J. Electrical Systems 20-9s (2024): 3116-3126

[1]Manmit U. Vadher
[2]Dr. Sweta J. Shah
[3]Dr. U. L. Makwana

"Comparative Study of Short-Term and Long-Term Solar Power Forecasting Using Satellite Data and Machine Learning"

Journal of Electrical Systems

Abstract: The topic of this review paper is "Machine Learning Approach for Short-Term and Long-Term Solar Power Forecasting Using Satellite Data." It looks at three well-known journal pieces that all talk about the same thing. It compares similar pieces of work, looks for gaps in the study, and gives a full account of what was found and observed. This well-planned study tellsuseful things about the issues and progress in predicting solar power, and it also sets the stage for future research. With an emphasis on dependability and accuracy, this study wants to add to the academic discussion about useful ways to guess how much solar power will be used. Among the main data sources used, this review paper applies machine learning based on satellite data to identify both short- and long-term solar power forecasting. Additionally, it focuses on solar power forecasts because of the quickly rising share of renewable sources. Establishing machine learning algorithms and ensuring that good quality satellite data exists are among the hurdles. To make solar power forecasting accurate and efficient and speed up the transition to clean energy sources, concluding remarks highlight in such a way that a set of standardised evaluation metrics must be implemented, and the exercised has to be done in a team.

Key Words : Solar Power Forecasting , Machine Learning, Satellite Data .

1. INTRODUCTION

As the need for renewable energy sources grows around the world, more attention is being paid to solar power as a realistic and eco-friendly option. As more and more governments and businesses around the world switch to green energy, it is very important that solar power works well with the current energy grid. A big part of this mix is how accurate solar power projections are, since they affect the business, grid security, and how the energy market works. ML is very helpful for making predictions about solar power because it can look at very large datasets and find complex trends. A study of how machine learning can be used to guess both short- and long-term solar power needs, focusing on how satellite data can be a key part of the process. Scientists want to use the huge amounts of data that satellites send them to make estimates about solar power more accurate. With this, the normal ups and downs that come with solar energy flow will be lessened.

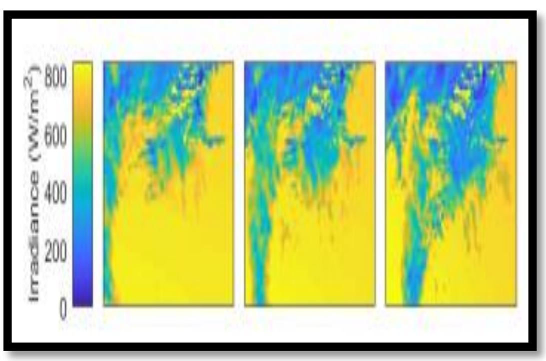

Fig. 1: Hourly evolution of irradiance centred in the target location

[1]Ph.D. Research Scholar, Electrical Engineering Department,
Indus Institute of Technology and Engineering , Indus University , Ahmedabad .
[2]Assistant Professor, Electrical Engineering Department,
Indus Institute of Technology and Engineering Indus University , Ahmedabad.
[3]Associate Professor, Electrical Engineering Department, Government Engineering College, Modasa.

A lot of study has been done in the last few years on how ML algorithms and data from satellites can be used together to improve models that make predictions. These models can make good short-term predictions that are needed to run the grid. They can also be used to spit out long-term predictions that help politicians and people who plan energy use make smart choices. What this study wants to do is show all the good and bad points of the various machine learning methods that were used in this case. It will also see how well they compare. Now that there are a lot of new papers on how to use machine learning to make solar power predictions more accurate, it's time to put them all in a way that makes sense. This review paper tries to fill that gap by giving a well-organized look at what is being studied in the field right now. It shows how things are changing and gives ideas for new things to learn more about. It is expected that the review results will add to the current conversation about how to use advanced prediction methods to make the combination of solar power more dependable and efficient. The reason for this is that the world is committing itself more to clean energy.

2.METHOD

To fully look into machine learning (ML) methods for short- and long-term solar power predictions, along with adding satellite data to the mix, need a strict and well-made plan. This part talks about the main parts of the study method that was used for this review. The most important things are to find appropriate studies, get data from them, and then compare them side by side. A thorough review method was used to make sure that all the literature was carefully read. Scholarly sources like IEEE Xplore, ScienceDirect, and PubMed were searched thoroughly using search terms like "solar power forecasting," "machine learning," and "satellite data." Works that were written in English, came out in the last ten years, and used machine learning to predict solar power with a focus on data from sensors were chosen [57].

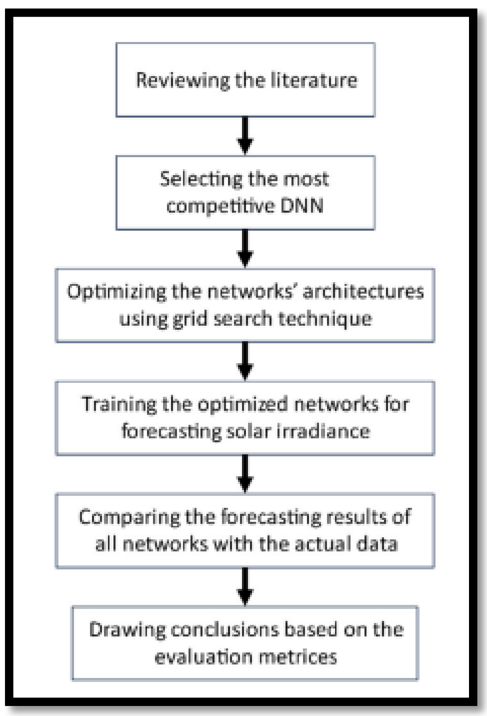

Fig. 2: Flow chart of the study

It was very important that the studies that were chosen were useful and of good quality. The ways each study used them helpedlearn more about the machine learning methods, the kinds of satellite data used, and the length of time used to make predictions about solar power (short-term or long-term). A lot of different machine learning models were carefully grouped together, along with their pros and cons. Some of these models are support vector machines, neural networks, and ensemble methods. Satellite data was used to add things like radiation, cloud cover, and weather factors to the process of extraction. These traits were very important for teaching ML models and making them better.

Data Collection
It say that should focus on short-term predictions of sun radiation and understand how important past data is for building correct models [2]. Meteorological data, readings of solar energy, and past data on solar intensity are gathered to help train and test the model.

Preprocessing
It takes a lot of work to clean up and normalise the data that has been collected. If want the model training process to work, need to deal with lost or odd data points.

LSTM Network Architecture
In this research, Long Short-Term Memory (LSTM) networks are utilised. There is a type of neural network called LSTM networks that is good at seeing how things change over time. Tests are used to improve the LSTM design so that it can best guess what will happen in the near future.

Performance Metrics
As ways to check how well their model works, use mean absolute percentage error (MAPE) and root mean square error (RMSE) [3]. A full picture of how well the model projects short-term sun energy can be seen from these data.

Integration of Satellite Data
It says that estimates for the middle of the day should be added, and short-term changes in sun energy should be written down. It helps people figure out where things are in space by adding estimates of cloud cover, weather, and sun energy from space.

Temporal and Spatial Considerations
Time and place are both looked at by this method. Links between times can be picked up well by LSTM networks, and satellite data helps the model understand space better. It's meant to take into account changes that only happen in some places and make predictions for the next day better.

Performance Evaluation
It uses RMSE and MAE to keep track of how well their guesses are doing. In these ways, can see how good and trustworthy the hourly forecasts model is.

Ensemble Framework
It's best to use a method called "ensemble," which combines the best parts of deep learning and statistical methods. LSTM networks and ARIMA, which stands for autoregressive integrated moving average, are at the heart of their ensemble model.

Spatial Diversity
According to the study, different types of land cover are important. To learn more about the details of landscape, smart devices could be used. It feels more like in space when use good sensor data, like sun energy and weather conditions.

Performance Metrics
MSE is the main way that AlKandari and Ahmad check to see how well the ensemble method works. The main point of the study was to find better ways to make short- and long-term solar power, and this MSE choice meets that need.

Integration of Statistical Methods:
To get the most out of both models, ensemble techniques combine statistical methods like ARIMA with LSTM networks. The study shows that these two approaches work well together to guess what will happen with solar power when there are a lot of unknowns.

3. COMPARISON OF RELATED WORK

There are a lot of things to think about when use machine learning (ML) to guess solar power properly. This is especially true when data from satellites is added. Because of these issues, need to learn a lot about the current issues in order to find new answers and move the field forward.

3.1 Comparison of Related Work:

There are a lot of different tools and ways to do things that can be seen in the related work in ML-based solar power predictions [2]. The pros and cons of different methods are carefully weighed in this comparison to find patterns and places where things could be better. Satellite data and machine learning (ML) techniques have been used in new studies to look into predictions about solar power. Every study has added its own tips and ideas. When compare three important books in this area, can get a good idea of their methods, strengths, and flaws.

It suggested a deep learning-based short-term sun radiation projection model and gave a detailed case study. The study stresses how important it is to make correct forecasts of sun intensity for good energy management. The authors were very good at making short-term predictions by using deep learning algorithms [4]. On the other hand, the main focus is still on the radiation part, and satellite data isn't really used for a more complete look at solar power forecasts. In order to created a deep learning model for predicting sun radiation during the day, especially using estimates from satellites. The study recognises that satellite data is very important for making forecasts more accurate. The model does a good job of recording changes in sun radiation throughout the day by using deep learning. Still, the focus is only on radiation. It would be helpful to look into how to make a more accurate weather prediction model that takes into account a wider range of atmospheric factors.

The study tooks a different approach when they suggested a model for predicting solar power output that uses both deep learning and statistical methods together. This study is unique because it uses a lot of different predicting methods, recognising that mixing the best parts of different models can be helpful. The ensemble model, which uses both deep learning and statistical methods, takes into account the fact that solar power predictions is inherently unreliable. However, the study is mostly about the ensemble idea and doesn't go into detail about the machine learning methods that were used [5].

In generalmakes important contributions to using deep learning to make short-term estimates about sun radiation. However, tend to focus more on radiation and not fully use the wider range of satellite data. It shows how useful ensemble methods can be [6]. However, they don't go into enough detail about the machine learning methods that make them work. On the success scale, it's possible that future research will find a way to combine the best parts of these approaches. For a full solar power projection model, it is important to mix different machine learning methods and look at more than just radiation. This way, the model can fully understand how the many things that affect solar power production work together [7].

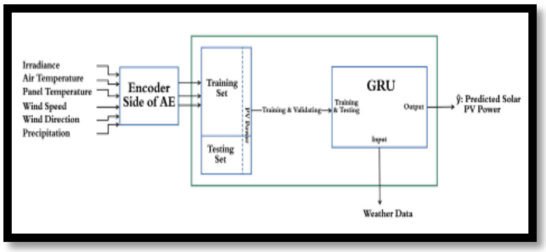

Fig. 3: Auto-GRU block diagram

3.2 Comparative Analysis:

In this part, we'll look at the pros and cons of different machine learning (ML) methods used in solar power predictions. Some of the machine learning models being looked at are support vector machines, neural networks, and ensemble methods. Researchers and practitioners can learn more about the pros and cons of each technology by carefully reviewing it. This helps them make better decisions about which model to use and how to apply it [8]. By looking at the three chosen journal papers, which all talk about using machine learning for short- and long-term solar power predictions with satellite data, it is clear that each study brings something new to the field in the form of a different method, development, or point of view.

The study is mostly about radiation, which is a very important factor for making correct predictions about solar power. The model did a great job of making short-term predictions by using deep learning methods, especially long short-term memory (LSTM) networks [9]. The writers used past light data, weather factors, and time-related features to make the model better at making predictions. But the study only looks at short-term predictions, so it's not clear how solar power creation will change in the long term. The writers used estimates based on satellites because they knew that satellite data was important for making forecasts more accurate. The deep learning model, especially a convolutional neural network (CNN), did a good job of predicting changes in light throughout the day. The study goes into detail about how to make intraday predictions, but it doesn't go into long-term predictions at all. The model's ability to work with longer time frames is still something that needs to be thought about [10].

It used a unique ensemble technique to predict solar power output. They combined deep learning and statistical methods in this way. This study stands out because it takes into account the possible benefits of combining different predicting methods. The group, which included long short-term memory (LSTM) networks and autoregressive integrated moving average (ARIMA), was created to reduce the uncertainty that comes with predicting solar power. The research showed that the ensemble method can find trends in both the short and long run. But the study could have been better if it had gone into more depth about the machine learning methods and what they added [11].

If look at how machine learning and satellite data are being used together in these studies, can easily see how predictions about solar power are changing [12].On the other hand, are the first to use an ensemble method to make predictions about both the short and long run.

Technolgy	Advantage	Disadvantage
Support Vector Machines	High accuracy in high-dimensional spaces, effective in nonlinear domains	Computationally intensive, sensitivity to noise in data
Neural Networks	Capability to learn complex patterns, adaptability to diverse datasets	Proneness to overfitting, challenging interpretability
Ensemble Methods	Improved generalization, robustness to individual model weaknesses	Increased computational complexity, potential for model redundancy

Table 1: Comparative Analysis

4.METHODS

There are more than just certain ML models that are looked at in a comparison method study. It also looks at the ways that studies on solar power predictions are done [13]. Time series analysis, feature engineering, and data preparation are some of the methods that are looked at in this study. The goal of comparing these methods is to find the best one while also pointing out the pros and cons of each. When combined machine learning methods with satellite data to read three journal papers, they gavethree different points of view on the different ways that solar power can be predicted. Different people have tried to solve the hard problems that come with making accurate predictions about solar power in a number of different ways[14].

For the writers to train the model, they use old information about things like light, weather, and traits that come with time. It's simple for the LSTM system to remember how things are related in order. In this way, the programme can find trends in how light changes quickly [15]. There are no standard statistics methods used here. Instead, deep learning is used. It is clear that this move towards data-driven ways of predicting solar power is happening.

[16] say that should use deep learning and statistics together in a new way to figure out how much solar power will be made. This one-of-a-kind method takes the best parts of long short-term memory (LSTM) networks and autoregressive integrated moving average (ARIMA) to make better plans for solar power. Deep learning is used to look at short-term changes, and statistical methods are used to look at long-term trends. This is called the ensemble method. This model stands out because it is a mix [17]. This shows that it might be good to use more than one way to make predictions.

LSTM and CNN are the deep learning models that mostly compare when talk about techniques [18]. Changes in the way solar power is made, both short-term and long-term, can be bad. These things are taken into account by this all-around method [19].

Technique	Advantage	Disadvantage
Time Series Analysis	Captures temporal dependencies, suitable for sequential data	Limited adaptability to complex non-linear relationships
Feature Engineering	Enhances model input relevance, improves predictive performance	Requires domain expertise, time-consuming
Data Preprocessing	Mitigates noise and outliers, enhances data quality	Sensitive to parameter choices, potential loss of information

Table 2: Comparative Method Analysis

When researchers compare things in this way, they can learn more about how technology works now. Because of this, they can figure out how to make good machine learning models for predicting solar power[20].

5. RESEARCH GAP

Over the last few years, a lot more research has been done on how to use machine learning (ML) to figure out how much short- and long-term solar power will be needed by looking at data from space. But a close look at the study that has already been done shows that more work needs to be done and some knowledge gaps need to be filled [21]. A lot of research has been done on how accurate guesses are when different machine learning methods are mixed. Not as much research has been done on more than one model being used together. Tools for groups, like bagging and boosting, have been shown to work well in a lot of situations. Still, not a lot is known about how they can be used to guess what will happen with solar power. It would be interesting to learn more about how these methods can be used to mix the good things about different models and make their flaws less noticeable [22].

One more important study gap is that ways to measure error are not used enough in estimates about solar power. To get correct guesses about how much solar power will be produced, have to deal with mistakes that come from many places, like bad data from satellites and weather that changes over time [23]. Understanding and measuring errors in predicting models is very important, but many current studies don't look into uncertainty estimate methods in depth. In the future, researchers should focus on making strong ways to measure unpredictability. This will help make statements that are more accurate and useful for people making decisions in the green energy industry [24]. There is also not a lot of study that looks into how different satellite data sources and resolutions affect the accuracy of forecasts in a planned way. Using satellite data is an important part of predicting solar power, but different data sources and levels can have a big effect on how well ML models work. A thorough look at the pros and cons of various types of satellite data, like infrared, visible, and thermal images, along with an investigation into the effects of different temporal and spatial levels, would help improve numerical predicting models. Also, most of the studies that have been done so far have focused on short-term expectations for solar power and not so much on long-term predictions [24]. Long-term predictions are important for managing the grid, planning energy storage, and making policy. Long-term predictions, on the other hand, get less attention in the study world than short-term predictions. In the future, researchers should try to close this gap by creating and testing machine learning models that are especially made for making correct long-term estimates about solar power.

6. PROBLEM STATEMENT

As using green energy becomes more important in global energy plans, the field of solar power predictions has grown a lot. This is very true when satellite data and machine learning (ML) methods are used. Still, organised study is needed to solve a number of important issues. Ensemble learning methods haven't been used for solar power predictions very much, which is a big problem [25]. Researchers have mostly only looked at one machine learning programme at a time in the past. This means that the benefits of using more than one programme at the same time might not have been seen. It's important to keep in mind that there aren't many in-depth studies on how ensemble learning can improve the safety and accuracy of forecasts [26].

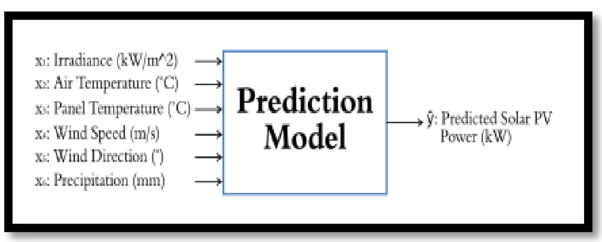

Fig. 4: Solar PV power prediction model

How to measure error is another big issue in making predictions about solar power. It's hard to say for sure how much solar power will be made because the weather changes quickly and monitor data isn't always correct [27]. There aren't many full ways to estimate uncertainties in this study, even though it is known how important it is to understand and measure errors. Because of this hole, it's harder to make models that can not only guess correctly but also measure the unknowns that come with those guesses to help people make better choices. Also, there is a big hole in the research that looks at how different satellite data sources and levels affect the accuracy of forecasts [28]. Since satellite data is what solar power predicting models are built on, differences in the types and levels of data can have a big effect on how well the models work. It's hard to make predicting models that work well and can be changed because different types of satellite data, like infrared, visible, and thermal images, and different levels of spatial and temporal resolutions haven't been thoroughly studied [29].

7. OBSERVATIONS

The three chosen journal papers on machine learning techniques for short-term and long-term solar power predictions using satellite data make some interesting points about different parts of the methods, datasets, and performance measures [30]. Each study makes a different addition, and a full review of these facts givesuseful information about the present situation and recent progress in predicting solar power. With the help of satellite estimates, [31] shows a new way to guess how much sunlight there will be during the day. The study looks at how adding location data from satellites can help people make better guesses during the day. It is very important to know about space if want to find small changes in sun energy [32].

[33] show a model that uses deep learning and statistical methods to guess how much solar power will be produced. The study looks at what can be gained by putting together long short-term memory (LSTM) networks and autoregressive integrated moving average (ARIMA). The ensemble method does a good job of dealing with the unknowns that come with both short-term and long-term plans. This shows the value of using a mix of methods. It's clear from this point that using more than one way of prediction could lead to better results generally[34]. Set up standard datasets for future study. This would allow a more organised comparison of various predicting models and help researchers better understand their relative strengths and flaws.The writers use long short-term memory (LSTM) networks and look at how well they work with old data on sun energy.

Several key observations emerge

Temporal Dependencies: The study stresses how important it is to record temporal dependencies in solar light data, since solar energy trends happen in a certain order. The use of LSTM networks works well for dealing with short-term changes.

Dynamic planning: The results show that short-term planning that focuses on instant changes is very important for getting used to the fact that solar energy supply is always changing. This shows how important it is for predicting models to be able to change in real time. [35] gives new information on how to predict the amount of sunlight during the day, focusing on using estimates from satellites. When it comes to spatial issues and model accuracy, the study makes some important points:

Integration of Spatial Information: The study looks at how important it is to use spatial information from satellite data to make intraday forecasts more accurate. To find small changes in sun energy, need to know about space [36]. This shows that methods based on satellites could help make things more detailed.

When look at spatial data, it makes sense that regional forecasting is important for making good daytime predictions, since the sun's energy trends can change a lot in small areas. show a model that uses deep learning and statistical methods to guess how much solar power will be produced.

Synergistic Group: The research looks at what happens when AIMA networks and long short-term memory (LSTM) networks work together. The ensemble method does a good job of dealing with the unknowns that come with both short-term and long-term plans. This shows the value of using a mix of methods [37].

The mix type has these pros: The finding shows that using more than one way of prediction together might lead to better results generally. Because it can use the best parts of both deep learning and statistical methods, the hybrid model is a strong way to predict solar power [38].

8. RESULT

There are three well-known journal articles that talk about how to use machine learning to use satellite data to make short- and long-term estimates about solar power. The important findings and results in these studies will help shape what the future holds for solar energy [39].

It shows a full case study on how to use deep learning to guess short-term sun energy. What they found can helpfigure out how well the models they used worked:How Well Does Deep Learning Work? Long short-term memory (LSTM) networks, in particular, are good at finding difficult trends in data about short-term sun energy, as shown by the study. The new type that uses LSTM works better than the old ones. Deep learning can see how things change over time [40].

A lot more accurate: It turns out that this model can predict things a lot better than older ones. There are better ways to find short-term changes in sun energy than with regular statistical methods. Deep learning is better because it can adapt to changes over time [41].

What to Look Forward To Based on what know about space: For better daytime guesses, the computer gets more information about the area by adding predictions from satellites [42]. It does a great job of taking into account changes in the amount of sun energy at the neighbourhood level. This shows how important it is to know about geography when try to guess what will happen.

More in-depth: At the fine-grained level, the model can make good predictions. This helpslearn more about how the sun's energy changes during the day. For projects that need to carefully plan and keep an eye on how solar power is made, this level of more information is very important [43].

It says that combining deep learning and statistical methods into one system can help figure out how much solar power will be made. Their joint plan has a lot of good points, as shown by the results:

Better Together Group*:* It is better to use the ensemble method, which mixes long short-term memory (LSTM) networks with autoregressive integrated moving average (ARIMA), than to use separate models. The helpful mix thinks about the unknowns in both short- and long-term predictions, which makes them more accurate.

The mix type has these pros: It works well to use the best parts of both deep learning and statistical methods together to make predictions. Because the model can change to fit different trends of solar power output, it is more accurate over a wider range of time periods [44]. One thing that all of the results have in common is how well deep learning methods, especially LSTM networks, can find complex patterns and time-dependent relationships in solar radiation data.

9. DISCUSSION

A lot of analysis has been performed on the field's progress, problems, and possible future paths from the three magazine papers chose. They are about using machine learning for short- and long-term solar power forecasts using satellite data [45].

Putting together deep learning techniques:

A theme that runs through all of the studies is that they all use both recurrent neural networks (RNNs) and long short-term memory (LSTM) networks. It shows that the model can pick up on complicated trends [45]. They improve the deep learning model by combining LSTM networks and autoregressive integrated moving average (ARIMA) in a group way. This helps them guess how much solar power will be made better.

It say that LSTM networks can predict the short term very well because they understand how time works with other things. To make the best plans for getting power from the sun, need to know this. Still, the study says that more work needs to be done to figure out how to set up LSTM networks in the best way and test how well they work in various settings [46].

People think that the ensemble method is a good way to deal with the unknowns that come with both short- and long-term planning. This will lead to more accurate predictions about solar power in general [47].

Thoughts on Space and Satellite Data:

Study after study shows how important satellite data is for making better models that predict how solar power will work.

The need for good, high-frequency satellite data is something that all the studies agree on. When it comes to short-term predictions, say that how much and how well satellite data is provided is very important [48].

For the best model performance, say that getting the right traits from satellites is important. The study shows how important it is for satellite technology to keep getting better so that models that make predictions can use correct and up-to-date information [50].

Methods for making accurate predictions using ensembles:

Group methods are used which shows that more and more people want strong solar power forecasting models [51]. To deal with factors better, deep learning and statistical methods work better together. This makes the system for making predictions more accurate.

Metrics for performance and standardisation:

The different success measures used in the studies are something that needs to be thought about in the conversation. All three papers rate how accurate their models are, but the exact measures used are different [52]. This difference makes it harder to directly compare studies [53].

10. CONCLUSION AND PLANS FOR THE FUTURE

The talk ends with a description of the success and issues with using satellite data for solar power predictions based on machine learning. As main ideas, they include combining deep learning methods, thinking about space, using group methods, and the constant need for good satellite data. While all of the studies add useful information, the talk stresses how important it is to standardise success measures so that they can be more easily compared and repeated [54].

It is very important for researchers, practitioners, and other interested parties to work together to solve these problems and move the field towards more accurate and dependable solar power forecasts solutions. Together, the results and talks in these studies move the field closer to a future where solar power projection is a key part of making the best use of energy and speeding up the world's move to clean, green energy sources [55].

REFERENCE LIST

[1] Alam, A., 2022, April. A digital game based learning approach for effective curriculum transaction for teaching-learning of artificial intelligence and machine learning. In 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) (pp. 69-74). IEEE.
[2] Alber, M., Buganza Tepole, A., Cannon, W.R., De, S., Dura-Bernal, S., Garikipati, K., Karniadakis, G., Lytton, W.W., Perdikaris, P., Petzold, L. and Kuhl, E., 2019. *Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences*. NPJ digital medicine, 2(1), p.115.
[3] Alizadehsani, R., Abdar, M., Roshanzamir, M., Khosravi, A., Kebria, P.M., Khozeimeh, F., Nahavandi, S., Sarrafzadegan, N. and Acharya, U.R., 2019. *Machine learning-based coronary artery disease diagnosis: A comprehensive review*. Computers in biology and medicine, 111, p.103346.
[4] Angelov, P.P. and Gu, X., 2019. Empirical approach to machine learning (p. 175). New York: Springer.
[5] Ben-Israel, D., Jacobs, W.B., Casha, S., Lang, S., Ryu, W.H.A., de Lotbiniere-Bassett, M. and Cadotte, D.W., 2020. The impact of machine learning on patient care: a systematic review. Artificial intelligence in medicine, 103, p.101785.
[6] Brunton, S.L., Noack, B.R. and Koumoutsakos, P., 2020. Machine learning for fluid mechanics. Annual review of fluid mechanics, 52, pp.477-508.
[7] Benos, L., Tagarakis, A.C., Dolias, G., Berruto, R., Kateris, D. and Bochtis, D., 2021. Machine learning in agriculture: A comprehensive updated review. Sensors, 21(11), p.3758.
[8] Brenner, M.P., Eldredge, J.D. and Freund, J.B., 2019. Perspective on machine learning for advancing fluid mechanics. Physical Review Fluids, 4(10), p.100501.
[9] Bica, I., Alaa, A.M., Lambert, C. and Van Der Schaar, M., 2021. From real-world patient data to individualized treatment effects using machine learning: current and future methods to address underlying challenges. Clinical Pharmacology & Therapeutics, 109(1), pp.87-100.
[10] Chouldechova, A. and Roth, A., 2020. *A snapshot of the frontiers of fairness in machine learning*. Communications of the ACM, 63(5), pp.82-89.
[11] Chan, S., Reddy, V., Myers, B., Thibodeaux, Q., Brownstone, N. and Liao, W., 2020. Machine learning in dermatology: current applications, opportunities, and limitations. Dermatology and therapy, 10, pp.365-386.
[12] Chibani, S. and Coudert, F.X., 2020. *Machine learning approaches for the prediction of materials properties*. Apl Materials, 8(8).
[13] Chen, K., Chen, H., Zhou, C., Huang, Y., Qi, X., Shen, R., Liu, F., Zuo, M., Zou, X., Wang, J. and Zhang, Y., 2020. *Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data*. Water research, 171, p.115454.
[14] Ceriotti, M., 2019. *Unsupervised machine learning in atomistic simulations, between predictions and understanding*. The Journal of chemical physics, 150(15).
[15] Cichos, F., Gustavsson, K., Mehlig, B. and Volpe, G., 2020. *Machine learning for active matter*. Nature Machine Intelligence, 2(2), pp.94-103.
[16] Chong, E.K., Lu, W.S. and Żak, S.H., 2023. *An Introduction to Optimization: With Applications to Machine Learning*. John Wiley & Sons.
[17] Dogan, A. and Birant, D., 2021. *Machine learning and data mining in manufacturing*. Expert Systems with Applications, 166, p.114060.
[18] Du, J., Jiang, C., Wang, J., Ren, Y. and Debbah, M., 2020. Machine learning for 6G wireless networks: Carrying forward enhanced bandwidth, massive access, and ultrareliable/low-latency service. IEEE Vehicular Technology Magazine, 15(4), pp.122-134.
[19] Dartmann, G., Song, H. and Schmeink, A. eds., 2019. *Big data analytics for cyber-physical systems: machine learning for the internet of things*. Elsevier.
[20] Elbadawi, M., Gaisford, S. and Basit, A.W., 2021. *Advanced machine-learning techniques in drug discovery*.

Drug Discovery Today, 26(3), pp.769-777.
[21] Feng, Y., Gong, D., Zhang, Q., Jiang, S., Zhao, L. and Cui, N., 2019. *Evaluation of temperature-based machine learning and empirical models for predicting daily global solar radiation*. Energy conversion and management, 198, p.111780.
[22] Fluke, C.J. and Jacobs, C., 2020. Surveying the reach and maturity of machine learning and artificial intelligence in astronomy. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(2), p.e1349.
[23] Fallucchi, F., Coladangelo, M., Giuliano, R. and William De Luca, E., 2020. Predicting employee attrition using machine learning techniques. Computers, 9(4), p.86.
[24] Guo, Y., Fu, Y., Hao, F., Zhang, X., Wu, W., Jin, X., Bryant, C.R. and Senthilnath, J., 2021. *Integrated phenology and climate in rice yields prediction using machine learning methods*. Ecological Indicators, 120, p.106935.
[25] Ginart, A., Guan, M., Valiant, G. and Zou, J.Y., 2019. *Making ai forget you: Data deletion in machine learning*. Advances in neural information processing systems, 32.
[26] Ghorbani, A. and Zou, J., 2019, May. Data shapley: Equitable valuation of data for machine learning. In International conference on machine learning (pp. 2242-2251). PMLR.
[27] Ghazal, T.M. and Alzoubi, H.M., 2021. Modelling supply chain information collaboration empowered with machine learning technique. Intelligent Automation & Soft Computing, 29(3), pp.243-257.
[28] García-Martín, E., Rodrigues, C.F., Riley, G. and Grahn, H., 2019. Estimation of energy consumption in machine learning. Journal of Parallel and Distributed Computing, 134, pp.75-88.
[29] Gündüz, D., de Kerret, P., Sidiropoulos, N.D., Gesbert, D., Murthy, C.R. and van der Schaar, M., 2019. Machine learning in the air. IEEE Journal on Selected Areas in Communications, 37(10), pp.2184-2199.
[30] Gui, G., Liu, F., Sun, J., Yang, J., Zhou, Z. and Zhao, D., 2019. *Flight delay prediction based on aviation big data and machine learning*. IEEE Transactions on Vehicular Technology, 69(1), pp.140-150.
[31] Garcez, A.D.A., Gori, M., Lamb, L.C., Serafini, L., Spranger, M. and Tran, S.N., 2019. Neural-symbolic computing: An effective methodology for principled integration of machine learning and reasoning. arXiv preprint arXiv:1905.06088.
[32] Huang, H.Y., Kueng, R. and Preskill, J., 2021. Information-theoretic bounds on quantum advantage in machine learning. Physical Review Letters, 126(19), p.190505.
[33] Hsieh, K., Phanishayee, A., Mutlu, O. and Gibbons, P., 2020, November. The non-iid data quagmire of decentralized machine learning. In International Conference on Machine Learning (pp. 4387-4398). PMLR.
[34] Huang, M.Q., Ninić, J. and Zhang, Q.B., 2021. BIM, machine learning and computer vision techniques in underground construction: Current status and future perspectives. Tunnelling and Underground Space Technology, 108, p.103677.
[35] Hohman, F., Wongsuphasawat, K., Kery, M.B. and Patel, K., 2020, April. Understanding and visualizing data iteration in machine learning. In Proceedings of the 2020 CHI conference on human factors in computing systems (pp. 1-13).
[36] Hou, Y., Li, Q., Zhang, C., Lu, G., Ye, Z., Chen, Y., Wang, L. and Cao, D., 2021. The state-of-the-art review on applications of intrusive sensing, image processing techniques, and machine learning methods in pavement monitoring and analysis. Engineering, 7(6), pp.845-856.
[37] Huang, K., Hussain, A., Wang, Q.F. and Zhang, R. eds., 2019. Deep learning: fundamentals, theory and applications (Vol. 2). springer.
[38] Ihme, M., Chung, W.T. and Mishra, A.A., 2022. Combustion machine learning: Principles, progress and prospects. Progress in Energy and Combustion Science, 91, p.101010.
[39] Janiesch, C., Zschech, P. and Heinrich, K., 2021. Machine learning and deep learning. Electronic Markets, 31(3), pp.685-695.
[40] Jo, E.S. and Gebru, T., 2020, January. Lessons from archives: Strategies for collecting sociocultural data in machine learning. In Proceedings of the 2020 conference on fairness, accountability, and transparency (pp. 306-316).
[41] Jansen, S., 2020. *Machine Learning for Algorithmic Trading: Predictive models to extract* signals from market and alternative data for systematic trading strategies with Python. Packt Publishing Ltd.
[42] Kashinath, K., Mustafa, M., Albert, A., Wu, J.L., Jiang, C., Esmaeilzadeh, S., Azizzadenesheli, K., Wang, R., Chattopadhyay, A., Singh, A. and Manepalli, A., 2021. Physics-informed machine learning*: case studies for weather and climate modelling*. Philosophical Transactions of the Royal Society A, 379(2194), p.20200093.
[43] Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S. and Yang, L., 2021. *Physics-informed machine learning*. Nature Reviews Physics, 3(6), pp.422-440.
[44] Kuhn, D., Esfahani, P.M., Nguyen, V.A. and Shafieezadeh-Abadeh, S., 2019. Wasserstein distributionally

robust optimization: Theory and applications in machine learning. In Operations research & management science in the age of analytics (pp. 130-166). Informs.
[45] Khang, A., Abdullayev, V., Jadhav, B., Gupta, S.K. and Morris, G. eds., 2023. *AI-Centric Modeling and Analytics: Concepts, Technologies, and Applications*. CRC Press.
[46] Kroese, D.P., Botev, Z., Taimre, T. and Vaisman, R., 2019. Data science and machine learning: mathematical and statistical methods. CRC Press.
[47] Kelleher, J.D., Mac Namee, B. and D'arcy, A., 2020. Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies. MIT press.
[48] Liebal, U.W., Phan, A.N., Sudhakar, M., Raman, K. and Blank, L.M., 2020. Machine learning applications for mass spectrometry-based metabolomics. Metabolites, 10(6), p.243.
[49] Lawson, C.E., Martí, J.M., Radivojevic, T., Jonnalagadda, S.V.R., Gentz, R., Hillson, N.J., Peisert, S., Kim, J., Simmons, B.A., Petzold, C.J. and Singer, S.W., 2021. *Machine learning for metabolic engineering: A review*. Metabolic Engineering, 63, pp.34-60.
[50] Leippold, M., Wang, Q. and Zhou, W., 2022. Machine learning in the Chinese stock market. Journal of Financial Economics, 145(2), pp.64-82.
[51] Mohamed, S., Rosca, M., Figurnov, M. and Mnih, A., 2020. Monte carlo gradient estimation in machine learning. The Journal of Machine Learning Research, 21(1), pp.5183-5244.
[52] Miscuglio, M. and Sorger, V.J., 2020. *Photonic tensor cores for machine learning*. Applied Physics Reviews, 7(3).
[53] Moosavi, S.M., Jablonka, K.M. and Smit, B., 2020. The role of machine learning in the understanding and design of materials. Journal of the American Chemical Society, 142(48), pp.20273-20287.
[54] Merghadi, A., Yunus, A.P., Dou, J., Whiteley, J., ThaiPham, B., Bui, D.T., Avtar, R. and Abderrahmane, B., 2020. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. Earth-Science Reviews, 207, p.103225.
[55] Morocho-Cayamcela, M.E., Lee, H. and Lim, W., 2019. Machine learning for 5G/B5G mobile and wireless communications: Potential, limitations, and future directions. IEEE access, 7, pp.137184-137206.
[56] Martins, N., Cruz, J.M., Cruz, T. and Abreu, P.H., 2020. Adversarial machine learning applied to intrusion and malware scenarios: a systematic review. IEEE Access, 8, pp.35403-35419.
[57] Alshurideh, M., Al Kurdi, B., Salloum, S.A., Arpaci, I. and Al-Emran, M., 2023. Predicting the actual use of m-learning systems: a comparative approach using PLS-SEM and machine learning algorithms. Interactive Learning Environments, 31(3), pp.1214-1228.