iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XI **Month of publication:** November 2022

DOI: https://doi.org/10.22214/ijraset.2022.47538

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

Machine Learning for Wind Energy Analysis and Forecasting

Zahid Mushtaq[1], Baljit Kaur[2], Aamir Yousuf Malik[3]
[1]*M. Tech Scholar, Department of Electrical Engineering, IET Bhaddal Technical Campus, Ropar, India*
[2]*Assistant professor, Department of Electrical Engineering, IET Bhaddal Technical Campus, Ropar, India*
[3]*M. Tech Scholar, Department of Electrical Engineering, IET Bhaddal Technical Campus, Ropar, India*

Abstract: ***Better prediction designs for the contributes directly of renewables are essential to reducing the requirement for controlling the energy associated with traditional power plants. For the extremely variable wind power output to be successfully integrated into the power system, a reliable forecast is especially crucial. For a wind inclusion study in the western United States, we use data from the National Renewable Energy Laboratory (NREL) and focus on summary wind power projection in this chapter. Our approach immediately derives functional connections from data, in contrast to physical procedures that depend on exceedingly complex differential equations. By reprising the regression model as a regression issue, we explore several regression strategies including regression models, k-nearest peers, and support vector regression. In our experiments, we look at forecasts for both individual wind turbines and entire wind parks, proving the viability of a machine learning technique for short-term wind power projection.***
Keywords: NREL, Wind Energy, Machine Learning

I. INTRODUCTION

Due in part to public spending [1] and increased consciousness of climate change, rapid advances in the techniques used to gather renewable energy have led to an increase in the share of such sources relative to traditional ones (e.g. fossil fuels). Turbines, which can be onshore (on land) or offshore, are especially used to harvest wind energy (at sea). Off shore wind farms will become increasingly prevalent for a number of reasons, notably stronger and more consistent ocean winds, larger units being easier to ship and build, less visual interference, and minimizing potential conflicts of interest [2].. However, maintaining offshore wind turbines is expensive since it ensures that they function well for the entirety of their lifespan (often 20–25 years). Condition monitoring (CM) is the practice of keeping an eye on a wind turbine's parts to spot variations in performance that might be the first signs of a problem. It should go without saying that employing robust CM to find defects before they happen could reduce operations and maintenance (O&M) costs significantly [4]. CM approaches have made use of analyses of certain operational metrics and traits (e.g. vibration analysis, strain measurement, thermography and acoustic emissions). Recent improvements in big data management, machine learning (ML), transmitter and signal processing systems, and computational resources have created opportunities for integrated and in-depth CM analysis tools, where various data types can support informed, dependable, cost-effective, and improvement — in CM. This system investigates modern ML-based methods to CM for wind turbines (from 2011 forward). Using the search terms "wind turbine structural health monitoring transformation" and "wind turbine analyzed classification," news stories were retrieved from Google Scholar and filtered by year (>2011), access, citations, and relevance; some papers published earlier than 2011 are included due to their previous importance.

Due to its high production unpredictability, renewable energy is expanding quickly and creating a number of problems. Accurate prediction is a key technology for the efficient integration of wind power into the grid since it enables the design of reserve plants, battery loading strategies, and consumer scheduling.

There are two types of models available for prediction jobs. The bulk of forecasting techniques, like [3], are based on physical models with numerical weather simulations. Both short-term and long-term hourly forecasts are made using these models. The second class of prediction methods includes machine learning algorithms, which have gained in popularity recently. Since these analysis methods get their functional correlations directly through observed, they are sometimes referred to as statistics models. Since they may be used for predictions with a temporal horizon spanning from seconds to hours, they are helpful for balancing the electrical grid with its numerous authorities.

In this work, daily wind speed data were utilized to predict wind power using ml techniques. We chose to utilize daily wind power projections since wind speed is inherently unstable and unpredictable.

