REPUBLIC OF TURKEY
ISTANBUL GELISIM UNIVERSITY
INSTITUTE OF GRADUATE STUDIES

Department of Electrical and Electronics Engineering

MACHINE LEARNING TECHNIQUES FOR SOLAR POWER OUTPUT PREDICTING

Master Thesis

Samaa Yahya Dawood AL_MAYYAHI

Supervisor

Asst. Prof. Dr. Mahmoud HK. ALDABABSA

Istanbul – 2023

THESIS INTRODUCTION FORM

Name and Surname	**:** Samaa Yahya Dawood AL_MAYYAHI
Language of the Thesis	**:** English
Name of the Thesis	**:** Machine Learning Techniques For Solar Power Output Predicting
Institute	**:** Istanbul Gelisim University Institute of Graduate Studies
Department	**:** Electrical and Electronics Engineering
Thesis Type	**:** Master
Date of the Thesis	**:** 20/07/2023
Page Number	**:** 69
Thesis Supervisors	**:** Prof. Dr. Mahmoud H. K. ALDABABSA
Index Terms	**:** Solar power forecasting, Long Short-Term Memory (LSTM), Nadam optimizer, Time series analysis, Meteorological parameters
Turkish Anstract	**:** Güneş Enerjisi Çıkışını Tahmin Etmek İçin Makine Öğrenimi Teknikleri
Distribution List	**:** 1. To the Institute of Graduate Studies of Istanbul Gelisim University 2. To the National Thesis Center of YÖK (Higher Education Council)

Signature

Samaa Yahya Dawood AL_MAYYAHI

REPUBLIC OF TURKEY
ISTANBUL GELISIM UNIVERSITY
INSTITUTE OF GRADUATE STUDIES

Department of Electrical and Electronics Engineering

MACHINE LEARNING TECHNIQUES FOR SOLAR POWER OUTPUT PREDICTING

Master Thesis

Samaa Yahya Dawood AL_MAYYAHI

Supervisor
Asst. Prof. Dr. Mahmoud HK. ALDABABSA

Istanbul – 2023

DECLARATION

I hereby declare that in the preparation of this thesis, scientific ethical rules have been followed, the works of other persons have been referenced in accordance with the scientific norms if used, there is no falsification in the used data, any part of the thesis has not been submitted to this university or any other university as another thesis.

Samaa Yahya Dawood AL_MAYYAHI

…/…/2023

TO ISTANBUL GELISIM UNIVERSITY

THE DIRECTORATE OF GRADUATE EDUCATION INSTITUTE

The thesis study of Samaa Yahya Dawood ALI AL-MAYYAHI titled as Machine Learning Techniques For Solar Power Output Predicting has been accepted as MASTER in the department of Electrical-Electronic Engineering by out jury.

Director ______________________________

Asst. Prof. Dr. Mahmoud H. K. ALDABABSA

(Supervisor)

Member ______________________________

Asst. Prof. Dr. Yusuf Gürcan SAHIN

Member ______________________________

Asst. Prof. Dr. Mohammed SALEM

APPROVAL

I approve that the signatures above signatures belong to the aforementioned faculty members.

... / ... / 20..

Prof. Dr. Izzet GUMUS

Director of the Institute

SUMMARY

Accurate solar power forecasting is necessary for solar power facilities to operate dependably and effectively. This thesis suggests a novel approach for making long-term forecasts regarding solar power output using Long Short-Term Memory (LSTM) neural networks and the Nadam optimizer. Despite the challenges brought on by the sun's erratic beams, the objective is to provide accurate projections for the system's design and operation.The research's initial phase involves using a variety of unique methodologies to compare and contrast the findings from numerous LSTM models with those from more traditional time series models like ARIMA and SARIMA. The suggested LSTM model using the Nadam optimizer generates more accurate predictions when compared to traditional methods. To increase the system's accuracy and dependability, the impact of climatic factors on solar power forecasting is being researched.The proposed approach results in a number of significant improvements. By taking into account the sensitivity of SPV output power to a variety of environmental conditions, it first presents a novel viewpoint on SPV power forecasting. By comparing the method to other widely used SPV power forecasting methodologies, it also confirms the technique's effectiveness. The recommended method facilitates in the prediction of mitigating factors like solar irradiance and SPV module efficiency in addition to increasing forecast accuracy.The limitations of the suggested technique are also highlighted in the research, including the need for suitable training data and the difficulties in managing the LSTM's forget gate's memory. Future research may look at different neural network topologies and integrate more input parameters to further boost prediction accuracy.Finally, this thesis presents a novel approach for long-term solar power forecasting that combines LSTM with the Nadam optimizer. The findings have ramifications for solar power system optimization, design, and operation, and they also help with the creation of solar power forecasting algorithms. Accurate estimates of solar power output enable improved system architecture, increased dependability, and increased economic viability.

Key Words: Solar power forecasting, Long Short-Term Memory (LSTM), Nadam optimizer, Time series analysis, Meteorological parameters

ÖZET

Güneş enerjisi santrallerinin güvenilir ve verimli bir şekilde çalışması için, doğru güneş enerjisi tahmini gereklidir. Uzun Kısa Süreli Bellek (LSTM) sinir ağlarını ve Nadam iyileştiriciyi kullanan bu tez, güneş enerjisi çıkışı hakkında uzun vadeli tahminler yapmak için yeni bir yöntem önermektedir. Amaç, düzensiz güneş ışınlarının neden olduğu zorluklara rağmen, sistemin tasarımı ve işleyişi için doğru tahminler vermektir. Araştırmanın ilk adımı, birçok LSTM modelinden elde edilen sonuçları, bir dizi farklı teknik kullanarak ARIMA ve SARIMA gibi daha geleneksel zaman serisi modellerinden elde edilen sonuçlarla karşılaştırmak ve karşılaştırmaktır. Geleneksel yaklaşımlarla karşılaştırıldığında, Nadam optimize edici kullanılarak önerilen LSTM modeli daha güvenilir tahminler üretir. Ek olarak, sistemin hassasiyetini ve güvenilirliğini artırmak için iklimsel unsurların güneş enerjisi tahmini üzerindeki etkisi incelenir. Önerilen yöntemle birkaç önemli ilerleme kaydedilmiştir. İlk adım olarak, SPV çıkış gücünün çeşitli iklimsel faktörlere duyarlılığını dikkate alarak SPV güç tahminine yeni bir bakış açısı sunar. İkincisi, tekniğin başarısını diğer popüler SPV güç tahmin stratejileriyle karşılaştırarak doğrular. Önerilen teknik yalnızca tahmin doğruluğunu iyileştirmekle kalmaz, aynı zamanda güneş ışınımı ve SPV modülü verimliliği gibi hafifletici faktörlerin tahminine de yardımcı olur. Makale ayrıca önerilen stratejinin, uygun eğitim verilerinin gerekliliği ve LSTM'nin unutma kapısının hafızasını kontrol etmenin zorluğu gibi uyarılarının altını çiziyor. Tahmin doğruluğunu daha da artırmak için, gelecekteki çalışmalar diğer sinir ağı topolojilerini incelemeyi ve daha fazla girdi parametresi eklemeyi içerebilir. Son olarak, bu tez, LSTM'yi Nadam optimizer ile birleştirerek uzun vadede güneş enerjisi tahmini için orijinal bir yöntemi açıklamaktadır. Sonuçlar, güneş enerjisi sistemlerinin tasarımı, işletimi ve optimizasyonu için çıkarımlara sahiptir ve güneş enerjisi tahmin algoritmalarının geliştirilmesine katkıda bulunur. Daha iyi sistem tasarımı, daha yüksek güvenilirlik ve daha ekonomik fizibilite, güneş enerjisi çıkışının kesin tahminleriyle mümkün kılınmıştır.

Anahtar kelimeler: Güneş enerjisi tahmini, Uzun Kısa Süreli Bellek (LSTM), Nadam optimize edici, Zaman serisi analizi, Meteorolojik parametreler.

TABLE OF CONTENTS

CHAPTER THREE
METHODOLOGY

CHAPTER FOUR
RESULTS AND ANALYSIS

LIMITATIONS AND FUTURE WORK

ABBREDIVATIONS

SPV	:	Solar Photovoltaic
LSTM	:	Long Short-Term Memory
ARIMA	:	Auto Regressive Integrated Moving Average
SARIMA	:	Seasonal AutoRegressive Integrated Moving Average
RMSE	:	Root Mean Square Error
MSE	:	Mean Square Error
NN	:	Neural Network
SGD	:	Stochastic Gradient Descent
RMSprop	:	Root Mean Square Propagation
Adam	:	Adaptive Moment Estimation
Adamax	:	Adaptive Moment Estimation with Infinity Norm
Adagrad	:	Adaptive Gradient Algorithm
Adadelta	:	Adaptive Delta
Ftrl	:	Follow-the-Regularized-Leader
DNN	:	Deep Neural Network
GRU	:	Gated Recurrent Unit
MAE	:	Mean Absolute Error
MAPE	:	Mean Absolute Percentage Error
LSTM	:	Long Short-Term Memory
RNN	:	Recurrent Neural Network

LIST OF TABLES

LIST OF GRAPHICS

LIST OF FIGURES

INTRODUCTION

Moving away from energy sources that use fossil fuels and toward ones that are more sustainable and better for the earth is a challenge the world has never faced before. This change is being driven by the pressing need to lessen the effects of climate change and make sure that future generations will have a sustainable future. Specifically, this change is being pushed by how quickly things need to change. Solar photovoltaic (PV) systems are quickly becoming one of the most exciting innovations in the area of renewable energy. But because these systems only work sometimes and in different ways, adding them to the energy grid poses a number of important problems.

Recent study has looked into how artificial intelligence (AI) and machine learning (ML) could be used to improve how renewable energy systems are run, controlled, monitored, maintained, and how well they work. This study was done to answer the questions that were brought up. ML has made a promise to predict how much power PV systems will make, which is important for adding them to the energy grid. In this area, it's important to make correct predictions. With the help of correct forecasts, grid operators can improve reliability, cut costs, and make the most of renewable energy (Dehghani, et al., 2018) (Deng , Peng;, Zhang; , & Qian, 2018) (Mosavi, et al., 2019).

The goal of this thesis study is to find out if machine learning and, in particular, long short-term memory (LSTM) neural networks can be used to predict how much power photovoltaic (PV) systems will produce. The long short-term memory (LSTM) design is a type of neural network that has done well in time series forecasting. The goal of this project is to come up with a new way to predict weather that, when applied to a time horizon of 24 hours, can make reliable and accurate predictions of power output (Anwar, El Moursi, & Xiao, 2017).

As input data, past measurements of power output and sequences of meteorological data from before the prediction horizon will be used to see how well the proposed method works. In addition to this, a different set of data from an oracle weather forecaster will also be taken into account. The goal of this thesis research is to find the best way to predict the power output of PV systems by judging the performance of an LSTM model that has been trained on different kinds of input data. In particular, the different kinds of data input will be examined.

In the end, the goal of this thesis project is to make an addition to the development of more accurate and reliable solar PV power forecasting systems. These ways of doing things can help make it easier to add green energy to the power grid.

CHAPTER ONE

GENERAL OVERVİEW

1.1. Background

Electricity is an essential aspect of modern life, and the demand for energy is continually growing. However, the use of conventional fossil fuel-based energy sources has resulted in significant environmental degradation, such as greenhouse gas emissions, which contributes to global warming. Renewable energy sources, such as solar photovoltaic (PV) systems, offer a sustainable solution to this challenge. Solar PV systems are an abundant and clean energy source that can contribute to reducing our dependence on conventional energy sources (Neacşa, Panait, Mureșan, Voica, & Manta, 2022). Accurate forecasting of solar PV power output is crucial for the efficient management and utilization of renewable energy resources. This thesis project aims to investigate the potential of machine learning techniques, specifically long short-term memory (LSTM) neural networks, for reliably predicting the power output of solar PV systems. By developing a novel forecasting approach using LSTM, this study seeks to contribute to the advancement of renewable energy systems and support the transition towards a more sustainable energy future.

Solar energy is a form of energy that is clean, renewable, and good for the environment, and it is simple to incorporate into existing power infrastructures. An ideal solar photovoltaic (PV) power projection method is very necessary if one wishes to achieve efficient grid functioning, effective energy management, and cost-effective scheduling. In their time on the market, prominent prediction approaches such as autoregressive integrated moving average (ARIMA), numerical weather prediction (NWP), artificial neural network (ANN), and hybrid methods have demonstrated only moderate levels of success. However, they can only be used for short-term projections, which may be sufficient for more archaic freestanding or tiny PV systems, but long-term forecasts are necessary for the operation of more contemporary grid-integrated PV systems. In light of this, a method that is both significantly improved and trustworthy is urgently required as the structure of the renewable power network continues to get more complicated. According to the findings of a comprehensive literature analysis that was carried out on the ways of forecasting, the majority of the strategies that are now in use continue to place their emphasis on outmoded approaches to solar photovoltaic (SPV) power projection. These approaches do not take into account the influence of the most important meteorological elements, which have a significant impact on the accuracy of the forecasts and result in inefficient monitoring, maintenance, and regulation of the power generated by renewable energy sources. The accuracy of long-term solar power forecasting has been the subject of a number of research studies, including those that made use of NN, ARIMA/SARIMA, NWP, LSTM, and hybrid models (Basurto, et al., 2019) , (Das, et al., 2018) (Khalid & Javaid, 2020) (Paliwal, Patidar, & Nema, 2020) (Santhosh,, Venkaiah, & Kumar, 2020) (Seyedmahmoudian, et al., 2018) (Sharadga, Hajimirza, & Balog, 2020) (Sharma, Sharma, Irwin, & Shenoy, 2011) (VanDeventer, et al., 2019) (Vaziri, et al., 2021). Some of these studies have been published. (Behera, Majumder, & Nayak, 2018). Behera et al. (2018) employed the extreme learning machine (ELM) and its modifications like particle swarm optimization-extreme learning machine (PSO-ELM), craziness particle swarm optimization-extreme learning machine (CRPSO-ELM) and accelerate particle swarm optimization-extreme learning machine (APSO-ELM) (Behera, Majumder, & Nayak, 2018). The study that conducted used the best-first search algorithm with forwarding selection for the variable selection algorithm (Rana, Koprinska, & Agelidis). Using NN modeling and error metrics analysis, Sonia Leva et al. (2019) developed a persistence model to estimate the output power of a BIPVS. This model was generated by (de Paiva, Pimentel, Marra, de Alvarenga, & Muss, 2019) (Sonia Leva et al). Recurrent neural networks (RNN) linked with multi-time-horizon predictions were used to provide short-time-horizon SPV power forecasts by (Mishra & Palanisamy, 2018). Mishra et al. (2020) (Mishra & Palanisamy, 2018). These forecasts had a time horizon of between one and four hours. Kardakos et al. (2020) developed an NWP model for the power prediction of a grid-connected photovoltaic plant by applying the SARIMA model to the ANN in order to forecast the solar insolation as a consequence of numerous inputs (Chen, Duan, Cai, & Liu, 2021) (Kardakos, et al., 2013). Developed a 24-hour forward forecasting model that made use of a self-organized map (SOM) for the purpose of weather data categorization (Chen, Duan, Cai, & Liu, 2021). On the other hand, Lee et al. (2021) created two long-short-term memory (LSTM) models that had three hidden layers in each of the models (Lee & Kim, 2019). Nasser et al. (2021) developed, trained, and evaluated five LSTM models for hour ahead SPV forecasting of a system located between Aswan and Cairo, Egypt (Nasser & K, 2019). These models included varying inputs, different types of LSTM, and a number of layers.

1.2. Problem Statement

Especially in the context of current plants with bigger capacity and their integration into the grid, solar photovoltaic (SPV) power forecasting plays an essential part in the effective integration and operation of SPV plants. This is especially true when it comes to newer plants. However, the majority of the research that has been done on SPV power forecasting has been on methodologies with a limited time horizon. These methods are not sufficient to meet the requirements of current SPV plants. As a consequence of this, there is an urgent requirement to move the focus towards methodologies that provide a power prediction over a long-term horizon for SPV systems.

In order to fill this vacuum in research, it is vital to conduct an in-depth comparison of the methodologies of predicting short and long time horizons. The literature that is currently available on such comparisons is, unfortunately, quite limited, which further impedes the development and deployment of effective SPV power forecasting systems.

In addition, traditional techniques of forecasting have a tendency to become less accurate as the number of steps rises, which limits the usefulness of these approaches for future power projection. This constraint presents a substantial problem for dependable and accurate long-term power forecasting, in particular when taking into consideration the dynamic nature of SPV systems.

In addition, the vast bulk of research done in this area has mostly focused on SPV plants that have limited capacities, omitting to take into account the unique issues that are presented by plants with bigger capacities. Existing methods of forecasting, when applied to bigger plants, result in a considerable rise in the root mean square error (RMSE), which indicates the inadequacy of present approaches for handling the complexity of higher-capacity SPV systems. RMSE stands for root meaning square error.

In addition, the short period of meteorological and power data records, as well as the possibility of errors in these records, presents a barrier to their application in forecasting. Because the data gathering only lasts for a short period of time, it is impossible to accurately capture the whole spectrum of weather conditions and fluctuations in power output. As a result, the accuracy and reliability of forecasting models are negatively impacted.

In light of these research deficiencies, there is an immediate need to create SPV power forecasting approaches that are competent, efficient, and capable of meeting the long-term horizon forecasting needs of current SPV plants. These methods are crucial for the efficient installation and operation of real-world SPV plants, and they make possible both an improvement in system design and an efficient incorporation into the larger-scale grid.

1.3. Aim

This research project's objective is to create an enhanced method for long-term solar photovoltaic (PV) power forecasting by employing a long short-term memory (LSTM) model, and then to assess that method's effectiveness in comparison to a number of alternative time-series and neural network models utilizing a variety of optimizers. In the context of integrating renewable energy sources into the power grid, the goal is to improve the accuracy of solar photovoltaic (PV) power forecasts and offer a reliable approach that can be used for better system planning and management. These goals are intended to be accomplished simultaneously.

1.4. Importance Of The Study

This project's relevance rests in the significance of accurate solar PV power forecasting for the effective integration and operation of renewable energy sources, notably solar power, inside the power grid. This is where the project's significance resides. The following is a list of the most crucial reasons why this project should be done:

1. Improvements to Existing Grid Integration Accurate solar photovoltaic power forecasting methods are an absolute necessity for the successful incorporation of solar energy into the existing electrical grid. The capacity of grid operators and energy

planners to more effectively control the unpredictability and intermittency of solar power output is made possible by accurate predictions. This helps to provide a steady and predictable supply of electricity.

2. Enhancement of System Planning: Accurate long-term solar PV power forecasting contributes to improvements in system planning for the deployment of renewable energy sources. It gives vital insights into future patterns of power generation, which enables stakeholders to make educated decisions regarding capacity growth, infrastructure expenditures, and energy management techniques.

3. Accurate Prediction of Solar PV Power output: This project helps improve resource allocation since it makes accurate predictions of solar PV power output. Participants in the energy market, legislators, and investors may all make use of the projections to more effectively manage resources, strike a balance between supply and demand in the energy market, and maximize income streams.

4. Cost Reduction: There are various different ways in which accurate solar PV power forecasting leads to cost reduction. It makes it possible to make efficient use of solar electricity, which in turn lessens the need for traditional sources of backup power and helps keep energy imbalances to a minimum. As a consequence of this, operating costs are reduced, improvements are made to energy trading, and overall grid efficiency is improved.

5. The Penetration of Renewable Energy Sources Accurate solar PV power forecasting is an essential component in the process of increasing the percentage of renewable energy sources in use. This initiative fosters confidence among grid operators, investors, and consumers by giving accurate projections. This confidence helps to facilitate the greater deployment of solar power and supports the transition to an energy system that is more environmentally friendly.

6. Benefits to the Environment: The successful application of dependable solar PV power forecasting techniques makes it possible to include a greater percentage of clean and renewable energy into the overall energy mix. This helps to reduce greenhouse gas emissions, lessen our reliance on fossil fuels, and provides support for global efforts to address climate change and work toward a more sustainable future.

In conclusion, the significance of this study rests in the possibility that it will raise the precision of solar photovoltaic power forecasting, which will ultimately result in greater grid integration, improved system planning, optimum resource allocation, cost reduction, increasing penetration of renewable energy sources, and environmental advantages.

1.5. Objectives

The following list illustrates all of this project's goals:

1. Construct a model based on long-term and short-term memory (LSTM) The project's primary objective is to build and implement an LSTM model that is especially customized for long-term solar PV power forecasting. The LSTM model was selected because of its capacity to properly manage time-series data and to capture long-term dependencies, which led to an improvement in the accuracy of forecasting.

2. Assess the performance of several optimizers for LSTM The purpose of this project is to evaluate the effectiveness of various optimizers in combination with the LSTM model. These optimizers include Nadam, RMSprop, Adam, Adamax, SGD, Adagrad, Adadelta, and Ftrl. The goal of the project is to improve the accuracy of solar PV power forecasts by identifying the optimizer that is the most effective, and this will be done by evaluating the effects of several optimizers.

3. Evaluate the suggested LSTM model in relation to other time-series models and other neural network models the purpose of this project is to evaluate the proposed LSTM model in relation to other time-series models and different neural network models. This comparison will give insights into the relative performance of the LSTM-based technique, as well as its superiority, for long-term solar PV power forecasting.

4. Evaluate the increase in forecasting accuracy the purpose of this project is to measure the improvement in forecasting accuracy achieved by the LSTM model with the Nadam optimizer in comparison to other models. The purpose of the study is to demonstrate that the technique that has been presented is superior by carrying out an exhaustive assessment, which will include comparisons against models such as the autoregressive integrated moving average (ARIMA), the seasonal autoregressive integrated moving average (SARIMA), and other neural network models.

5. Evaluate the suggested approach the purpose of this research is to evaluate the proposed LSTM model with Nadam optimizer using real-world data from a solar PV power system that has a 250.25 kW installed capacity and is situated at MANIT Bhopal in the Indian state of Madhya Pradesh. During the validation step, the performance of the model will be evaluated in a real-world environment to see whether or not it is successful for solar PV power forecasting.

6. Showcase the practical application of the suggested methodology for improved system planning and management: The project's goal is to demonstrate the practical applicability of the methodology being offered. The project aims to show the relevance and utility of the created model in assisting decision-making processes linked to the integration of renewable energy sources and grid management by giving accurate long-term solar PV power estimates.

The development of an LSTM-based forecasting model is one of the overarching goals of this project, as are the evaluation of various optimizers, comparison with previously developed models, quantification of forecasting accuracy improvement, validation with real-world data, and demonstration of practical applicability for system planning and management.

1.6. Contributions

It is necessary for the design and management of solar power facilities to have accurate forecasts of solar electricity. In the present investigation, a unique method for long-term solar power forecasting was developed. This method is based on LSTM and makes use of the Nadam optimizer. Several alternative methods were used to conduct an analysis of the LSTM models, ARIMA models, and SARIMA models to see which one produced the best results. The accuracy of the forecasts produced by the suggested method was significantly better than those produced by current techniques. The accuracy and dependability of the system were both improved as a result of this study's investigation of the influence of climatic elements on solar power forecasting. The following is a condensed summary of the most important contributions made by this study:

• The presentation of a unique approach for SPV power forecasting that is based on LSTM and uses the Nadam optimizer. This technique takes into consideration the fluctuations in SPV output power with regard to meteorological data.

• Solar power forecasts using eight different LSTM models, in addition to ARIMA and SARIMA models; evaluation and comparison of these models.

• A comparison of the proposed method with other SPV power forecasting methods that are commonly used to determine whether or not the suggested method is effective.

• The prediction of mitigating factors such as solar irradiance, solar photovoltaic module efficiency, and other meteorological characteristics, which improves the system's accuracy and dependability.

1.7. Outline

The thesis will be organized in a systematic manner, beginning with an introduction that provides background information and defines the goals of the study. The literature review aims to offer an in-depth summary of the pertinent literature and identify any gaps in the previous study. The deep learning strategy utilizing the LSTM model will be described in the methods chapter, along with data collection, preprocessing, feature engineering, and performance measures. In the section titled "Results and Analysis," the collected findings will be presented and discussed, focusing on areas of excellence, shortcomings, and opportunities for growth. Lastly, the conclusion will include a summary of the research findings, an evaluation of the difficulties associated with project management, and recommendations for further research. The thesis will also include a list of references, documenting the sources utilized throughout the study.

CHAPTER TWO

LITERETURE REVIEW

The purpose of this literature review is to give background information and an in-depth study of the current state-of-the-art methods in machine learning and time-series analysis for solar PV power forecasting, which is covered in Chapter 2. The purpose of this chapter is to analyze the advantages and disadvantages of current methods, as well as to spot research gaps. To this end, a literature assessment will be done, with an emphasis on studies that use machine learning methods to solar PV power forecasting. The important performance indicators used to evaluate the accuracy of forecasting models will also be discussed in this chapter, along with an introduction to the core principles of machine learning and time-series analysis. The information presented in this chapter will provide the groundwork for the next several chapters, which will focus on the creation of a revolutionary deep learning technique for solar PV power forecasting.

2.1. Photovoltaic Solar Power

A photovoltaic (PV) system is a collection of solar modules, each of which is made up of solar cells units that are able to convert the energy contained within solar radiation into electricity that can be used. When compared to the non-renewable energy sources that have been routinely utilized in the past, these systems are distinguished by their absence of greenhouse gas emissions and pollution, which earns them the reputation of being ecologically beneficial (Fraas, 2010). However, the performance of photovoltaic (PV) systems that are linked to the grid is affected by a variety of variables. To arrive at an accurate estimate of the amount of power that may be generated by a PV system, it is necessary to take into consideration a number of elements, including the state of the surface, the amount of solar irradiance, the level of radiation intensity, and the amount of cloud cover. In addition, solar cells are affected by the temperature of the surrounding air since the conversion efficiency of solar cells diminishes as the temperature increases. Because of this, the power output of a photovoltaic (PV) system is not a straightforward linear function of the amount of solar irradiation; rather, it is impacted by the random nature of a number of other climatic conditions (Cîrstea, Martiş, Cîrstea, Constantinescu-Dobra, & Fülöp, 2018). It is essential to have a thorough understanding of these complexities and to factor them in if one wants to effectively evaluate and anticipate the performance of PV systems. Researchers are able to construct accurate forecasting models by taking into account the aforementioned parameters. These models reflect the complexities of PV system functioning, which in turn improves the dependability and efficiency of solar energy output.

2.1.1. Technology review

Around the year 1839, when Edmond Becquerel discovered the photovoltaic effect (jobim & junior, 2014). There was a rise in the number of people interested in photovoltaic solar energy. The first solar cell was manufactured in 1876, but for a considerable amount of time, advancements in technology were limited to research carried out outdoors. It wasn't until perhaps around the year 1960 that this sort of technology first began to be manufactured on an industrial basis (jobim & junior, 2014) (J, 2021).

The advancement of photovoltaic technology was catalyzed by two important milestones that played an essential part in the process. The necessity for alternative power sources to provide electricity to remote areas was the impetus behind the first major accomplishment. The second significant event was the so-called "space race," which made use of solar technology to supply energy to the many pieces of equipment that were housed in space (jobim & junior, 2014).

The price of oil skyrocketed in the 1970s, which further pushed the development of photovoltaics forward. Countries such as the United States made investments in the subject as a result of their desire to find new sources of energy. The Asian market, and China in particular, started making significant investments in the manufacturing of solar modules in the early 2000s and eventually took the lead in this industry in 2009 (jobim & junior, 2014).

The expansion of the solar sector between 2003 and 2014 reached an annual rate of 54.2%, as reported by CRESESB (Manual de Engenharia Para Sistemas Fotovoltaicos – 2014 – Solenerg Energia Solar Fotovoltaica, n.d.). This expansion was largely fueled by incentives to promote the utilization of renewable energy sources as an alternative to fossil fuels. China leads the globe in terms of installed capacity for solar electricity, followed by the United States and Japan (Ponprathom & Teekasap, 2022). As of 2016, the total amount of photovoltaic power that was installed globally reached 294GW. The price of solar modules continues to be a barrier to broad deployment (jobim & junior, 2014). Despite the fact that production levels have grown. To find a solution to this problem, a number of nations have instituted financial incentives for research and technical development, and the governments of these nations are providing financial assistance as well.

The manufacture of solar cells and modules has not been a primary emphasis of technological advancement in photovoltaics in Brazil, as opposed to the more theoretical studies that have taken place at research institutions. However, Brazil has made large expenditures in the construction of solar systems, primarily for the purpose of delivering electricity to regions that are separated from the national grid, which is referred to as the National Interconnected System (SIN). The Brazilian regulatory organization known as Aneel issued Normative Resolution number 482/2012 in the year 2012, which prescribed standards for the development of solar systems that were linked to the distribution network. The installation of tiny and micro producers was given a significant boost as a result of this resolution.

By the year 2017, Brazil had a total of 438.3 megawatts (MW) of solar electricity installed, which was spread throughout 15.7 thousand projects associated to these systems (J, 2021). Tax breaks, assistance programs from government agencies and institutions like the Brazilian Development Bank (Banco Nacional do Desenvolvimento, BNDES), and incentives for micro and distributed mini-generation systems are some of the ways the country hopes to further encourage the expansion of solar installations. Solar power is expected to account for 9 percent of Brazil's entire national energy supply by the year 2050 (J, 2021), according to recent projections.

The rise in interest in photovoltaic generating may be attributed to the improvements in system efficiency that have been made, in addition to the incentives for research and development. In the 1950s, the efficiency of solar panels was around 5%, and the cost of producing one watt peak of electricity was $1.785. However, the efficiency of modules available today has grown to around 15%, and their costs are at around $1.20 per watt peak (J, 2021). Silicon is by far the most common material employed in the production of photovoltaic cells, accounting for about 95% of all cells manufactured on a global scale (jobim & junior, 2014) (J, 2021).This is because it is readily available, has a low price, and has production techniques that are already well-established. Monocrystalline silicon, polycrystalline silicon, and thin silicon film are the three primary varieties of photovoltaic cells that are available for purchase on the commercial market (jobim & junior, 2014).

2.1.2. Operation Of Photovoltaic System

It is necessary to have an understanding of both solar radiation and irradiance (Hersch & Zweibel, 1982). In order to comprehend the operation of a photovoltaic (PV) cell as well as the functioning of a PV panel. Solar radiation is the term used to describe the solar energy that is sent to Earth in the form of electromagnetic radiation. This solar energy can either be directly or diffusely received by the Earth. Diffuse radiation is the light that indirectly impacts a surface after being reflected and diffracted (Hersch & Zweibel, 1982). Direct radiation relates to the light that strikes a horizontal surface in a direct manner in a straight line of progression.

Irradiance, which is a quantification of solar radiation and shows the power of solar radiation in relation to the area, is expressed as watts per square meter. When evaluating the efficiency of PV panels, it is usual practice to use a reference value of one thousand W/m2 (jobim & junior, 2014).

As can be seen in Figure 1, the fundamental structure of a photovoltaic cell consists of two distinct types of semiconductor materials known as N-type and P-type, as well as electrical contacts that serve to complete the circuit and allow current to flow (Hersch & Zweibel, 1982). The generation of electricity by a photovoltaic cell is accomplished by applying the photovoltaic principle, which states that the cell should convert the solar energy that it receives into electric current.

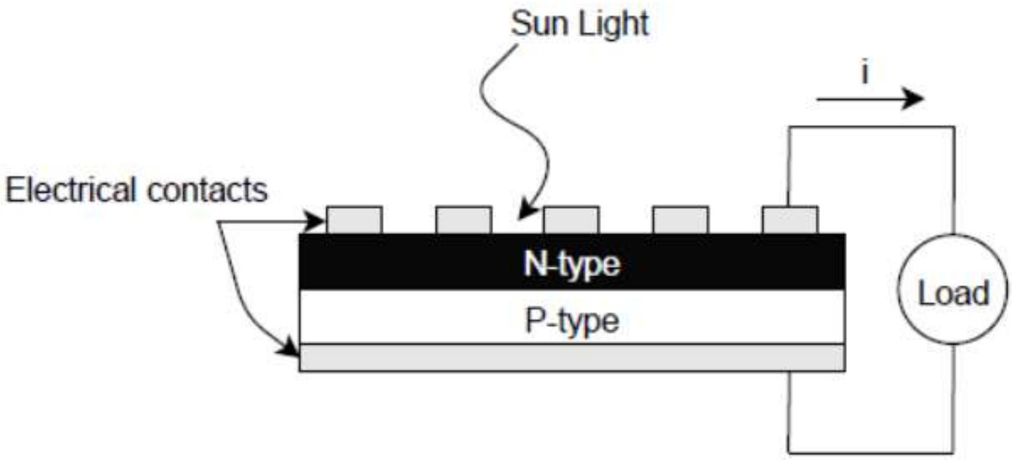


Figure 1. Photovoltaic cell model

When light reaches the surface of crystalline silicon, it has the potential to be absorbed, which then causes changes in the material's electrical characteristics. If there is enough energy, the electrons in the valence layer of the crystal will get excited, and they will be able to flow freely throughout the crystal. The excitation of electrons results in the creation of holes, and due to the attraction and mobility of valence electrons, it is possible for both electrons and holes to travel through the crystal. The electrons and holes in a system become more unsettled as the radiation's energy level rises (Hersch & Zweibel, 1982).

A potential barrier is required in order to complete the transformation of this excitement into an electric current. Due to the fact that the barrier is responsible for separating free electrons from holes, the photovoltaic cell will have an excess of electrons on one side of the cell, while the other side will have an excess of holes. This difference in potential generates an electric field, which may be used to generate current provided the right conditions are met. The potential barrier is produced at the intersection of a positively doped silicon shell (P-type) and a negatively doped silicon shell (N-type), both of which are doped with boron. The process of negatively doping the material results in the introduction of atoms that have five valence electrons, whereas the process of positively doping the material results in the introduction of atoms that have one fewer electron in their valence shell. It is the N-type material that ends up carrying the majority of electrons, while the P-type material ends up carrying the majority of holes (Hersch & Zweibel, 1982).

When N-type and P-type materials are brought into contact with one another, electrons from the N-type material fill the holes in the P-type material. This results in clusters of positive charges being created on the N side of the junction, and clusters of negative charges being created on the P side. An electrical barrier will build as the movement of charges continues; this barrier will prohibit any more charges from moving in. The potential barrier that was discussed before (Hersch & Zweibel, 1982). is created when this equilibrium is reached.

When light strikes a material with an N-type atomic structure, electrons get separated from their associated holes, and holes speed toward the barrier, where they attempt to recombine with the negative charges found on the P-type side. In a similar fashion, extra electrons are generated in the N-type material whenever light strikes the P-type material. When the cell is connected to an electric circuit, an electric current is produced when the electrons leave the N-type material, travel through the circuit, and then recombine with the holes on the P-type side of the material. This process takes place when the cell is in operation. The quantity of light energy that is absorbed by the cell as well as the energy of the electrons that are created is proportional to the amount of current that is produced (Hersch & Zweibel, 1982).

2.2. AI & The Deep Learning Model

Neurons are the processing units that are a part of a neural network's typical architecture. Neurons are coupled with one another. These neurons, when given input, create an output value depending on an activation threshold according to the value of the input. Neural networks are able to approximate a broad variety of functions, frequently including nonlinear ones, because to the intricate networks of linked neurons that they are built from. The number of stages of computation that are chained together to create a neural network determines the depth of the network. Each step of computation consists of one or more neurons operating in parallel. Deep learning may be thought of as an introduction to the notion of chaining together many levels of computation (Hao, Zhang, & Ma, 2016).

An approach to artificial intelligence that makes use of deep learning, which is a subset of neural networks, is described here. The study conducted in the field of neuroscience provides the basis for the ideas and fundamental concepts that underpin it. It is essential to keep in mind that the objective of deep learning is not to imitate the precise operations of the human brain. Instead, it takes its cues from a limited number of abstract and high-level ideas that are connected to the fundamental operation of the human brain. These ideas are used as the foundation for the construction of effective deep learning models, which makes it possible to automate difficult activities and derive meaningful patterns from enormous datasets (Deep Learning, n.d.).

2.3. Classification Of Forecasting Methods

Utilizing previous patterns included within the data that is now accessible, forecasting is an extremely important part of predicting what will happen in the future. It is a helpful statistical tool that may be used to estimate a variety of characteristics and can vary from short-term forecasts for the next few minutes to long-term estimates spanning many years. In the context of solar photovoltaic (SPV) power forecasting, the selection of a suitable forecasting approach is contingent on a number of parameters including the size of the PV plant, the required forecasting horizon, the geographical location, and the existence of other climatic fluctuations.

It is essential to decide which technique of forecasting is best appropriate for any given set of conditions in order to successfully manage the risks that are connected with predicting. In the next sections, an in-depth investigation into the various categorizations of SPV forecasting methodologies is provided. These categories offer insights into the many methodologies that are accessible for SPV power forecasting. As a result, they enable researchers and practitioners to make educated judgments based on the specific requirements of their individual situations.

By going deeper into the complexities of various forecasting systems, it is possible to evaluate their benefits and drawbacks, which paves the way for the selection of the method that is the most suitable for producing accurate and trustworthy SPV power forecasts. This thorough grasp of various forecasting approaches ensures that the inherent uncertainties and problems connected with SPV power forecasting are successfully handled, which ultimately leads to enhanced decision-making and optimum performance in the renewable energy industry. Moreover, this understanding ensures that the inherent uncertainties and challenges associated with SPV power forecasting are properly controlled.

2.3.1. Horizon forecasting

The forecasting horizon, which refers to the time period for which the SPV power production is forecasted, is one of the most important factors to consider when selecting the most suitable approach for forecasting. A visual depiction of the many types of solar PV forecasting methodologies based on their individual time horizons is shown in Figure 2. This picture is based on the information that is mentioned in Table 1, which can be found here.

These ways of forecasting can be generically characterized as short-term, medium-term, or long-term approaches, with each one catering to certain timelines and goals. Short-term forecasting approaches concentrate on making forecasts for the immediate or very near future, often spanning from a few minutes to a few hours into the future. The prediction horizon is extended to days or weeks with medium-term forecasting, providing insights into slightly longer-term power generation trends than short-term forecasting does. Lastly, long-term forecasting approaches enhance strategic planning and policymaking in the SPV sector by providing estimations for the months, years, or even decades into the future.

The characteristics, methodology, and applications of each forecasting approach are broken down in further detail within Table 1, which provides an overview of the various time ranges. Stakeholders are able to make educated judgments on the most appropriate forecasting technique based on their particular needs and goals provided they have a solid grasp of the distinctive characteristics of each methodology and the way in which they compare to one another.

The evolution of precise and efficient power projections is facilitated by this exhaustive categorization and explanation of solar PV forecasting systems based on time horizons. Improved resource allocation, grid management, and overall performance optimization of solar photovoltaic systems are all made possible as a result of this feature's ability to simplify the process of selecting the forecasting approach that is most suited for the various operating circumstances.

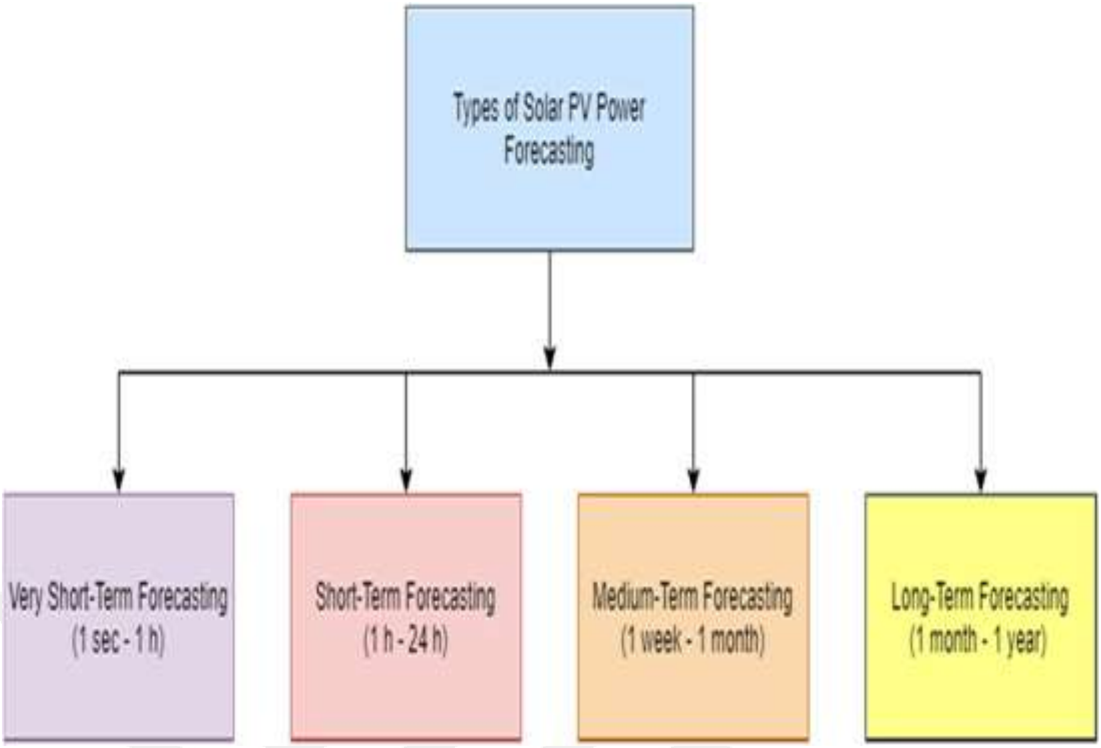


Figure 2. Time-horizon categorization of solar PV predictions

Table 1. Analyzing the time-horizon forecasting of solar PV generation.

Type of solar PV power forecasting	Time-horizon	Applications
Long-term	1 month–1 year	Helps authorities in planning the generation, transmission, and distribution of electricity along with the structuring and operation of electricity markets.
Medium-term	1 week–1 month	Unit commitment decisions, planning, and maintenance scheduling of the power system.
Short-term	1–24 h	Grid security, power reserve management.

Very short-term	1 s–1 h	Power and voltage regulation, real-time electricity dispatch, and grid stability.

2.3.2. Historical data-based

As can be seen in Figure 3, the subject of solar photovoltaic power forecasting has witnessed the emergence of a great deal of different ways of forecasting. Table 2 presents a comprehensive description of various strategies based on their features, which may be utilized for the purpose of navigating the varied terrain that is presented.

The selection of the proper approach, taking into consideration the data that is available and the time horizon that is wanted, is essential to the production of accurate forecasting results. Because every technique of forecasting has its own advantages and disadvantages, it is essential to match the compatibility of the approach to the particular criteria that are being asked of it in the process of predicting.

It is possible for practitioners to align the forecasting approach with the features of the data if they give careful consideration to the available data sources, such as historical data on solar PV power, data on climatic conditions, and other relevant elements. When choosing a technique for forecasting, it is important to take into account the desired time horizon, which can be short-term, medium-term, or long-term.

Table 2 serves as a helpful reference for decision-makers, researchers, and practitioners, allowing them to analyze and select the forecasting approach that is best suited to meet their individual requirements. It helps in the search of precise and dependable solar PV power forecasts by providing insights into the methodology, algorithms, and application of each method.

This exhaustive categorization of forecasting methodologies makes a contribution to the development of the discipline by presenting a methodical overview of the many possibilities that are now accessible. It gives stakeholders the ability to make educated decisions, which in turn leads to improved energy management, greater forecasting accuracy, and optimal exploitation of solar photovoltaic resources.

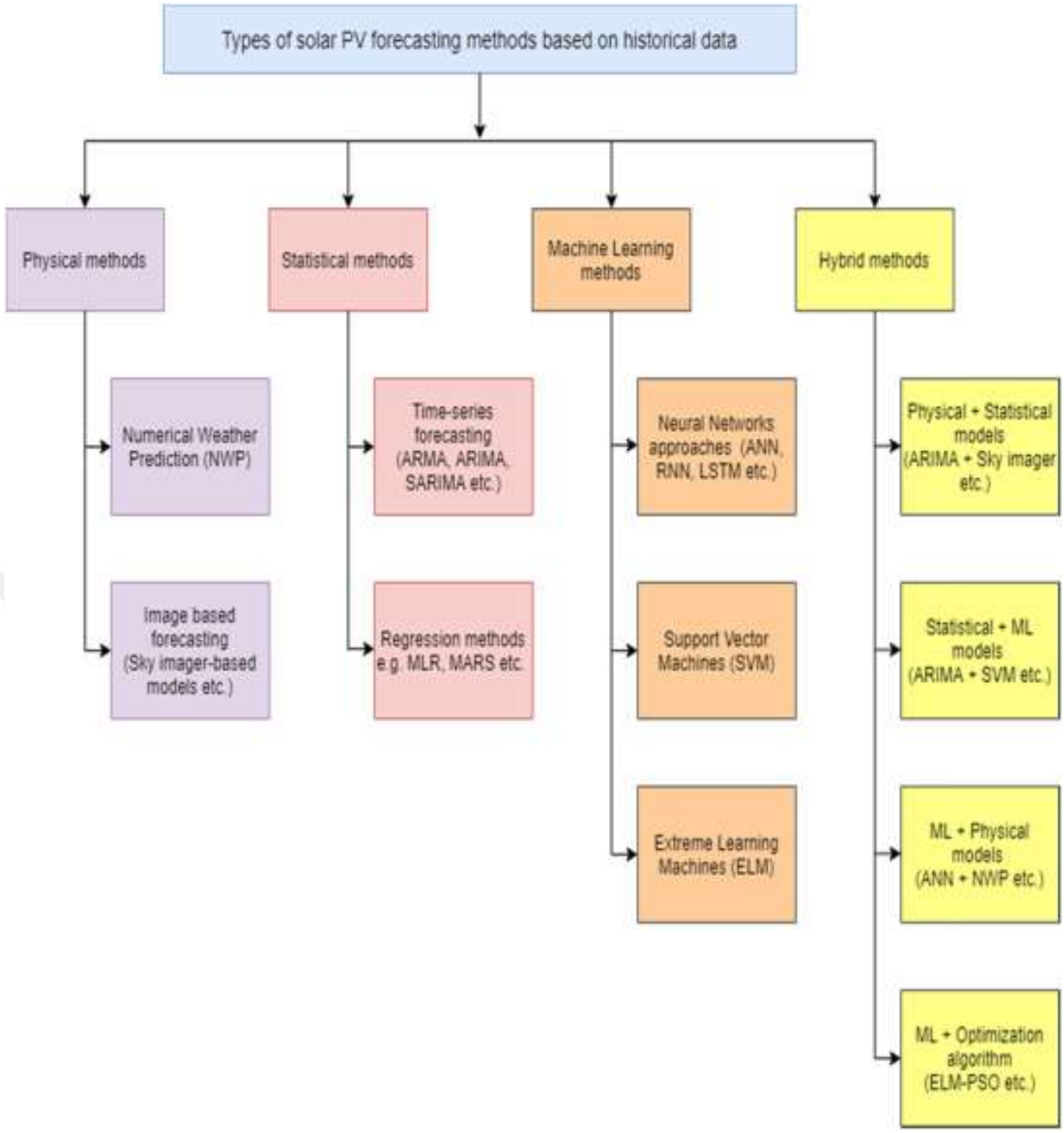


Figure 3. Solar photovoltaic projections are categorized using past data

Table 2. Methods for predicting the output of solar photovoltaic (PV) panels are surveyed.

Forecasting approach	Advantages	Disadvantages
Machine Learning models	- The suggested forecasting system can account for many variables that affect PV power output, which is a major plus.	- Alarge training dataset and optimal training process are needed to ensure the model can

	- The approach can improve its predictive abilities over time by learning from data through the training phase and applying a knowledge-based system. - The method's ability to represent complicated linkages and fluctuations, as captured by its ability to capture significant non-linearities in PV power ge It may also be used in large-scale systems to improve PV power production estimates.	learn from many data patterns and generalize well to new inputs. - A complex architecture and many modeling challenges must be considered to reflect the data's underlying patterns and linkages. - The technique requires more computer power and longer training cycles to properly maintain and use prior data to make accurate forecasts, which increases forecasting performance.
Physical models	- Optimal for cases where there is a scarcity of data and no access to past records. - Allows for the development of forecasting variables to be used in statistical models. - Especially helpful for making long-term predictions, especially when contrasted to satellite-based systems.	- Extremely sensitive to rapid changes in the values of climatic variables; - Requires detailed solar PV models and precise local observations. - Constant recalibration is required due to difficulties in gathering accurate physical input data.
Statistical models	- For accurate short-term forecasting, it makes use of publicly available meteorological data and, in most cases, outperforms physical models. - Its ease of use makes it a realistic option for use in forecasting.	- Accuracy is dependent on a large amount of input data from the past. - Less accurate in the long run when making predictions. - Cannot reliably capture sporadic patterns in input variables.

Hybrid models	- Created with the express goal of improving the efficacy of physical and statistical methods. - Outperforms any other technique for physical modeling. - Increases the precision and dependability of future projections.	- The temporal dynamics of historical PV data are often ignored, and more computing resources and effort are required to integrate several approaches. - To effectively manage various data sets, more RAM is required.

2.4. State-Of-Art Of Forecast Methods

As was noted before, a number of research have been carried out to investigate various strategies for predicting the output of solar electricity (Antonanzas, et al., 2016). These investigations included in-depth literature evaluations, in which every prediction approach was broken down and illustrated with examples. In addition, fundamental ideas such as the prediction horizon and methods for conducting analyses of the methodologies were established. An intriguing idea that was presented in (Antonanzas, et al., 2016). was the skill score (ss), which was used to measure the performance of the model in comparison to more straightforward reference models.

There are research papers that focus on a single method or compare many ways to discover which one is the best fit for a certain assignment. In addition to studies that examine a variety of forecasting strategies, there are also research papers that analyze various forecasting methodologies. For example (Brano, Ciulla, & Falco, 2014). Investigated a single approach, whereas (Almonacid, Rus, Higueras, & Hontoria, 2011). carried out research that compared an artificial neural network (ANN) to three different physical methods: the Osterwald, Araujo-Green, and Single Diode Models. The researchers used irradiance and temperature data as inputs when applying these approaches to various photovoltaic systems. As a result, they were able to derive typical V-I curves as outputs for varying input values measured over the course of a year.

The comparative analysis that was done in (Almonacid, Rus, Higueras, & Hontoria, 2011). Revealed that the ANN technique had lower prediction errors (ranging from 6% to 8%) in comparison to the other methods (which ranged from 6%

to 30%). The capacity of artificial neural networks (ANNs) to take into account a wider variety of losses inherent in solar systems, such as temperature, irradiance, and panel inclination angle, is what the researchers believe is responsible for the superior performance of ANNs. On the other hand, traditional approaches merely took into account temperature losses, which made them far less accurate.

It was also stated in (Dumitru, Gligor, & Enachescu, 2016). That ANN algorithms were used in order to anticipate the amount of electricity generated by solar panels. In this comparison research, the performance of two neural networks, namely the Multilayer Perceptron and the Elman Neural Network, was tested for this specific job. Both of these networks used a method called Backpropagation for their training. The values of the photovoltaic system's historical output power were used as the basis for the input data, which was modeled as a time series. A strategy known as a moving window was utilized, in which the input was a window containing t samples and the output was determined to be the sample that was t samples plus one. The researchers were pleased with their findings, which showed an error rate of around 0.5 percent, and they proposed adding climatic parameters to the variables used in the analysis as a way to further increase accuracy.

The accuracy of two different types of dynamic neural networks, the Focused Time Delay Neural Network and the Distributed Time Delay Neural Network, was compared in another study that was published under the reference number (Al-Messabi, Li, El-Amin, & Goh, 2012). In a manner analogous to that described in (Dumitru, Gligor, & Enachescu, 2016), the input data consisted entirely of historical readings of power output taken from the system's time series and were analyzed with a moving window. Surprisingly, this study did not require historical meteorological data from the installation location, such as irradiance and temperature. This finding suggests that this technique was more relevant in real-world settings since it did not rely on hypothetical conditions. On the other hand, this runs counter to the claim that was made in (Dumitru, Gligor, & Enachescu, 2016), which said that the addition of climatic data improves the effectiveness of forecasting systems.

(Ding, Wang, & Bi, 2011). Utilized a feed-forward neural network (MLP) in order to anticipate the generation of a solar system with a 24-hour window. This research was published in (Ding, Wang, & Bi, 2011). They were able to directly estimate the power production by making use of the system's past data on the generation of power. Both the use of the improved backpropagation algorithm, which addresses limitations of the standard backpropagation algorithm, and the use of the Similar Day Selection Algorithm, which finds historical data from days with similar climatic conditions to enhance forecasting accuracy, particularly on non-sunny days, are two noteworthy aspects of this research. The improved backpropagation algorithm addresses the limitations of the standard backpropagation algorithm.

The research that was carried out by (Brano, Ciulla, & Falco, 2014). Consisted of a comparative study between three distinct artificial network topologies that were applied to the prediction of the generation produced by photovoltaic systems. These topologies were as follows: a multilayer perceptron with one hidden layer; a recurrent neural network multilayer perceptron; and a gamma memory artificial neural network. The findings provided convincing evidence that these networks can be useful for forecasting solar generation, with prediction errors of less than 1% of the total. The maximum power level that was recorded by the actual PV system. In addition, correlation analysis was utilized as a preprocessing technique for the input data. This allowed for the identification of the factors that had the largest influence on the output variable (the amount of solar power generated).

Carried out an intriguing study that was published in , which went beyond only forecasting the output power of PV systems. They came up with a solution to the problem of maximizing the functioning of a microgrid that consisted of a reversible hydropower plant and a solar system. The forecasting system was put to use in order to plan the utilization of the hydroelectric plant, which served as a mechanism for the storage of energy and absorbed power fluctuations caused by the output of photovoltaics. The reliability of the distribution network was preserved by utilizing the hydroelectric plant either to store extra energy or to deliver electricity in the event that the PV system had a malfunction.

(Dolara, Grimaccia, Leva, Mussetta, & Ogliari, 2015). Offered a hybrid strategy that combined a parametric methodology and a statistical method. This hybrid approach was called a hybrid approach. The parametric method included the use of the Clear Sky Solar Radiation Model (CSRM), which represented the irradiance that was incident on the location where the PV system was being installed when there was no cloud cover. After that, a neural network was applied to the problem of predicting the amount of energy generated. According to the findings of the research, when the performance of the hybrid approach was compared to that of employing an ANN (MLP) by itself, the hybrid method produced superior results. It underlined how important it is to accurately predict weather conditions in order to make forecasting more efficient.

(Ogliari, Grimaccia, Leva, & Mussetta, 2013). Created yet another hybrid prediction technique by integrating Dynamic Genetic Swarm Optimization (GSO) with artificial neural networks (ANNs) that were trained using the Backpropagation algorithm. This hybrid technique was devised with the intention of improving forecast accuracy by overcoming the constraints imposed by the Backpropagation algorithm. When compared with just employing ANNs, the findings indicated much better performance.

In a nutshell, the following are some of the most important takeaways from the findings that were presented:

• The manner in which input data in forecasting systems are processed has a substantial impact on the accuracy of the results.

• Similar Day Selection Algorithm, such as that which is provided by the Similar Day Selection Algorithm, helps to increase the accuracy of weather forecasts.

• In predicting problems, hybrid systems that integrate several methodologies perform exceptionally well.

• The relevance of doing correlation analysis between variables cannot be overstated since the selection of suitable input variables for the forecasting model is of the utmost importance.

2.5. Time series

Time series analysis involves the examination and interpretation of statistical data collected at regular intervals. In this type of data, each row is arranged in chronological order, establishing a clear relationship between the data and the time of collection.

Measurements in a time series can be captured at various regular intervals, such as hourly, daily, monthly, or yearly. When dealing with future values that are unknown until a specific date, it becomes necessary to estimate new data points at the corresponding time in the time series (Box, Jenkins, & Reinsel, 2008). This estimation ensures the continuity of the data and enables the application of prediction techniques.

Understanding the pattern exhibited by the data is crucial in comprehending the nature of the time series in the past and anticipating its behavior in the future. The data pattern serves as a valuable guide for selecting an appropriate prediction method.

Time series analysis offers a robust framework for examining historical patterns and making predictions about future trends. By discerning recurring patterns, trends, seasonality, and other relevant characteristics within the data, one can make informed decisions and forecasts in diverse fields, including finance, economics, and weather forecasting, among others.

2.5.1. Time Series Data Structure

A time series is a collection of one or more data points that have a distinct structure that is defined by trend, seasonality, and noise. A time series can include as little as one data point or as many as millions of data points. A continuous observation of data demonstrating an overall increase or decline might be characterized as a trend. It is not always present in every dataset, but it is typical to detect a declining trend in one portion of the data and a rising trend in another section of the data. Although this phenomenon does not always occur, it is common to observe it when it does. For example, due to the broad use of technology, there has been a discernible rise in the amount of daily mobile shopping. A similar pattern can be seen with the overall quantity of energy produced, which has been exhibiting an upward trend due to the expanding number of solar panels that have been placed. During the warmer months, when there is less need for heating, energy consumption goes down, but during the colder months, it exhibits an upward tendency.

A pattern that is repeated over a certain amount of time is said to having seasonality. Similar to trend, it is not an essential quality, but it is something that may be seen in certain datasets. An illustration of a yearly pattern is the rise in the amount of power consumed during the winter months, followed by a fall in that amount during the summer months. This cycle repeats again each year. Another illustration of this phenomenon is the rise and fall, on a daily basis, of the number of times users log on to a certain website, with the peak occurring in the morning and the lowest point being in the evening. Depending on how frequently data is collected, seasonality can manifest itself over a wide range of time periods, from seconds to minutes to even years. When trying to anticipate future values in a time series, having an understanding of seasonality is absolutely necessary.

The term "noise" refers to irregularities in the data that cannot be explained by either a trend or seasonal patterns. It is a result of circumstances that cannot be predicted and persists even after trend and seasonal factors have been taken into account. For example, despite the fact that the price of gold tends to go up over the summer months, there may be some circumstances, such as an economic crisis, in which the price of gold goes down at that time. This variation that is not consistent with the pattern that was anticipated is an example of noise.

Data from time series can be organized into univariate or multivariate categories. Only the information gathered at the specified times is taken into consideration in a univariate dataset. On the other hand, a multivariate dataset contains supplementary information that was gathered concurrently with the observed data. For instance, solar energy generation numbers by themselves are an example of univariate data since they incorporate both time and the data that was seen. On the other hand, a multivariate dataset includes additional pertinent data that either directly or indirectly contributes to the data that was seen at the same time. In the context of the generation of solar energy, a multivariate dataset would contain not only the solar generating values but also meteorological information, the cloud rate, the rain rate, and the locations of the various power producing facilities. Visualizing univariate datasets is often simpler compared to multivariate ones.

Understanding the features of time series, such as trend, seasonality, and noise, as well as the difference between univariate and multivariate datasets, is essential for properly assessing and forecasting future values in a variety of industries, including finance, economics, energy, and other related areas.

2.5.2. Time series terminologies

Trend refers to a long-term change in the mean level of a time series. However, defining what constitutes "long-term" can be challenging. In certain cases, climatic variables exhibit cyclic variations over extended periods, such as 50 years. If one only has access to 20 years of data, this long-term oscillation would appear as a trend. However, with several hundred years of data, the long-term oscillation would become more evident.

A stationary time series is characterized by the absence of systematic changes in its mean (first moment), variance (second moment), and strictly periodic variations. Most of the probability theory of time series focuses on stationary time series. Consequently, time series analysis often involves transforming non-stationary series into stationary ones to leverage this theory. For instance, removing trends and seasonal variations from a dataset allows modeling the remaining variation using a stationary stochastic process.

A different filter is a specific type of filter that is particularly useful for removing trends. The first-order difference is commonly employed and often sufficient to achieve apparent stationarity. In some cases, a second-order difference may be necessary for specific datasets. However, over-differencing should be avoided as it amplifies the variance.

Autocorrelation is a vital indicator of time series properties. Sample autocorrelation coefficients (rk) measure the correlation between observations at different lags. These coefficients provide insights into the underlying probability model that generated the data.

A correlogram, which plots the autocorrelation coefficients against the corresponding lags, is a valuable tool for interpreting a set of autocorrelation coefficients. Correlograms aid in model identification and are particularly useful for selecting the most appropriate type of Autoregressive Integrated Moving Average (ARIMA) model to represent the observed time series. When dealing with stationary series, the correlogram is compared to the theoretical autocorrelation functions of different ARMA processes to determine the most suitable one. For instance, the correlogram of a Moving Average (MA) process of order q exhibits a distinctive cutoff at lag q.

Stationary series often display short-term correlations, which are characterized by a relatively large initial autocorrelation coefficient (rt) followed by a few subsequent coefficients that gradually decrease. Autocorrelation coefficients for longer lags tend to be approximately zero.

In the case of a completely random time series, autocorrelation coefficients at non-zero lags are approximately zero for large sample sizes (n). For a random time series, the autocorrelation coefficients (rk) approximately follow a normal distribution with a mean of zero and a variance of 1/n. Consequently, if a time series is random, around 95% of the autocorrelation coefficients are expected to fall within the range of $\pm 2/\sqrt{n}$, which is within two standard deviations of the true value of zero.

Outliers can significantly impact the correlogram of a time series. It is advisable to address outliers before commencing formal analysis to mitigate their effects.

2.6. Sequence Modeling

In this part, the idea of sequence modeling is broken down, and several architectural approaches that have been established to solve sequence modeling issues are investigated. The recurrent neural network, often known as an RNN, is an example of a popular type of architecture (Deep Learning, n.d.). Even though recurrent neural networks (RNNs) are not the major focus of this paper, it is vital to appreciate the underlying ideas behind RNNs in order to understand the reason behind the deep learning model that was chosen for this study: long short-term memory neural networks (LSTM). A concise summary of RNNs and their limitations is provided in sections 2.3.2 and 2.3.3, respectively, of this thesis in order to offer correct context for LSTM within the framework of this thesis. This is done so in order to provide adequate context for LSTM within the framework of this thesis. In the following sections, a more extensive description of LSTM neural networks will be provided, illuminating their design and operation in the process.

2.6.1. Optimization

It is essential to have a solid understanding of optimization and loss functions (Deep Learning, n.d.) before getting into the complexities of RNNs and LSTMs. Standard procedures for assessing and evaluating deep learning models focus on the models' overall performance in respect to an error function, which is also referred to as a loss function. The objective of a loss function is to supply a quantifiable measurement of the degree to which a particular deep learning algorithm accomplishes a given job, which will be represented by the letter T. In the case of regression tasks, the output of a high value by the loss function implies that the deep learning model is not well optimized and struggles to represent the given dataset in an appropriate manner. On the other hand, a low value output indicates that the deep learning model is highly optimized and successfully captures the patterns and relationships included within the dataset. This is indicated by the fact that the model has a low value. The selection of an appropriate loss function is essential because it directs the learning process of the deep learning model. This gives the model the ability to alter its parameters and minimize the loss while it is being trained, thus it is essential that this function be chosen carefully. The goal of the deep learning algorithm is to enhance its performance and produce better outcomes for the task that has been presented to it by iteratively improving the model with the loss function that has been chosen.

2.6.2. Recurrent Neural Networks

Recurrent neural networks (RNNs), a prominent form of neural network that is particularly intended for processing sequential input, are frequently used in sequence modeling. This is because RNNs are one of the few types of neural networks. RNNs are capable of capturing sequential relationships within the data, which is something that typical feed-forward neural networks are unable to do since they regard each input as independent. RNNs are advantageously suited for applications such as natural language processing, speech recognition, and time series analysis as a result of this feature.

The incorporation of recurrent connections among hidden units is the basic idea that underpins RNNs. These connections make it possible for information to be maintained and passed on from one time step to the next. Because of this, the network is able to keep a recollection of previous inputs and incorporate them into the prediction or output it is producing at the moment. A fundamental RNN design is depicted in Figure 4. In this architecture, each hidden unit is connected to both the previous hidden state and the current input. These connections are shown to be bidirectional. The RNN performs processing on the input at each time step and provides an output as a result. This output may then be utilized for further prediction or may be sent back into the network for later time steps.

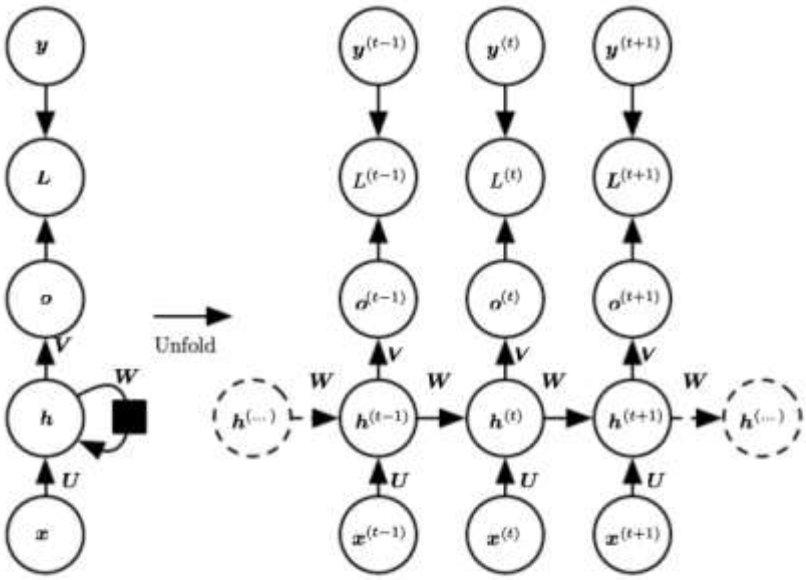


Figure 4. A representation of a basic RNN

RNNs are excellent at modeling sequential data that displays temporal dependencies, such as predicting the next word in a phrase or projecting future values in a time series. This is made possible by the use of recurrent connections, which are utilized by RNNs. However, classic RNNs have an issue known as "vanishing gradient," which occurs when the gradient signal weakens with time and makes it difficult to capture long-term dependencies. This makes it difficult to train RNNs.

In order to circumvent this constraint, more sophisticated RNN topologies have been created, such as the Long Short-Term Memory (LSTM) network. LSTMs are equipped with memory cells and several other gating mechanisms, which enable them to selectively store and retrieve information. As a result, they are able to seize and remember significant data across longer sequences. These architectural innovations provide a solution to the problem of vanishing gradients and make it possible to describe complicated sequences in a more accurate manner.

In conclusion, recurrent neural networks (RNNs) are an effective method for sequence modeling because of their capacity to deal with sequential dependencies. However, recent developments such as LSTM have further increased their capacities. These developments have resulted in greater long-term memory as well as improved performance across a wide range of sequence-based activities.

2.6.3. LSTM Neural Networks

The application of recurrent neural networks (RNNs), which show promise in principle but are generally prevented from being used in practice owing to inherent constraints, has potential. To be more specific, problems with bursting and disappearing gradients reduce their usefulness in long-term sequence modeling. Long-Short-Term Memory (LSTM) neural networks are a specific architecture that was built to solve these difficulties. LSTM stands for "long short-term memory." The LSTM algorithm, which was developed in 1997 (Hochreiter & Schmidhuber, 1997). By Hochreiter and Schmidhuber, was the first to introduce the idea of a memory block that might include one or more memory cells.

Alongside the memory cell, the LSTM memory block's first iteration of design called for the inclusion of a pair of gating devices known as an input gate and an output gate. The LSTM architecture was improved by (Gers, Schmidhuber, & Cummins, 1999). In the year 2000 by the addition of a gate that is now often referred to as the forget gate. Figure 5 presents a graphical depiction of a single LSTM memory block, highlighting the fundamental elements that make up this type of memory. Three vector inputs, namely s(t1)i, h(t1)i, and x(t)i, are subjected to changes within the context of this composition. These transformations are made possible via a combination of sigmoid units, vector addition units, and vector multiplication units.

Figure 6 is an illustration of the recurrent nature of a single LSTM unit. It highlights how the internal state vector s, and the hidden unit vector h are converted or conserved. The purpose of this illustration is to convey the temporal dependencies that occur across time. This part goes more into the underlying structure of the LSTM memory block, illuminating its separate components and demonstrating how LSTM efficiently handles the problem of disappearing and bursting gradients in the process of doing so.

To gain a grasp of how the LSTM system overcomes the obstacles presented by long-term sequence modeling, it is essential to comprehend the complexities of the LSTM architecture, including its gating mechanisms and memory cells. In the following sections, an in-depth investigation will be conducted into the inner workings of the LSTM, aiming to provide a greater understanding of its capacity to acquire and remember essential information for lengthy periods of time.

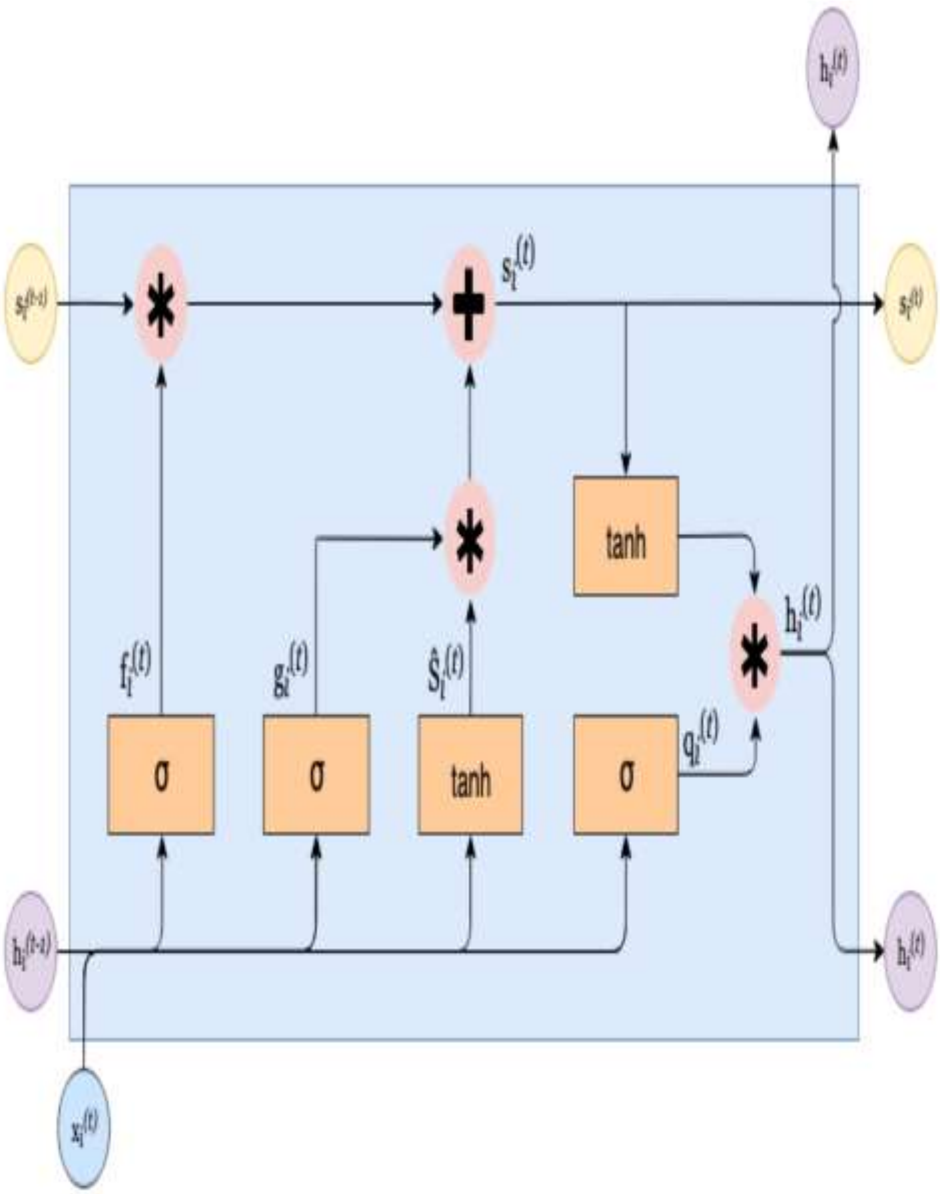


Figure 5. A standard LSTM memory block

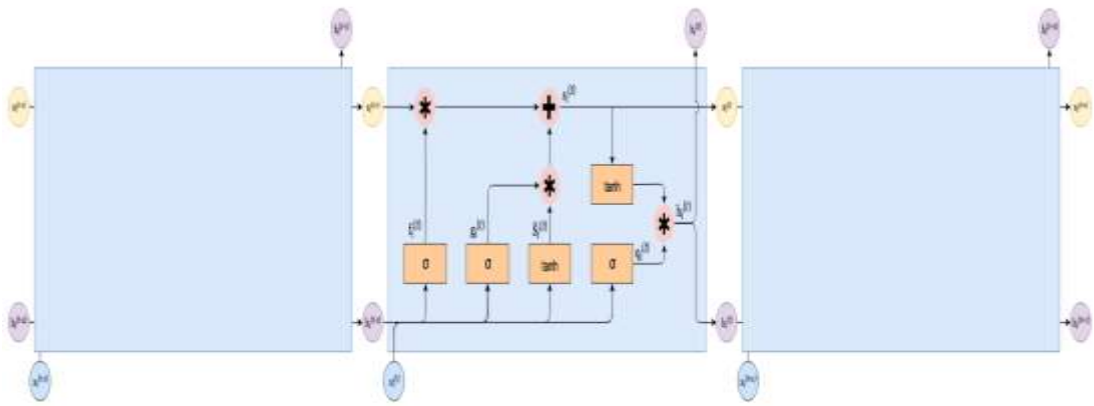


Figure 6. temporal recurrence of a LSTM memory block

2.6.4. Cell state

The capacity of the LSTM memory block to sustain and transmit a modified or maintained state, which is more frequently referred to as memory, over time steps is one of the most important elements of this type of memory. This is accomplished by including a cell state within the memory block, which may be altered by a variety of components including the input gate, output gate, and forget gate. This allows the memory block to function as intended. Figure 7 provides a graphical representation of the cell state in the form of the horizontal connection.

The cell state goes through a series of operations that are carried out within the LSTM memory block i. These operations include pointwise vector multiplications and vector additions. An operation known as pointwise vector multiplication is performed on the prior cell state using the output from the forget gate as the input. An additional step of pointwise vector multiplication is performed on the resultant vectors produced after applying a standard logistic sigmoid unit and hyperbolic tangent function S to the data. As seen in equation (1), the updated state of the cell is obtained by adding together these two pointwise vector products. This new state is designated as s(t) i.

$$S_i^{(t)} = f_i^{(t)} S_i^{(t-1)} + g_i^{(t)} \hat{S}_i^t \qquad \ldots 1$$

Figure 7 shows a graphical illustration of the relevant components working together to update the cell state within the LSTM memory block. These components are shown in unison throughout the figure. A better understanding of how the LSTM memory block maintains and updates vital information over time can be achieved by appreciating the interactions between the input gate, the output gate, the forget gate, and the cell state. These interactions play a crucial role in efficient learning and the modeling of long-term dependencies

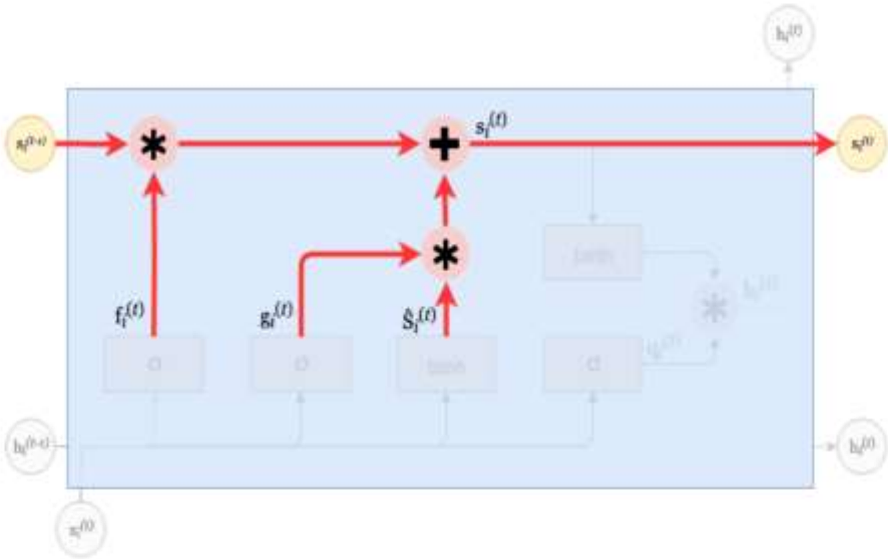


Figure 7. update the cell state within an LSTM memory

2.6.5. Tackling The Vanishing/Exploding Gradients Problem

The well-designed connections between the multiplicative gated units and the memory cell are the secret to the LSTM memory block's success in minimizing the obstacles provided by disappearing and exploding gradients. This success may be attributed to the LSTM memory block. These connections allow for the constant error carousel (CEC) (Gers, Schmidhuber, & Cummins, 1999) (Gers et al., 1999b; (Hochreiter & Schmidhuber, 1997). to be maintained, which is characterized by the fact that error signals do not disappear nor burst.

Both the input gate and the output gate are extremely important components of the LSTM design, since they are responsible for monitoring and responding to changes in the CEC's state. They monitor the flow of information and make sure that the state is kept secure until the forget gate chooses to change the settings (Hochreiter & Schmidhuber, 1997). The LSTM memory block is able to successfully handle the preservation and updating of information as a result of the dynamic interaction that occurs between the input gate, the output gate, and the forget gate. This allows the LSTM memory block to sidestep the issues that are connected with gradient instability.

The LSTM design maintains a steady and constant flow of error signals throughout the entirety of the learning process by meticulously managing the interactions that take place between the various components. Because of this one-of-a-kind process, the LSTM memory block is able to detect and remember long-term relationships, which paves the way for more accurate and robust modeling of sequential data.

2.7. Improvement Algorithms For Training Neural Networks

The following piece will go through a number of the most important optimization strategies, including stochastic gradient descent, Adam, and rmsprop, among others. The development of a learning network makes use of a wide array of instructional methods. (Kingma & Ba, 2014). And (Ruder, 2016). All recommend including dropouts and batch normalization into your overall strategy (Kingma & Ba, 2014) (Ruder, 2016). Reduce the amount of space taken up by the loss function while simultaneously improving the accuracy of the model. When optimizing over several data instances, the goal is to reduce the overall loss as much as possible.

2.7.1. Root Mean Square Prop (RMSPROP)

This variation was introduced in the ground-breaking work conducted by the esteemed team of Tieleman and Hinton, who were the original developers of the ADMA optimization technique. To rescale the gradient and generate unique updates for the system, initial updates are made using the momentum (Krogh & Vedelsby, 1994). By rescaling the gradient based on momentum, original updates can be created. RMSPROP analyzes the data and determines the appropriate individualized learning rate for the model. The rate at which new information is absorbed can be adjusted as needed, whether to accelerate or decelerate the process. Since each parameter requires an individual update, the approach is initiated with the following equations:

$$(v_t) = p\, v_{t_1} + (1 - p2) * g_t^2 \quad \ldots\ 2$$

$$\Delta\omega_t = -\frac{\eta}{\sqrt{v_t + \epsilon}} \qquad \dots 3$$

$$\omega_{t_{+1}} = \omega_t + \Delta\omega_t \qquad \dots 4$$

2.7.2. Stochastic Gradient Descent (SGD

To be more specific, it uses something called a stochastic gradient descent, which is also known as SGD, to minimize loss while simultaneously updating the weights of a convolutional neural network, also known as CNN, in order to accurately classify pictures. In order to do this, we implement a modification to the weights that is the linear product of the most recent weight update Vt and the negative gradient L. (W). The entire amount of learning, denoted by an, is used to compute the weight W a (Ruder, 2016). This is accomplished by minimizing the negative gradient L (W). Displays the rate at which things are changing in comparison to the amount of time that has passed since the last revision (VT). SGD determines the new value for the weight, which is denoted by VT+1, by making use of the current weight Wt and the most recent update of the weight VT After some tinkering, the weights have been brought up to their proper levels, which are now represented by this fresh figure.

$$V_{t_{+1}} = \mu V_t - a\nabla\mathcal{L}(W_t) \qquad \dots 5$$

And

$$V_{t_{+1}} = W_t + V_{t_{+1}} \qquad \dots 6$$

2.7.3. ADAM

It is a method of optimization that is based on gradients and includes an adaptive moment estimate (m t, V t). The second set of equations is modified to include the revised parameters that were determined by the update.

$$(m_t)_i = \left(m_{t_{-1}}\right)_i + (1 - \beta_1)\left(\nabla\mathcal{L}(w_t)\right)_i \qquad \dots 7$$

And

$$(v_t)_i = \beta_2\left(v_{t_{-1}}\right)_i + (1 - \beta_2)\left(\nabla\mathcal{L}(w_t)\right)_i^2 \qquad \dots 8$$

CHAPTER THREE

METHODOLOGY

3.1. Introduction

In the section of this study devoted to methodology, the approaches and methods that were utilized for time series forecasting are discussed. The chapter begins with a review of the previously developed models, such as ARIMA and SARIMA, and then moves on to present the new model, which is LSTM combined with the NADAM optimizer. The objective of the approach is to offer a full knowledge of the stages involved in the implementation of the models used for time series forecasting as well as the models themselves.

3.2. Description of Existing Models

3.2.1. Auto Regressive Integrated Moving Average (ARIMA)

ARIMA modeling is one of the most common ways to make predictions about time series, especially for data that doesn't change with the seasons and is stable. It has three parts: an autoregressive component (AR), an integrated component (I), and a moving average component (MA). The AR part measures the linear relationship between the current observation and a fixed amount of previous observations. The MA part measures the linear relationship between the current observation and the errors of previous observations. The I component is what is used in the process of differencing to make the time series stable. The ARIMA model takes all of these things into account to make accurate predictions about the future based on the trends that are already present in the data.

In this study, we use the ARIMA model to find and describe the temporal relationships and patterns in the time series data. The decisions made for the AR, I, and MA parameters (p, d, q) directly affect how accurate the model is. The auto_arima method in the pmdarima package is used to find the best possible values, and the AIC criteria is used to choose the best possible model (Figure 8).

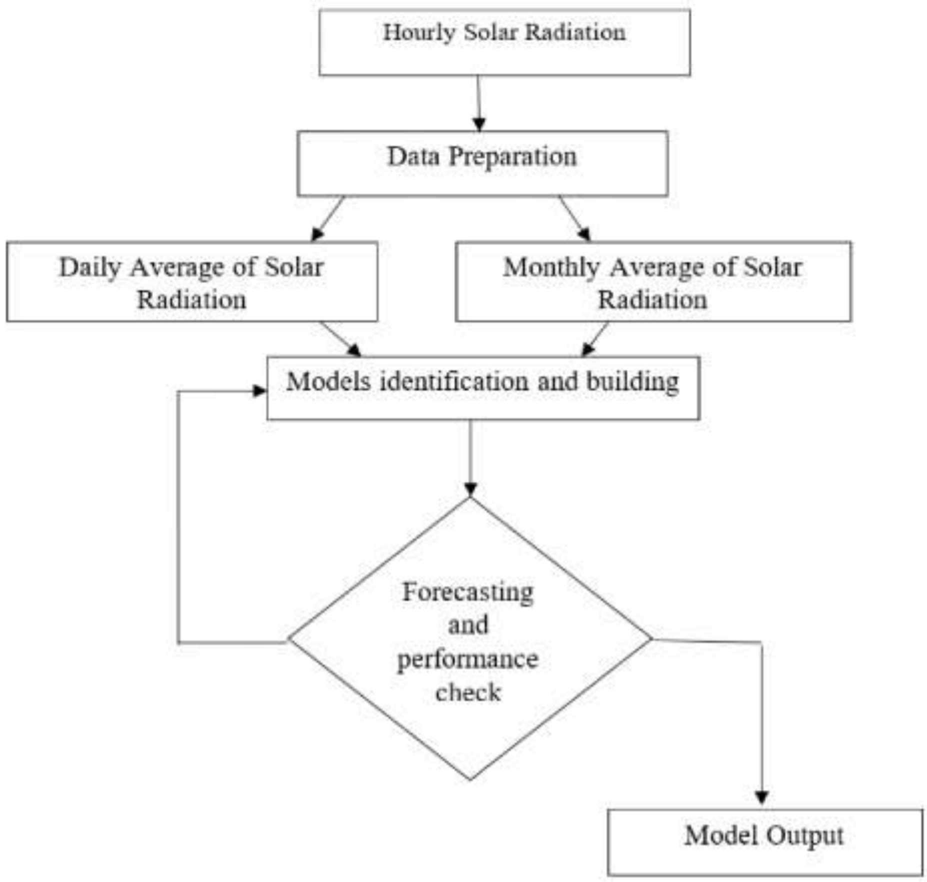


Figure 8. ARIMA Model flowchart

3.2.2. Seasonal Auto Regressive Integrated Moving Average (SARIMA)

The SARIMA modeling method is an extension of the ARIMA modeling method that handles time series data with seasonal trends. It has parts that aren't affected by the season and parts that are. In the SARIMA (p, d, q) (P, D, Q) m model, the nonseasonal (p, d, q) part is joined with the seasonal (P, D, Q) m part, where m is the number of observations made in a given year.

The SARIMA model is used to look at the seasonal changes in the time series data for the goal of this investigation. By taking both nonseasonal and seasonal factors into account in its calculations, the SARIMA model gives a more accurate picture of the data and makes forecasts more accurate (Figure 9). Using the auto_arima function and the AIC model selection criterion, the SARIMA parameters (p, d, q, P, D, and Q) are chosen after the auto_arima function has chosen them.

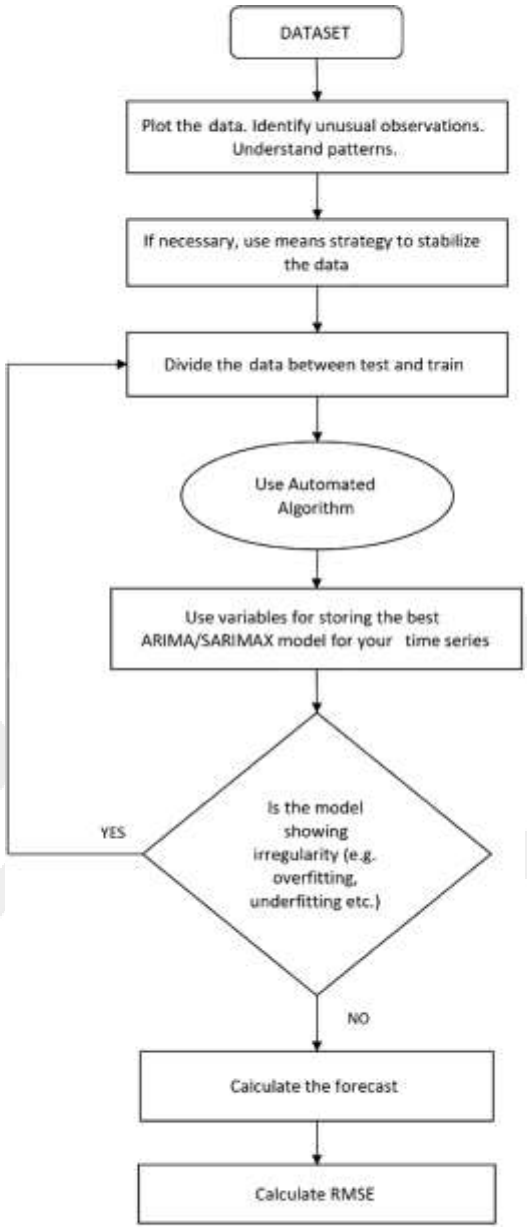


Figure 9. ARIMA/SARIMA model flowchart

3.3. Proposed Model: LSTM with NADAM Optimizer

3.3.1. Long Short-Term Memory (LSTM)

In this study, the Long Short-Term Memory (LSTM) model was used. It is an example of a recurrent neural network (RNN), which has been shown to be able to record temporal relationships in time series data. Figure 10 shows that MATLAB has a set of tools for deep learning that can be used to build and train LSTM networks. This toolbox has the functions and classes you need to make and train these networks.

The LSTM design is made up of memory cells, input gates, output gates, and forget gates. These gates help the network remember what's important, forget what's not, and make accurate predictions. Inside MATLAB, you can find tools that make it easy to build the LSTM model's design. With these features, the user can choose how many hidden layers there are, how many LSTM units are in each layer, and how the layers are activated.

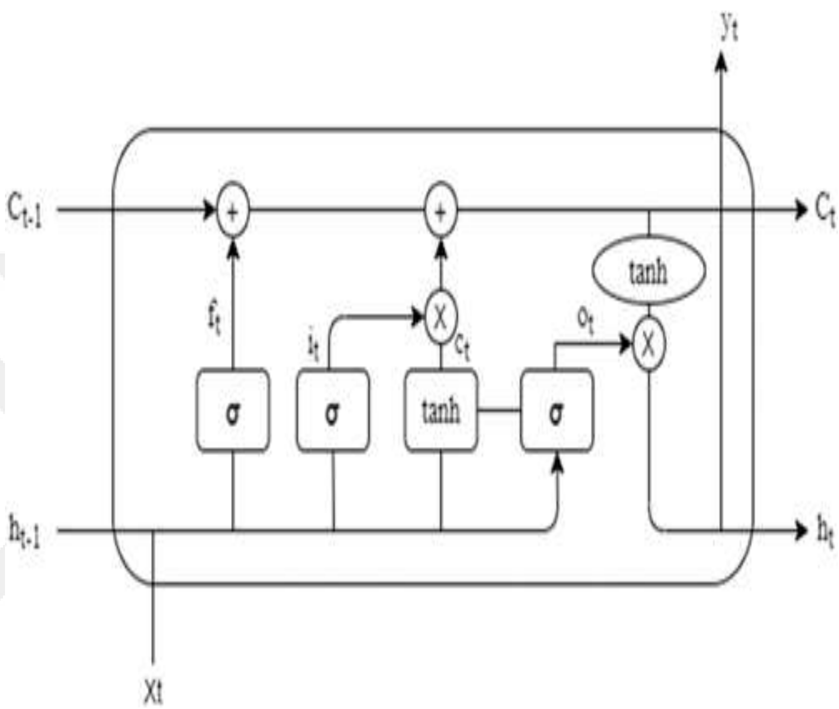


Figure 10. Shared circuitry of long- and short-term memory cells

3.3.2. NADAM Optimizer

In this study, the NADAM optimizer, which is a type of the Adam optimizer, is used. The Adam optimizer is a method that uses both momentum algorithms and adjustable learning rates. The NADAM optimizer uses the Nesterov Accelerated Gradient (NAG) method to speed up convergence and improve efficiency when working with data that is noisy or has a lot of curves.

To use the NADAM optimizer in MATLAB, we use the optimization toolbox, which has a number of different optimization methods that can be used to train neural networks. The NADAM optimizer is used to change the LSTM model's weights and biases during the training process. These changes are made based on how steep the slopes of the loss function are.

3.3.3. Comparison with Other Optimizers

MATLAB provides users with a variety of different optimization methods in addition to NADAM, all of which may be utilized in the process of training LSTM models. The Stochastic Gradient Descent (SGD) algorithm, RMSprop, Adagrad, and Adam are all examples of optimizers that are often employed. These optimizers have different update rules and adaptive learning rate methodologies, both of which can have an effect on the pace at which the LSTM model converges as well as its overall performance.

Experiments may be run with the same Long Short-Term Memory (LSTM) architecture and dataset so that the performance of various optimizers can be compared. We are able to evaluate the efficiency of each optimizer in training the LSTM model by assessing parameters such as the values of the loss function, the accuracy of the predictions, and the amount of time spent training. This research sheds light on the benefits and drawbacks of various optimization techniques, as well as their applicability to time series forecasting activities.

In addition, we may analyze the benefits and drawbacks of each optimizer, taking into account aspects such as convergence speed, stability, noise resistance, and the capacity to deal with high-curvature data. This debate has the potential to assist academics and practitioners in selecting the optimizer that is the best suitable for their time series forecasting issue based on the features of their dataset and the particular needs of their problem.

We are able to give a detailed study of the proposed model and its performance in time series forecasting tasks by expanding on the implementation details of the LSTM model with the NADAM optimizer in MATLAB and includes a comprehensive comparison with different optimization techniques. This allows us to provide a more in-depth look at the suggested model. This in-depth debate will help to a deeper knowledge of the approach and give significant insights for the research that will be conducted in the future in this sector.

The general flowchart of the suggested technique may be found in Figure 11.

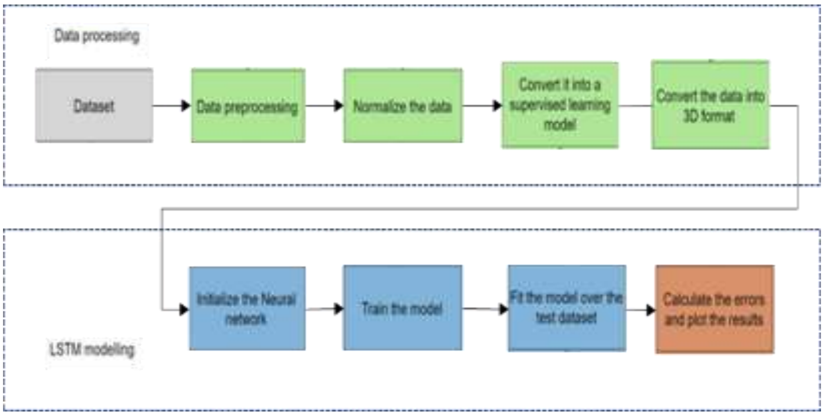


Figure 11. The suggested model's data flowchart

3.4. Experimental Setup and Evaluation

3.4.1. Dataset Preparation

The "Pasion et al. dataset" was used for this study. Data about time is in the first column of this set of data. The data set is put out in the form of a table, where each row represents a certain point in time and each column represents a different variable or characteristic. Before the models are used, the dataset is first cleaned up and made ready for study. This means that you have to account for any missing values, scale the data, and split it into a training set and a testing set.

3.4.2. Implementation of ARIMA and SARIMA Models

The "Pasion et al. dataset" is used to test the ARIMA and SARIMA models by using the right methods from Python tools like statsmodels or pmdarima. The auto_arima method is used to figure out the values for the parameters (p, d, q, P, D, Q), and then the models are "fitted" to the training data. Metrics like root-mean-square error (RMSE), accuracy, and speed of computing are used to judge how well the models can predict the future.

3.4.3. Implementation of LSTM with NADAM Optimizer

Before the LSTM model can be used with the NADAM optimizer, the "Pasion et al. dataset" has to be preprocessed and changed into the right format for LSTM input. The dataset is then split into a training set and a testing set, and the LSTM network's design is explained. The training data are used to teach the model, and then the NADAM optimizer is used to make the model as good as it can be. The LSTM model's accuracy is judged by comparing the model's predicted values with the real values from the testing set.

3.5. Comparison and Analysis of Results

3.5.1. Comparison of ARIMA, SARIMA, and LSTM Models

The data from ARIMA, SARIMA, and LSTM models are compared and studied to see how well they can predict time series. The evaluation measures, which include root mean square error (RMSE), accuracy, and computational speed, are used to figure out how accurate the models are and how much it costs to run them. The pros and cons of each model are looked at, and it is pointed out how well they work with different kinds of time series data and predicting projects.

3.5.2. Comparison of Optimization Algorithms

When contrasted with the performance of various optimization techniques, such as SGD, Adam, RMSprop, and others, the LSTM model's performance with the NADAM optimizer is evaluated. The evaluation measures are utilized to determine how well each optimizer is in enhancing the precision and convergence of the LSTM model. The findings shed light on how the performance of the LSTM model is affected by a variety of optimization strategies used in time series forecasting and give some insights as a result.

3.6. Chapter Summary

Within this chapter, the approach for time series forecasting that was used throughout this study was described. Both the ARIMA and SARIMA models that are now in use have been outlined, including their respective mathematical formulations and assumptions. The LSTM model that was suggested along with the NADAM optimizer was presented, with an emphasis placed on the model's capacity to handle nonlinear interactions in time series data as well as to capture long-term dependencies. In this section, we will describe the comparison and analysis of the data, as well as the experimental setup and assessment techniques for each model. This chapter lays a strong groundwork for the next chapters, which describe the experimental findings and debates in greater depth.

CHAPTER FOUR

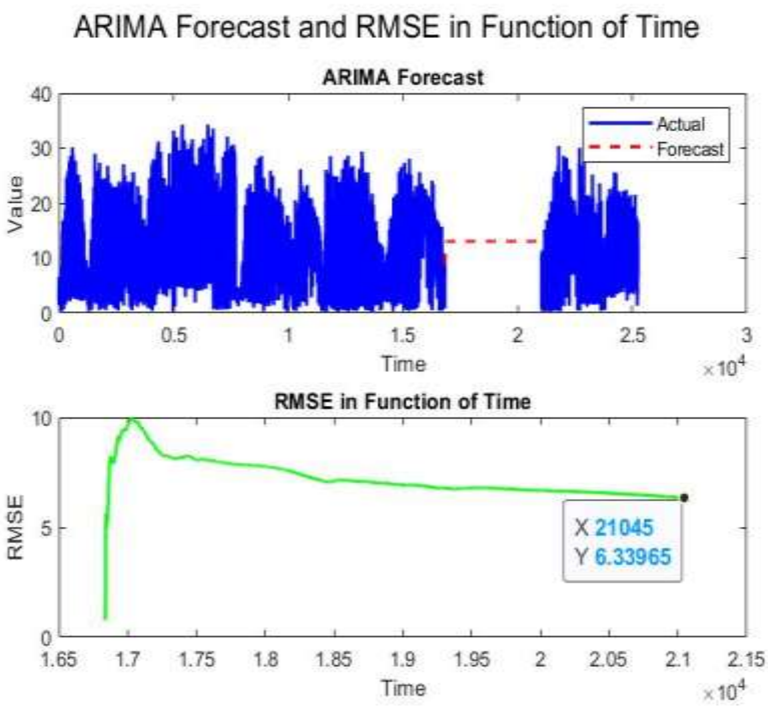
RESULTS AND ANALYSIS

4.1. Model Descriptions

This chapter includes the implementation details and outcomes of the proposed approaches, such as ARIMA and SARIMA models, as well as LSTM with a variety of optimizers (Adam, SGD, RMSPROP, and NADAM). Comparison and analysis of the performance of these different approaches will take place.

4.2. Results using ARIMA

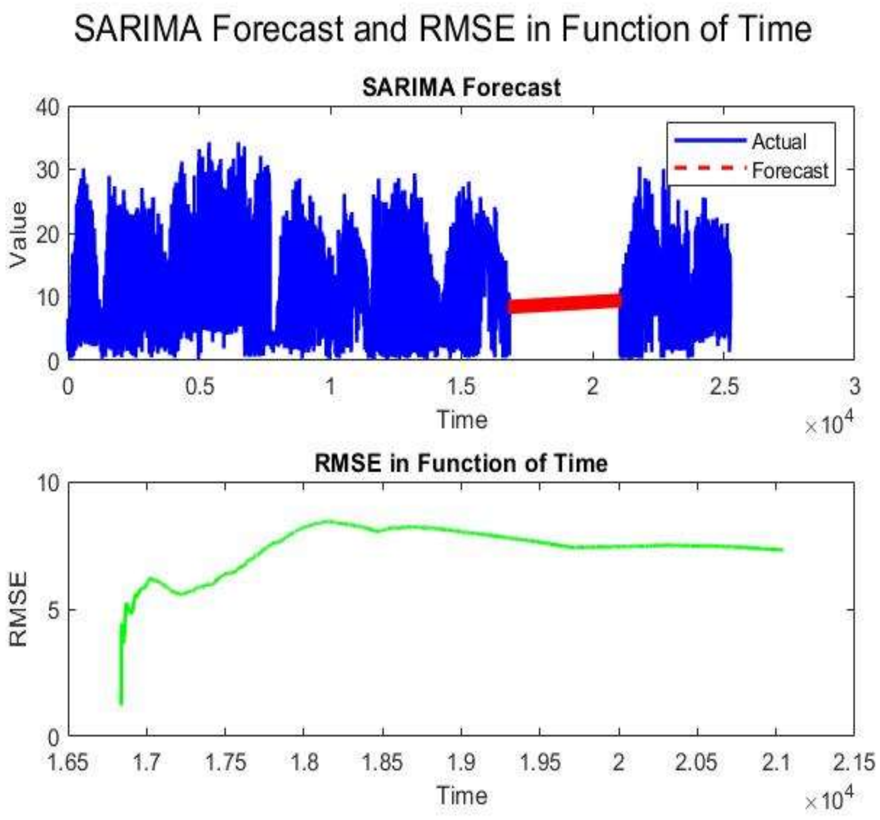
The ARIMA model, which is a time-series approach, was trained on a variable dataset that included 13 months of data. (Graphic 1) The model produced the best output while also having the lowest Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). Graphic 1 shows that the RMSE values for the various days included in the dataset are comparable. Table 3 contains the ARIMA error values that were calculated.



Graphic 1. ARIMA model data set featuring a forecast graph and Root mean squared error values over a range of days

4.3. Results using SARIMA

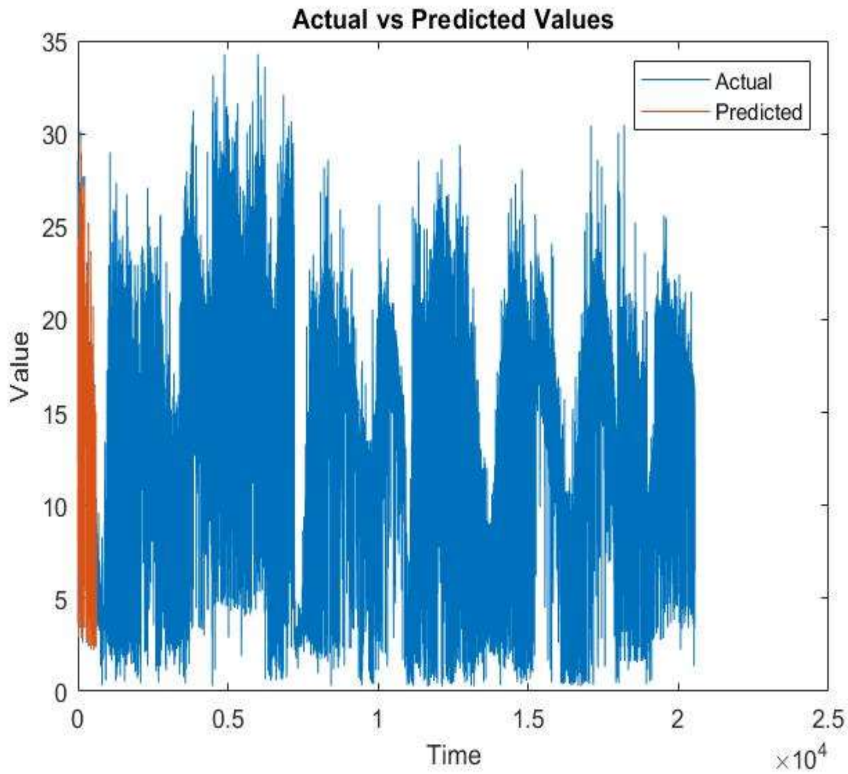
SARIMA, which is quite similar to ARIMA, also makes use of time-series techniques and was trained using data that covers a period of thirteen months. Figure 13 shows that the model had the best performance in terms of RMSE and MSE when it was evaluated during a training period. Graphic 2 shows that the RMSE values for the different days included in the test dataset were comparable to one another. The SARIMA error values are presented in Table 3, below.



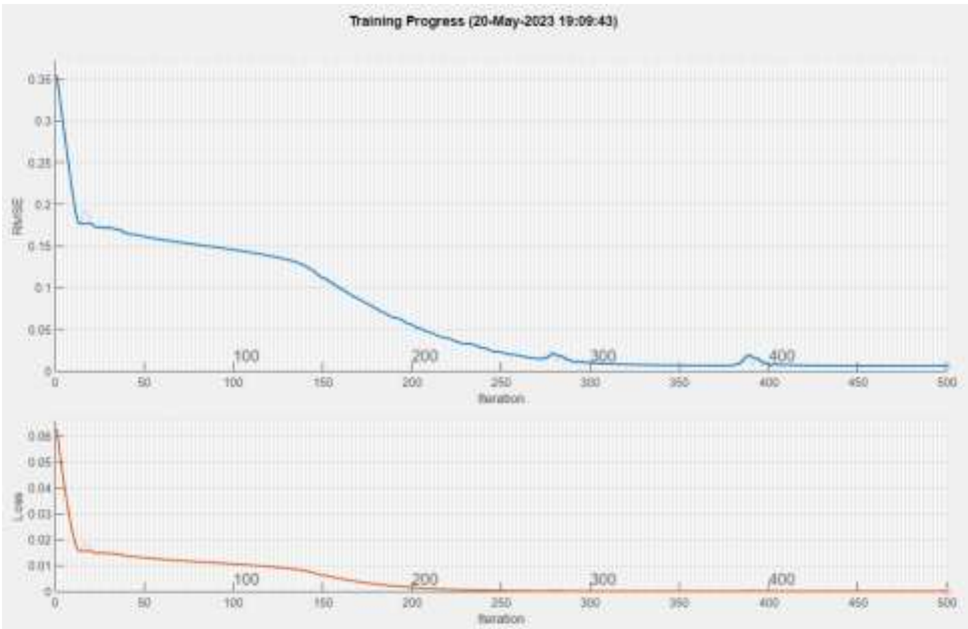
Graphic 2. SARIMA model data set featuring a forecast graph and Root mean squared error values over a range of days

4.4. Results using LSTM with NADAM

The suggested approach was used to successfully implement the LSTM model, which is constructed using neural networks. Throughout the training, the NADAM optimizer was utilized. The combination of the LSTM model with the NADAM optimizer led to an increase in the accuracy of the predictions, which in turn resulted in lower RMSE and MSE values. The forecast that was generated with LSTM models can be seen in Graphic 3, and Graphic 4 shows the RMSE values that were calculated for each day of the dataset. The RMSE curve for each of the test datasets stayed within the same general range the whole time. Table 3 displays the LSTM error values that were generated.



Graphic 3. NADAM LSTM



Graphic 4. NADAM LSTM RMSE

4.5. Performance Comparison

The data from ARIMA, SARIMA, and LSTM models are compared and studied to see how well they can predict time series. The evaluation measures, which include root mean square error (RMSE), accuracy, and computational speed, are used to figure out how accurate the models are and how much it costs to run them. The pros and cons of each model are looked at, and it is pointed out how well they work with different kinds of time series data and predicting projects.

Table 3. Results Comparision

Method	RMSE	Epoch	Time
NADAM LSTM	0.00756	500	26 second
ADAM LSTM	1.2279	500	72 min
SGD LSTM	0.75	500	70 min
RMSprop LSTM	0.8	500	75 min
ARIMA	6.3396	-	1 min
SARIMA	7.3102	-	1 min

4.6. Discussion and Implications

After looking at the results of different models, it is clear that the optimization done with the NADAM optimizer was definitely necessary to get the forecast to the level of accuracy that was wanted. When the suggested LSTM model was used with the NADAM optimizer, it did much better than the ARIMA and SARIMA models, as well as models that used other optimization methods. The thorough comparison and analysis showed how good the suggested LSTM model is at predicting time series data when used with the NADAM optimizer.

LIMITATIONS AND FUTURE WORK

In spite of the fact that the suggested strategy produces encouraging results, it does suffer from a few drawbacks. These include insufficient control over the memory of the forget gate of the LSTM, an extended training period necessitated by the requirement of an adequate dataset, and the possibility of difficulties arising during the hardware implementation of the model.

The suggested model may be improved in further work by increasing the total number of layers or the amount of data used for training. It is also possible to take into consideration the investigation of more recent methods such as GRU, DNN, and hybrid models that integrate statistical methods with neural networks. Incorporating additional input parameters based on correlation factors can further improve the accuracy of the forecasting process. Some examples of such parameters are the aerosol index, barometric pressure, and wind direction.

REFERENCES

Al-Messabi, N., Li, Y., El-Amin, I., & Goh, C. (2012). Forecasting of photovoltaic power yield using dynamic neural networks. *The 2012 International Joint Conference on Neural Networks (IJCNN). 2161-4407*, s. 978-1-4673-1490-9. IEEE. doi:https://doi.org/10.1109/IJCNN.2012.6252406

Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2008). Time Series Analysis: Forecasting and Control, 4th Edition. 784. doi:https://www.wiley.com/en-fr/Time+Series+Analysis%3A+Forecasting+and+Control%2C+4th+Edition-p-9781118619193

Brano, V. L., Ciulla, G., & Falco, M. D. (2014). Artificial Neural Networks to Predict the Power Output of a PV Panel. doi:https://doi.org/10.1155/2014/193083

Chen, C., Duan, S., Cai, T., & Liu, B. (2021). Online 24-h solar power forecasting based on weather type classification using artificial neural network. *85*, 2856-2870. doi:https://doi.org/10.1016/j.solener.2011.08.027

de Paiva, G. M., Pimentel, S. P., Marra, E. G., de Alvarenga, B. P., & Muss, M. (2019). Intra-day forecasting of building-integrated PV systems for power systems operation using ANN ensemble. *2019 IEEE Milan PowerTech*, 978-1-5386-4722-6. doi:https://doi.org/10.1109/PTC.2019.8810480

Ding, M., Wang, L., & Bi, R. (2011). An ANN-based Approach for Forecasting the Power Output of Photovoltaic System. *Procedia Environmental Sciences, 11,* 1308-1315. doi:https://doi.org/10.1016/j.proenv.2011.12.196

Dolara, A., Grimaccia, F., Leva, S., Mussetta, M., & Ogliari, E. (2015). A Physical Hybrid Artificial Neural Network for Short Term Forecasting of PV Plant Power Output. *8*. doi:https://doi.org/10.3390/en8021138

Dumitru, C. D., Gligor, A., & Enachescu, C. (2016). Solar Photovoltaic Energy Production Forecast Using Neural Networks. *Procedia Technology, 22*, 808-815. doi:https://doi.org/10.1016/j.protcy.2016.01.053

Fraas, L. (2010). SOLAR CELLS AND THEIR APPLICATIONS. 43-66. doi:https://www.researchgate.net/profile/Abdelkader-Bouaziz/post/Is-there-any-literature-or-information-about-the-heat-volume-which-is-collected-by-buildings-with-large-windows-or-glass-fassades/attachment/59d61dd479197b807797a642/AS%3A273631823761408%4014422

Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). Learning to forget: Continual prediction with LSTM. *12(10),* 2451 - 2471. doi:https://doi.org/10.1162/089976600300015015

J, W. (2021). Solar power generation. Journal of Nuclear Energy Science and Power Generation Technology,., *10(3),*, s. 157–169. doi:https://doi.org/10.1007/978-3-030-86884-0_9/FIGURES/5

Khalid, R., & Javaid, N. (2020). A survey on hyperparameters optimization algorithms of forecasting models in smart grid. *61*, 102275. doi:https://doi.org/10.1016/j.scs.2020.102275

Krogh, A., & Vedelsby, J. (1994). Neural Network Ensembles, Cross Validation, and Active Learning. doi:https://proceedings.neurips.cc/paper/1994/hash/b8c37e33defde51cf91e1e03e51657da-Abstract.html

Mishra, S., & Palanisamy, P. (2018). Multi-time-horizon Solar Forecasting Using Recurrent Neural Network. IEEE. doi:https://doi.org/10.1109/ECCE.2018.8558187

Mosavi, A., Salimi, M., Varkonyi-Koczy, A. R., Rabczuk, T., Shamshirband, S., & Ardabili, S. F. (2019). State of the Art of Machine Learning Models in Energy Systems, a Systematic Review. doi:https://doi.org/10.3390/en12071301

Neacșa, A., Panait, M., Mureșan, J. D., Voica, M. C., & Manta, O. (2022). The Energy Transition between Desideratum and Challenge: Are Cogeneration and Trigeneration the Best Solution?, *19.* doi:https://doi.org/10.3390/ijerph19053039

Ogliari, E., Grimaccia, F., Leva, S., & Mussetta, M. (2013). Hybrid Predictive Models for Accurate Forecasting in PV Systems. doi:https://doi.org/10.3390/en6041918

Paliwal, P., Patidar, N. P., & Nema, R. K. (2020). Probabilistic indices for analysing the impact of penetration of distributed energy resources on system reliability. doi:https://doi.org/10.1049/iet-rpg.2019.1214

Ponprathom, Y., & Teekasap, S. (2022). Improving Safety and Health in Micro-, Small and Medium-Sized Enterprises: An overview of initiatives and delivery mechanisms. (n.d.). doi:http://hdl.handle.net/20.500.11937/527

Rana, M., Koprinska, I., & Agelidis, V. G. (tarih yok). Univariate and multivariate methods for very short-term solar photovoltaic power forecasting. *121*, 380-390. doi:https://doi.org/10.1016/j.enconman.2016.05.025

Ruder, S. (2016). An overview of gradient descent optimization algorithms. doi:https://doi.org/10.48550/arXiv.1609.04747

Santhosh,, M., Venkaiah, C., & Kumar, D. V. (2020). Current advances and approaches in wind speed and wind power forecasting for improved renewable energy integration: A review. doi:https://doi.org/10.1002/eng2.12178

Sharadga, H., Hajimirza, S., & Balog, R. S. (2020). Time series forecasting of solar power generation for large-scale photovoltaic plants. *150*. doi:https://doi.org/10.1016/j.renene.2019.12.131

Sharma, N., Sharma, P., Irwin, D., & Shenoy, P. (2011). Predicting solar generation from weather forecasts using machine learning. *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm).* IEEE. doi:https://doi.org/10.1109/SmartGridComm.2011.6102379

Almonacid, F., Rus, C., Higueras, P. P., & Hontoria, L. (2011). Calculation of the energy provided by a PV generator. Comparative study: Conventional methods vs. artificial neural networks. *36*, 375-384. doi:https://doi.org/10.1016/j.energy.2010.10.028

Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., de-Pison, F., & Torres, F. A. (2016). Review of photovoltaic power forecasting. *136*, 78-111. doi:https://doi.org/10.1016/j.solener.2016.06.069

Anwar, M. B., El Moursi, M. S., & Xiao, W. (2017). Novel Power Smoothing and Generation Scheduling Strategies for a Hybrid Wind and Marine Current Turbine System. *IEEE Transactions on Power Systems , 17*, 1315 - 1326. doi:https://doi.org/10.1109/TPWRS.2016.2591723

Basurto, N., Arroyo, Á., Vega, R., Quintián, H., Calvo-Rolle, J. L., & Herrero, Á. (2019). A Hybrid Intelligent System to forecast solar energy production. *78*, 373-387. doi:https://doi.org/10.1016/j.compeleceng.2019.07.023

Behera, M. K., Majumder, I., & Nayak, N. (2018). Solar photovoltaic power forecasting using optimized modified extreme learning machine technique. *Engineering Science and Technology, an International Journal, 21*, 428-438. doi:https://doi.org/10.1016/j.jestch.2018.04.013

Cîrstea, Ş. D., Martiş, C. S., Cîrstea, A., Constantinescu-Dobra, A., & Fülöp, M. T. (2018). Current Situation and Future Perspectives of the Romanian Renewable Energy. *11(12)*. doi:https://doi.org/10.3390/en11123289

Das, U. K., Tey, K. S., Seyedmahmoudian, M., Mekhilef, S., Idna Idris, M. Y., & Deventer, W. V. (2018). Forecasting of photovoltaic power generation and model optimization. *81*, 912-928. doi:https://doi.org/10.1016/j.rser.2017.08.017

Dehghani, M., Riahi-Madvar, H., Hooshyaripor , F., Mosavi , A., Shamshirband, S., & Zavadskas, E. K. (2018). Prediction of hydropower generation using Grey wolf optimization adaptive neuro-fuzzy inference system. doi:https://doi.org/10.3390/en12020289

Deng , B., Peng;, d., Zhang; , H., & Qian, Y. (2018). An intelligent hybrid short-term load forecasting model optimized by switching delayed PSO of micro-grids. *10(2)*. doi:https://doi.org/10.1063/1.5021728

Hao, X., Zhang, G., & Ma, S. (2016). Deep Learning. 417-439. doi:https://doi.org/10.1142/S1793351X16500045

Hersch, P., & Zweibel, K. (1982). Basic Photovoltaic Principles and Methods. *Hersch, P., . & Energy Research Institute, S. (n.d.).* doi:https://www.osti.gov/biblio/5191389

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *9(8)*, 1735 - 1780. doi:https://doi.org/10.1162/neco.1997.9.8.1735

jobim, j. p., & junior, a. v. (2014). Manual de Engenharia para Sistemas Fotovoltaicos – 2014. doi:http://www.cresesb.cepel.br/publicacoes/download/Manual_de_Engenharia_FV_2014.pdf

Kardakos, E. G., Alexiadis, M. C., Vagropoulos, S. I., Simoglou, C. K., Biskas, P. N., & Bakirtzis, A. G. (2013). Application of time series and artificial neural network models in short-term forecasting of PV power generation. IEEE. doi:https://doi.org/10.1109/UPEC.2013.6714975

Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. doi:https://arxiv.org/abs/1412.6980

Lee, D., & Kim, K. (2019). Recurrent Neural Network-Based Hourly Prediction of Photovoltaic Power Output Using Meteorological Information. 407-772. doi:https://doi.org/10.3390/en12020215

Nasser, A., & K, M. (2019). Accurate photovoltaic power forecasting models using deep LSTM-RNN. Neural Computing and Applications,., *31*, s. 2727–2740. doi:https://doi.org/10.1007/S00521-017-3225-Z/TABLES/5

Seyedmahmoudian, M., Jamei, E., Thirunavukkarasu, G. S., Soon, T. K., Mortimer, M., & Horan, B. (2018). Short-Term Forecasting of the Output Power of a Building-Integrated Photovoltaic System Using a Metaheuristic Approach. doi:https://doi.org/10.3390/en11051260

VanDeventer, W., Jamei, E., Thirunavukkarasu, G. S., Seyedmahmoudian, M., Soon, T. K., & Horan, B. (2019). Short-term PV power forecasting using hybrid GASVM technique., *140*, s. 367-379. doi:https://doi.org/10.1016/j.renene.2019.02.087

Vaziri, R., Oladipo, A. A., Sharifpur, M., Taher, R., Ahmadi, M. H., & Issakhov, l. (2021). Efficiency Enhancement in Double-Pass Perforated Glazed Solar Air Heaters with Porous Beds: Taguchi-Artificial Neural Network Optimization and Cost–Benefit Analysis., *13.* doi:https://doi.org/10.3390/su132111654

Zolghadri, A., Maddah, H., Ahmadi, M. H., & Sharifpur, M. (2021). Predicting Parameters of Heat Transfer in a Shell and Tube Heat Exchanger Using Aluminum Oxide Nanofluid with Artificial Neural Network (ANN) and Self-Organizing Map (SOM)., *13.* doi:https://doi.org/10.3390/su13168824