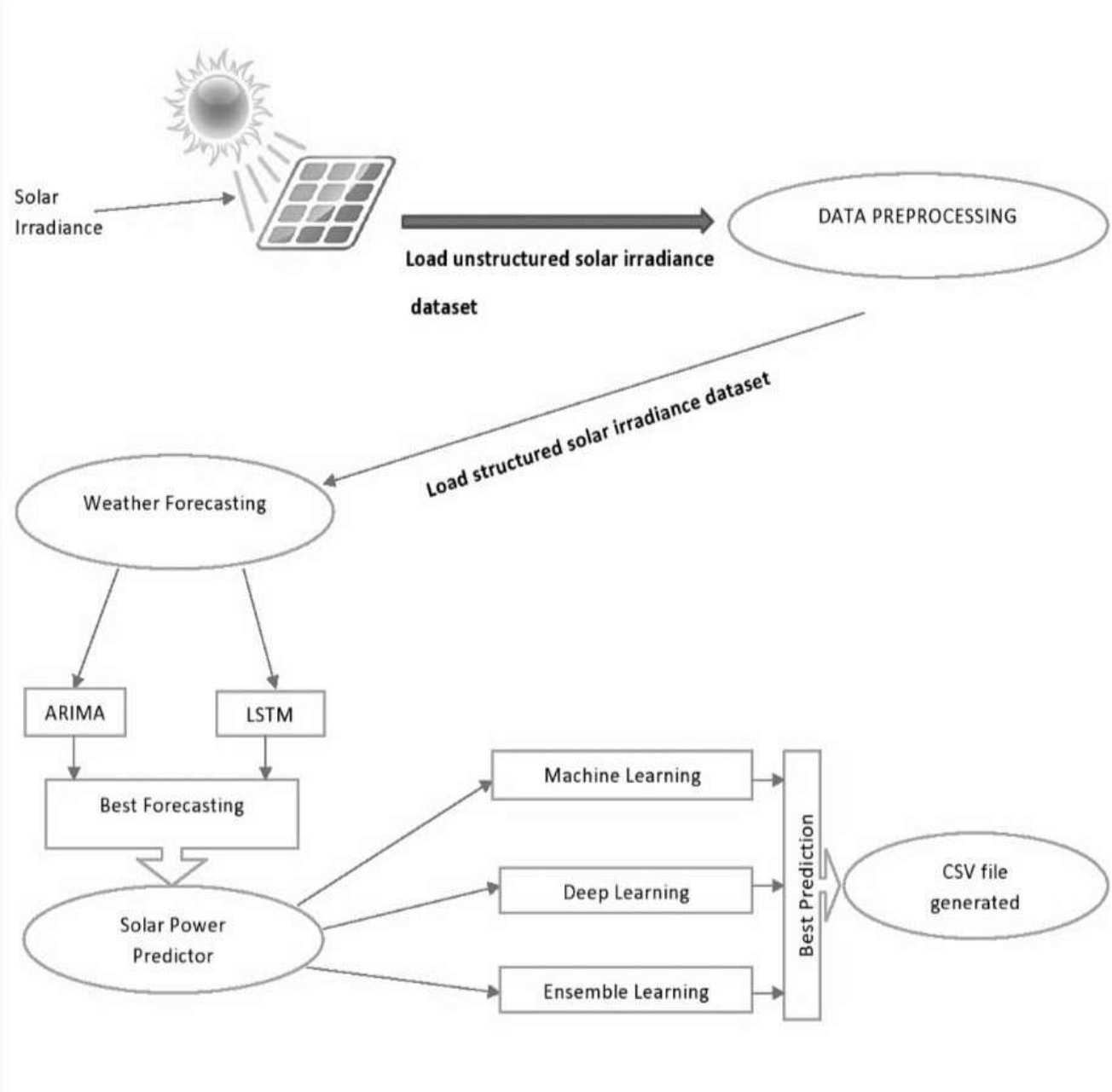

Figure 1: Proposed System Architecture

The snapshot result of comparative study:

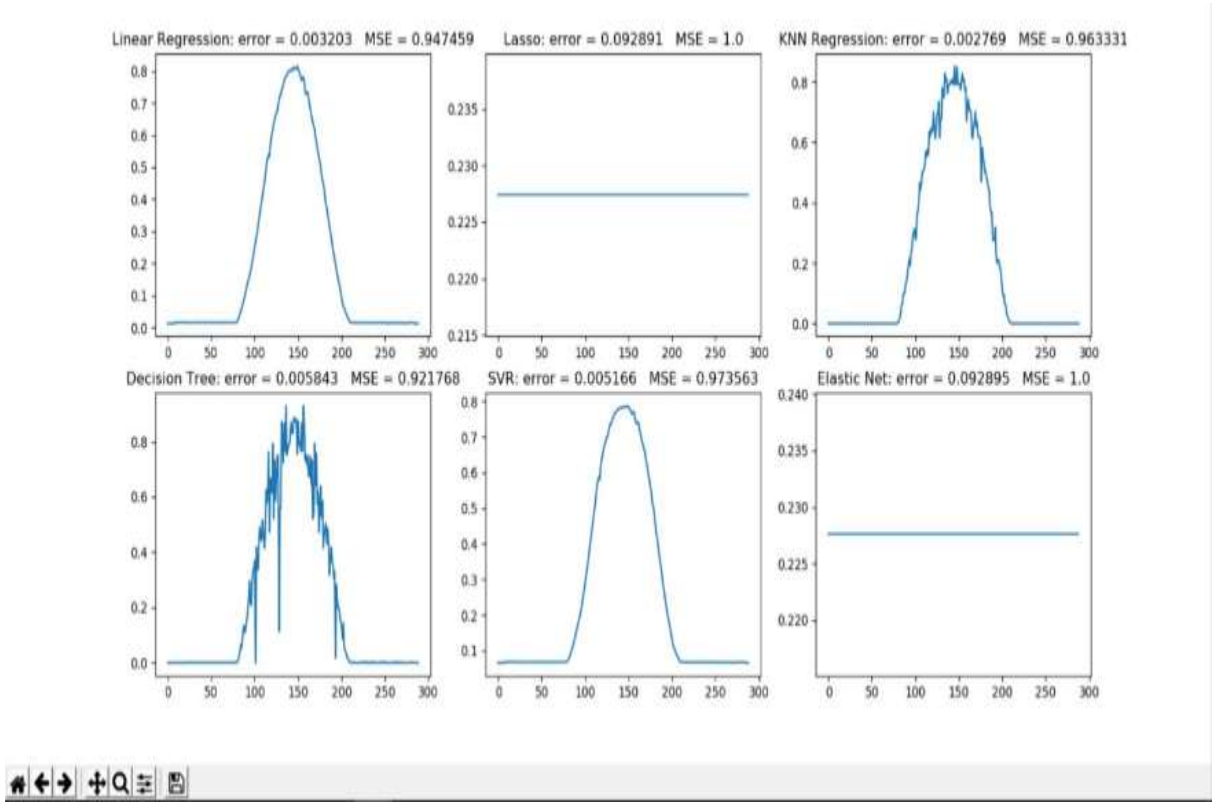

Figure 2: Machine Learning Results

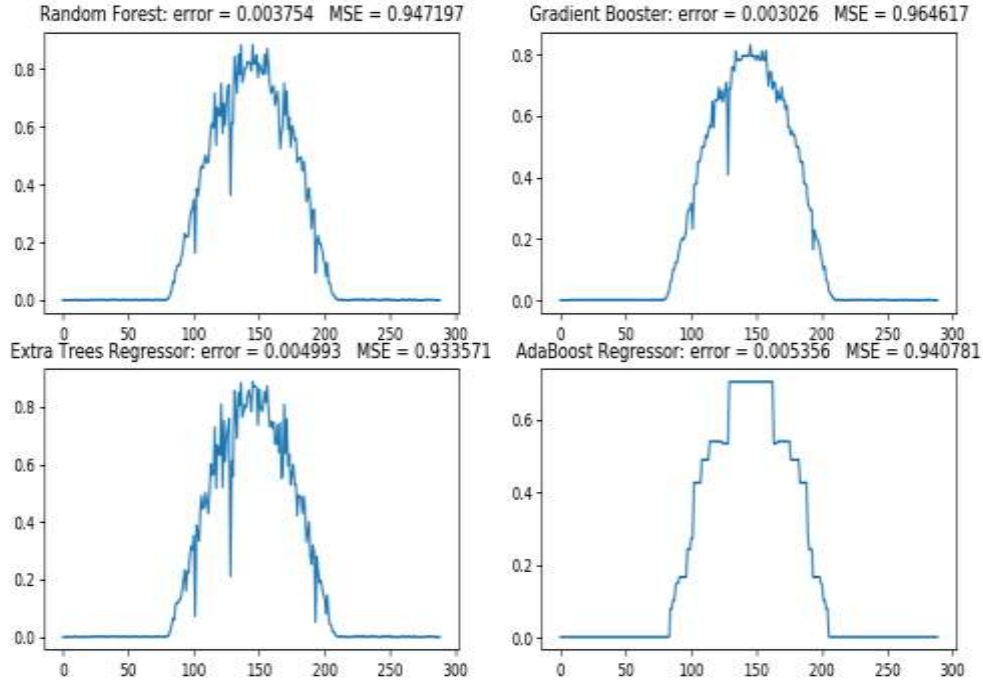

Figure 3: Ensemble Learning Results

6. CONCLUSION

This software application will act as motivation to use alternate means of energy generation which is the need of the hour. It supports Solar Power Plants to efficiently generate electricity and monetize the generated energy accordingly. It will not only help and facilitate the use of Solar Power Generated electricity but also help to estimate the energy generated per day depending on the climatic conditions. This forecasting of weather along with the energy generated will give accurate energy generation prediction. This energy prediction and thus be observed and optimum energy generation plans can be made to support cities. This software will also help Solar Power Plants to monetize their power generation from a business as well as social point of view for the greater good. It will encourage little steps taken to use cleaner energy sources and alternates to Non-Renewable energy sources for a saver planet by reducing the harmful effects of Global Warming. This tool provides efficient comparison among all the provided algorithms to give the most accurate results. It will provide a comparative study on the dataset of user's choice to give predicted or forecasted results, whichever may be required.

7. ACKNOWLEDGEMENT

I would like to take this opportunity to express my heartfelt gratitude towards people whose help and coordination has made this project a success. This research was supported by Computer Engineering branch of Alamuri Ratnamala Institute of Engineering and technology. We thank our guide Vikky Jambhulkar who provided insight and expertise that greatly assisted the research.

