OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG
FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGY
LENA

Chair of Electric Power Networks and Renewable Energy Sources

Master Thesis

Energy Prediction Using Artificial Intelligence for Solar and Wind Power

Submitted on July 21, 2023
Surya Chandra Karanam
Student Number: 230344

First Examiner: Prof. Dr.-Ing. habil. M. Wolter

Second Examiner: Prof. Dr.-Ing. A. Lindemann

Supervisor: M.Sc. T. Woldu

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG | EIT | FAKULTÄT FÜR ELEKTROTECHNIK UND INFORMATIONSTECHNIK | LENA

Master's thesis

for Mr. Surya Chandra Karanam

Topic: Energy Prediction Using Artificial Intelligence for Solar and Wind Power

Task:

The renewable energies are not efficiently used due to lack of accurate forecast of power generation. The reasons for inaccurate power forecast could be change in the climatic conditions. There is a huge possibility of economic loss when the plant failed to generate the planned energy. In such cases, non-renewable resources are used, and it costs more than the renewable sources. Using artificial intelligence in the prediction can decrease the chances of not meeting the load demanded. A machine learning model which predicts the power that can be generated by a plant is essential to avoid such economic losses.

Therefore, the objective of this thesis is to use the renewable sources more efficiently by accurately estimating the power generation using artificial intelligence. The goal is to predict the power that can be generated by each plant in the next few weeks accurately. A suitable artificial intelligence (AI) model for prediction the power generation must be identified in literature research, and its structure and fundamental functionality have to be described. The forecast parameters shall be identified and a datasheet has to be created to train them in the AI model. To accurately forecast the upcoming power generation of each renewable power plant, at least a one-year power generation data by each plant under different conditions shall be used as input.

The study has to cover the following points:

- Literature research on power plants and parameters affecting the total energy produced
- Preparing a dataset to train the machine learning model
- Analysis of different algorithms using ML and Deep learning in python
- Compare the results and consider the best algorithm
- Documentation of results.

Magdeburg,	01.02.2023	
Date of edition:	06.03.2023	
Date of submission:	24.07.2023	Prof. Dr.-Ing. habil. M. Wolter Task tutor
Supervisor:	M.Sc. T. Woldu	
1st examiner:	Prof. Dr.-Ing. habil. M. Wolter	
2nd examiner:	Prof. Dr.-Ing. A. Lindemann	Prof. Dr.-Ing. R. Leidhold Chairman examination board

Declaration of Independence

I assure that I have written this master thesis independently without external help and used only the given sources and aids. Literally or meaning points, which were obtained from other works are labeled with their source.

Karanam, Surya Chandra

Magdeburg, July 21, 2023

Abstract

Solar and wind power are the primary sources of renewable energy worldwide. However, the amount of energy they generate depends on several environmental factors, such as temperature, irradiance, humidity, wind speed, and wind direction. Accurately predicting these conditions is a significant challenge. In addition, the integration of renewable energy sources such as solar and wind power into energy grids has created challenges for the accurate forecasting of energy supply and demand. Increasing production costs of non-renewable energy sources make it essential to forecast electricity supply to the grid in a cost-efficient way. While plant capacity is estimated mathematically, climatic factors often affect output. Artificial intelligence (AI) can help predict weather conditions and improve the accuracy of power forecasts. Additionally, a machine learning (ML) model can identify the relationship between environmental factors and energy output considering the effects of clouds and other losses in a given region.

Firstly, this involves predicting the climate accurately and then estimating power output in the second step. Efficient power scheduling can then be planned, ensuring that renewable resources are used effectively. This thesis uses a variety of ML models in the first stage to make meteorological predictions. The most accurate algorithm is then selected to forecast future weather patterns. Based on this forecast, the thesis estimates the amount of power that solar plants and wind farms will generate every hour for any selected time period.

Overall, this thesis explores the use of AI in power forecasting for solar plants and wind farms. The thesis examines the types of machine learning algorithms that can be used to determine the energy produced by solar plants and wind farms, and the data requirements and limitations of these algorithms.

Acknowledgement

First and foremost, I would like to thank Prof. Wolter for accepting my thesis proposal.

Next, I extend my heartfelt thanks and gratitude to my advisor, Mr. T. Woldu, for his invaluable contribution to my research journey. His expertise and guidance were essential in shaping the outcome of the thesis and providing me with the direction I needed to successfully complete it.

I would like to express my love and gratitude to my parents Adisiva, Parimala Kumari, and my sister Pravallika Deepak Kumar for their unwavering support, love, and encouragement throughout my academic journey. The thesis would not have been possible without their love and support.

I would also like to acknowledge the support of my friends, who provided me with motivation and encouragement during these challenging times.

Finally, I am incredibly grateful for the transformative experience I had at Otto-von-Guericke University Magdeburg, where I not only honed my academic and professional skills but also grew as a person and shaped my future aspirations.

Thank you all once again.

Contents

List of Figures

List of Tables

List of Symbols

Constants

ρ	Air density in kg/m^3
A_t	Altitude in m
C_p	Power coefficient
P	Pressure in kPa
PR	Performance ratio in %
V	Wind speed in m/s
η	Efficiency in %
A_s	Area of a solar panel in m^2
A_w	Swept area of a wind turbine blade in m^2
G	Solar irradiance in W/m^2
S_p	Solar power in kW
T	Temperature (°C)

List of Abbreviations

Glossary

AC	Alternating Current
AI	Artificial Intelligence
DC	Direct Current
GHI	Global Horizontal Irradiance
LSTM	Long-Short Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MPP	Maximum Power Point
MVTPL	Mytrah Vayu Tungabhadra Private Limited
NWP	Numerical Weather Prediction
POA	Plane of Array
PV	Photovoltaic
RFR	Random Forest Regression
SVM	Support Vector Machine
SVR	Support Vector Regression
XGB	Extreme Gradient Boosting

1 Introduction

Solar power has emerged as an effective solution to meet the growing demand for clean, renewable energy worldwide. It involves the conversion of sunlight into electricity using Photovoltaic (PV) panels. This technology offers numerous advantages, including its abundance, affordability, and zero-emission footprint. The widespread adoption of solar power has significantly reduced greenhouse gas emissions, contributing to a more sustainable future. However, there are challenges to its adoption, particularly the intermittent nature of solar power due to factors like cloud cover and weather patterns. Additionally, integrating solar power into existing energy infrastructures presents logistical obstacles.

Likewise, wind power has become a prominent source of clean and renewable energy. By harnessing the kinetic energy of the wind, wind turbines generate mechanical energy, which is then converted into electricity. The amount of power generated depends on factors such as wind speed, air density, and turbine size. Despite its benefits, the adoption of wind power also faces challenges, including intermittency and the need for advanced energy storage technologies. To address these challenges, experts have developed grids that combine both solar and wind power. These grids consist of solar plants and wind farms, working together to provide a more reliable and consistent energy supply. By leveraging the complementary nature of solar and wind energy sources, the combined output becomes more stable and reliable than relying on either source alone. Moreover, the integration of solar and wind power helps to balance the grid, enhancing energy efficiency and reducing overall costs. Accurate forecasting of energy supply and demand is crucial for efficient grid management and the optimal utilization of renewable energy sources.

In recent years, Artificial Intelligence (AI) and machine learning algorithms have proven instrumental in addressing power forecasting challenges for solar plants and wind farms. These technologies leverage historical data, taking into account factors such as temperature, wind speed, humidity, and solar irradiance to predict energy output with greater accuracy. By analyzing patterns and relationships between external factors and energy output, machine learning models can be trained to provide precise forecasts based on current and projected weather conditions. The application of AI in power forecasting not only improves grid efficiency but also enables better energy distribution and storage management, reducing reliance on traditional fossil fuel-based sources.

In conclusion, solar and wind power are pivotal in driving the transition towards a sustainable energy future. Overcoming challenges through grid integration and leveraging AI for power forecasting can enhance the efficiency and utilization of renewable energy sources, leading to a cleaner and more sustainable energy landscape.

1.1 Thesis Motivation

The motivation behind undertaking this thesis lies in the interest in the intersection of AI and the energy sector. In future, AI is poised to play a critical role in transforming various industries, and the energy sector is no exception. By leveraging advanced algorithms and machine learning techniques, there is an opportunity to revolutionize power forecasting and pave the way for a more efficient and sustainable energy landscape.

One of the key motivations driving this research is the realization that developing a tailored Python script for plant-specific power forecasting holds immense potential. By customizing the forecasting model to the characteristics and parameters of individual power plants, the accuracy can be significantly enhanced in the case of energy forecasts. This, in turn, empowers plant operators to better manage the supply and demand dynamics of the grid. The ability to balance the grid effectively can lead to improved energy efficiency, reduced costs, and minimized reliance on traditional energy sources.

Moreover, by delivering a comprehensive model capable of comprehending weather patterns and plant behavior, plant operators can gain a deeper understanding of the overall losses in the plant. This understanding is vital in identifying inefficiencies, optimizing performance, and mitigating any potential shortcomings. By using the power of AI and machine learning, it is possible to unlock valuable insights that drive continuous improvement in the energy sector. By addressing these motivations, this thesis aims to contribute to the broader goal of achieving a sustainable and intelligent energy future.

1.2 Thesis Objective

The objectives of this thesis are to:

1. Understand the factors that influence the overall energy production of solar and wind plants.
2. Identify the key parameters among these factors and utilize them to train a machine learning model.
3. Experiment with different machine learning algorithms and determine the most effective model for power forecasting.
4. Develop a model capable of accurately forecasting energy and predicting the energy supply to a grid system for user-specified dates.

1.3 Thesis Structure

This thesis is organized in seven chapters, each contributing to the exploration of AI in energy forecasting for solar plants and wind farms. The following provides an overview of the chapters:

The first chapter serves as an introduction to the thesis, presenting the underlying motivation behind the research and outlining the study's objectives. This initial chapter establishes the foundation for the subsequent sections.

Second chapter explores the seamless integration of art and science in these renewable energy sources. The focus is on elucidating the problem statements and challenges encountered in power forecasting for solar parks, wind farms, and hybrid systems.

Third chapter provides an overview of the methodologies and algorithms previously employed for power forecasting in India. Additionally, a proposed solution is introduced, taking into account the influencing parameters and losses associated with solar and wind energy systems.

Chapter 4 considers a case study of a real-time system, with specific emphasis on a particular solar plant and wind farm. The chapter outlines the plant specifications and parameters monitored by the plant authority, along with the power generation observed under varying meteorological conditions. The methodology employed is illustrated through block diagrams.

Chapter 5 presents the process of data extraction and processing is explained, with a particular focus on implementing machine learning techniques using Python. A comprehensive explanation of the Python script developed for energy forecasting is provided, along with a step-by-step guide for executing the script and utilizing the code.

Chapter 6 presents the forecasting results and a comparison with historical plant datasets. The outcomes of the energy forecasting model are thoroughly examined, considering the hourly rate of power generation from the solar plant and wind farm to the power grid for user-specified dates. These results shed light on the effectiveness and accuracy of the developed machine learning model.

The last Chapter summarizes the key findings and insights derived in the research. The implications of the study are discussed, highlighting the potential for future work and outlining steps to enhance existing methodologies in energy forecasting.

In summary, this thesis provides a comprehensive exploration of the application of AI in energy forecasting for solar plants and wind farms. Each chapter contributes to the understanding of machine learning algorithms, data requirements, and limitations within the context of renewable energy systems. The thesis concludes with a multi-layered machine learning model designed for energy forecasting, while also discussing future research possibilities and potential improvements to existing methodologies.

2 Art of Science for Solar and Wind Energy

Renewable energy resources have gained significant popularity as an alternative to non-renewable resources due to their negative environmental impacts. The key advantage of renewable energy resources lies in their sustainability and abundant availability. Consequently, there has been a growing global interest in developing and implementing renewable energy technologies. Different types of renewable energy resources exist, each with their own unique benefits and drawbacks. This section provides an overview of solar and wind energy resources [1].

- Solar energy stands out as one of the most widely adopted renewable energy resources. It relies on the sun's radiation, converting it into electricity through photovoltaic cells composed primarily of silicon. These cells absorb solar energy and transform it into electricity. Solar energy offers cleanliness, safety, and abundant availability, making it an attractive option for power generation worldwide.
- Wind energy represents another popular renewable energy resource derived from the movement of air molecules. Wind turbines convert wind energy into electricity, utilizing rotating blades that respond to wind currents. This rotational movement is then converted into electrical energy via generators. Wind energy proves particularly suitable in areas characterized by high wind speeds, such as offshore locations or mountainous regions.

In addition to the aforementioned renewable energy resources, several emerging technologies hold promise for the future. These include wave energy, tidal energy, and solar thermal energy. Although still in the early stages of development, these technologies exhibit potential to become significant contributors to the renewable energy landscape in the coming years.

Overall, renewable energy resources are essential components for achieving a sustainable energy future. Each resource possesses distinct advantages and limitations, making them suitable for power generation in various regions worldwide. Additionally, ongoing investments and advancements in emerging technologies will further enhance the significance of renewable energy in meeting the world's energy needs.

2.1 Solar Energy Influencing Factors

The power generated by a single solar cell is dependent on the photons that reach the P-layer of the cell. However, direct measurement of photons emitted by the sun is not sufficient for calculating panel output, as they must pass through multiple layers of

the atmosphere which results in diffusions in the troposphere and stratosphere. The overall energy produced by solar parks relies heavily on environmental conditions in the troposphere. Additionally, physical and electrical parameters play crucial roles in determining power generation and overall system efficiency. Through extensive research and experiments, scientists have identified key parameters that significantly impact electron extraction from a solar panel. This section provides an understanding of these influencing parameters and their mathematical expressions.
Three main categories are environmental, physical, and electrical which can affect power generation in a solar plant. Solar energy has become a vital source of renewable energy due to its abundance and availability. A solar panel's power output depends on various factors discussed in the previous section. Understanding these influencing factors is essential for designing efficient and effective solar power systems. This section examines the relationship between solar power and the factors that affect it.

2.1.1 Environmental Parameters

1. **Solar Irradiance:**

   Solar radiation refers to the energy from the sun that reaches the earth and is a critical environmental parameter affecting the power output of a solar plant. Solar radiation is known as solar irradiance and is usually measured in watts per square meter (W/m^2). It is influenced by factors such as the time of the day, angle of incidence of the sun's rays, cloud cover, season, and weather conditions [2]. Higher solar irradiance generally leads to increased power output, while lower levels can result in reduced output. Cloudy or overcast days significantly reduce the available sunlight, resulting in lower power output. Parameters similar to solar radiation, such as solar irradiance, plane of array POA, and global horizontal irradiance GHI, accurately measure the solar energy reaching the panel through diffusion and reflection. Solar irradiance is affected by the angle of incidence of the sun rays, which varies with the sun position in the sky and atmospheric conditions. The angle of incidence represents the angle between the sun ray and a line perpendicular to the solar panel. The solar panel receives the maximum amount of solar radiation when the angle of incidence is zero. The relationship between solar power and solar irradiance can be mathematically expressed using the following formula:

   $$S_p = G \cdot \eta \cdot A \cdot PR$$

   Solar Irradiance depends on solar radiation, temperature and pressure:

$$G = S \cdot \left(1 + 0.33 \cdot \left(\frac{T}{273}\right)\right) \cdot \left(1 + 0.2 \cdot \left(\frac{P}{101.325}\right)\right)$$

Here, G is the solar irradiance in watts per square meter,
PR is the performance ratio,
P represents the pressure in kilo pascals,
S is the solar radiation,
A is the panel surface area in square meters,
η is the panel efficiency

Solar Irradiance versus Power Output

Irradiance measures the solar energy received per unit area on a surface. The power generated by a solar panel is directly proportional to irradiance. The maximum power output of a solar panel is achieved at the maximum power point (MPP), a specific irradiance level. The MPP varies depending on factors such as temperature, cell type, and technology. As irradiance increases, the power output of the solar panel follows an almost linear increment until reaching the MPP. Beyond the MPP, power output decreases due to efficiency losses and increased temperatures resulting from higher irradiance levels.

Why Lower Irradiance Might Cause High Power?

It may seem counterintuitive that lower levels of irradiance can sometimes lead to higher power output. A solar panel is influenced by various factors and one such factor is the maximum power point (MPP). It describes the point where the panel generates the maximum power given specific conditions. At low levels of irradiance, the MPP shifts to lower voltage and higher current, resulting in higher power output. This is because, at low irradiance, the panel cannot generate sufficient voltage to overcome losses and reach its maximum power output. By lowering the voltage, the panel can operate at a higher current and generate more power [3]. However, as irradiance continues to decrease, power output eventually decreases as well. This occurs when the current also decreases and the voltage is no longer able to compensate for the losses. Additionally, high panel temperatures can decrease power output regardless of the irradiance level. While lower irradiance can sometimes lead to higher power output, it is crucial to consider various factors that affect the MPP and power output and optimize the panel's operating conditions for maximum efficiency.

Why High Irradiance Might Give Lower Power Output?

While it is generally true that higher irradiance leads to higher power output in a solar panel, there are cases where this relationship doesn't hold. This is due to the temperature coefficient, which describes how changes in temperature affect the panel's performance. When a solar panel is exposed to high irradiance, it absorbs more energy and generates more electricity. However, the panel's temperature also increases, negatively impacting its efficiency. As a result, the cell performance decreases, which results in a decrease in power [3].

The temperature coefficient measures how much the panel's efficiency decreases with each degree celsius increase in temperature. For example, a panel with a temperature coefficient of -0.5 %/°C will experience a 0.5 % efficiency decrease for every degree celsius rise in temperature. In certain cases, the decrease in efficiency caused by higher temperatures can outweigh the increase in power output from higher irradiance. This is especially true in hot climates or when operating at high temperatures. In such conditions, the temperature coefficient can cause a decrease in power output despite high irradiance levels. It is vital to consider that the temperature coefficient is not the only factor affecting output power. The angle of the sun, reflectivity of surrounding surfaces, panel orientation, and shading also play significant roles. Therefore, considering all these factors is crucial when designing and optimizing solar power systems.