The approaches and algorithms used for prediction accuracy are unable to produce accurate and pleasing results, particularly for long-term forecasting scenarios. Regression analysis approaches are used because predicting consistent wind power levels might be difficultLeast Absolute Shrinkage Selector Operator (LASSO), k Nearest Neighbor (kNN), eXtreme Gradient Boost (XGBoost), Random Forest (RF), and Support Vector Regression (SVR) were the regression approaches used in this study. The hourly wind speed dataset was used to calculate the daily mean wind speed, and the daily total wind power was calculated using both the daily wind direction and the standard error. The proposed method was evaluated on a variety of prospective sites to see if the engines could produce respectable outcomes in each instance. In order to further demonstrate the effectiveness of the suggested machine continuing to learn renewable power forecasting, we assessed them against wind direction observations from four different sites.... This study showed that by using a base location mode, machine learning algorithms may be used to decide if it makes sense to set up wind farms in a site that is unidentified.

II. OBJECTIVES

1) Various machine learning methods were employed to forecast future wind power.
2) Daily wind speed, daily point difference of monthly wind speed, and daily wind and solar data were given to machine learning. A total of five machine learning techniques were used for long-term wind power modeling.
3) The study supports building a new wind farm in an unproven location.

III. LITERATURE REVIEW

Catalao et al. [1] trained a three-layered backpropagation network for short-term wind power prediction using the Levenberg-Marquardt approach, outperforming the tenacity model and ARIMA methods. A neural network ensemble approach for wind power prediction was also developed by Han et al. [7]. Focken et al. [4] examined how regional smoothing effects lowered the prediction inaccuracy of combined power generation in the context of wind turbine clustering. Electrical engineers Pöller and Achilles [16] looked studied the possibility of fusing many wind turbines into a single power. The visuo wind power forecast method that serves as the foundation of our investigation was initially introduced in [10], with a more thorough explanation appearing in [11].
. In [20], we provided a description of our turbine preselection method. Since the optimization model is challenging to solve, we provided an evolutionary gray box approach for effective feature selection, which relates to a selection of suitable turbines. In [6], we proposed an ensemble method for SVR, in which a large number of SVRs' predictions are combined to create a potent classifier from small, randomly selected chunks of training data. We examined this problem individually since wind power ramps are difficult to integrate into the system [12]. Ramp prediction is addressed as an SVM-based classification problem. Recursive feature selection demonstrates that this method is affected by the number of nearby turbines.

IV. METHODOLOGY

The main goal is to automatically compute models using data rather than subjective judgments. These models will seek to extract as much underlying information about the data as they can because they are trained on real-world facts. It is being utilized in a wide range of industries, including security, botany, chemistry, visual effects, and even medicine. An application of machine learning that is well recognized is the automated identification of handwritten numerals (see Figure 2.1). Greyscale handwritten digits that have been scanned must be recognized and assigned a number 0 to 9. Although the samples to be recognized may have been produced by an entirely different and formerly unidentified individual, a sizable historical library of handwritten digits from diverse persons is used to train the model. Similar to how people can understand the numerals and letters written by an unknown person, the ability to generalize is a crucial requirement for machine-learning models..
The early performance of machine learning across a wide range of applications demanded enough computing power [2]. The speed of computer processors, main memory, and storage has only made handling larger data volumes simpler.

Figure 1. Illustration of handwritten numbers taken from the MNIST database [71]. The categorisation is challenging since the same digit written by multiple writers might change significantly. The capacity for generalization is a key criterion for an effective classifier.

Datasets actually exist. Machine learning models can also resolve problems that individuals are unable to. A new development in computer science called "big data" seeks to overcome the challenge posed by rapidly growing data quantities [127]. This is especially true in the energy and smart grids sectors, where a rising number of sensors need to be regularly managed [20].

For a variety of applications, there are different machine learning techniques that vary in terms of model construction, accuracy, usability, and computational cost [11, 39, 79]. The three different categories of machine learning algorithms are supervised, unsupervised, and semi-supervised learning. Finding the structure of a dataset made up of patterns without label information is the aim of unsupervised learning. Finding groups with similar patterns is made easier, in particular, with the use of clustering algorithms.