8. REFERENCES

[1] D. K. Chaturvedi, Isha. "Solar Power Forecasting: A Review" .International Journal of Computer Applications (0975 – 8887)
Volume 145 – No.6, July 2016

[2] Andre Gensler, Janosch Henze, Bernhard Sick, "Deep Learning for Solar Power Forecasting – An Approach Using Autoencoder and LSTM Neural Networks" IEEE International Conference on Systems, Man, and Cybernetics.

[3] Kivanc Basaran, Akın Özçift, Deniz Kılınç . "A New Approach for Prediction of Solar Radiation with Using Ensemble Learning Algorithm".Arabian Journal for Science and Engineering.

[4] Azhar Ahmed Mohammed, Waheeb Yaqub and Zeyar Aung. "Probabilistic Forecasting of Solar Power: An Ensemble Learning Approach".Springer International Publishing Switzerland 2015 R. Neves-Silva et al. (eds.), Intelligent Decision Technologies, Smart Innovation, Systems and Technologies 39.

[5] Afan Galih Salman, Yaya Heryadi, Edi Abdurahman and Wayan Suparta . "Weather Forecasting Using Merged Long Short Term Memory Model (LSTM) and Autoregressive Integrated Moving Average (ARIMA) Model".Afan Galih Salman et al. / Journal of Computer Science 2018.

[6] Kuk Yeol Bae, Han Seung Jang and Dan Keun Sung . "Hourly Solar Irradiance Prediction Based on Support Vector Machine and Its Error Analysis".IEEE Transactions on Power Systems.

[7] Shahida Begum, Reshma Banu, Ali Ahamed and Parameshachari BD. "A Comparative Study on Improving The Performance Of Solar Power Plants Through IOT and Predictive Data Analytics" . 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques.

[8] Mohamed Abuella, Badrul Chowdhury . "Solar Power Forecasting Using Support Vector Regression".American Society for Engineering Management International AnnualConference,2016.