Understanding the relationship between irradiance and power output is essential for optimizing the performance of solar plants and accurately forecasting power output. By accurately predicting irradiance levels, it becomes possible to estimate the power output of solar plants and plan energy production accordingly.

2. **Temperature:**

   Temperature is another crucial environmental parameter that affects the efficiency of the panels. As the temperature of the panel increases, its efficiency decreases. Solar panels operate most efficiently at moderate temperatures, typically between 25-35 degrees celsius. At higher temperatures, the electrical resistance of the panel increases, leading to reduced power output. The temperature coefficient represents the percentage decrease in power output per degree celsius increase above a reference temperature, often 25°C [3]. There are two types of temperatures to be monitored in a plant.

   Module Temperature The temperature of solar panels, known as module temperature, significantly affects their performance and efficiency. Higher temperatures can cause a decrease in energy conversion efficiency. Therefore, accurate measurements of module temperature are crucial for power forecasting. By incorporating module temperature in the forecasting models, adjustments can be made to the predicted power output, accounting for temperature variations and their impact on solar panel performance.

Ambient Temperature Ambient temperature, or the temperature of the surrounding environment, also influences the efficiency of solar panels. High ambient temperatures can lead to increased module temperatures and reduced energy conversion efficiency. Therefore, monitoring and considering ambient temperature in power forecasting is important for more precise predictions of the actual power output of solar systems.

3. **Global Horizontal Irradiance (GHI):** GHI measures the solar radiation that reaches the surface, encompassing both direct and diffuse components on a horizontal plane. The output power of a solar panel is directly proportional to GHI, as it serves as the primary energy source for electricity generation [4]. However, the relationship between GHI and solar power output is not always linear. Factors like temperature and shading can cause non-linearities, resulting in efficiency losses at higher GHI levels. Overheating due to excessive GHI can reduce efficiency and power output. Careful monitoring of panel angle, orientation, temperature, and shading is essential for optimal power output.
4. **Plane of Array (POA) Irradiance:** POA irradiance specifically measures the irradiance reaching the panel surface, perpendicular to its orientation. This parameter is also directly influenced by the amount of POA irradiance received [5]. Accurately predicting POA irradiance is crucial for forecasting solar power output, considering factors such as panel angle, orientation, shading, and atmospheric conditions.
5. **Humidity** Humidity is the measure of water vapor present in the air and can also affect solar panel efficiency by reducing the sunlight that reaches the panel [6]. However, for the purpose of this thesis, the effect of humidity is neglected due to data availability constraints.

2.1.2 Physical Parameters

1. **Panel orientation and tilt:** The orientation of the solar panel refers to its direction relative to the sun, while the tilt represents the angle between the panel and the ground. The tilt affects the amount of irradiance received by the panel. When the panel is tilted, the irradiance decreases, reducing energy production. Optimal performance requires solar panels to face the sun directly. The ideal orientation and tilt depend on the panel's location. In the northern hemisphere, the panel must be

facing south to receive most possible energy, while in the southern hemisphere, a north-facing orientation is preferred. The tilt angle is determined by the location's latitude, with a tilt angle close to 0 degrees at the equator and closer to 90 degrees at higher latitudes. The tilt angle can be calculated using the formula [6]:

$$\text{Tilt angle} = \text{Latitude} + 15^{\circ} \text{ during winter}$$
$$\text{Tilt angle} = \text{Latitude} - 15^{\circ} \text{ during summer}$$

2. **Shading:** Shading from trees, buildings, or other objects can diminish the amount of sunlight reaching the panel, leading to a decrease in power output. Even minimal shading can significantly impact the panel's power output.

3. **Age and degradation:** Solar panels experience efficiency loss over time due to wear and tear and exposure to the elements. Aging panels can become damaged, dirty, or less effective at converting sunlight into electricity.

4. **Dust and pollution:** Dust and pollution in the air can hinder the amount of sunlight reaching solar panels. When dust and pollutants settle on the panels, they reduce the penetration of sunlight, thereby decreasing efficiency. This effect is particularly notable in areas with high air pollution levels, such as cities or industrial regions.

2.1.3 Electrical Parameters

1. **Series and parallel connection:** The electrical connections between photovoltaic cells influence the voltage and currents produced by the solar plant. Series connections maintain the same voltage across equipment, while parallel connections increase the voltage while maintaining a constant current. The topology and interconnections of the solar plant are subjective and depend on the overall power grid system. A combination of series and parallel connections is often chosen to balance grid requirements.

2. **Electrical equipment:** The type of inverter, wiring, and other electrical components can affect the performance of the solar plant. These parameters impact the efficiency of converting DC power generated by the panels to AC power which is later fed to the grid. All the inverters that are improperly sized or configured can lead to power losses and reduced plant efficiency.

These five parameters—solar irradiance, POA irradiance, GHI, module temperature, and ambient temperature—were selected for power forecasting due to their direct impact on solar power output. Incorporating these parameters into forecasting models enables more accurate predictions, taking into account the specific conditions, orientation, and efficiency of solar panels. By considering these parameters, reliable power forecasts can be made, supporting efficient energy planning and management. In summary, numerous factors influence the power output of solar panels. Sunlight levels, panel angle and orientation, temperature, and shading are crucial factors to consider for optimizing power output. However, monitoring all these factors can be complex and time-consuming. To streamline the approach, this thesis focuses on five major parameters with significant effects on power generation. These parameters will be explained in the subsequent section.

2.1.4 Unveiling the Losses in Solar Power Generation

A solar power plant harnesses photovoltaic (PV) modules to convert solar radiation into electricity, but there are various losses that reduce overall power generation efficiency. Understanding the sources of system's loss is essential for optimizing plant design and operation. This section discusses the different types of losses occurring in solar power generation [7].

1. **Module temperature losses:** The temperature of the solar panel significantly affects power generation efficiency. As the module's temperature increases, its efficiency decreases.
2. **Wiring and connection losses:** Electrical connections between solar panels and inverters can introduce losses due to resistance in the wiring. These losses are proportional to the wiring length and the current flowing through it.
3. **Mismatch losses:** Mismatch losses occur when solar panels of different types or orientations are used together, leading to a decrease in overall system efficiency.
4. **Inverter losses:** The inverter, responsible for converting DC power to AC power, can introduce losses if not properly sized or configured.
5. **Shading losses:** Shading losses arise when solar panels are partially or fully shaded by objects such as trees or buildings, resulting in decreased system efficiency.
6. **Dust and soiling losses:** The dust and pollutants accumulated on the panel surface stops the sunrays and reduces power generation efficiency.
7. **Reflection and transmission losses:** Reflection and transmission losses occur when solar radiation is reflected or transmitted through the panel. The use of anti-reflective coatings can mitigate these losses.

Understanding the various losses in a solar power plant is crucial for estimating power output and ensuring maximum efficiency. By identifying these losses, plant operators can increase profitability and contribute to the transition toward sustainable energy.

Calculating each loss is time-consuming and requires a deep understanding of multiple factors. However, these losses generally follow certain patterns based on the environmental parameters mentioned above, and these patterns can be identified using machine learning. This thesis aims to streamline the estimation of overall losses, providing faster and accurate results.

2.2 Wind Energy Influencing Factors

Various factors can impact the power generation of a wind farm, including environmental, physical, and electrical parameters. Gaining an understanding of these parameters and their interactions is important for optimizing solar plant performance and maximizing renewable energy generation. These parameters can be mathematically modeled to describe their effects on power generation.

2.2.1 Environmental Parameters

1. **Wind Speed:** Wind speed is a crucial factor influencing wind power output. Higher wind speeds result in increased kinetic energy, leading to higher power production. The relationship between them follows a nonlinear pattern, adhering to the power curve that is defined by the wind turbine manufacturer. Generally, the output power of a turbine is directly proportional to the cube of the wind speed. The power output can be estimated using the formula [8]:

   $$P = 0.5 \cdot \rho \cdot A \cdot C_\mathrm{p} \cdot V^3$$

   In this representation,
   P represents Power
   ρ represents air density
   A represents rotor swept area
   C_p represents power coefficient
   V represents wind speed

2. **Wind direction:** Wind direction significantly impacts wind turbine efficiency and performance. Optimal wind direction aligns the rotor blades with the incoming wind, maximizing power production. Wind turbines are designed to operate within specific wind direction ranges for optimal performance [9]. Wind rose analysis is a technique used to determine wind direction frequency and distribution in a specific location over time. This analysis provides valuable information on prevailing

wind directions and their corresponding frequencies, aiding wind power project developers in turbine placement and orientation. Wind direction can vary over time, and wind turbines have tolerance ranges for wind direction variations. If variations fall within the range, turbines can adjust their orientation for optimal energy capture. However, significant variations beyond the range may require turbines to shut down or operate at reduced capacity to prevent potential damage.

3. **Turbulence intensity:** Turbulence intensity refers to wind speed and direction variations caused by atmospheric conditions, obstacles, and wind shear. Higher turbulence intensity can increase loads on wind turbines, affecting their performance and power output. Turbines are designed to handle specific turbulence intensity levels, exceeding which may lead to reduced efficiency and increased maintenance requirements.
4. **Relative humidity:** Relative humidity measures the water vapor in the air compared to the maximum capacity at a given temperature, indirectly influences wind behavior. Higher relative humidity indicates saturated air, potentially resulting in stable atmospheric conditions and less energetic wind patterns [6]. The specific effects on wind power generation are complex and depend on local climate and geography. Relative humidity impacts air density, which, in turn, affects power output. As relative humidity increases, air density decreases, leading to decreased power output.
5. **Air density:** Air density directly affects wind power generation and it is directly proportional to turbine output. Factors such as altitude, humidity, and temperature influence air density [10]. As air density decreases, wind turbine power output decreases as well. Standard air density is considered to be 1.225 kg/m^3 but decreases with increasing altitude.

2.2.2 Physical Parameters

Turbine characteristics: Wind turbine design and specifications significantly impact power output. Parameters like rotor diameter, hub height, blade design, and generator type influence turbine performance. Modern turbines optimize aerodynamics and components for maximum power production.

Wind farm layout: Efficient wind farm layouts consider wind direction patterns. Careful arrangement of turbines minimizes negative effects such as wake and turbulence from upstream turbines on downstream ones. Proper spacing reduces interference and power losses due to wake effects.

Mechanical and electrical losses: Mechanical losses like friction in bearings, gearbox, and generator, along with electrical losses from transformers, cables, and converters, affect overall turbine efficiency. Efficient component design and maintenance practices minimize these losses to maximize power generation.

Wake effects: Operating wind farms experience wake effects, leading to reduced wind speeds and turbulence downstream. Downwind turbines are affected. Optimized turbine layouts and spacing minimize wake effects, enhancing overall power output.

Rotor diameter: Rotor diameter directly impacts wind energy capture. Power generated is proportional to the square of rotor diameter. Larger rotors increase swept area, capturing more wind energy. However, larger rotors require robust support structures and may have space or height limitations.

Blade angle: Blade angle determines power extraction from the wind. Wind turbine control systems automatically adjust blade angles based on wind speed and direction to optimize power output.

Wind Shear: Wind shear refers to wind speed and direction variations over a distance. It affects turbine performance by causing uneven blade loading. Anemometers placed at different heights measure wind shear, enabling adjustment of blade angles for optimized power output.

Blade pitch: Blade pitch determines the angle of attack of turbine blades relative to wind direction. It influences rotor aerodynamic efficiency and power output. Maintenance and Operation: Regular maintenance ensures efficient turbine operation and prevents breakdowns. Optimizing turbine operation maximizes efficiency and power output.

In summary, the turbine output is influenced by factors such as wind speed, air density, blade size, turbine efficiency, blade angle, wind shear, and wake effects. Understanding and measuring these factors optimize wind energy systems for maximum power output and efficiency. By analyzing these parameters, wind energy operators can identify improvement opportunities and implement strategies to enhance wind energy production.

2.2.3 Electrical Parameters

Voltage and frequency: The grid voltage and frequency play a crucial role in wind turbine performance. Synchronization with the grid ensures stable power delivery. Deviations in grid parameters affect power output and may require strategies like reactive power compensation and voltage regulation.

Power factor: Power factor is the ratio of real power delivered and the apparent power of an electrical system. Wind turbines aim to operate at a unity power factor or a desired power factor defined by grid codes. Controlling the power factor optimizes power transfer and grid stability.

Turbine efficiency: Turbine efficiency measures the conversion of wind's kinetic energy into electrical energy. Factors such as turbine design, blade size, and material influence efficiency. It is typically expressed as a percentage, ranging from 30% to 50%, and depends on turbine design and component quality.

2.2.4 Unveiling the Losses in Wind Power Generation

A wind farm harnesses wind energy to generate electricity, but various factors can lead to losses that reduce turbine efficiency. These losses can be classified into mechanical and electrical categories. Let's explore the different types of losses in a wind farm and their impact [11].

Mechanical losses result from friction within turbine components and aerodynamic losses caused by wake effects and turbulence. Key mechanical losses include bearing, gearbox, generator, and blade and hub losses. These losses are quantified using a mechanical power loss coefficient representing the percentage of mechanical power lost due to friction and aerodynamics.

Electrical losses occur due to resistance in electrical components like cables, transformers, and inverters. Significant electrical losses include generator, transformer, cable, and inverter losses. These losses are expressed using an electrical power loss coefficient (Ce), representing the percentage of electrical power lost due to resistance.

Total losses in a wind farm combine mechanical and electrical losses. To enhance overall efficiency, it is crucial to minimize losses. Strategies such as proper maintenance, lubrication, and alignment of turbine components reduce mechanical losses. Additionally, utilizing high-quality electrical components and optimizing wind turbine layout can minimize electrical losses.

Aerodynamic losses refer to energy losses obtained while converting the wind's kinetic energy to mechanical energy. These losses stem from various factors and impact wind power generation efficiency. Let's explore them in detail:

- Turbine blade drag: Blade drag occurs when wind encounters resistance on the turbine blade surfaces, resulting in energy loss. Factors like blade shape, surface roughness, and wind speed influence blade drag. Optimizing blade design with aerodynamic profiles minimizes turbulence, improves lift-to-drag ratios, and reduces drag forces.

- Turbulence and flow separation: Turbulent wind conditions, such as wind shears or gusts, induce irregular airflow around turbine blades, leading to increased drag and reduced energy extraction. Turbulence can also cause flow separation, detaching airflow from blade surfaces, resulting in decreased lift and increased drag. Measures like blade pitch control, rotor yaw control, and aerodynamic stall control mitigate turbulence effects and maintain stable airflow, minimizing losses.
- Tip and root losses: Blade tips and root sections experience unique aerodynamic effects contributing to energy losses. At blade tips, increased wind velocity caused by tip vortices generates higher drag and reduced efficiency. Near the blade root, where blades connect to the hub, increased turbulence and disturbances increase drag and energy loss. Blade designs and rotor configurations are optimized to mitigate these losses by minimizing vortex effects and improving airflow.

Wake effects occur when turbines extract energy from the wind, leading to reduced wind speeds and altered airflow patterns in the wake region behind turbines. Downstream turbines experience lower wind speeds, resulting in decreased energy capture and increased aerodynamic losses. Optimal wind farm layout design and turbine spacing minimize wake effects and associated losses.

Other losses include blade curvature losses, tip losses, and tower shadow losses, which can be estimated using empirical formulas or computational fluid dynamics (CFD) simulations. It's important to note that losses may vary depending on turbine design and operating conditions. Regular monitoring and maintenance are crucial to minimizing losses and optimizing energy production.

In summary, wind power losses occur due to mechanical friction, aerodynamic factors like blade drag, turbulence, flow separation, tip and root effects, wake effects, and tower shadow. Wind turbine design, blade profiles, rotor configurations, wind farm layout optimization, and advanced control mechanisms aim to mitigate these losses and enhance overall wind power generation efficiency.

In conclusion, Chapter 2 explores the influencing factors of solar and wind energy. The analysis encompasses various aspects, including environmental, physical, and electrical parameters, as well as losses in power generation. By comprehending these factors, a deeper understanding of the complexities involved in harnessing solar and wind energy can be achieved. The investigation of environmental parameters provides insights into the impact of weather conditions, such as solar irradiance and wind speed, on energy generation. Similarly, understanding the physical parameters, including the design and orientation of solar panels and wind turbines, is crucial for optimizing energy production.

Additionally, the examination of electrical parameters sheds light on voltage regulation, power conversion, and transmission aspects, contributing to efficient energy utilization. Furthermore, the exploration of losses associated with solar and wind power generation highlights the importance of identifying and mitigating factors such as shading, soiling, and system inefficiencies. By addressing these losses, the overall energy output can be maximized.

The outlined sections in this chapter cover all relevant aspects related to solar and wind energy influencing factors. This comprehensive analysis provides a solid foundation for the subsequent chapters of this thesis. It lays the groundwork for the exploration of the role of artificial intelligence in forecasting the generation of solar and wind energy, which will be the focus of Chapter 3.

3 Hybrid Systems and Forecast Methodologies

Transitioning to renewable energy sources is essential for mitigating the adverse impacts of climate change and reducing reliance on fossil fuels. The fields of solar and wind energy have witnessed remarkable advancements and widespread adoption as prominent renewable energy technologies. However, the intermittent nature of both solar and wind energy, resulting from weather patterns and diurnal variations, can limit their effectiveness individually. To overcome these limitations and enhance renewable energy generation, hybrid systems that combine solar and wind energy have emerged as a promising solution. By harnessing the complementary attributes of solar and wind resources, hybrid systems offer improved reliability and energy output. This article delves into the concept of hybrid systems, specifically those comprising solar and wind energy, and explores their advantages and potential applications.