The best possible reconstruction of the original patterns is maintained while image compression algorithms give a mapping to a reduce space, or patterns with fewer properties than the original patterns. This study focuses on the supervised environment, even if semi-supervised methods are an attractive extension of supervised approaches when there is limited data available. Here, labels are available and may be utilized to build inference and simulation methods.

The structure of this section is as follows. The idea of supervised learning and its formalization are introduced in Section 2. Figure 1 displays the k-nearest neighbors regression.

A. Supervised Learning

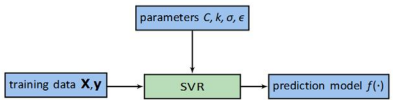


Figure 1 An illustration of the supervised machine learning environment. An method using the labeled training dataset and various parameter values computes a model f.

In this example, an SVR model is trained using the parameters penalty C, kernel k, kernel bandwidth, and loss sensitivity g..

B. Supervised Learning

In this project, we limit ourselves to the task of supervised learning. The goal is to use labels assigned to patterns in the training set to predict a label for an unlabeled test pattern. When discrete labels are applied to the task, it is referred to as classification. When employing continuous values, the task is referred to as regression. The time series prediction problem in this study is a regression problem, hence the focus of this chapter is on regression techniques.

A recent description and comparison of several existing techniques are provided by Fernandez-Delgado et al. in their article from the year 2008 (p. 28). They evaluate 17 distinct algorithms families and 179 different classifiers. While some algorithms are motivated by statistics, some by symbolic AI, and yet others by biological principles. The ideal algorithm might sometimes depend on the practitioner's preferences or experience. The optimum strategy also heavily depends on the available data. C4.5According to the IEEE International Conference on Data Mining (ICDM) in December 2006, support vector machine and k-Means are two of the top 10 data mining algorithms1. Algorithms include (SVM), Apriori, EM, PageRank, AdaBoost, k-NN, Naive Bayes, and CART [123]. In recent years, owing to the idea of "Deep Learning," neural networks have become a popular machine learning technique.

SVR, k-NN, and DTs are our main tools for this endeavor. The k-NN and DT algorithms are discussed in this chapter and are important for comprehending machine learning methods. The regression form of SVM is taught in Chapter 5, which deals with ensembles of support vector regressors for wind power prediction. For further details on machine learning ensembles, see Chapter 4.

C. k-Nearest Neighbors

The well-known k-nearest neighbors (k-NN) technique is a relatively simple yet effective solution for regression and classification problems [11, 39, 121]. It is a well-known machine learning technique that has been applied in a number of real-world settings, such as computer vision [98], astronomy [35, 44], computer graphics, biology, and physics.

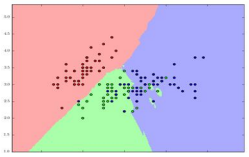


Figure 2 (a) Classification with k = 15.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

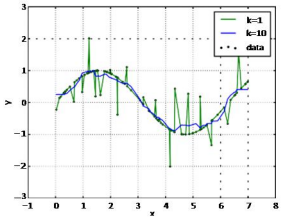


Figure 4.) Regression with k = 1 and k = 1

1) Nearest Neighbors Model

The distance metric-based search for similar patterns in the training dataset is the foundation of the k-NN model. The Euclidean distance, which is referred to as the distance between two points, is the most widely used choice. When using this metric, it is essential to normalize the feature values to a number between 0 and 1. If not, as illustrated in [121], the characteristics' relative relevance would depend on the scale at which they were measured. As an alternative, this behavior might be exploited to develop an adaptive distance measurement, such as by learning a general Mahalanobis distance [77, 120]. This academic discipline is known as distance metric learning..

The k nearest neighbors w.r.t. (x, xj) are searched from all patterns in the training set x X when making a prediction for a test pattern pattern xj. The proposed method averages the label information of the k closest neighbors with regressed when dealing with jobs that need

The efficacy of the k-NN model for classification or regression depends on the choice of coefficient k. Figure 2.3 displays one case (b). If k is set too low, the perceptron underpredicts test patterns while being overfit to the training set. The job of model selection is described in Section 2.4.

