SOLAR ENERGY FORECASTING

Student 1: Akanshi Jain
Email: Akanshijain1234@gmail.com **& Contact:** 9826393848
Student 2: Jatin Sadhwani
Email: Jatinsadhwani.1234@gmail.com **& Contact:** 7566647818
Student 3: Sanil Sarathe
Email: Sanilsarathe76@gmail.com **&Contact:** 6260102074
Supervisor: Mr. Amit Mittal Sir
Email: amittal@ietdavv.edu.in **and Contact:** 9993315549

Computer Engineering

Keywords: Solar Energy Forecasting using Machine Learning Techniques.

1. INTRODUCTION

Time series forecasting is a technique for the prediction of events through a sequence of time. The technique is used across many fields of study. The techniques predict future events by analyzing the trends of the past, on the assumption that future trends will hold similar to historical trends.
Time series forecasting is performed in a variety of applications including:

- Weather forecasting
- Earthquake prediction
- Pattern recognition
- Signal processing
- Control engineering, etc.

Time series forecasting is sometimes just the analysis of experts studying a field and offering their predictions. In many modern applications, however, time series forecasting uses computer technologies, including:

- Machine Learning
- Neural Network, etc.

In response to many factors, including economics, global warming and climate change, electric power generation has been shifting towards renewable sources. There is no doubt that solar energy is one of the chief sources of renewable energy. Solar energy can be directly converted into electrical energy using photovoltaic (PV) cells. Extensive research and increasing use of the PV cells to generate utility level electricity is driving the drop in price that can be easily seen by the price difference of PV panels between 2010 and 2018, which decreased by 74% during that period [1].

2.Literature Survey

PV power generation is dependent on many factors, such as weather conditions and PV module temperature. Natural variation in climatic conditions can vary these factors changing the PV power generation. Since PV power generation is variable, intermittent, and nonlinear, an accurate solar power forecasting method is required to operate the power system reliably and stably and to ensure quality power production.
There are three types of solar power forecasting: physical methods, statistical methods, and hybrid methods. The machine learning method also comes under the statistical models. Physical models are numerical weather prediction (NWP), sky imagery, and satellite imaging model. Statistical methods are more accurate for a short time horizon (1 to 6 hours), whereas physical models are suitable for the long term.

3. PROBLEM DOMAIN

In recent years, the rapid boost of variable energy generations particularly from wind and solar energy resources in the power grid has led to these generations becoming a noteworthy source of uncertainty with load behavior still being the main source of variability. Generation and load balance is required in the economic scheduling of the generating units and in electricity market trades. Energy forecasting can be used to mitigate some of the challenges that arise from the uncertainty in the resource. Solar power forecasting is witnessing a growing attention from the research community.

4. SOLUTION DOMAIN

A. What makes Time Series Special?

As the name suggests, TS is a collection of data points collected at **constant time intervals**. These are analyzed to determine the long-term trend so as to forecast the future or perform some other form of analysis. But what makes a TS different from say a regular regression problem? There are 2 things:

1. It is **time dependent**. So, the basic assumption of a linear regression model that the observations are independent doesn't hold in this case.
2. Along with an increasing or decreasing trend, most TS have some form of **seasonality trends**, i.e. variations specific to a particular time frame. For example, if you see the sales of a woolen jacket over time, you will invariably find higher sales in winter seasons.

B. Loading and Handling Time Series in Pandas

Pandas has dedicated libraries for handling TS objects, particularly the **datatime64[ns]** class which stores time information and allows us to perform some operations really fast.

C. How to Check Stationarity of a Time Series?

A TS is said to be stationary if its **statistical properties** such as mean, variance remain **constant over time**. But why is it important? Most of the TS models work on the assumption that the TS is stationary. Intuitively, we can say that if a TS has a particular behavior over time, there is a very high probability that it will follow the same in the future. Also, the theories related to stationary series are more mature and easier to implement as compared to non-stationary series.

D. How to make a Time Series Stationary?

Though stationarity assumption is taken in many TS models, almost none of practical time series are stationary. So statisticians have figured out ways to make series stationary, which we'll discuss now. Actually, its almost impossible to make a series perfectly stationary, but we try to take it as close as possible.

E. Estimating & Eliminating Trend

One of the first tricks to reduce trend can be **transformation**. For example, in this case we can clearly see that the there is a significant positive trend. So we can apply transformation which penalize higher values more than smaller values. These can be taking a log, square root, cube root, etc.

F. Moving average

In this approach, we take average of 'k' consecutive values depending on the frequency of time series. Here we can take the average over the past 1 year, i.e. last 12 values.

G. Eliminating Trend and Seasonality

The simple trend reduction techniques discussed before don't work in all cases, particularly the ones with high seasonality. Let's discuss two ways of removing trend and seasonality:

1. **Differencing** – taking the difference with a particular time lag
2. **Decomposition** – modeling both trend and seasonality and removing them from the model.