The integration of solar plants and wind farms in hybrid systems presents several advantages over standalone solar or wind systems. Firstly, hybrid systems bolster overall reliability by mitigating the intermittent nature of renewable energy generation. While solar energy gives peak outputs in daylight hours, wind energy tends to be consistent all over the day and night. By synergistically utilizing both energy sources, hybrid systems provide a stable and reliable power output. Usually, one of the main parameters monitored by plant operators is the capacity factor representing the ratio between actual energy to the maximum output, compared to standalone solar or wind systems. This is made possible by the complementary generation patterns of both plants, resulting in a more consistent overall energy production.

3.1 Hybrid Systems

The integration of both the plants in hybrid systems has emerged as a promising approach to maximize renewable energy generation. Hybrid systems harness the complementary nature of solar and wind energy, enhancing system reliability, optimizing capacity factors, and promoting optimal resource utilization [12]. This subsection delves into the design considerations and plant specifications associated with hybrid systems. Furthermore, it highlights the systematic utilization of AI. The findings underscore the potential of hybrid systems to significantly contribute to the transition towards a sustainable energy future.

The successful implementation of hybrid systems relies on proper design and integration strategies. Sizing and optimization play pivotal roles in effectively utilizing solar and wind resources. Accurate estimation of solar and wind resources, coupled with load analysis, facilitates determining the appropriate capacity and configuration for the hybrid system. Advanced modeling techniques, including simulation and optimization algorithms, aid in achieving the optimal balance between solar and wind capacities. Grid integration is another critical consideration as it enables efficient management and distribution of energy. Integrating energy storage technologies like batteries or pumped hydro storage within hybrid systems allows excess energy to be stored when the energy is generated more than anticipated and later supplied when there is lower power generated, thereby further enhancing system reliability and stability.

3.1.1 Complementary Characteristics of Hybrid Systems

Solar and wind power plants possess complementary characteristics that make them ideal for integration within hybrid systems. These characteristics include [12]:

Seasonal and daily variation: Solar energy generation reaches its peak during daylight hours when sunlight is abundant, while wind energy generation tends to be more consistent throughout the day and can peak during nighttime. This seasonal and daily variation allows for a more balanced power generation profile when solar and wind plants are combined within a hybrid system. By integrating both energy sources, a hybrid system ensures a consistent power supply throughout the day and year, reducing reliance on a single energy source.

Geographic distribution: Solar resources tend to be more abundant in regions closer to the equator, while wind resources are more widely distributed globally. This geographic distribution offers the potential to develop hybrid systems in diverse locations, capitalizing on the available solar and wind resources in different regions. By combining solar and wind plants, hybrid systems can leverage the strengths of each technology and maximize energy generation.

Complementary generation profiles: Solar power generation typically peaks during the summer months, while wind power generation can be more pronounced during the winter. This complementary generation profile aligns well with the seasonal variations in energy demand. Integrating solar and wind plants within a hybrid system enables a more balanced and stable power output throughout the year, reducing the impact of seasonal fluctuations in energy production.

Mitigating intermittency: Solar and wind energy sources exhibit inherent intermittency due to factors such as cloud cover, nighttime hours, and variable wind speeds. By combining these two renewable sources in a hybrid system, the intermittent nature of each technology can be mitigated. When one source experiences lower generation, the other source can compensate, resulting in a more reliable and consistent power supply.

3.1.2 Disadvantages of Hybrid Systems

While hybrid systems comprising solar and wind energy offer numerous advantages, it is important to acknowledge and address their potential disadvantages. Some of the disadvantages associated with hybrid systems include [13]:

Complexity and higher initial costs: Integrating multiple energy sources within a hybrid system requires complex design, control, and monitoring systems. This complexity can lead to higher installation and maintenance costs compared to standalone solar or wind systems. Additionally, the need for additional equipment such as power converters, energy storage systems, and grid integration components can further increase the initial costs.

Land and space requirements: Hybrid systems typically necessitate more significant land or space compared to individual solar or wind installations. Combining solar panels and wind turbines within a single system requires adequate space to accommodate both technologies, which can be challenging in densely populated areas or locations with limited available land.

Variable output and system control: Efficient control algorithms are essential for managing the variable output from solar and wind sources within hybrid systems. Balancing power generation from different sources and optimizing energy flow to the grid or storage systems can be complex. Ensuring stable and reliable operation under varying weather conditions and fluctuating energy demands requires advanced control strategies and real-time monitoring.

Weather dependency: The performance of hybrid systems is influenced by weather conditions, such as solar irradiation levels and wind speed. Simultaneous decreases in solar and wind resources can affect overall energy generation. Careful planning and consideration of backup power sources or energy storage systems are necessary to ensure a continuous power supply.

Maintenance and technical challenges: Hybrid systems involve multiple components, including solar panels, wind turbines, inverters, batteries, and control systems. Each component requires maintenance and monitoring to ensure optimal performance. The complexity of the system can pose challenges in terms of troubleshooting, repair, and overall system management.

Addressing these disadvantages necessitates thorough planning, proper system design, advanced control algorithms, and regular maintenance. Despite the challenges, hybrid systems still offer significant benefits and can play a pivotal role in achieving a sustainable and resilient energy future.

3.2 Power Forecasting Methodologies for Power Forecasting

Accurate power forecasting is crucial for optimizing the operation and management of solar plants, wind farms, and hybrid systems in India. By accurately predicting power generation, operators can effectively plan and schedule energy dispatch, grid integration, and maintenance activities. Various methodologies have been employed to forecast power output, combining historical data, weather information, and advanced modeling techniques. This section explores the methodologies previously and currently used for power forecasting in solar, wind, and hybrid systems, along with their performance evaluations.

3.2.1 Solar Power Forecasting Methodologies

Different forecasting technologies used are as following [14]:

1. Clear sky model:

   The clear sky model uses solar radiation data, cloud cover observations, and atmospheric conditions to estimate solar power output. It calculates the expected power generation based on the available solar energy and the system's characteristics. However, this method does not consider local weather patterns and can result in significant deviations from the actual power output.

2. Numerical Weather Prediction (NWP):

   NWP models utilize fundamental atmospheric variables like temperature, wind speed, and others to forecast solar power generation. These models incorporate sophisticated algorithms and mathematical equations to simulate the behavior of the atmosphere and predict solar irradiance. NWP models provide improved accuracy compared to clear sky models by considering local weather conditions.

3. Statistical models:

   Statistical approaches utilize historical data of solar power generation and corresponding weather parameters to develop forecasting models. Simple regression techniques, like multivariate linear regression and polynomial linear regression are commonly employed to establish relationships between weather variables and solar power output. These models can capture complex interactions and dependencies, leading to more accurate forecasts.

4. Satellite imagery and data assimilation:

   Satellite imagery, including cloud cover and solar radiation measurements, combined with data assimilation techniques, are utilized to improve solar power forecasts. Data assimilation integrates real-time observations with numerical models, resulting in more precise predictions by continuously updating the model inputs based on the latest measurements.

3.2.2 Wind Power Forecasting Technologies

Different wind forecasting technologies are as following [15]:

1. Numerical Weather Prediction (NWP): Similar to solar power forecasting, NWP models are extensively used for wind power forecasting. These models analyze atmospheric conditions, including wind speed, wind direction, temperature, and pressure, to estimate wind power generation. NWP models have advanced over time, incorporating higher resolutions and more sophisticated algorithms, leading to improved accuracy in wind power forecasts.

2. Time-Series Models: Time-series models utilize historical wind power data to capture patterns and trends. Autoregressive Integrated Moving Average (ARIMA) and autoregressive conditional heteroskedasticity (ARCH) models are commonly used to predict wind power based on past observations. These models can handle seasonality and account for temporal dependencies, making them suitable for short-term wind power forecasting.

3. Ensemble Forecasting: Ensemble forecasting combines multiple forecasts from different models or model configurations to enhance prediction accuracy. It takes advantage of the diversity of models and accounts for the inherent uncertainties in weather forecasting. Ensemble forecasting methods, such as the ensemble mean and weighted average, provide more robust wind power forecasts.

The performance of power forecasting methodologies is evaluated based on metrics like mean absolute error, accuracy, and correlation coefficient (R). These metrics assess the model performance by comparing the predicted power output with the actual measurements. The performance of each methodology can vary depending on factors such as geographical location, system configuration, and data availability.

Hybrid systems encompassing both solar and wind energy sources can benefit from a combination of the methodologies mentioned above. Power forecasting in hybrid systems involves integrating forecasts for solar and wind power outputs, considering their respective contributions and intermittency. This requires advanced modeling techniques that combine weather forecasts for solar irradiance, wind speed, and other relevant parameters. Statistical models, machine learning algorithms, and hybrid modeling approaches are employed to predict the power output of hybrid systems accurately.

Continuous research and development are undertaken to improve power forecasting methodologies in India. Advancements in machine learning, data assimilation techniques, and the availability of high-resolution weather data contribute to enhanced forecasting accuracy. The integration of real-time measurements, artificial intelligence, and ensemble forecasting approaches hold promise for further improving the performance of power forecasting in solar, wind, and hybrid systems, ultimately facilitating the efficient integration of renewable energy into any power grid.

3.3 Problem Statement for Solar Energy

As the demand for solar power continues to rise, understanding the influence of weather conditions on solar power generation becomes increasingly important. Weather plays an important role in the overall efficiency of solar power systems. Therefore, comprehending local weather patterns in areas where solar panels are installed is crucial. Unfortunately, weather conditions can be highly unpredictable, making it challenging to accurately forecast the amount of solar power generated at any given time. Factors such as cloud cover, rain, snow, and wind all impact the amount of sunlight reaching the solar panels and, consequently, the power output.

Among these factors, cloud cover has a substantial influence on solar power generation. When clouds obscure the sun, the solar radiation reaching the panels decreases significantly, resulting in reduced power generation. The extent to which cloud cover affects solar power depends on factors such as cloud thickness, type, and time of day. Thin, high-altitude clouds may have a minimal impact, while thick, dark clouds can considerably diminish solar power generation. Rain and snowfall also affect solar power output. Rainfall reduces the amount of sunlight reaching the panels, resulting in decreased power generation. Snow accumulation on the panel surface exacerbates this reduction by further blocking sunlight and reflecting it away, leading to a decrease in power output.

Furthermore, wind can impact solar power generation by carrying dust and debris that settle on the solar panels, obstructing sunlight and diminishing power output. Wind can also induce vibrations in the panels, potentially causing microcracks in the solar cells, thereby reducing their efficiency over time. It is worth noting that while weather conditions can significantly affect solar power generation, various technologies, such as solar trackers, can help mitigate the impact of adverse weather conditions. Solar trackers dynamically orient the panels to follow the sun's movement across the sky, thereby maximizing sunlight exposure even under suboptimal weather conditions.

In summary, weather conditions exert a substantial influence on the efficiency and reliability of solar power systems. Cloud cover, rain, snow, and wind all impact the amount of sunlight reaching the panels and, consequently, the power output. Understanding the effects of weather on solar power generation is vital for designing and maintaining efficient solar power systems. One potential methodology to address the overall losses associated with solar energy production is the utilization of machine learning (ML), which will be further explained in subsequent sections.

3.4 Problem Statement for Wind Energy

Wind power plays a significant role in the renewable energy landscape, contributing a substantial share to the global energy mix. However, the performance and predictability of wind turbines heavily rely on weather conditions, which can be highly unpredictable [3]. Wind speed and direction exhibit significant variability, posing challenges in accurately forecasting energy generation at specific times.

Weather conditions are influenced by several factors, including atmospheric pressure, temperature, humidity, and local geography. Wind speed and direction can fluctuate rapidly over short distances and timeframes, making it difficult to forecast wind energy output accurately. Sudden changes in wind speed, for instance, can cause turbines to stall or shut down, reducing overall efficiency and output. To address these challenges, wind energy companies employ a range of tools and techniques to predict and manage weather-related risks.

One approach involves utilizing sophisticated weather modeling software that simulates and forecasts wind speeds and directions based on various data inputs, such as satellite imagery, meteorological observations, and historical wind patterns. These models enable operators to anticipate changes in wind conditions and adjust turbine settings accordingly, optimizing performance and minimizing downtime.

Another strategy involves using advanced sensors and control systems to monitor real-time wind conditions. This allows operators to adjust turbine settings dynamically according to the wind speed and direction. Turbines may be equipped with pitch control systems that adjust the blade angles or yaw control systems that orient the entire turbine to capture the maximum amount of wind energy. Despite these efforts, weather-related variability remains a significant challenge, especially in regions with highly unpredictable weather patterns.

Sudden changes in wind speed and direction can lead to a phenomenon known as "turbine wake," where wind turbines create a low-pressure zone behind them. This reduces wind speed and diminishes the efficiency of downstream turbines. Additionally, extreme weather events like hurricanes, typhoons, and thunderstorms can cause severe damage to wind turbines, resulting in costly repairs and downtime. While modern wind turbines are designed to withstand extreme weather conditions, unexpected damage or failure can still occur. Wind energy companies have developed various tools and techniques to manage weather-related risks. However, ongoing research and innovation in this field are essential to ensure the continued growth and success of the wind energy industry.

3.5 Proposed Solution with AI

To address the problem of unpredictability in solar and wind energy generation caused by weather conditions, the integration of artificial intelligence (AI) has emerged as a powerful solution for accurate energy forecasting. This proposed solution focuses on utilizing a combination of machine learning and deep learning algorithms to forecast relevant environmental parameters. The specific algorithms chosen for this study are Support Vector Regression, Random Forest Regression, Extreme Gradient Boosting, and Long Short-Term Memory (LSTM).

The objective is to determine the most suitable algorithm for forecasting each environmental parameter by conducting a thorough exploration. The performance of these models is evaluated by comparing their predictions with historical data. By integrating the best-performing algorithms into a single model, the overall forecast performance can be significantly enhanced.

The integration of AI brings valuable capabilities for analyzing data from diverse sources such as weather sensors, satellite imagery, and historical records. Leveraging this data, AI models generate precise weather forecasts that play a crucial role in optimizing the operation of renewable energy systems and managing energy output.

In the domain of solar energy, accurate prediction of solar irradiance is essential for forecasting solar energy generation. AI-based models utilize historical data, weather forecasts, and environmental information to estimate the amount of solar radiation reaching the panels. This knowledge enables optimization of solar panel operation, adjustments to panel tilt and orientation, and accurate predictions of energy generation at specific times. Studies have demonstrated that AI-based solar irradiance forecasting models can enhance the accuracy of solar energy generation forecasts by up to 30

Similarly, AI-based models are deployed in wind energy to forecast wind speed, direction, and turbulence, which are critical factors in predicting wind energy generation. These models analyze data from weather sensors, satellite imagery, and historical records to generate precise wind forecasts. The forecasts contribute to optimizing wind turbine operation, adjusting blade pitch angles, and predicting energy generation levels at specific time intervals. AI-based wind speed forecasting models have shown improvements of up to 10% in the accuracy of wind energy generation forecasts. Furthermore, AI algorithms facilitate the optimization of renewable energy system operation and enable more efficient energy output management. For instance, they can predict energy demand and adjust renewable energy system output accordingly, minimizing wastage and enhancing overall system efficiency. Additionally, AI models can detect faults and anomalies in renewable energy systems, leading to swift and effective maintenance and repairs.

In conclusion, the integration of AI in renewable energy systems provides a viable solution to address the unpredictability in energy generation within hybrid systems. AI-based models excel in accurate energy forecasting, optimization of renewable energy system operation, and efficient energy output management. These capabilities are pivotal for the successful integration of renewable energy systems into the power grid and for reducing the carbon footprint of the energy sector. The focus of this thesis is primarily on accurate predictions of solar irradiance, wind speed, and other relevant parameters. The selected machine learning algorithms, such as Support Vector Regression, Random Forest Regression, Extreme Gradient Boosting, and deep learning algorithm Long Short-Term Memory (LSTM), are utilized to forecast meteorological variables and improve energy forecasts.

4 Real-time Hybrid System and Specifications

This chapter introduces a case study of a hybrid system comprising a solar park and wind farm located in Kurnool, India. The case study focuses on understanding the integration of solar and wind energy sources and their impact on power output. Additionally, this section highlights how the data of important environmental parameters are monitored and obtained from an internal employee working at Azure Power, India. The parameters relevant to power output are shortlisted to provide valuable insights into the system's performance.

The hybrid system in Kurnool, India, combines the Ultra Mega Solar Plant and the MVTPL Wind Farm, both operating in the region. These two energy sources complement each other, maximizing the overall energy generation and improving the reliability of the system. The solar park harnesses solar energy through photovoltaic panels, while the wind farm utilizes wind turbines to convert wind energy into electricity.

To monitor and analyze the performance of the hybrid system, data on important environmental parameters are collected. These parameters include solar irradiance, plane of array (POA) irradiance, global horizontal irradiance (GHI), module temperature, ambient temperature, wind speed, and wind direction. These parameters have a direct influence on the power output of the solar park and wind farm. By monitoring and analyzing these parameters, insights can be gained into the factors affecting the energy generation and the system's overall efficiency.

The dataset used for this case study is obtained from an internal employee working at Azure Power, India. The employee has access to real-time and historical data collected from various monitoring systems within the hybrid system. The dataset provides valuable information on the environmental conditions, allowing for a comprehensive analysis of the system's performance. By shortlisting the important environmental parameters, this case study aims to provide a deeper understanding of how these factors impact the power output of the hybrid system. The selection of these parameters is based on their direct influence on energy generation and their significance in power forecasting and optimization.

In conclusion, this section introduces the case study of a hybrid system consisting of a solar park and wind farm in Kurnool, India. It highlights the data monitoring process, the influence of important environmental parameters on power output, and the dataset obtained from an internal employee at Azure Power, India. The subsequent sections will delve into the specifications of the Ultra Mega Solar Plant and MVTPL Wind Farm, providing a detailed analysis of each component. The utilization of machine learning in power and weather forecasting will be discussed in the following section, building upon the information presented in this chapter.

4.1 Ultra Mega Solar Park

The Kurnool Ultra Mega Solar Park, located in the Kurnool district of Andhra Pradesh, India, stands as one of the largest solar power parks globally. Spanning approximately 5,932 acres, this park was commissioned in 2017 and plays a significant role in India's renewable energy transition. With a total installed capacity of 1,000 megawatts (MW), it serves as a substantial source of clean energy. Within the park, there are multiple solar power plants, each with its own set of specifications [16].

One notable solar power plant within the Kurnool Ultra Mega Solar Park is the 1,000 MW solar power plant developed by Andhra Pradesh Solar Power Corporation Private Limited (APSPCL). This plant contributes significantly to the park's overall capacity and generates renewable energy on a massive scale.

The Kurnool Solar Park consists of several sections or blocks, managed by different entities. Prominent manufacturers and maintainers operating within the park include [16]:

SB Energy (Soft Bank Group): As a subsidiary of the Soft Bank Group, SB Energy holds a significant portion of the Kurnool Solar Park's area. They have developed solar power plants with a total capacity of 350 MW.

Greenko Group: Greenko Group, a leading renewable energy company in India, is another prominent player within the Kurnool Solar Park. They own and maintain solar power plants with a capacity of 500 MW.

Azure Power: Azure Power, renowned for its expertise in solar power development and operation, has a presence in the Kurnool Solar Park. They have developed and maintain solar power plants with a capacity of 100 MW.

Adani Green Energy: Adani Green Energy, a part of the Adani Group, has a stake in the Kurnool Solar Park. They own and operate solar power plants with a capacity of 100 MW.

ACME Solar: ACME Solar, a leading solar power developer in India, is also involved in the Kurnool Solar Park. They have developed and maintain solar power plants with a capacity of 100 MW.

The solar panels utilized in the Kurnool Ultra Mega Solar Park consist of poly crystalline silicon cells and possess the specifications shown in Table 1.

Table 1: Specifications of Solar Panels Used in Kurnool Ultra Mega Solar Park

Specification	Value
Nominal power rating	325 Wp
Module efficiency	16.76 %
Open circuit voltage (Voc)	46.22 V
Short circuit current (Isc)	9.32 A
Maximum power voltage (Vmp)	37.02 V
Maximum power current (Imp)	8.80 A
Operating temperature range	-40°C to +85°C

Monitoring and Data Analysis: The plant operations within the Kurnool Ultra Mega Solar Park are monitored through the Azure Power website1. This platform provides real-time information and allows for efficient management of the solar plant's performance. The overall metrics and filters present in the dashboard in the homepage, as shown in 4.1 [16]

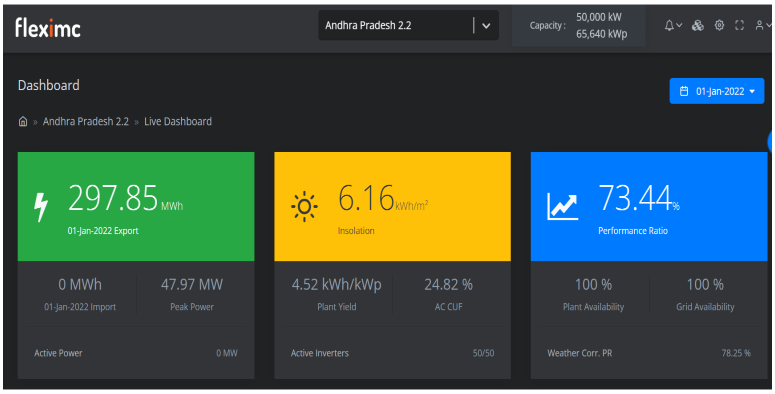

Figure 4.1: Azure Power Portal - Home Page

Specifically, this thesis focuses on meteorological parameters that influence the total plant output and the power generated at different instances. Figure 4.2 [16] shows all important parameters that are selected for ML model training.

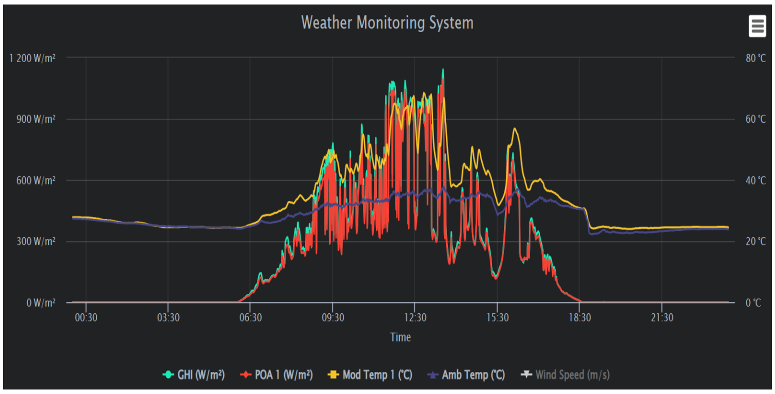

Figure 4.2: Azure Power Portal - Weather Monitoring Dashboard

Similarly, another dashboard in the same portal presents the power generated for a solar irradiance at any time instant. It is illustrated in 4.3 [16].

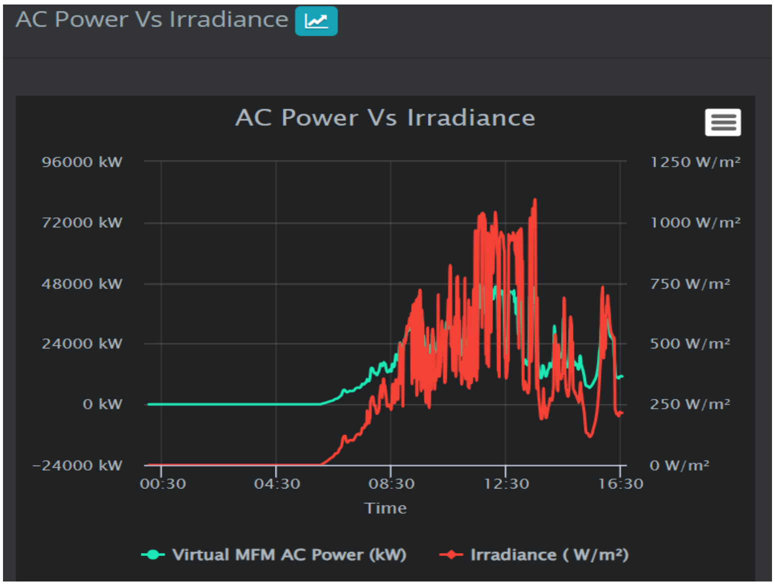

Figure 4.3: Azure Power Portal - Irradiance vs Power

For this case study, the dataset from the Kurnool Solar Plant spans from January 1, 2022, to January 31, 2023. A sample of minute-level data is presented in Table 2.

Table 2: Sample Minute Level Data

Time	Virtual MFM AC Power (kW)	Irradiance (W/m^2)
02/04/2022 08:00	12252.3	248.8
02/04/2022 08:01	12429.7	252
02/04/2022 08:02	12647.1	255.84
02/04/2022 08:03	12808.2	257.92
02/04/2022 08:04	12978.3	262.4
02/04/2022 08:05	13251.3	266.88
02/04/2022 08:06	13482.9	271.04
02/04/2022 08:07	13715.1	275.2
02/04/2022 08:08	13928.7	279.84
02/04/2022 08:09	14125.4	283.52
02/04/2022 08:10	14385.1	288.16

In analyzing the data, the recorded ambient temperatures throughout the year are presented in Figure 4.4.

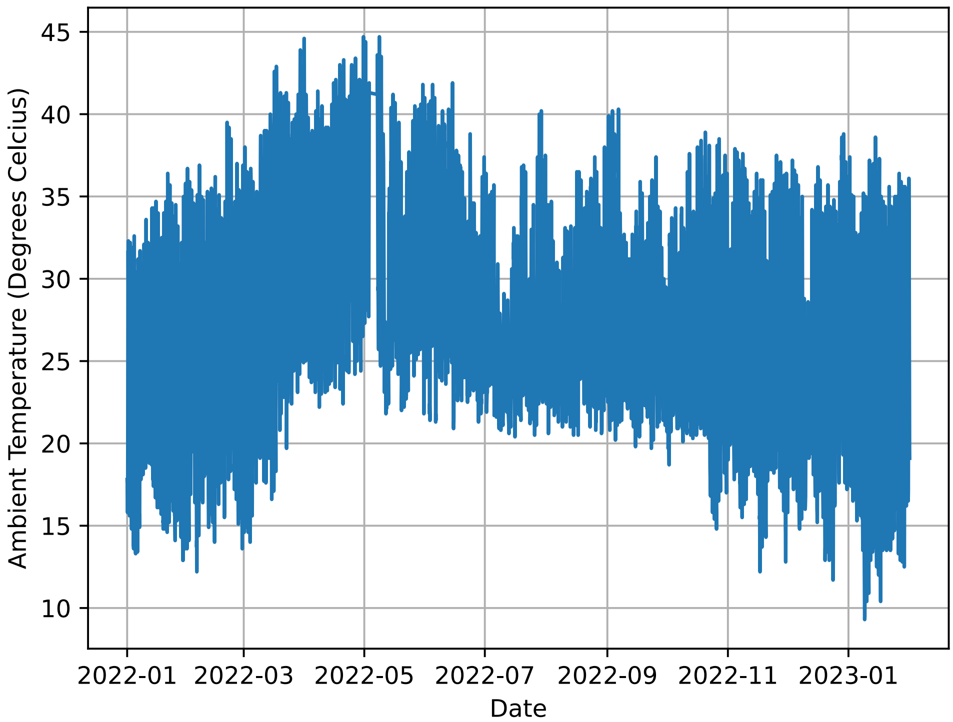

Figure 4.4: Solar Park Temperature from Jan 2022 to Dec 2023

Now, the minute level dataset is converted into hourly data by taking the average of the irradiance in an hour and the average of the power generated in that hour. Table 3 displays a sample of the hourly dataset.

Table 3: Sample Solar Data on January 27, 2023

Time	Virtual MFM AC Power (kW)	Irradiance (W/m^2)	GHI (W/m^2)	POA (W/m^2)	Mod Temp (°C)	Amb Temp (°C)
2023-01-27 05:00:00	0	0.64	0.32	0.64	12.6	12.9
2023-01-27 06:00:00	0	0.64	0.32	0.64	12	13.2
2023-01-27 07:00:00	454.9	21.76	16.32	21.76	12.7	13
2023-01-27 08:00:00	10969.5	270.08	194.72	270.08	18.9	16.1
2023-01-27 09:00:00	27003.6	574.72	437.44	574.72	37.5	26
2023-01-27 10:00:00	37585	788.96	630.72	788.96	47	29.2
2023-01-27 11:00:00	44082.6	960.16	784.32	960.16	52.9	31.4
2023-01-27 12:00:00	47753.6	1062.08	877.6	1062.08	54.6	32.4
2023-01-27 13:00:00	46064	1054.08	869.76	1054.08	56.8	35
2023-01-27 14:00:00	44079.3	991.52	809.92	991.52	54.5	34.6
2023-01-27 15:00:00	35808	808	637.44	808	51.4	35.9
2023-01-27 16:00:00	25621	559.2	416.32	559.2	43.4	34.4
2023-01-27 17:00:00	9854.3	264.8	181.12	264.8	34.8	31.3
2023-01-27 18:00:00	434.4	17.6	12.96	17.6	25	25.4
2023-01-27 19:00:00	0	0.48	0.32	0.48	19.8	21.4

The relationship between temperature and output power is not directly reliant. Instead, it varies based on the level of irradiance present at the given time. The ambient temperature serves as a determining factor for the module temperature, consequently impacting the efficiency of the panel. Therefore, the combined influence of ambient temperature, module temperature, and irradiance assumes a critical role in estimating the overall power generated by the plant. As previously defined, GHI and POA provide measurements of the solar energy received by the panel surface and the inclination angle of the panel, albeit indirectly. Remarkably, through pattern analysis, the tilt angle and orientation do not require explicit specification to the forecasting algorithm, as it can autonomously extract this information.

Based on the analysis conducted, the following five parameters - ambient temperature, module temperature, plane of array (POA), global horizontal irradiance (GHI), and irradiance - are considered as significant features for training the model to predict power output. These parameters exhibit crucial influences on the performance of the solar power plant.

By considering these parameters, the model can capture the environmental conditions and their impact on energy generation accurately. The inclusion of these features enhances the forecasting capability and facilitates optimal energy management within the solar power plant. In the next section, the utilization of machine learning techniques in power and weather forecasting will be discussed, building upon the analysis and insights obtained from the Kurnool Solar Power Plant case study.

4.2 Mytrah Vayu Tungabhadra Private Limited

MVTPL (Mytrah Vayu Tungabhadra Private Limited) is a renowned renewable energy company in India that operates multiple wind farms across the country. One of their notable wind farms is situated near the Tungabhadra river in the Kurnool district of Andhra Pradesh. This wind farm has made significant contributions to the renewable energy sector in India. With a total installed capacity of 500 MW, it harnesses the power of several wind turbines strategically installed along the river. During the installation process, the management shortlisted three types of wind turbines based on their suitability for the location. Factors such as the height above sea level, ground strength to withstand wind tower forces, and regional wind speed characteristics played a role in determining the placement of each turbine. The MVTPL wind farm features a range of wind turbines, including the Suzlon 2.1 MW-S-111, GE 1.7 MW-103, and GE 2.3 MW-11 models. These turbines are classified as doubly fed induction generator horizontal axis wind turbines with variable rotor speed. The turbines have a rated power of 1700 KW and operate at rotation speeds ranging from 10 to 17.14 rpm [17].

The wind farm comprises a total of 97 towers, with an average air density of 1.225 kg/m^3. The annual average wind speed at the site is recorded at 7.5 m/s, and the wind shear coefficient is 0.2. The cut-in wind speed, at which the turbines start operating, is 3 m/s, while the cut-out wind speed, indicating the maximum wind speed at which the turbines can safely operate, is 20 m/s. The rated wind speed, at which the turbines produce their maximum power output, is 9.4 m/s. Additionally, the turbines are designed to withstand a survival wind speed of 52.5 m/s.

The rotor of the wind turbines has a diameter of 103 m and a length of 50.2 m. Each turbine is equipped with three blades, resulting in a swept area of 8332 m^2. The orientation of the rotor is upwind, and the direction of rotation is clockwise. The generators within the wind turbines have a rated power output of 1745 KW and operate as 4-pole, 3-phase systems. The rated voltage is 690 V, and the frequency is 50 Hz. The towers at the MVTPL Wind Farm have a height of 79.7 m, providing a suitable elevation for harnessing wind energy efficiently. The wind farm consists of 30 units of the GE 1.7 MW turbines, 44 units of the GE 2.3 MW turbines, and 23 units of the Suzlon 2.1 MW-S-111 turbines. These specifications and details define the characteristics and capabilities of the wind turbines at the MVTPL Wind Farm, providing valuable insights into their performance and contribution to renewable energy generation. The operating frequency is consistently maintained at 50 Hz. Table 4 provides detailed specifications of the wind turbines categorized by turbine model [17].

The specifications of the wind turbines at the MVTPL wind farm in Kurnool are as follows:

Table 4: Wind Farm Specifications

Turbine Model	GE-1.7	GE-2.3	S-111
Number of Turbines	30	44	23 m/s
Rated Power	1.7 MW	2.3 MW	2.1 MW
Rotor Diameter	100 m	107 m	110 m
Swept Area	7854 m^2	8992 m^2	9500 m^2
Tower Height	80 m	80 m	90 m
Cut-In Wind Speed	3 m/s	3 m/s	3 m/s
Cut-Out Wind Speed	25 m/s	25 m/s	25 m/s

The wind turbines at the MVTPL wind farm are equipped with a variable speed generator and pitch-controlled blades. They utilize a gearbox with a step-up ratio of 1:97 to increase the rotor's rotational speed to the generator's rated speed. The generator output is then transformed to the grid voltage level using a step-up transformer. To monitor and control the operation of the wind turbines, the wind farm utilizes a supervisory control and data acquisition (SCADA) system. The SCADA system collects real-time data on wind speed, power output, and other operating parameters of the turbines. This data is instrumental in optimizing the wind farm's operation and facilitating predictive maintenance. For the case study, the plant parameters and data are obtained from an employee working at the MVTPL wind farm. The dataset used in the analysis covers the period from January 1, 2022, to December 31, 2022. A sample of dataset stored for every 10 minutes is shown in Table 5.

Table 5: Sample of Minute Level Wind Dataset

Date/Time	**LV Active Power (kW)**	**Wind Speed (m/s)**	**Theoretical Power Curve (KWh)**	**Wind Direction (°)**
01 01 2018 00:00	380.04	5.31	416.32	259.99
01 01 2018 00:10	453.76	5.67	519.91	268.64
01 01 2018 00:20	306.37	5.21	390.90	272.56
01 01 2018 00:30	419.64	5.65	516.12	271.25

Other required parameters are pulled from different sources in the same dashboard and integrated into a single dataset using basic Pandas operations in Python. The single dataset formed with all the influencing parameters converted to hourly format is presented in Table 6.

Table 6: Sample Wind Data on December 29, 2022

Date-Time	Wind Speed (kmph)	Relative Humidity (%)	Temperature (°C)	Pressure (kPa)	Air Density (Kg/m3)	Energy (MWh)
2022-12-29 00:00:00	16.70142	59.01	20.1	100.78	1.1972	79511.78
2022-12-29 01:00:00	14.40055	60.3	20.3	100.83	1.19698	50959.92
2022-12-29 02:00:00	15.09242	60.15	20	100.93	1.19939	58781.63
2022-12-29 03:00:00	20.1125	60.72	19.5	101.01	1.20239	139459.5
2022-12-29 04:00:00	18.1817	60.59	19.3	101	1.2031	103088.39
2022-12-29 05:00:00	16.18654	60.14	16.5	101.01	1.21485	73449.58
2022-12-29 06:00:00	15.51076	60.9	17	101.04	1.21311	64536.29
2022-12-29 07:00:00	17.07149	59.75	14	101.07	1.22615	86968.6
2022-12-29 08:00:00	17.48983	59.63	12.5	101.15	1.23356	94085.27
2022-12-29 09:00:00	17.90817	59.09	12.4	101.23	1.23497	101114.77
2022-12-29 10:00:00	17.95644	60.89	12.4	101.27	1.23546	101975.06
2022-12-29 11:00:00	18.31042	60.61	12	101.24	1.23683	108245.41
2022-12-29 12:00:00	19.56544	60.84	18.9	101.22	1.20737	128918.08

In conclusion, several parameters, including wind speed, air density, pressure, ambient temperature, module temperature, plane of array (POA), global horizontal irradiance (GHI), and irradiance, are considered as significant features for training the model to predict power output. These parameters have a direct impact on the performance of the wind turbines and the overall power generation. By incorporating these features, the model can effectively capture the environmental conditions and their influence on energy production.

In the next section, the utilization of machine learning techniques for power and weather forecasting will be discussed, building upon the analysis and insights obtained from the MVTPL Wind Plant case study.

4.3 Overview of Implementation with AI

The implementation of Artificial Intelligence (AI) is crucial in understanding weather patterns and analyzing plant output patterns. However, to enable AI to perform accurate predictions, it requires access to historical data and a labeled dataset that clearly specifies the corresponding weather conditions and the power generated. Therefore, the initial step involves creating a structured dataset that serves as a foundation for developing statistical or mathematical models.

To accomplish this, a real-time dataset is obtained with the assistance of an employee working at Azure Power in India. This dataset comprises solar plant information from January 2022 to January 2023. Additionally, the influential factors mentioned in sections 4.3 and 3.3 are considered for analyzing the behavior of the plant. The power generated by the plant is recorded every 10 minutes using measuring devices for key parameters such as Ambient Temperature, Module Temperature, Plane of Array (POA), Global Horizontal Irradiance (GHI), and Solar Irradiance. It's important to note that the sensors placed at the plant station may not capture the irradiance accurately for panels located far away from the station. To address this, the data is modified and aggregated to provide hourly information for our case study. This enables a comprehensive understanding of the plant's performance and facilitates intraday and day-ahead forecasts.

The preprocessed dataset is then fed into a machine learning model to extract insights. Performance evaluation plays a pivotal role in assessing the accuracy of the model. Initially, the objective is to predict weather conditions using suitable machine learning or deep learning algorithms. The trained model is subsequently utilized to forecast the weather for the last 30 days, 15 days, and 7 days at an hourly rate, with comparisons made against existing values. Several metrics, including accuracy, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R2Score, are evaluated to optimize the model's performance through data preprocessing and hyperparameter tuning. Furthermore, a similar model with certain modifications is employed to predict the energy produced by the hybrid system and its integration with the grid.

In conclusion, the implementation of AI in the solar and wind energy hybrid system is a key aspect of this study. By utilizing historical data and a carefully curated dataset, machine learning models are developed and trained to forecast weather conditions and predict power output. The evaluation of model performance using various metrics ensures accuracy and enables optimization through data preprocessing and parameter tuning. The successful integration of AI in the hybrid system allows for enhanced prediction capabilities, optimized energy management, and improved efficiency in renewable energy generation.

Overall, the combination of solar and wind energy, coupled with the implementation of AI, holds immense potential for a sustainable and reliable energy future. Continued research and development in this field will further enhance the performance and effectiveness of hybrid systems, accelerating the transition to clean and renewable energy sources.

5 Power Forecasting with Four Layer ML Model

This chapter presents a comprehensive methodology for power forecasting using a machine learning (ML) model with four layers. The methodology aims to predict the power output of the hybrid solar and wind energy system discussed in the previous chapters. The chapter begins by explaining the architecture of the ML model through a detailed block diagram, highlighting the ten individual blocks that make up the model. Each block is then thoroughly explained, providing insights into its specific role in the forecasting process.

Furthermore, the chapter outlines the overall execution cycle of the Python program used for power forecasting. It starts with data preprocessing steps, ensuring the data is prepared and ready for model training. The procedure for training the ML models, including Support Vector Machine (SVM), Extreme Gradient Boosting (XGB), Random Forest Regression (RFR), and the deep learning model Long Short-Term Memory (LSTM), is described. The performance of each model is assessed, and the results are presented through comparison line graphs that illustrate the relationship between the original and predicted values.

Additionally, the chapter compares the power output predictions obtained from the ML model with historical predictions made by existing forecast techniques. This evaluation process aids in the selection of the best performing model for power forecasting. The chapter concludes by highlighting the chosen model and its readiness for forecasting. It also connects with the next chapter by explaining how users can utilize the provided Python code to obtain future power predictions for the entire hybrid system. The chapter further explores how plant operators can gain insights into the supply-demand curve and emphasizes the model's potential applicability to other power plants worldwide with minimal modifications. Finally, the drawbacks of the methodology are discussed, providing a comprehensive understanding of the limitations and areas for future improvement.

5.1 Architecture of Four Layer ML Model

The ML model consists of four different sections and each contributes to estimating overall power output of the entire hybrid system. The block diagram is presented in Figure 5.1.

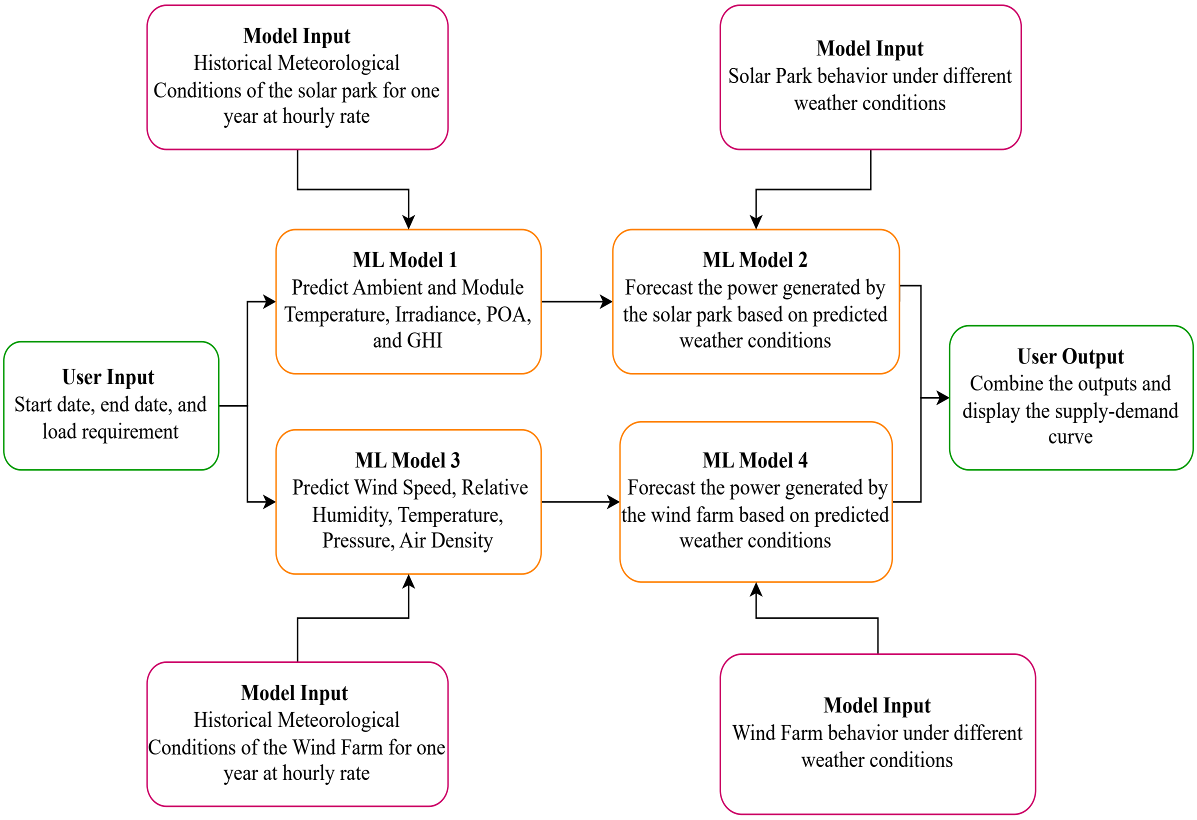

Figure 5.1: AI Model Architecture with Four Layers

The top layer in the block diagram is for Solar power forecast and the bottom layer is for Wind power forecast. Firstly, the user enters the start date, end date, and the load demanded in the selected days. This input is fed into model 1 and model 2, each pre-trained with the weather conditions at the plant location using best algorithm. The time series forecast results produced by these two blocks are fed to the other two blocks (model 3 and model 4) which are already pre trained with the plant behavior for any weather conditions. Finally, the power forecasting by both these models are combined and a presented to the user at the end.

5.2 Overview of Python Programming and Modeling

Python is one of the most popular programming languages used for machine learning. Its simplicity, readability, and extensive library support make it a preferred choice for ML practitioners. Python provides a rich ecosystem of libraries such as NumPy, Pandas, and Scikit-learn, which offer powerful tools for data manipulation, numerical computations, and machine learning algorithms. Google Colab is an online platform provided by Google that enables researchers and developers to write and execute Python code in a Jupyter Notebook environment. First, it provides free access to powerful hardware resources, including GPUs and TPUs, which are crucial for accelerating the training and inference processes of deep learning models. Second, Colab offers seamless integration with other Google services such as Google Drive, making it easy to access and manage datasets. Overall, Python and the Google Colab environment form a powerful combination for implementing and experimenting with machine learning algorithms.

A complete Python program developed to forecast the energy produced by the hybrid system is presented in Figure 5.2. The block diagram starts with taking inputs from the user about start and end date of the forecast along with the load requirements. The right branch of the block diagram is to find the best machine learning model to the case study based on the accuracy of each model. The left and center branch of the block diagram makes of the best ML model and predicts the weather conditions and their respective power generated. At the end, the AI model is used to predict future weather conditions and power generations.

The hyperlink to Python script is given in Appendix 1 and the code contains different sections to perform multiple functions and operations. In addition, the comments attached beside each line explain the function of that specific line or a code snippet. The user is free to make changes in the input section by entering the load demanded as per the requirements. The program goes through a series of predefined machine learning models to perform meteorological predictions and consequent energy produced by the hybrid system. Finally, it declares the future predictions and concludes whether the forecast meets the required demand or not.

Furthermore, the program used to develop these predefined machine learning models is integrated into another Google Colab file (See Appendix 1.1). This file is divided into three sub-sections which compare the results of different machine learning algorithms for solar and wind energy forecasts. However, the graphical representations of these results are discussed in next chapter in detail. Moreover, this chapter specifically compares all the results obtained and concludes the most suitable ML model.

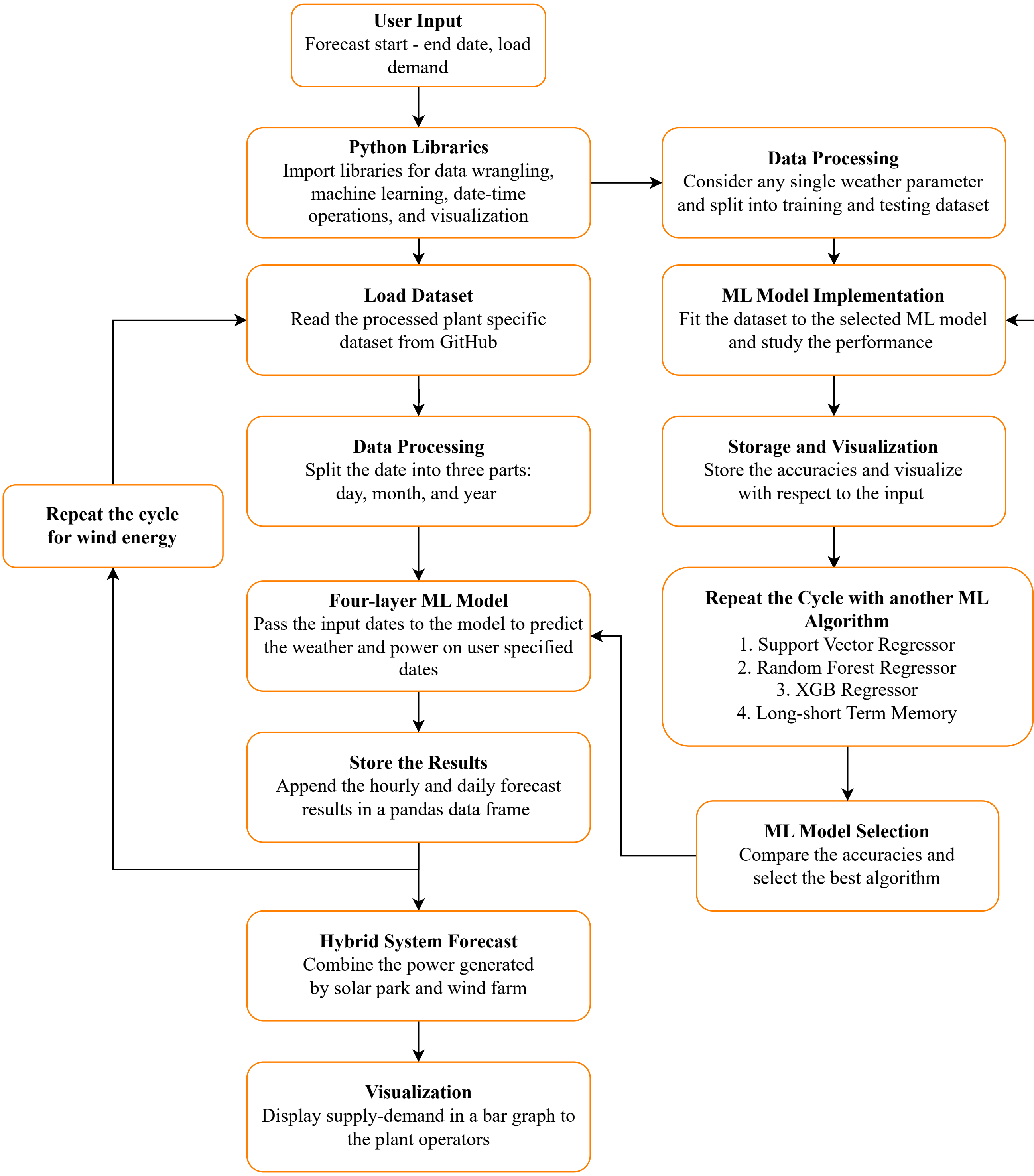

Figure 5.2: Block Diagram of Python Script Execution Cycle

5.3 Data Preprocessing Techniques

Data preprocessing is a crucial step in creating a high-quality dataset for power forecasting in the hybrid solar and wind energy system. The dataset is prepared by converting minute-level data to hourly data and then aggregating it to daily data. The following steps are undertaken in the data preprocessing phase:

1. Data Collection: Historical data on temperature, irradiance, and power output from the solar plant and wind farm is collected at an hourly rate. This data is obtained from sensors placed in the solar panels, weather stations near the solar plant, and wind turbines.

2. Data Cleaning: The collected data is carefully examined to identify and handle missing or corrupted data. Missing data can be imputed using techniques like mean, median, or interpolation. Removing duplicates and addressing inconsistencies is essential to ensure data integrity.

3. Data Integration: The temperature, weather information, irradiance, and power output data are integrated into a unified dataset. The timestamps are matched to create a synchronized dataset for analysis and modeling.

4. Data Normalization: To facilitate accurate model training, the dataset is normalized by scaling the temperature, irradiance, and power output values to a range between 0 and 1. Normalization ensures that all features contribute equally to the training process.

5. Feature Selection: Relevant features, including temperature, irradiance, and power output, are selected from the dataset for training the machine learning models. These features are known to have a significant impact on power generation.

6. Feature Engineering: Additional features are derived from existing features to enhance the predictive capability of the models. For instance, time of day or day of the week can be derived from timestamps to provide the models with more contextual information.

7. Data Splitting: The dataset is divided into training, validation, and testing sets. The training set is used to train the machine learning models, the validation set helps fine-tune the model hyperparameters, and the testing set is utilized to evaluate the model's performance.

By following these data preprocessing techniques, a clean and well-organized dataset is created, ready for training machine learning models to forecast power output accurately.

5.4 Four Layer ML Model Training

Machine learning models play a crucial role in power forecasting for the hybrid solar and wind energy system. This section explains the training process and performance evaluation of four different models: Support Vector Machines (SVM), Random Forest Regression (RFR), Support Vector Regression (SVR), and Long Short-Term Memory (LSTM) networks.

5.4.1 Support Vector Machine (SVM)

Support Vector Machines (SVM): SVM is a powerful algorithm for classification and regression tasks. In weather forecasting, SVM can be used to predict weather patterns based on historical data, such as temperature, humidity, wind speed, and atmospheric pressure. Similarly, SVM can be applied to energy forecasting, predicting energy demand based on factors like historical energy consumption, weather conditions, time of day, and season. SVM captures complex patterns and provides accurate predictions, aiding in efficient energy planning and distribution [18].

To train the SVM model, the data is preprocessed, ensuring a tabular dataset with numerical features and a continuous target variable. Categorical variables are encoded using techniques like one-hot encoding or label encoding, while missing values are handled through imputation or removal. Feature scaling is important for SVM to ensure that all features are on a similar scale. The model is then trained on the preprocessed data, finding the optimal hyperplane that separates different classes or predicts continuous values. During training, SVM captures the underlying patterns in the data, enabling accurate predictions. The trained SVM model can make predictions on new unseen data by mapping it to the feature space using the learned hyperplane. The performance of the SVM model is evaluated using various metrics such as accuracy, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R2Score. These metrics provide insights into the accuracy and reliability of the predictions. Additionally, a comparison line graph between the original and predicted values helps visualize the performance of the SVM model.

Support Vector Regression (SVR): SVR is an extension of SVM for regression tasks. It can be used for energy forecasting, predicting energy demand based on factors like historical energy consumption, weather conditions, and other relevant features. SVR operates by mapping the input data into a higher-dimensional feature space and finding a hyperplane that best fits the data while maximizing the margin [18]. The training and evaluation process of SVR is similar to that of SVM. The data is preprocessed, and the model is trained on the preprocessed data using appropriate kernels and hyperparameter values. The performance of the SVR model is evaluated using metrics such as MAE, MAPE, and R2Score, and the predicted values are compared to the actual values using a line graph.

The trained SVR model is used to predict the temperatures and wind speeds for the last 7 days. The results are provided in the form of line graph in Figure 5.3 and 5.4.

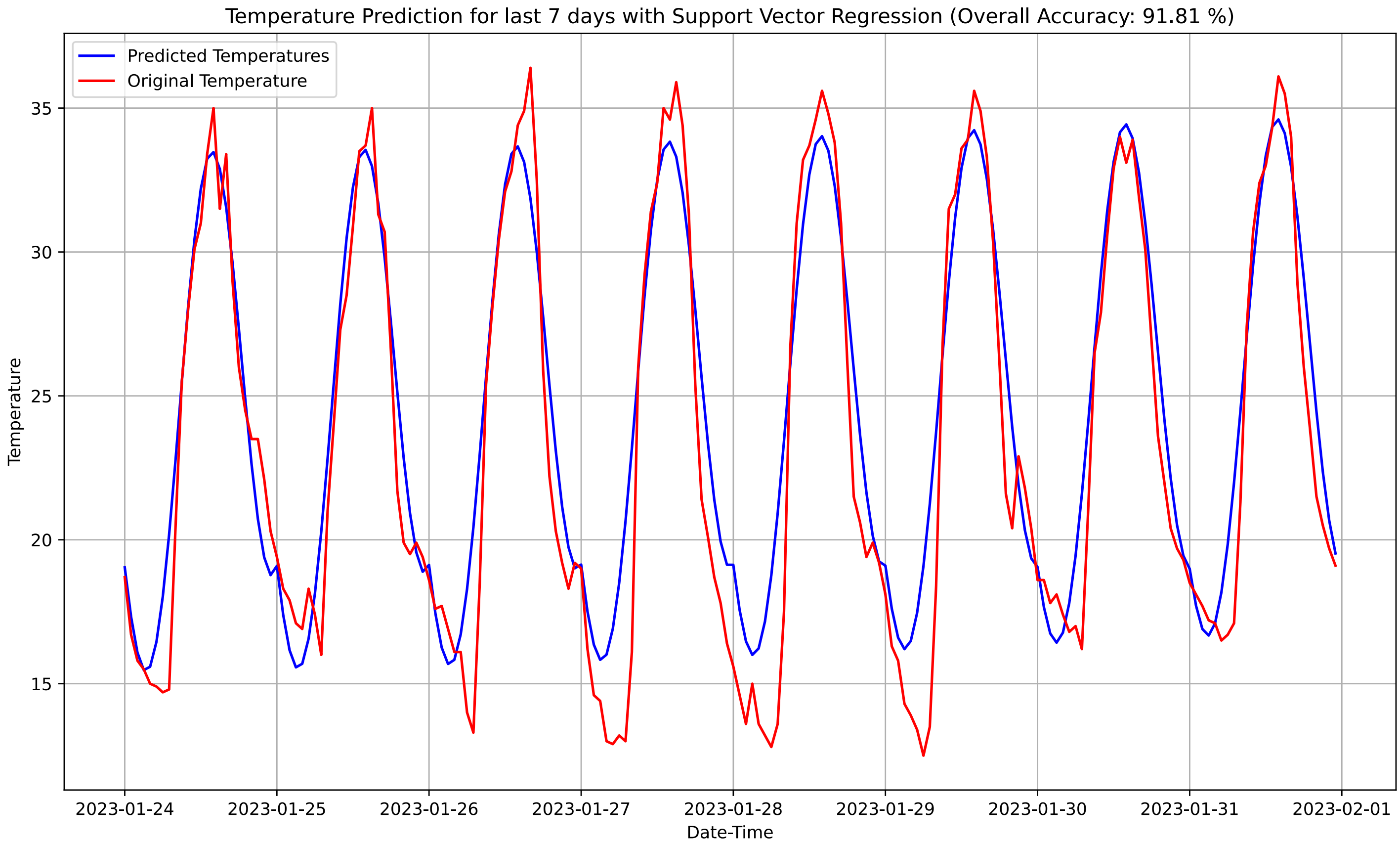

Figure 5.3: SVR Temperature Predictions - Last 7 Days

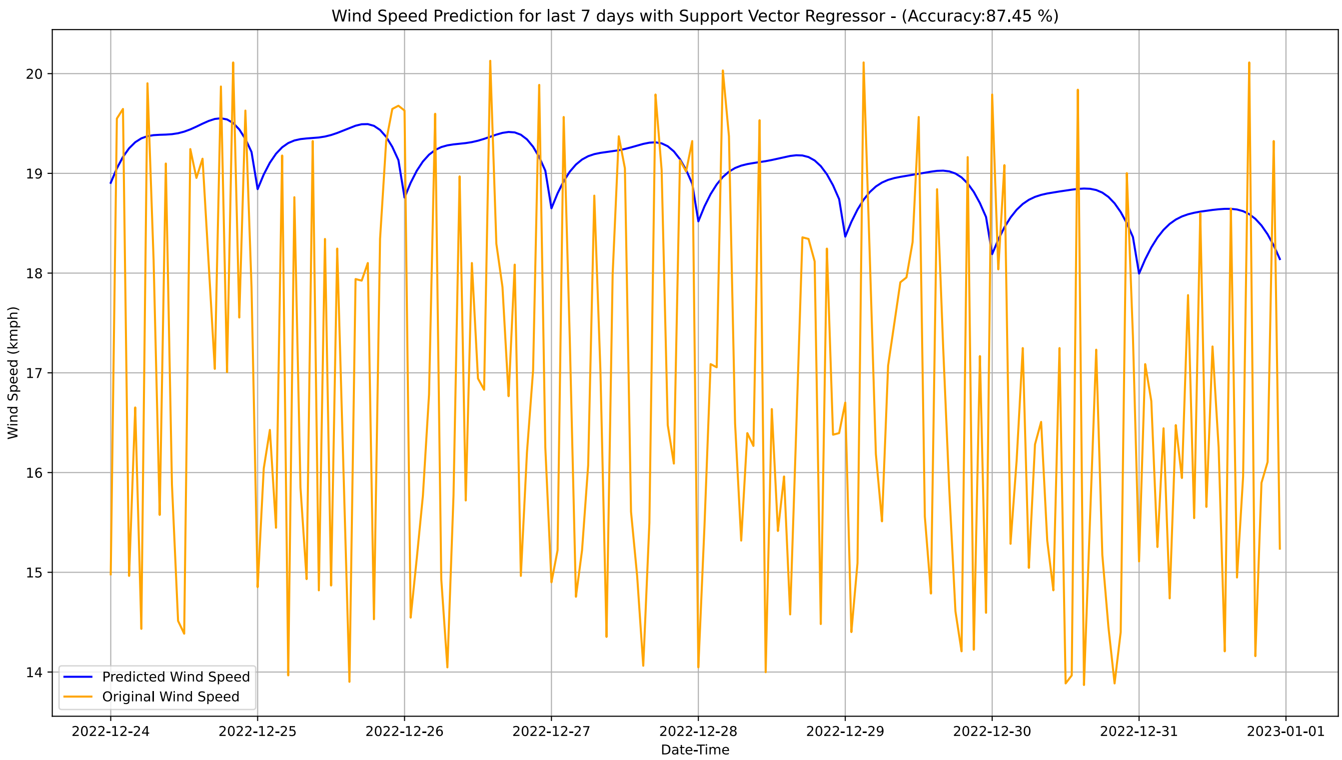

Figure 5.4: SVR Wind Predictions - Last 7 Days

Although the accuracy of temperature prediction is more than 91.81 %, but the wind speeds are predicted with 87.45 % accuracy. From the figure, it is evident that the relation between predicted and original wind speed is not accurate. However, there is a possibility of increasing the accuracy by using another algorithm. Further, three more techniques are explored with similar procedure in the next sections and they are assessed by their accuracy.

5.4.2 Random Forest Regression (RFR)

Random Forest Regression (RFR): Random Forest is an ensemble learning algorithm that combines multiple decision trees to make predictions. In weather forecasting, Random Forest can handle large amounts of meteorological data and make predictions on parameters like rainfall, cloud cover, or temperature. In energy forecasting, Random Forest can predict energy demand or generation based on historical data, weather conditions, economic factors, and other relevant features. To train the RFR model, the data is prepared by ensuring a tabular dataset with numerical features and a continuous target variable. Similar to SVM, categorical variables can be encoded, missing values can be handled, and feature scaling is not necessary for decision tree-based algorithms [19]. The RFR model is built by constructing an ensemble of decision trees using a technique called bagging. Each decision tree is trained on a bootstrap sample of the training data, and a subset of features is randomly selected at each node to determine the best split. The number of decision trees, known as the n-estimators hyperparameter, can be specified by the user. The trained decision trees in the ensemble are then used to make predictions. For regression tasks, the final prediction is often the average or weighted average of the predictions from all the trees. This aggregation process helps reduce the impact of individual noisy or overfitting trees, leading to a more stable and accurate prediction. The performance of the RFR model is evaluated using metrics such as MAE, MAPE, and R2Score. A comparison line graph between the original and predicted values illustrates the performance of the RFR model. The results showcase the ability of the RFR model to capture non-linear relationships and handle complex interactions between different variables.

The trained RFR model is used to predict the temperatures and wind speeds for the last 7 days. The results are provided in the form of line graph in Figure 5.5 and 5.6.

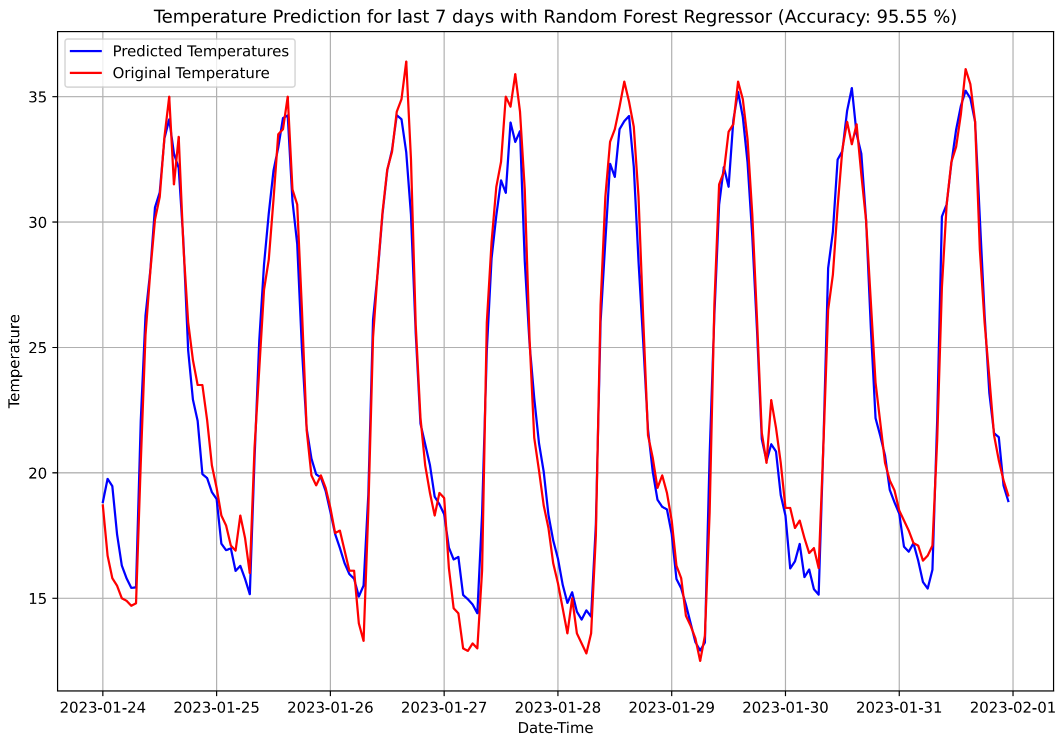

Figure 5.5: RFR Temperature Predictions - Last 7 Days

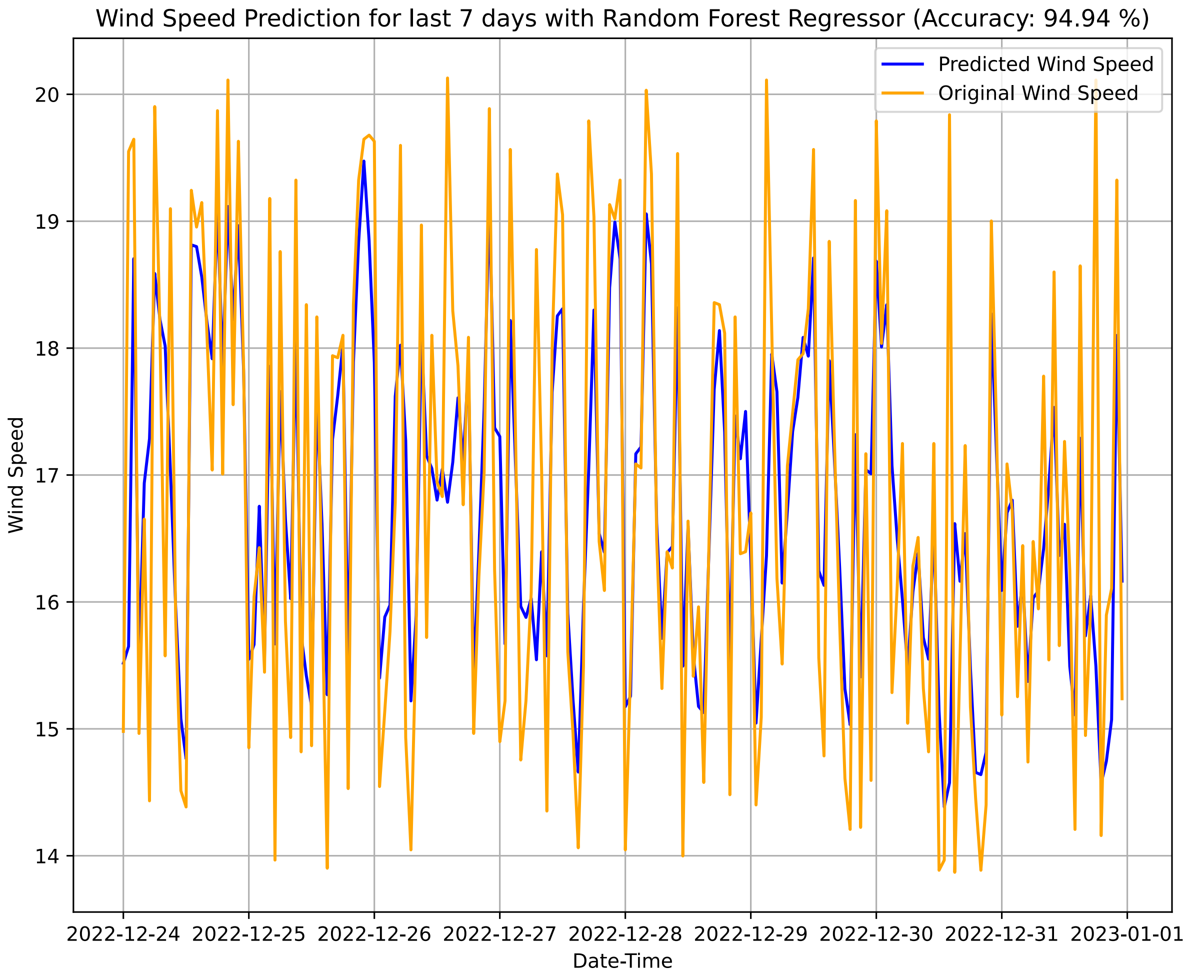

Figure 5.6: RFR Wind Predictions - Last 7 Days

5.4.3 Long-Short Term Memory

LSTM Networks: LSTM networks are a type of recurrent neural network (RNN) architecture designed to capture long-term dependencies in time series data. In weather forecasting, LSTM networks can handle the complexities of meteorological data and predict future weather conditions accurately. In energy forecasting, LSTM networks can predict energy demand or generation at high granularity, incorporating factors such as historical energy consumption, weather conditions, holidays, and special events [20]. The operation of an LSTM network involves memory cells that store and update information over time, with input, forget, and output gates controlling the flow of information. These gates enable LSTMs to retain relevant information over long sequences and effectively capture long-term dependencies. To train an LSTM network, the data is prepared in a suitable format, ensuring sequential input and a continuous target variable. The network architecture is defined, including the number of LSTM layers, the number of memory cells, and the output layer. The model is then trained on the prepared data, leveraging its ability to capture short-term fluctuations and long-term trends. The performance of the LSTM model is evaluated using metrics such as MAE, MAPE, and R2Score. A comparison line graph between the original and predicted values showcases the accuracy and effectiveness of the LSTM model in capturing complex temporal patterns. For our case study, the LSTM model is trained and evaluated using historical temperature data. The trained LSTM model is used to predict the temperatures and wind speeds for the last 7 days. The results are provided in the form of line graph in Figure 5.7 and 5.8.

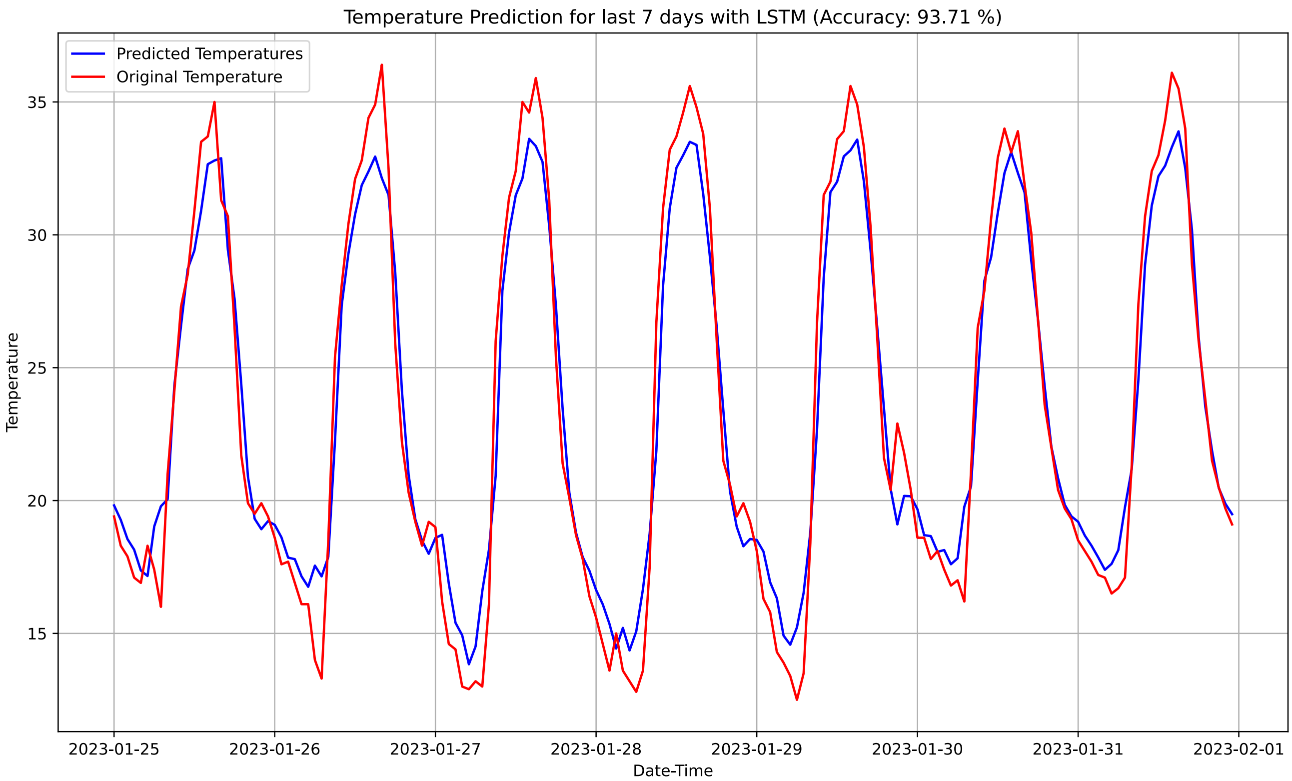

Figure 5.7: LSTM Temperature Predictions - Last 7 Days

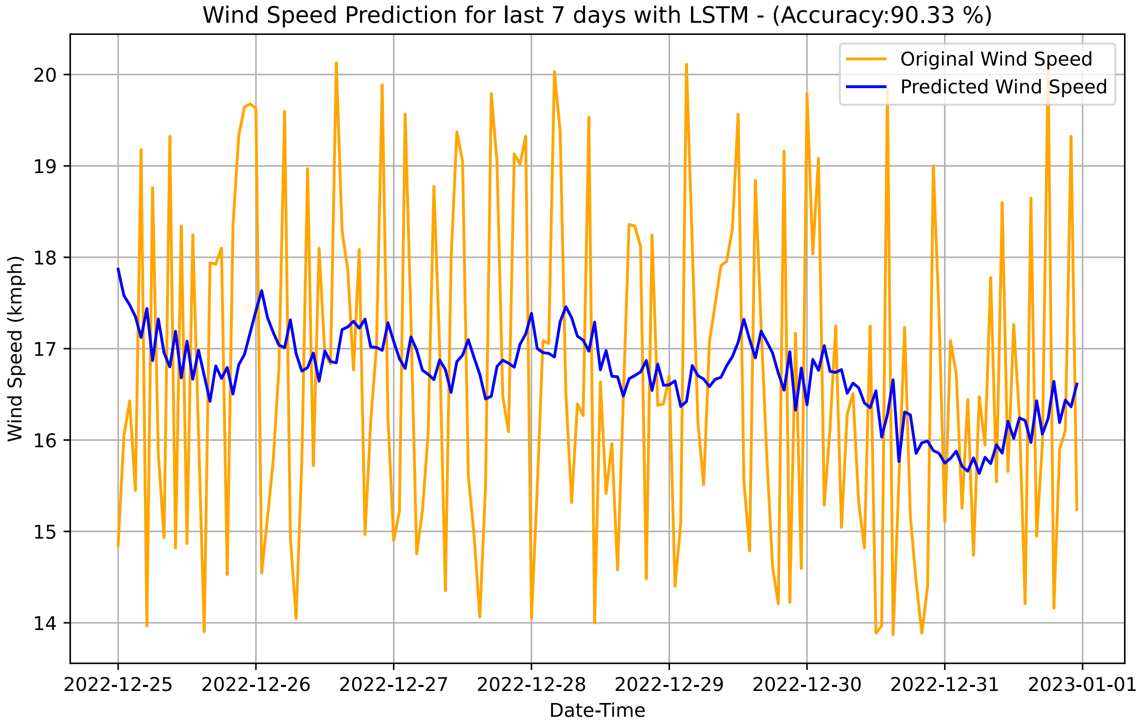

Figure 5.8: LSTM Wind Predictions - Last 7 Days

5.4.4 Extreme Gradient Boosting

XGB (Extreme Gradient Boosting) is a powerful machine learning algorithm widely used for forecasting tasks, including time series forecasting. XGB incorporates boosting, a technique that combines multiple weak prediction models (decision trees) into a strong predictive model. In the context of our case study on weather and energy forecasting, XGB can be applied to predict variables such as temperature, wind speed, and energy demand [21]. One key advantage of XGB is its ability to handle complex relationships and capture non-linear patterns in the data. It achieves this by recursively building decision trees, where each subsequent tree corrects the errors made by the previous trees. This iterative process allows the model to gradually improve its performance by focusing on challenging instances in the data. XGB also optimizes an objective function that quantifies the difference between predicted values and actual target values, allowing it to find the best combination of decision trees for accurate forecasts. In our case study, XGB Regression is utilized to forecast temperature and wind speed based on historical data. The model is trained using a tabular dataset that includes relevant features such as historical weather conditions, time of day, and season. Categorical variables are one-hot encoded, missing values are handled through appropriate techniques, and feature scaling is applied to ensure all features are on a similar scale. The training process involves building an ensemble of regression trees, where each tree is added sequentially to correct the errors of the previous trees. XGB optimizes an objective function, which includes a loss term and a regularization term, to find the best parameters for the model. Regularization techniques are applied to prevent overfitting and enhance the model's ability to generalize to unseen data. Once the XGB Regression model is trained, it can make predictions on new data by combining the predictions of individual trees. The trained XGB model is used to predict the temperatures and wind speeds for the last 7 days. The results are provided in the form of line graph in Figure 5.9 and 5.10. These predictions demonstrate the accuracy and effectiveness of the XGB Regression model in capturing the underlying patterns in the weather data.

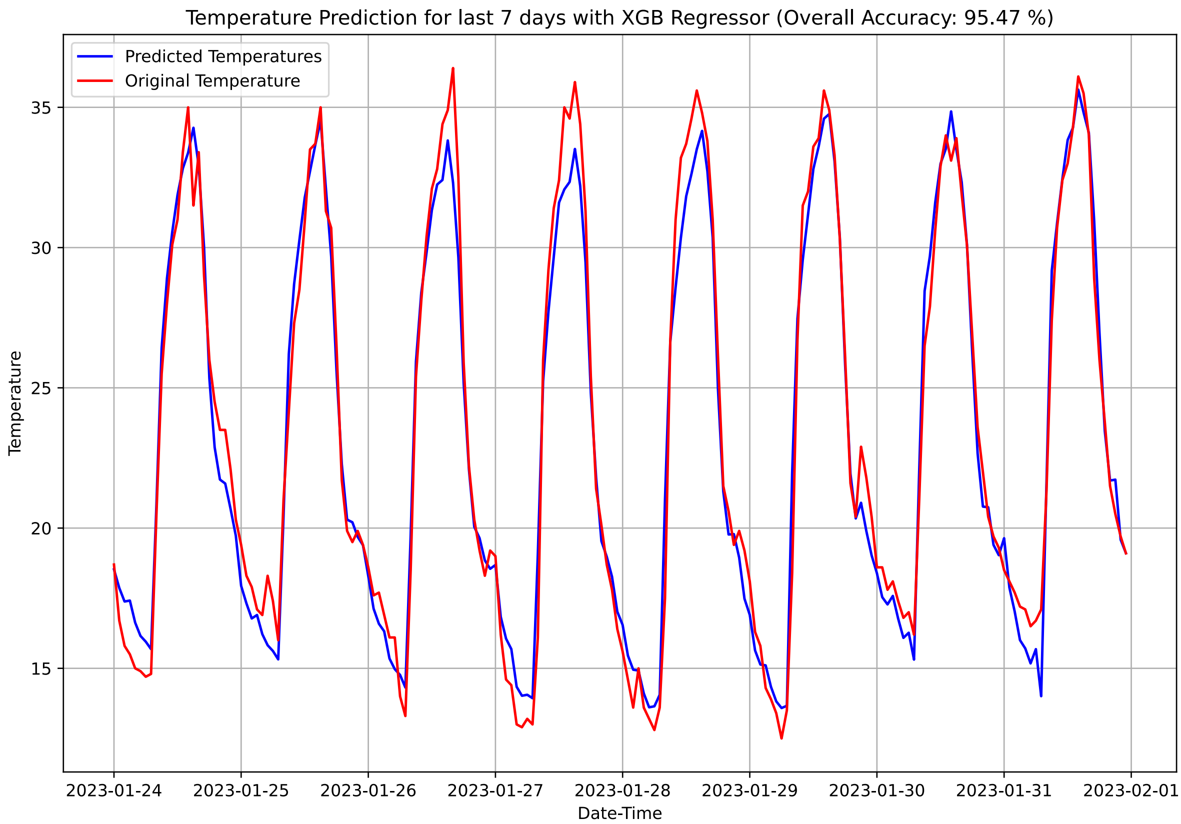

Figure 5.9: XGB Temperature Predictions - Last 7 Days

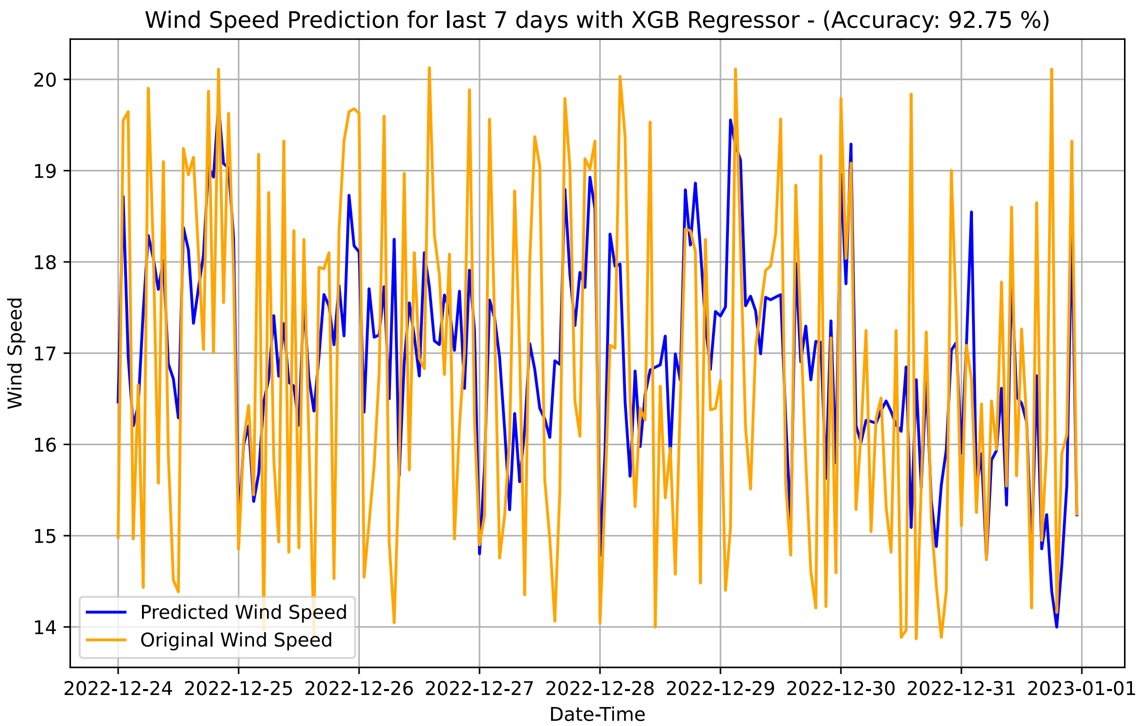

Figure 5.10: XGB Wind Predictions - Last 7 Days

In conclusion, XGB Regression is a valuable machine learning algorithm in weather and energy forecasting. It has been successfully applied in our case study to predict temperature and wind speed. The model's ability to capture complex relationships and its performance in forecasting make it a reliable choice for similar forecasting tasks in the future.

In summary, this section has explored three different machine learning algorithms and one deep learning algorithm. It provided an overview of the operation of each algorithm and their results on predicting the temperatures. However, Random forest regression has outperformed all other techniques in all parameters. Therefore, the AI model used in this project uses RFR to forecast the energy of the hybrid system. As this section focused mainly on meteorological predictions, next chapter provides the forecast results obtained and also guides on using the AI model to predict for user defined dates.

5.5 Performance Evaluation and Selection

This section provides a comparison of four machine learning models used for power forecasting, namely Support Vector Machines (SVM), Random Forest Regression (RFR), Long Short-Term Memory (LSTM) Networks, and XGB Regression. The evaluation focuses on the accuracy of these models in predicting power output, as well as their performance in forecasting temperature and wind speed for the last 7 days. The accuracies obtained for temperature and wind speed predictions are compared with existing temperature values and predictions from current forecast models. Moreover, the overall power output predictions of each model are also assessed and compared.

Upon careful evaluation and comparison, it becomes evident that the Random Forest Regression (RFR) model outperforms the other models in power forecasting. This model demonstrates higher accuracy and better generalization to unseen data, making it the preferred choice for forecasting power in the entire hybrid system.

The four-layer machine learning model, comprising SVM, RFR, LSTM, and XGB Regression, is now equipped with the Random Forest Regression model placed in all four layers. This configuration ensures accurate and reliable power forecasts by leveraging the strengths of multiple machine learning algorithms.

To summarize, the Random Forest Regression model emerges as the top performer in terms of accuracies for power forecasting, temperature prediction, and wind speed forecasting. With the selection of the Random Forest Regression model, the four-layer machine learning model is fully prepared to deliver precise power forecasts for the entire hybrid system.

In the next chapter, the practical implementation of the power forecasting model using Python will be explored. This chapter provides a detailed explanation of how to effectively utilize the Python script. Users will gain insights into the supply-demand curve and make informed decisions by inputting desired dates and obtaining accurate power forecasts. The Python script, designed with user-friendliness in mind, offers a valuable tool for power forecasting in various renewable energy settings. The developed power forecasting model not only contributes to the efficient operation of the hybrid system but also holds the potential for adaptation in similar plants worldwide with minimal modifications. By harnessing the capabilities of machine learning and artificial intelligence, the utilization of renewable energy resources can be enhanced, leading us towards a more sustainable and efficient energy future.

5.6 Historical Meteorological and Power Prediction

The four layer AI model is fed with the plant data set and it has formed a complex mathematical equations and decision trees in the back end. Now, the model is completely ready to make predictions. To test the overall performance of the model, random date range is selected and the model is used to predict the environmental parameters as well as the overall power generated. The results are joined with the original dataset that is fed to the system and the accuracy is calculated on different time instants. The accuracy can be assessed on hourly rate and daily rate. A sample of the observations is presented below.

5.6.1 Historical Solar Power Prediction

The model is asked to predict the energy produced by the Solar plant on "30.01.2023". The predicted temperature, irradiance, anticipated power along with the original power generated at that specific time and accuracy is presented in the Table 7.

Table 7: Solar Energy Predictions on 30 Jan, 2023 (Overall Accuracy: 98.23 %)

Date-Time	**Temperature (°C)**	**Irradiance (W/m^2)**	**Predicted Power (kW)**	**Original Power (kW)**	**Accuracy (%)**
2023-01-30 0:00:00	18.1059	0.48	0	0	100
2023-01-30 1:00:00	16.140429	0.64	0	0	100
2023-01-30 2:00:00	16.435549	0.64	0	0	100
2023-01-30 3:00:00	16.452076	0.64	0	0	100
2023-01-30 4:00:00	15.813996	0.64	0	0	100
2023-01-30 5:00:00	16.121477	0.48	0	0	100
2023-01-30 6:00:00	15.452744	0.64	0	0	100
2023-01-30 7:00:00	15.283107	19.57264	462.5142	473.5	97.679873
2023-01-30 8:00:00	21.17745	243.2288	8883.7316	8800.5	99.05424
2023-01-30 9:00:00	28.245851	535.11403	27215.8583	27098.7	99.567661
2023-01-30 10:00:00	29.66453	780.4144	37176.9911	37126.3	99.863463
2023-01-30 11:00:00	31.711647	951.98336	45544.3184	48513.5	93.87968
2023-01-30 12:00:00	32.822152	1044.11867	46986.1732	46782.1	99.563779
2023-01-30 13:00:00	34.432053	1028.04166	46407.3203	45982.3	99.075687
2023-01-30 14:00:00	34.43563	942.821912	43738.3884	42430.5	96.917575
2023-01-30 15:00:00	33.494699	703.15824	32316.562	34101.1	94.76692
2023-01-30 16:00:00	32.222487	509.89914	23072.2522	22595.4	97.889605
2023-01-30 17:00:00	30.103716	234.04368	9582.1382	8100.7	81.71222
2023-01-30 18:00:00	25.813858	14.90128	435.2622	445.5	97.701953
2023-01-30 19:00:00	22.850825	0.48	0	0	100
2023-01-30 20:00:00	21.503319	0.48	0	0	100
2023-01-30 21:00:00	20.716862	0.48	0	0	100
2023-01-30 22:00:00	19.162355	0.48	0	0	100
2023-01-30 23:00:00	18.804385	0.48	0	0	100

Similarly, the model is used to predict for a week starting from 24 Jan, 2023 to 31 Dec, 2023. The model successfully predicted the total power generated by the solar park with more than 90 % of accuracy on a day and with an overall accuracy of 95.64 % in that week. The results are presented in Table 8.

Table 8: Solar Energy Predictions from 24 Dec to 31 Dec, 2023 (Overall Accuracy: 93.29 %)

Date	Generated Power (kW)	Predicted Power (kW)	Accuracy (%)
2023-01-24	13449.26667	12965.30939	84.636596
2023-01-25	13779.57083	13214.79618	91.095212
2023-01-26	14215.75833	12836.0722	94.252859
2023-01-27	13737.925	13169.39718	96.837698
2023-01-28	13603.30833	13314.06384	96.369713
2023-01-29	13232.91667	12980.6794	95.69499
2023-01-30	13435.42083	13409.22959	98.236361
2023-01-31	11513.475	13315.09084	89.22192

5.6.2 Historical Wind Farm Prediction

The model is asked to predict the energy produced by the Wind Farm on "08.12.2022". The predicted wind speeds, Relative Humidity, Temperature, Pressure, Air density, anticipated energy along with the original energy produced at that specific time and the accuracy is presented in the Table 9

Table 9: Wind Energy Predictions on 08 Dec, 2022 (Accuracy: 90.61%)

Date-Time	Wind speed (kmph)	Relative Humidity (%)	Temperature (°C)	Pressure (kPa)	Air Density (Kg/m^3)	Predicted Energy (MWh)	Original Energy (MWh)	Accuracy (%)
2022-12-08 00:00:00	18.415681	60.71403	28.4671	101.66151	1.174676	41266.42709	44275.29	93.204194
2022-12-08 01:00:00	18.260364	63.31901	27.3646	101.07229	1.172299	39870.86885	35942.20	89.069481
2022-12-08 02:00:00	18.251193	63.70071	27.2244	101.10476	1.173217	39723.30680	35552.08	88.267278
2022-12-08 03:00:00	18.874487	63.17428	29.8243	101.61509	1.168751	44155.59095	51760.66	85.307241
2022-12-08 04:00:00	19.505714	65.43549	31.0011	101.37462	1.161294	47981.30423	51227.37	93.663415
2022-12-08 05:00:00	19.269175	64.70341	30.3395	101.67076	1.167525	46954.67088	43269.62	91.483515
2022-12-08 06:00:00	19.156239	61.03423	23.9398	101.30118	1.188768	47108.54885	51474.25	91.518670
2022-12-08 07:00:00	17.076945	60.20510	21.7487	101.34700	1.197670	33431.01973	36439.87	91.742972

Similarly, the model is used to predict for a week starting from 01 Dec, 2022 to 07 Dec, 2022. The model successfully predicted the total energy produced by the wind farm with more than 90 % of accuracy on a day and with an overall accuracy of 95.64 % in that week. The results are presented in Table 10.

Table 10: Wind Energy Predictions in Typical Week (Overall Accuracy: 95.64 %)

Date	Energy (MWh)	Original Energy (MWh)	Accuracy (%)
2022-12-01	855217.042	930775.7	91.88218408
2022-12-02	842553.4368	860231.27	97.94499063
2022-12-03	797262.459	827561.75	96.33872748
2022-12-04	763126.1382	843445.59	90.47722192
2022-12-05	842776.2096	872786.8	96.56152105
2022-12-06	687410.9306	676308.38	98.35835974
2022-12-07	850151.6117	867915.81	97.95323485

Therefore, the developed model is performing with around 95 % accuracy. Hence, the four layer ML model can be used to forecast the total power generated by the hybrid system. These results are used to guide the plant operators for power scheduling and to maintain the grid balance.

6 Forecasting Results and Limitations

In this chapter, the execution cycle of the Python script, which utilizes a four-layer machine learning model for power forecasting, will be explored. The script operates in a step-by-step manner to generate accurate power predictions based on user-defined dates and load demand. The following section delves into the process:

1. Data Input: The user provides input dates and load demand for which power forecasting is desired. The script prompts the user to enter the start date and end date, along with the corresponding load demand values.

2. Data Validation: The script performs data validation to ensure the input dates are in the correct format and fall within the available dataset range. It also validates the load demand values to ensure they are consistent and reasonable.

3. Model Loading: The script loads the trained Random Forest Regression model from the stored model file. As an ensemble model, Random Forest Regression requires more storage space compared to other models due to the multiple trees it forms at the back end.

4. Data Preprocessing: The input dates and load demand are preprocessed to align with the format required by the model. This involves transforming the dates into a suitable format and scaling the load demand values if necessary.

5. Power Forecasting: The preprocessed data is fed into the four-layer machine learning model. Each layer of the model sequentially processes the input data, passing it through the respective ML model (SVM, RFR, LSTM, XGB), and obtaining the power forecast for each layer.

6. Result Aggregation: The power forecasts from all layers are aggregated to obtain the final power forecast for the specified dates. This ensures that the model leverages the strengths of each ML algorithm and produces a comprehensive and accurate prediction.

7. Graphical Representation: The script generates a bar graph that visually represents the supply-demand curve based on the forecasted power and the provided load demand. This graph provides a clear visualization of the power availability and the load demand pattern over the specified dates.

The hyperlink to the Python script used in this project is presented in Appendix 1. The blocks replicates the block diagram and has four sections for each model and follows execution cycle described in Chapter 5. The user can open the link, click on *"Energy Forecast for the Hybrid System"* section and enter the required forecast dates and load demand on those dates in the following format available.

User_Start_date = datetime(2023, 7, 15, 0)
User_End_date = datetime(2023, 7, 21, 23)
Pload = [90, 80, 84, 75, 110, 97, 115]

After entering the required dates in *"year-month-day-hour)"* format, click on *"Run after"* in the toolbar to view the supply-demand curve on those days at the bottom. Furthermore, the exploration script of all ML models is presented in Appendix 2.
A sample output is presented in the bar graph when the user enters the start date as "15.07.2023" and end date as "21.07.2023", along with the load demand values [95, 90, 105, 110, 120, 125, 98]. The bar graph in 6.1 showcases the predicted power output and load demand, enabling the user to analyze the balance between supply and demand.

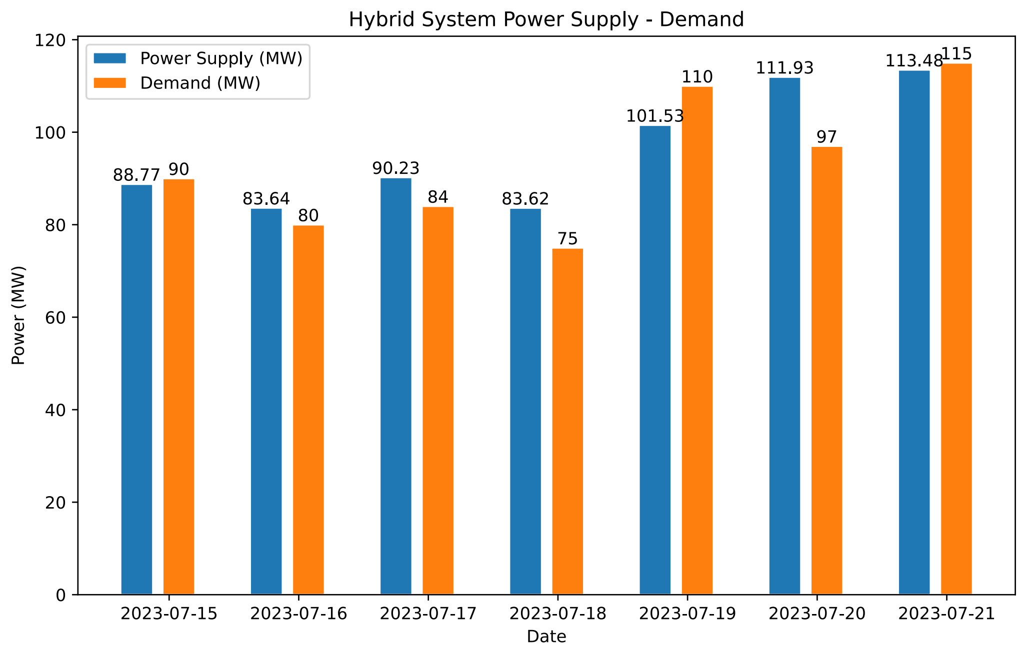

Figure 6.1: Supply-load Curve for User Defined Dates

6.1 Limitations

In this subsection, the drawbacks associated with the implemented power forecasting system using the four-layer machine learning model will be discussed. Despite the model's impressive performance, it is essential to consider certain limitations. Let's examine these drawbacks:

- Storage Space Requirement: The implemented model, which utilizes the Random Forest Regression algorithm, necessitates a substantial amount of storage space. This is attributable to the ensemble nature of the algorithm, which involves creating multiple decision trees in the backend. It is crucial to ensure the availability of sufficient storage resources to accommodate the model.
- Dependence on Historical Data: The four-layer machine learning model heavily relies on historical data to achieve accurate power forecasting. Optimal performance is typically attained when the model is provided with at least one year's worth of historical information. Insufficient historical data may impede the model's capacity to capture seasonality and other long-term patterns, resulting in diminished forecasting accuracy.

Despite these drawbacks, the implemented power forecasting system using the four-layer machine learning model holds significant potential for the renewable energy sector. It enables precise predictions and facilitates efficient resource management, thereby contributing to the advancement of renewable energy integration and sustainability.

In the next chapter, the thesis will be concluded by summarizing the key findings, discussing the implications of the developed power forecasting model, and providing recommendations for future research and enhancements.

7 Conclusions and Future Scope of Work

This thesis explored the implementation of a power forecasting system for a hybrid renewable energy system. It began by analyzing the components of the hybrid system, including the Ultra Mega Solar Plant and MVTPL Wind Farm. Then delved into the importance of accurate power forecasting and its impact on efficient resource management.

Later developed a four-layer machine learning model consisting of SVM, RFR, LSTM, and XGB algorithms. Each algorithm was trained and evaluated for its ability to forecast power output based on historical weather data and other influential factors. The results revealed that the Random Forest Regression (RFR) algorithm outperformed the other models, showcasing better accuracy and robustness in capturing complex nonlinear relationships.

The system allows users to interact by providing input dates and load demand, obtaining accurate power forecasts, and visualizing the supply-demand balance. The developed model has the potential to revolutionize the renewable energy sector, enabling efficient resource management, better decision-making, and a smoother integration of renewable energy into the grid.

The execution cycle of the Python script was explained, detailing the step-by-step functioning of the four-layer model. Users were provided with the capability to interact with the system by providing input dates and load demand, obtaining accurate power forecasts, and visualizing the supply-demand balance through graphical representations.

7.1 Implications of the Developed Model

The developed power forecasting model has significant implications for the renewable energy sector. Accurate power forecasts enable effective resource planning and allocation, leading to enhanced operational efficiency and reduced costs. Plant operators can make informed decisions regarding power generation, grid integration, and energy trading based on the forecasted output. Additionally, the model facilitates a deeper understanding of the relationship between weather patterns, environmental factors, and power generation. It allows for the identification of trends, seasonality, and other patterns that can inform future planning and optimization strategies.

7.2 Recommendations for Future Research and Improvements

While the developed power forecasting model demonstrates promising results, there are opportunities for further research and improvements. Some recommendations for future work include:

1. Incorporating Advanced Feature Engineering: Explore additional features and variables that could enhance the accuracy of power forecasts. This could involve integrating real-time data from sensors, satellite imagery, or other external sources.
2. Expanding the Hybrid System: Extend the scope of the model to encompass other renewable energy sources, such as biomass or hydroelectric power. Incorporating a wider range of energy generation technologies would provide a more comprehensive forecast and enable better resource management.
3. Integration of Hybrid Forecasting Approaches: Investigate the potential benefits of combining machine learning algorithms with physical or statistical models. Hybrid forecasting approaches that leverage the strengths of different methods could further improve the accuracy and reliability of power forecasts.
4. Scalability and Efficiency: Explore methods to optimize the storage space required for the ensemble models. Investigate techniques to reduce the computational complexity and improve the efficiency of the power forecasting system.

References

[1] E. Energy, “Types of renewable energy,” 2023. [Online; accessed 11-April-2023].

[2] D. o. A. E. C. o. E. T. P. S. U. Mohamed Amer Chaaban, Instructor, “Irradiance and pv performance optimization,” 2023. [Online; accessed 11-April-2023].

[3] M. S. Khan, V. Hegde, and G. Shankar, “Effect of temperature on performance of solar panels- analysis,” pp. 109–113, 2017.

[4] J. F. Jønler, F. Brunø Lottrup, B. Berg, D. Zhang, and K. Chen, “Probabilistic forecasts of global horizontal irradiance for solar systems,” *IEEE Sensors Letters*, vol. 7, no. 1, pp. 1–4, 2023.

[5] P. Babal, M. Korevaar, S. Franken, J. Mes, T. Bergmans, and K. Wilson, “Uncertainties in irradiance measurements of sensors to poarear of bifacial solar panels,” pp. 0959–0963, 2020.

[6] A. M. H. Suad Hassan Danook, Khamis Joir Jassim, “The impact of humidity on performance of wind turbine, case studies in thermal engineering,” vol. 14, 2019. [Online; accessed 11-April-2023].

[7] T. Tahseen, “12 types of losses in solar pv system,” 2022. [Online; accessed 9-April-2023].

[8] N. Nguyen, S. Almasabi, and J. Mitra, “Impact of correlation between wind speed and turbine availability on wind farm reliability,” *IEEE Transactions on Industry Applications*, vol. 55, no. 3, pp. 2392–2400, 2019.

[9] Y. Ding, H. Zhou, Z. Tan, Y. Chen, and J. Ding, “The influence of wind direction on short-term wind power prediction: A case study in north china,” in *IEEE PES Innovative Smart Grid Technologies*, pp. 1–5, 2012.

[10] Q. Wang, P.-z. Zhou, L. Zhang, and Z. He, “Effect of air density on the fatigue life of wind turbine blades,” in *2021 International Conference on Intelligent Transportation, Big Data Smart City (ICITBS)*, pp. 278–281, 2021.

[11] W. E. T. Facts, “Wind farm energy loss factors,” 2020. [Online; accessed 1-April-2023].

[12] M. Z. Malik, K. Zehra, I. Ali, Ubedullah, M. Ismail, A. Hussain, V. Kumar, M. Abid, and M. H. Baloch, “Solar-wind hybrid energy generation system,” in *2020 IEEE 23rd International Multitopic Conference (INMIC)*, pp. 1–6, 2020.

[13] S. Technologies, “Advantages and disadvantages of hybrid solar energy systems,”

2020. [Online; accessed 1-May-2023].

[14] A. M. Alam, Nahid-Al-Masood, I. Asif Razee, and M. Zunaed, "Solar pv power forecasting using traditional methods and machine learning techniques," in *2021 IEEE Kansas Power and Energy Conference (KPEC)*, pp. 1–5, 2021.

[15] M. Joubert, A. Dalton, and B. Bekker, "A categorisation wind power forecasting methodologies, highlighting emerging short-term forecasting methods," in *2021 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, pp. 1–6, 2021.

[16] H. Reddy, "Azure power," 2023. [Online; accessed 1-April-2023].

[17] V. S. P. Ltd., "Environmental and social impact assessment," 2016. [Online; accessed 10-April-2023].

[18] R. Nguyen, Y. Yang, A. Tohmeh, and H.-G. Yeh, "Predicting pv power generation using svm regression," in *2021 IEEE Green Energy and Smart Systems Conference (IGESSC)*, pp. 1–5, 2021.

[19] R. Zhang, L. Zhu, J. Ye, S. Kuai, X. Wang, T. Wang, and S. Li, "Pv prediction using hierarchical bilstm-rfr model considering meteorological factors," in *2022 5th International Conference on Electronics and Electrical Engineering Technology (EEET)*, pp. 120–125, 2022.

[20] E. N. Kajabad, S. V. Ivanov, and I. Khodnenko, "Lstm algorithm for forecasting events in changing electric consumption," in *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 565–571, 2020.

[21] C. Sheng and H. Yu, "An optimized prediction algorithm based on xgboost," in *2022 International Conference on Networking and Network Applications (NaNA)*, pp. 1–6, 2022.

A Appendix 1

Link to the four layer ML model for hybrid system forecast:
`https://colab.research.google.com/drive/1Pbm_KiQiBQrp2NUppNoMpk6_8mwjGeVT#scrollTo=RsgO0pDFk722`

Steps to Execute:

1. Open the link, click on *"Runtime"* and *"Runall"* in the toolbar.
2. Scroll to the bottom, open *"Energy Forecast for the Hybrid System"*
3. Enter the dates and load demand on those dates in the below format

   User_Start_date = datetime(2023, 7, 15, 0)
   User_End_date = datetime(2023, 7, 21, 23)
   Pload = [90, 80, 84, 75, 110, 97, 115]

   Then, click on *"Runtime"* and *"Run after"*

B Appendix 2

Link to the Python script used to compare four different machine learning algorithms:

`https://colab.research.google.com/drive/1LLVfJ-4C1c2AoQriE6Q9WqYTpm28e5kz#scrollTo=mfIPkEMrL4Gq`