2) Efficient Neighbor Search

Since there is no need for a training phase with the k-NN model, no learning is necessary. A lookup in the training dataset is the only source of information used to forecast a pattern xj. Thus, the technique is an algorithm for instance-based learning (see Witten [121]). The time required for a brute-force implementation for a training set X and a test set Xtest is O(|X| |Xtest|), although there have been major advances that have made more efficient computations possible. The development of complex geographic data structures is rather frequent. The most important trees are cover trees [10] or K-d trees [8, 39]. Normal k-d tree is the definition of a balanced binary tree: The root of the tree T represents all points, while its two descendants are (almost) equal-sized subsets. The points are divided into these subsets level-wise starting at the root (level I = 0). One uses the median for each node v at level I in dimension I mod d to divide the points of v into two groupings. The recursion ends when a node v corresponds to a singleton or when a user-defined recursion level is achieved. Building such a tree in O(n log n) time for n patterns is feasible since obtaining a median needs linear time.

Locality-sensitive hashing [3, 98], which seeks to compute (1 +)-approximations, is another acceleration technique (with high probability). The latter category of methods primarily handles learning challenges in feature spaces with high dimensions. Additionally, certain GPU-based solutions for speeding up closest neighbor queries have been suggested (GPU). These methods typically seek to provide a respectable speedup for medium-sized datasets, but they fall short for big datasets [16, 34]. Although extremely effective solutions using GPU-based data structures do exist, they are not necessarily appropriate [35, 44].

D. Decision Trees

Decision Trees are fundamental machine learning methods for regression and classification. Apart from their cheap processing costs, decision trees' main benefit is the model interpretability: A decision tree is frequently a binary tree, where each node stands for a judgment standard based on one t - way test feature. Each leaf node is provided with a label. The machine learning specialist can therefore readily grasp the decisions taken while navigating the tree as a consequence. [11, 39]. For building decision trees using a training dataset, there are several approaches available. Quinlan created the well-known ID3, C4.5, and its commercial sequel, the C5.02. We restrict ourselves to using Breiman's classification and regression trees (CART) technique in this study [14]. According to Wu et al., CART and C4.5 are both among the top 10 algorithms listed by the IEEE International Conference on Data Mining (ICDM) in January 2006. [123]

Despite being utilized in many different professions, decision trees are simple and have few possibilities. Decision trees, however, are essential for this study and modern machine learning in general due to their applicability.

While the CART algorithm was employed in bagging [12] and the successful RF [13], the C5.0 method makes advantage of boosting capabilities [30]. Modern research has found that these methods are among the most reliable and accurate models [28].

V. SYSTEM ARCHITECTURE

A number of machine learning methods, each of which focuses on a different subject. Technologies like classification and regression are frequently used for forecasting. In order to produce actual values from the input dataset, several of these approaches were derived from algorithms. A subclass of classification is referred known as "regression analysis.".

This section introduced the regression analysis methods used in this study: Support Vector Regression, XGBoost Regression, Random Forest Regression, and LASSO Regression (SVR). Due to their widespread usage and impressive performance in the literature, these techniques were chosen for regression concerns. Different theoretical foundations underlie these algorithms, and it would be helpful to know which foundation and approach is better at predicting wind power.

Numerous algorithm parameters affect each algorithm's performance and runtime. In order to determine the best parameters for each algorithm given our circumstances, we took a trial-and-error method. Reporting the parameter value at the conclusion of each algorithm stage, we ran algorithms with a variety of parameter values and used the best observed results.

A. LASSO regression

LASSO (Least Absolute Shrinkage Selector Operator) regression is a specialization of linear regression [40]. It is sometimes referred to as the shrinkage model since it balances the educated assumptions. Finding the subset of predictors with the lowest prediction error for a quantitative response variable is the aim of LASSO regression. The only way LASSO regression varies from linear and ridge regression is by parameterizing the coefficient of some features to zero..