H. Forecasting a Time Series

We saw different techniques and all of them worked reasonably well for making the TS stationary. Lets make model on the TS after differencing as it is a very popular technique. Also, its relatively easier to add noise and seasonality back into predicted residuals in this case. Having performed the trend and seasonality estimation techniques, there can be two situations:

1. A **strictly stationary series** with no dependence among the values. This is the easy case wherein we can model the residuals as white noise. But this is very rare.
2. A series with significant **dependence among values**. In this case we need to use some statistical models like ARIMA, ARMA, SARIMA to forecast the data.

5. SYSTEM DOMAIN

The project work on Time series Forecasting for Solar energy involves various tools, Technologies, platform and environment , Let us get some insight of them-

* Tools, environment and platforms-

--- Jupyter notebook- It provides to develop open-source software, open-standards, and services for interactive computing across dozens of programming languages. It is the basic tool and the working platform for the given research project work while working with Python on it.

--- Google Colaboratory - It allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education. This would be main environment for the code deployment.

* Technologies -

--- Machine Learning - It is the study of computer algorithms that can improve automatically through experience and by the use of data. Machine learning algorithms will help in building the model based on the sample data in order to make predictions or decisions without being explicitly programmed to do so. It would be the main technology used for the project work.

-- Time series analysis - In addition to machine learning another important technology would be the time series analysis which would be helping in the accomplishment of the project. It is a specific way of analyzing a sequence of data points collected over an interval of time. In time series analysis, analysts record data points at consistent intervals over a set period of time rather than just recording the data points intermittently or randomly. Time series data can be used for forecasting—predicting future data based on historical data.

--- Python and it's library - The whole code work would be done with using python language and its libraries.

Some of the libraries required would be Scikit-learn, Numpy, matplotlib, seaborn, pandas, etc.

6 APPLICATION DOMAIN

Industrial Application: Sun’s thermal energy is used in office, warehouse and industry to supply power. Solar energy is used to power radio and TV stations. It is also used to supply power to lighthouse and warning light for aircraft.

Transportation: Solar energy is also used for public transportation such as trolleys, buses and light-rails.

Solar Electricity: his is one of the solar energy applications that has gained a lot of momentum in recent years. As solar panel costs decline and more people become aware of the financial and environmental benefits of solar energy, solar electricity is becoming increasingly accessible. While still a very small percentage of the electricity generated in the U.S.

Solar Lighting: Solar lights have become ubiquitous and can be found everywhere from home landscaping and security lights to road signs and street lights. These solar lighting technologies for your home are inexpensive and readily available from basic to high-end designs everywhere from your local hardware store to online shopping websites such as Amazon.com.

Pool heating: Solar heating system can be used to heat up water in pool during cold seasons.

7. EXPECTED OUTCOME

After examining the performance metrics of some selected timeseries techniques (ARMA, ARIMA, SARIMA, PROPHET). Time series analysis is one of the most important aspect of data analytics for any large organization as it helps in understanding seasonality, trends, cyclicality and randomness in the sales and distribution and other attributes. It may be concluded that ARIMA will outperformed all other techniques in terms of mean absolute error and root mean absolute error. The ARMA technique perform for the 15-minutes, 30-minutes, 1-hour, and 2-hours forecasting horizons, although better than the Machine Learning techniques.

References:

The reference must be completely mentioned in the list and cite in the text of the synopsis above. There should be at least 10 latest references. Formats (it is just like IEEE style) are as under:

[1] Kesh Bahadur Pun, "SHORT TERM FORECASTING OF SOLAR POWER WITH MACHINE LEARNING AND TIME SERIES TECHNIQUES", December, 2020, pp.
[2] Authors' name(s),"Paper Title," Proceedings of Conference Name, Place, Date, pp.
[3] Authors' name(s), Book Title, Edition No. Publisher, Year, pp.
[4] **https://machinelearningmastery.com/time-series-forecasting-methods-in-python-cheat-sheet/**, "**11 Classical Time Series Forecasting Methods in Python (Cheat Sheet)**"

Appendix (-A, -B, -C etc.)- if any

This may include the dataset, specifications etc. used, if any on/with, which the project work is being carried out.

Notes:

1. The students shall write every thing in his own words. Use MSWord, 12 Point Times New Roman Font, Single Column and Laser Printer. Approximately one-inch margin on each side of the page be maintained.
2. Give illustration and tables wherever appropriate.
3. The synopsis has to be signed by every group member and endorsed by respective supervisor.
4. The external supervisor must be technically qualified having at least ME/ MTech Degree.
5. Student must maintain a diary in which the talk made with the guide should be noted date wise with guide(s) signature. Meeting the guide(s) ones in a week is compulsory. Students must bring the diary at the time of external viva.

6 The student shall not send the project work for conference/ publication without permission of HOD.