B. k Nearest Neighbor (kNN)

aims to classify test cases based on their feature closeness to the number of k class centers and is a well-known instance-based lazy learning classification algorithm [41,42]. Distance measures like Euclidean, Manhattan, and Minkowski are used to gauge proximity. It begins with k randomly chosen locations and classifies training samples based on how far they are from these k centers. Iterative processes are the most effective way to mimic where k's class centers are located. Then test cases are divided into groups based on how similar their characteristics are to those of these k class centers. The kNN regression technique, which is a regression variation of the kNN classification method, maps test results to given training inputs and their related outputs.

C. xGBoost regression

The gradient boosting decision tree approach has a more effective and parallel variation called eXtreme Gradient Boost (xGBoost) [43]. The xGBoost approach is hence rapid and scalable. The technique intends to minimize the differential equation by designing better decision trees. Regression is another use for the xGBoost method. Considering that it is quicker and more effective than other boosting methods, it can find ideal answers more quickly.

D. Random forest regression

A popular decision tree approach that produces several decision trees from a single dataset is Random Forest (RF) [44]. The RF divides the input parameters for the dataset into several parts, creates decision trees for each subset of the features, and utilizes the results of each decision tree to get a final determination. Using this method, the complex issue of several feature spaces is divided into simpler and easier to understand parts. Each tree in random forest is formed using a random vector of k, which is a subset of the dataset's feature space, and the training dataset.

E. Support vector regression

The Support Vector Machines (SVM) technique has a regression variant called Support Vector Regression (SVR) [45]. To categorize input datasets, the SVM algorithm builds a line, plane, or hyperplane for a one-, two-, or multidimensional input space. A regression function is sought for by nonlinear SVR from the input hyperplanes. The SVR algorithm is the one that is used the most frequently. The SVR method attempts to fit a plane from the input variables within - distance using training data examples.

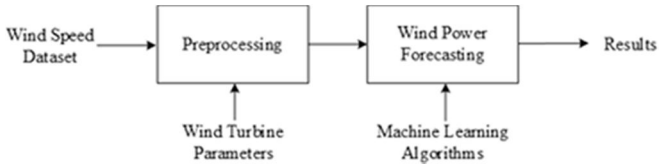


Figure 5 Model Block Diagram

Based on the daily mean wind speed and standard deviation, machine learning algorithms were used in the current study to anticipate the daily total wind power. The initial dataset included data on wind speed on an hourly basis. The daily mean wind speed and standard deviations were calculated from the hourly wind speed readings. Additionally, daily total wind power values were derived from hourly wind speed readings, and hourly wind power values were derived from hourly wind speed values. Following that, 4 years of daily mean wind speed, standard deviation, daily total wind power, and the following year's prediction were used to train the machine learning algorithms.

VI. SIMULATION AND RESULTS

The dataset is presented, and then the effectiveness of machine learning techniques for wind power forecasting is examined. The approaches are evaluated using the R2 values, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) metrics..

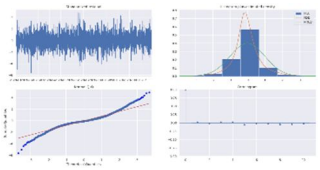


Figure 6 MAE and RMSE Matrix

predicting the next 24 steps

In [16]:

```
# u can change steps variable to any number of steps u would like to predict in the future
steps = 24
steps_index = []
for i in range(1, (steps+1)):
    steps_index.append(df.index[-1] + pd.DateOffset(hours=i))
```

In [17]:

```
forecast = model_auto.predict(n_periods=steps, exogenous=df[['wind_speed']][-steps:])
for i, val in enumerate(forecast):
    if val < 0:
        forecast[i] = val * -1
```

A. Ploting the predicted values

In [18]:

```
model_predict = pd.DataFrame(forecast, index=steps_index, columns=['next_' + str(steps) + '_steps'])
plt.figure(figsize=(20,10))
plt.plot(df,color='blue')
plt.plot(model_predict,color='red')
plt.show()
```

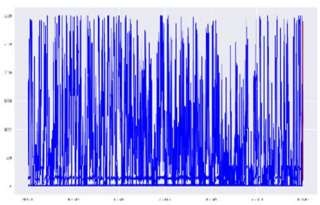


Figure 7. Predicted Values Plot

VII. CONCLUSION

In particular, classification techniques were used to predict the values of provided wind speed observations. The daily mean wind speed values were computed using hourly wind speed data, and the standard deviation was utilized to get the daily total wind power. The proposed method was tried in several places to see if the algorithms could produce results that were suitable for the training site. Long-term daily total wind power predictions were proven to be accurate using the xGBoost, SVR, and RF algorithms. RF is the most effective of these approaches, having an R2 score of 0.995 and an MAE of 7.048. The linear LASSO algorithm is by far the weakest one. On the other hand, the LASSO R2 value, which is 0.862, is

The research's key result is that by using the wind power model for a base site, machine learning algorithms may be used to decide if it makes sense to build wind farms in uncharted regions.

REFERENCES

[1] Jung J, Broadwater RP. Current status and future advances for wind speed and power forecasting. Renew. Sustain. Energy Rev. 2014;31:762–77. https://doi.org/ 10.1016/j.rser.2013.12.054. URL: http://www.sciencedirect.com/science/article/ pii/S1364032114000094.

[2] Wang X, Guo P, Huang X. A review of wind power forecasting models. Energy Procedia 2011;12:770–8. https://doi.org/10.1016/j.egypro.2011.10.103. URL: http://www.sciencedirect.com/science/article/pii/S1876610211019291.

[3] Chang W-Y. A literature review of wind forecasting methods. J. Power Energy Eng. 2014;2:161–8. https://doi.org/10.4236/jpee.2014.24023. URL: https://www. scirp.org/journal/PaperInformation.aspx?PaperID=44881.

[4] Samuel AL. Some studies in machine learning using the game of checkers ii – recent progress. IBM J. Res. Dev. 1967;11(6):601–17. https://doi.org/10.1147/rd.116. 0601.

[5] Alpaydin E. Introduction to Machine Learning. MIT Press; 2009.

[6] Rajagopalan S, Santoso S. Wind power forecasting and error analysis using the autoregressive moving average modeling. IEEE Power Energy Society General Meeting 2009;2009:1–6. https://doi.org/10.1109/PES.2009.5276019.

[7] Abdelaziz A, Rahman MA, El-Khayat M, Hakim MA. Short term wind power forecasting using autoregressive integrated moving average modeling. Proceedings of the 15th International Middle East Power Systems Conference (MEPCON'12). 2012. pp. 208:1–208: 6.

[8] Cadenas E, Rivera W, Campos-Amezcua R, Heard C. Wind speed prediction using a univariate ARIMA model and a multivariate NARX model. Energies 2016;9(2). https://doi.org/10.3390/en9020109. URL: http://www.mdpi.com/1996-1073/9/ 2/109.

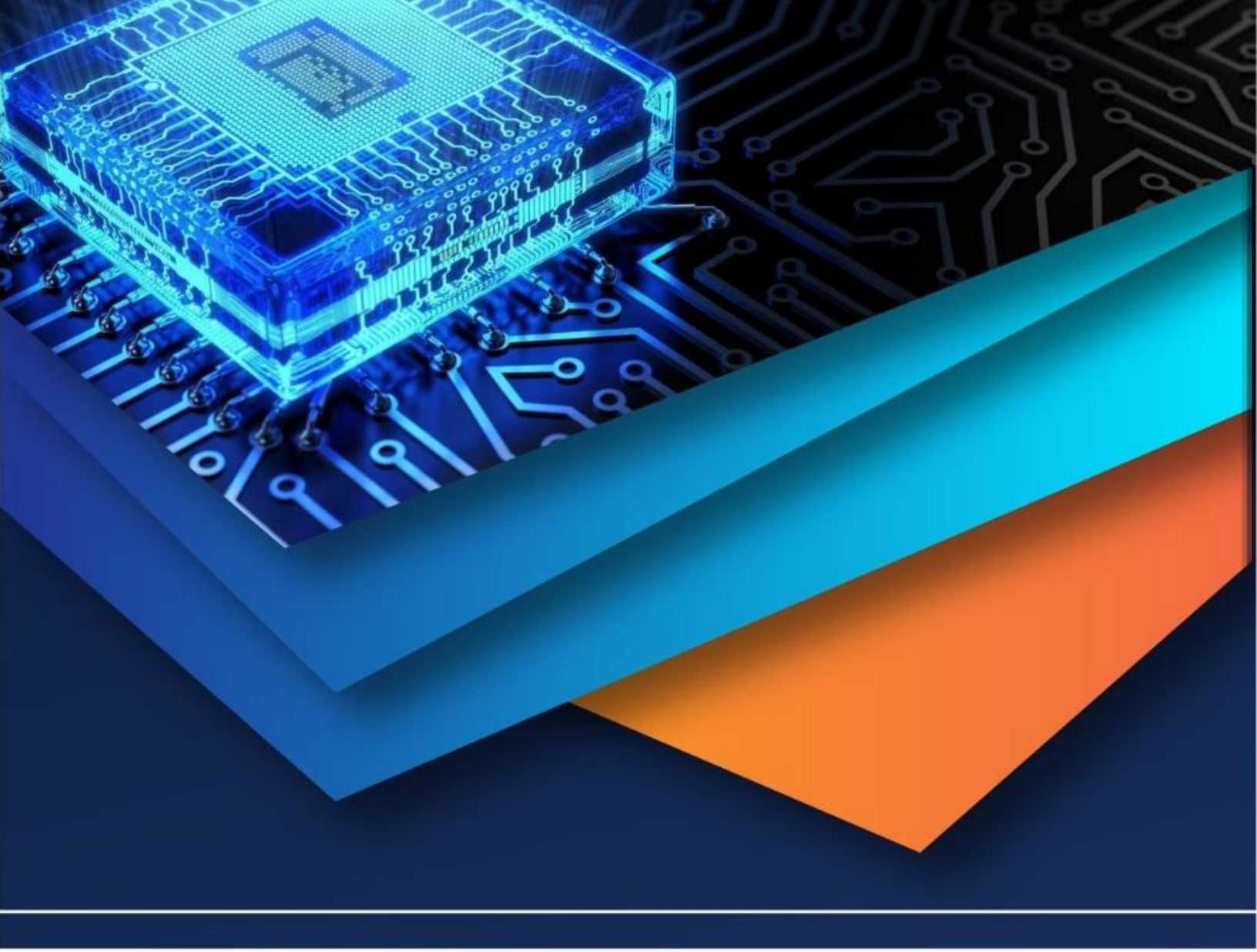
[9] Dowell J, Pinson P. Very-short-term probabilistic wind power forecasts by sparse vector autoregression. IEEE Trans. Smart Grid 2016;7(2):763–70. https://doi.org/ 10.1109/TSG.2015.2424078. [10] Lima JM, Guetter AK, Freitas SR, Panetta J, de Mattos JGZ. A meteorological-statistic model for short-term wind power forecasting. J. Control Autom. Electr. Syst. 2017;28(5):679–91. doi: 10.1007/s40313-017-0329-8.

[10] Wang J, Zhou Q, Zhang X. IOP Conf. Ser.: Earth Environ. Sci. 2018;199:022015https://doi.org/10.1088/1755-1315/199/2/022015.

[11] Robles-Rodriguez C, Dochain D. Decomposed threshold armax models for short- to medium-term wind power forecasting. IFAC-PapersOnLine 2018;51(13):49–54. https://doi.org/10.1016/j.ifacol.2018.07.253. 2nd IFAC Conference on Modelling, Identification and Control of Nonlinear Systems MICNON 2018.http://www. sciencedirect.com/science/article/pii/S240589631831005X.

[12] Pearre NS, Swan LG. Statistical approach for improved wind speed forecasting for wind power production. Sustainable Energy Technol. Assess. 2018;27:180–91. https://doi.org/10.1016/j.seta.2018.04.010. URL: http://www.sciencedirect.com/ science/article/pii/S221313881730512X.

[13] Sideratos G, Hatziargyriou ND. An advanced statistical method for wind power forecasting. IEEE Trans. Power Syst. 2007;22(1):258–65. https://doi.org/10.1109/ TPWRS.2006.889078.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY