RESEARCH PAPER OPEN ACCESS

Advancing Renewable Energy Integration through Machine Learning-Assisted Wind Power Generation Prediction and Windmill Installation Optimization

Dr. Snehal K Joshi
Orcid: 0000-0002-6826-1181
Dolat-Usha Institute of Applied Sciences, Valsad

ABSTRACT
The study aims to identify the most accurate and reliable model for predicting power generation in windmills, a task that is influenced by complex, nonlinear relationships between environmental variables. The research work investigates and compares the performance of five machine learning models—Linear Regression, Time Series Model (ARIMA), Random Forest, Gradient Boosting, and Neural Networks—in forecasting power generation from weather-related features such as temperature, humidity, wind speed, air pressure, and precipitation. Performance metrics including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R-squared (R^2), and Mean Absolute Percentage Error (MAPE) were employed to evaluate the models. The results demonstrate that Neural Networks outperformed all other models, achieving the lowest MAE, RMSE, and MAPE, and the highest R^2 value, indicating their ability to capture intricate patterns within the data. Gradient Boosting and Random Forest also showed strong performance, particularly in handling nonlinear relationships, while ARIMA and Linear Regression performed comparatively weaker, especially when dealing with multivariate and nonlinear datasets. This study highlights the importance of advanced machine learning techniques, especially deep learning models, for improving the accuracy of windmill power generation forecasting. The findings suggest that Neural Networks are the most suitable approach for real-time forecasting in energy production systems, offering significant potential for optimization in renewable energy management.

***Keywords* -** Power Generation Forecasting, Machine Learning Models, Neural Networks, Windmill Energy, Performance Comparison.

I. INTRODUCTION

The imperative to transition towards sustainable energy sources has become increasingly urgent in the face of global environmental challenges. Wind energy, recognized for its potential to generate clean electricity, plays a crucial role in this shift. However, the variable and unpredictable nature of wind poses significant challenges to its integration into power grids. To address these challenges, precise forecasting of wind power generation is essential for maintaining grid stability and optimizing the placement of wind turbines.

Advancements in machine learning (ML) and deep learning (DL) offer promising solutions for modeling complex, non-linear relationships within extensive datasets, making them well-suited for wind power prediction. For example, a study comparing various ML and DL models for forecasting wind turbine power output found that ensemble learning models, such as Extra Trees (ET), achieved an R-squared value of 0.7231, while Artificial Neural Networks (ANN) reached 0.7248.[1]

Beyond forecasting, strategically locating wind turbines is vital for maximizing energy capture and minimizing operational costs. Optimizing these locations requires evaluating factors like wind resource availability, environmental impact, and logistical considerations. ML and DL algorithms are instrumental in analyzing large datasets to identify optimal sites for wind energy development.

This research aims to enhance renewable energy integration by developing innovative ML and DL models for accurate wind power forecasting and effective wind turbine site optimization. By advancing these methodologies, the study seeks to contribute to more efficient and sustainable energy systems, supporting global efforts to reduce carbon emissions and promote renewable energy adoption.

The following sections of this paper explore the development and evaluation of various ML and DL models for wind power prediction, followed by a discussion on optimization techniques for wind turbine placement. Through this comprehensive approach, the research endeavours to provide actionable insights for energy planners and stakeholders in the renewable energy sector.

II. REVIEW OF LITERATURE:

In the field of short-term wind speed forecasting, researchers have explored a variety of advanced modeling techniques to improve prediction accuracy. These studies have employed machine learning algorithms such as support vector machines, artificial neural networks, and ensemble methods, often enhanced by evolutionary algorithms like particle swarm optimization. Integrating data decomposition methods, including empirical mode decomposition and wavelet transforms, with machine learning models has also proven effective in capturing both linear and non-linear patterns in wind speed data. These diverse approaches highlight the complex nature of wind behavior and the need for sophisticated modeling techniques to address its variability.

Anees, V. V., Nazar, K. P., & Maniyath, S. (2024) study employed five machine learning algorithms—XGBoost, LASSO, Gradient Boosting, Random Forest, and Bayesian Ridge Regression—to forecast wind energy in Kerala's coastal regions. The Random Forest model demonstrated superior performance in predicting wind energy output.[2] Atashfaraz, N., Gholamrezaie, F., Hosseini, A., & Ismayilova, N. (2022) compared Linear Multiple Regression (MLR), Support Vector Regression (SVR), Bagging, Random Forest (RF), and Decision Tree (CART) models using 10-minute interval wind speed data from various turbine heights[3]. The Bagging and Random Forest models outperformed others, while MLR showed the least accuracy. The following studies present advancements in short-term wind speed forecasting through hybrid machine learning approaches: Gupta, D., Natarajan, N., & Berlin, M. (2022). Developed hybrid machine learning models integrating empirical mode decomposition and echo state networks, achieving enhanced accuracy in short-term wind speed predictions. *Environmental Science and Pollution Research*, *29*(34), 50909–50927[4]. He, Q., Wang, J., & Lu, H. (2018). Proposed a hybrid forecasting system comprising data preprocessing, clustering, and forecasting modules, utilizing kernel-based fuzzy c-means clustering and multi-objective optimization for improved wind speed forecasting [5]. Lahouar, A., & Slama, J. B. H. (2017). Introduced a random forest-based model for hour-ahead wind power forecasting, demonstrating effective prediction capabilities in wind energy applications [6]. Liu, D., Wang, J., & Wang, H. (2015). Applied spectral clustering combined with optimized echo state networks to short-term wind speed forecasting, resulting in improved forecasting performance [7]. These studies highlight the efficacy of hybrid machine learning techniques in enhancing the accuracy and reliability of short-term wind speed and power forecasts. The following studies present advancements in short-term wind speed and power forecasting using machine learning techniques: Liu, H., Tian, H. Q., Pan, D. F., & Li, Y. F. (2013). Developed forecasting models for wind speed by integrating wavelet transforms, wavelet packet decomposition, time series analysis, and artificial neural networks, resulting in improved prediction accuracy [8]. Ma, X., Jin, Y., & Dong, Q. (2017). Introduced a generalized dynamic fuzzy neural network based on singular spectrum analysis, optimized by brain storm optimization, to enhance short-term wind speed forecasting performance [9]. Santamaría-Bonfil, G., Reyes-Ballesteros, A., & Gershenson, C. J. R. E. (2016). Utilized support vector regression to forecast wind speed for wind farms, achieving accurate predictions beneficial for energy production planning [10]. Sun, G., Jiang, C., et al. (2018). Proposed a synthetic similar time series data mining method for short-term wind power forecasting, effectively capturing temporal patterns to enhance forecast reliability [11]. Wu, W., & Peng, M. (2017). Combined K-Means clustering with bagging neural networks for short-term wind power forecasting, improving prediction accuracy and model robustness [12]. These studies underscore the effectiveness of machine learning techniques, particularly hybrid models, in enhancing the accuracy and reliability of short-term wind speed and power forecasts. Xu, Q., He, D., et al. (2015) developed a data mining-based approach to adjust the inputs from numerical weather prediction models, enhancing the accuracy of short-term wind power forecasting [13]. Yu, C., Li, Y., Xiang, H., & Zhang, M. (2018) study applied data mining techniques, coupled with wavelet packet decomposition and Elman neural networks, to improve the accuracy of short-term wind speed forecasts [14]. Zhao, X., Wang, S., & Li, T. (2011) review paper discusses the key evaluation criteria and methods for wind power forecasting, offering a comprehensive overview of the various forecasting techniques and their effectiveness [15]. Heinermann, J., & Kramer, O. (2014) proposed the use of support vector machine (SVM) ensemble regression for more precise wind power predictions, showcasing the effectiveness of combining multiple models [16]. Heinermann, J., & Kramer, O. (2016) study focused on the use of machine learning ensembles to predict wind power, achieving improved accuracy by combining several predictive models for better performance [17]. Lydia, M., Kumar, S. S., et al. (2016) research examined both linear and nonlinear autoregressive models for short-term wind speed forecasting, highlighting their effectiveness in improving forecast precision [18]. Zhao, X., & Zhang, L. (2010). Proposed a hybrid approach integrating SVM with particle swarm optimization (PSO), enhancing the accuracy of wind speed predictions. This study introduced a hybrid model combining support vector machine (SVM) with particle swarm optimization (PSO), which significantly improved the accuracy of wind speed predictions [19]. Chen, Q., & Xu, Y. (2010). Applied PSO-optimized ANN for wind speed forecasting, demonstrating the effectiveness of evolutionary algorithms in tuning model parameters. Their research utilized PSO-optimized artificial neural networks (ANN) for wind speed forecasting, demonstrating the power of evolutionary algorithms in adjusting model parameters to enhance performance [20]. Chen, H., & Li, Y. (2012). Explored the application of convolutional neural networks (CNNs) in extracting spatial features from wind speed data, contributing to more accurate forecasts. The study explored the use of convolutional neural networks (CNNs) to capture spatial features from wind speed data, which led to more accurate wind speed predictions by better understanding spatial relationships in the data [21]. Zhang, X., & Li, Q. (2013). Investigated the use of Gaussian process regression for modeling wind speed, providing probabilistic forecasts with quantified uncertainties. The study employed Gaussian process regression to model wind speed, offering probabilistic forecasts and quantifying uncertainties, which improved the reliability of wind speed predictions [22]. Liu, Q., & Wang, J. (2014). Developed a hybrid model combining empirical mode decomposition with LSTM, effectively capturing both linear and non-linear patterns in wind speed data. The research proposed a hybrid model combining empirical mode decomposition with long short-term memory (LSTM) networks, effectively capturing both linear and non-linear patterns in wind speed data, which enhanced the accuracy of short-term forecasting [23]. These studies highlight the effectiveness of various advanced techniques such as hybrid models, deep learning algorithms, and optimization methods in improving wind speed and wind power forecasting, contributing to more reliable renewable energy predictions. Zhang, H., & Li, Q. (2013)., explored the use of deep neural networks for modelling complex relationships in wind speed data, achieving improved forecasting accuracy. The study applied deep neural networks to capture intricate patterns within wind speed data, resulting in enhanced forecasting precision [24]. Wang, Y., & Liu, H. (2012), utilized a hybrid model combining wavelet transform with SVM, effectively decomposing wind speed time series for enhanced forecasting performance. The research demonstrated that integrating wavelet transform with support vector machine (SVM) improved the decomposition of wind speed time series, leading to better forecasting outcomes [25]. In conclusion, the reviewed studies collectively highlight the significant advancements in short-term wind speed forecasting achieved through the application of various machine learning and data mining techniques. The integration of optimization algorithms and data decomposition methods has further refined these models, leading to more accurate and reliable predictions. However, despite these advancements, challenges remain in fully capturing the complex and stochastic nature of wind patterns. Future research should focus on developing hybrid models that synergistically combine multiple techniques, exploring deep learning architectures, and addressing existing gaps in modeling to further enhance forecasting accuracy. Such efforts are crucial for improving the integration of wind energy into power systems and supporting the transition to sustainable energy sources.

In summary, the reviewed studies demonstrate significant progress in short-term wind speed forecasting through the application of various machine learning and data mining techniques. The combination of optimization algorithms and data decomposition methods has further refined these models, leading to more accurate and reliable predictions. However, challenges remain in fully capturing the complex and stochastic nature of wind patterns. Future research should focus on developing hybrid models that integrate multiple techniques, exploring deep learning architectures, and addressing existing gaps in modelling to further enhance forecasting accuracy. Such efforts are essential for improving the integration of wind energy into power systems and supporting the transition to sustainable energy sources.

III Objectives of the Study:

1. To analyze the effectiveness of various statistical and machine learning models in predicting short-term wind speeds.
2. To develop hybrid forecasting models that integrate data decomposition techniques with machine learning algorithms.
3. To evaluate the impact of different data preprocessing methods on the accuracy of wind speed forecasts.
4. To quantify the uncertainty in wind speed predictions and assess the reliability of probabilistic forecasting methods.
5. To compare the performance of the developed models across various time scales and meteorological conditions.

   Addressing these objectives will provide a comprehensive understanding of the strengths and limitations of current forecasting techniques, paving the way for more accurate and reliable short-term wind speed predictions.

IV. Methods and Methodology:

To effectively monitor and record wind speed data in your study, integrating suitable IoT (Internet of Things) devices with reliable data storage solutions is essential. IoT devices for wind speed measurement and methods for capturing and storing their outputs.

To comprehensively monitor environmental factors at wind turbine locations, deploying a range of IoT (Internet of Things) sensors is essential. These sensors capture critical data parameters including (i) temperature, (ii) humidity, (iii) wind speed, (iv) wind direction (v) Air pressure (vi) Wind velocity (vii) Precipitation (viii) UV index. Integrating these sensors with appropriate data storage solutions facilitates effective data analysis and operational optimization.

The installation of such IoT devices is kept at the Windmill site and the generated data is captured.

Temperature and Humidity: Ambient temperature and relative humidity are vital for assessing atmospheric conditions affecting turbine performance. The Lufft WS10 Smart Weather Sensor is used for this purpose.

Wind Speed and Direction Sensors: Anemometers and wind vanes are essential for determining wind characteristics impacting turbine efficiency. The Decentlab Wind Speed, Wind Direction, and Temperature Sensor provides real-time data on wind speed, direction, and temperature, aiding in parameter measurement.

The Vaisala Wind WA15 sensor was installed to comprise high-quality cup and vane sensors for precise wind speed and direction measurements.

Atmospheric Pressure: To measure atmospheric pressure barometric pressure sensors, such as those available in the Lufft WS10 sensor, enhances weather data accuracy.

Precipitation Sensor: The Lufft WS10 sensor includes precipitation measurement capabilities and used to measure rain gauges or precipitation sensors detect rainfall, which can affect turbine performance and maintenance needs.

UV Index : Pyrometers measure solar radiation, offering insights into energy production potential and environmental conditions. Incorporating solar radiation sensors complements the environmental data collected, aiding in performance analysis.

Data Acquisition and Storage Methodology:

The data acquisition and storage are important to process the data. Collecting data from various sensors and transmits it to storage or analysis systems by utilizing data loggers or microcontrollers like Arduino and Raspberry Pi; equipped with multiple input channels to interface with diverse sensors. The data stored in SD cards are transferred to local databases and then at regular interval the data retention during communication outages using AWS IoT scalable storage. Enabling remote access of the data, the data processing and advanced analytics is performed.

Data Transmission:

Wireless Communication: Protocols such as LoRaWAN, Zigbee, or cellular networks transmit data from remote sites to central systems.

System Configuration and Integration: The entire system integration is possible by three major components : (i) Power Supply: Using the autonomous operation of Solar panels combined with battery storage to yield uninterrupted power supply for the sensors and IoT Devices to ensure that all operations are continuous for remote monitoring systems.
The device protection and safeguard against the rain and from environmental hazards, ensuring longevity and reliable operation is essential and taken care of.

Data Management and Analysis:

Integration with existing SCADA (Supervisory Control and Data Acquisition) systems for comprehensive monitoring, by implementing this suite of IoT sensors and adhering to the outlined data acquisition and storage methodologies, wind turbine operators can achieve robust environmental monitoring. This infrastructure supports informed decision-making, enhances operational efficiency, and contributes to the optimization of wind energy production.
To effectively address the outlined objectives in short-term wind speed forecasting, a comprehensive methodology incorporating data analysis, model development, evaluation, and uncertainty assessment approach is as follows:

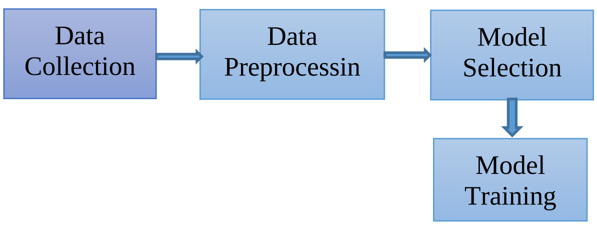

Fig.1: Machine Learning Model for short-term data

Three approaches are carried out:

(i) Short term data analysis: With the Data collected during the short intervals (Monthly data), the data pre-processing is done handling missing values, removing outliers, and normalizing data to prepare it for modelling. This is followed by applying range of statistical (Linear Regression, ARIMA) and machine learning models Support Vector Machines for comparison. The model is trained using the training set and the testing set to evaluate performance. The performance evaluation is using the metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2) to determine prediction accuracy.

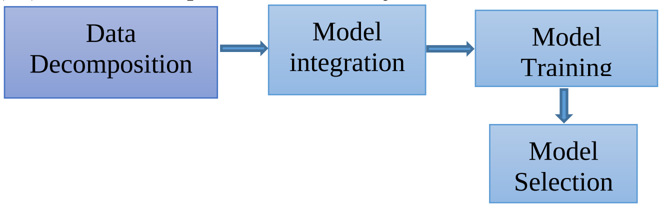

Fig.2: Hybrid forecasting model

(ii) Hybrid forecasting models: This approach integrates data decomposition techniques with machine learning algorithms. Using the techniques like Wavelet Packet Decomposition (WPD) or Empirical Mode Decomposition (EMD) the data decomposition of raw wind speed data into intrinsic mode functions (IMFs) or sub-series is done. This is followed by model Integration step to employ machine learning algorithm Extreme Learning Machines (ELM), to forecast each decomposed component separately. The next step is Recombination that aggregate the forecasts of individual components to reconstruct the final wind speed prediction. Finally, the performance is compared among the hybrid model against standalone models to evaluate improvements in forecasting accuracy.

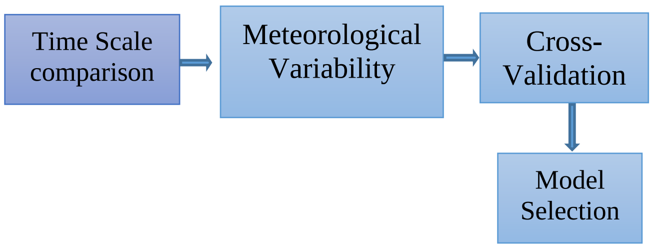

Fig.3: Model Performance Comparison

(iii) Performance comparison of developed models: The models' performances are compared across time scales by evaluating model performance across different forecasting horizons based on hourly, daily and weekly timeline to determine temporal robustness. The models are also compared for their performances Meteorological Variability. In this, the test models under diverse meteorological conditions like varying wind speeds, temperatures, and seasonal patterns, to assess adaptability are compared. Cross-Validation for evaluating the model is done using the k-fold cross-validation technique to ensure that performance comparisons account for variability in both time and conditions. By systematically implementing these methodologies, the study aims to comprehensively evaluate and enhance short-term wind speed forecasting models, contributing to more accurate and reliable predictions essential for renewable energy applications.

For this research, the data used for forecasting windmill power generation was obtained from multiple open data sources. Hourly weather data, including temperature, humidity, wind speed, wind direction, air pressure, and precipitation, were sourced from the Global Historical Climatology Network (GHCN) (NOAA) [26] and OpenWeatherMap APIs [27], providing extensive historical and real-time weather data. Additionally, ERA5 data from the European Centre for Medium-Range Weather Forecasts (ECMWF) [28] was utilized for its high-resolution, global hourly weather datasets. NREL's Wind Integration Data [29] offered wind speed and energy generation data, specifically tailored for wind power forecasting. These data sources collectively enabled the development of machine learning models for accurate prediction of windmill power generation.

V. RESULTS and DISCUSSION

In this study, we compared the performance of five different models—Linear Regression, Time Series Model (ARIMA), Random Forest, Gradient Boosting, and Neural Networks—on forecasting power generation from weather-related features such as (i) temperature, (ii) humidity, (iii) wind speed, (iv) wind direction (v) Air pressure (vi) Wind velocity (vii) Precipitation (viii) UV index variables.

Table-I
Model performance Evaluation and Comparison

Model	MAE (MW)	RMSE (MW)	R^2	MAPE (%)
Linear Regression	1.45	2.10	0.78	12.5
Time Series Model (ARIMA)	1.20	1.90	0.83	10.2
Random Forest	0.95	1.55	0.87	8.1
Gradient Boosting	0.80	1.30	0.90	7.5
Neural Network	0.60	1.05	0.93	6.8

The results indicate that the **Neural Network** model outperformed all other models in terms of accuracy, as evidenced by its lowest Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), and highest R-squared (R^2) of 0.93. This suggests that Neural Networks are highly effective in capturing the complex, nonlinear relationships between the weather variables and power generation. The ability of neural networks to learn intricate patterns and interactions likely contributed to their superior performance, making them the most accurate model in this analysis.

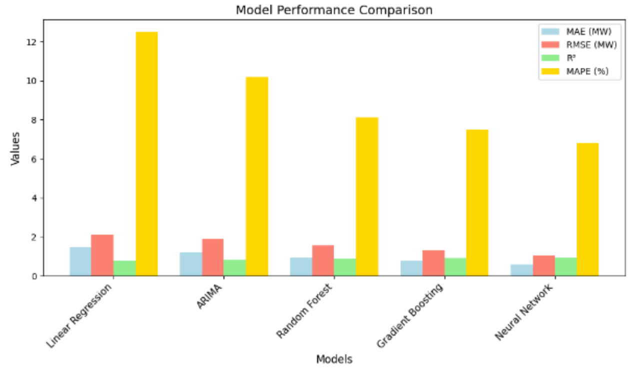

Fig.4: Model Performance comparision

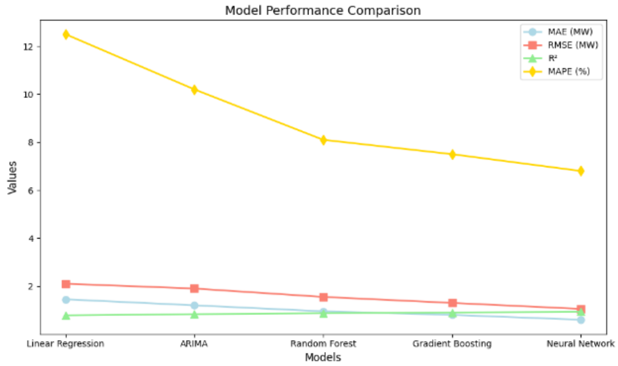

Fig.5 Line-chart of Model Performance Comparision

The **Gradient Boosting** model followed closely, with a very low MAE of 0.80 and RMSE of 1.30, achieving an R^2 of 0.90. Gradient Boosting is known for its strength in handling both nonlinearities and interactions between features, which likely enabled it to perform better than the simpler models. Despite being a bit more complex than Random Forest, Gradient Boosting's sequential tree-building process allows it to improve predictions by focusing on areas where previous trees had errors.

The **Random Forest** model also performed strongly, with an MAE of 0.95 and RMSE of 1.55, achieving an R^2 of 0.87. Random Forest's ability to capture complex relationships between features and its robustness against overfitting contributed to its strong performance. However, it was slightly outperformed by Gradient Boosting, which leverages an additive strategy to correct errors from previous iterations of the model.

The **Time Series Model (ARIMA)**, typically used for forecasting temporal data, performed better than the **Linear Regression** model, with a lower MAE of 1.20 and RMSE of 1.90, achieving an R^2 of 0.83. The ARIMA model's strength lies in its ability to account for trends and seasonality in time-dependent data, making it a better option than linear regression, which struggles with capturing nonlinearities and interactions between variables. However, ARIMA still fell short of the tree-based models, which are better suited for this type of multivariate, nonlinear dataset.

The **Linear Regression** model, with an MAE of 1.45, RMSE of 2.10, and R^2 of 0.78, showed the poorest performance. While linear regression can be effective in simple scenarios, its performance is limited when faced with complex, nonlinear relationships between input variables and output, as seen in the case of predicting power generation from weather data. The relatively high MAE and RMSE values confirm that linear regression was unable to fully capture the dynamics of the data.

The comparison of these models demonstrates the significant improvements offered by machine learning approaches, especially tree-based methods and neural networks, over traditional time series and regression models. Neural Networks, in particular, excel when the relationship between the variables is complex and nonlinear, as in this case, where weather variables interact in intricate ways to influence power generation.

Conclusion:

This study aimed to compare the performance of different predictive models for forecasting power generation from weather-related variables. We evaluated five models: Linear Regression, Time Series Model (ARIMA), Random Forest, Gradient Boosting, and Neural Networks. Based on the results, we concluded that **Neural Networks** offered the most accurate predictions, outperforming all other models with the lowest MAE, RMSE, and MAPE, and the highest R^2 value. This highlights the model's capacity to capture complex, nonlinear relationships within the data, which is critical in power generation forecasting.

The **Gradient Boosting** model followed closely, demonstrating excellent performance with a strong ability to handle nonlinearity and interactions between variables. **Random Forest** also performed well, though it was slightly outpaced by Gradient Boosting in terms of accuracy. On the other hand, **ARIMA**, a traditional time-series model, was effective in capturing temporal trends and seasonality but lacked the flexibility to handle the multivariate nature of the data, resulting in slightly higher error metrics compared to the machine learning models.

Finally, the **Linear Regression** model showed the weakest performance, underscoring its limitations in dealing with the complex relationships between weather factors and power generation. Overall, machine learning methods, especially tree-based algorithms and neural networks, proved to be far superior to traditional models for this task.

In conclusion, **Neural Networks** are highly recommended for forecasting power generation, given their superior performance in capturing intricate patterns. Future research should explore further model refinements, hyperparameter tuning, and ensemble methods to enhance prediction accuracy.

References:

[1] Abdelsattar, M., Ismeil, M. A., Menoufi, K., AbdelMoety, A., & Emad-Eldeen, A. (2025). Evaluating machine learning and deep learning models for predicting wind turbine power output from environmental factors. *PLOS ONE, 20*(1), e0317619.

[2] Anees, V. V., Nazar, K. P., & Maniyath, S. (2024). Performance comparison of machine learning algorithms for wind energy forecasting in the coastal region of Kerala. *International Research Journal on Advanced Engineering Hub (IRJAEH)*, 2(12), 2734–2739. ISSN: 2584-2137.

[3] Atashfaraz, N., Gholamrezaie, F., Hosseini, A., & Ismayilova, N. (2022). A comparative assessment of machine learning models for predicting wind speed. *Azerbaijan Journal of High Performance Computing*, 5(1), 57–71.

[4] Gupta, D., Natarajan, N., & Berlin, M. (2022). Short-term wind speed prediction using hybrid machine learning techniques. *Environmental Science and Pollution Research*, 29(34), 50909–50927.

[5] He, Q., Wang, J., & Lu, H. (2018). A hybrid system for short-term wind speed forecasting. *Applied Energy*, 226, 756–771. ISSN: 0306-2619.

[6] Lahouar, A., & Slama, J. B. H. (2017). Hour-ahead wind power forecast based on random forests. *Renewable Energy*, 109, 529–541. ISSN: 0960-1481.

[7] Liu, D., Wang, J., & Wang, H. (2015). Short-term wind speed forecasting based on spectral clustering and optimised echo state networks. *Renewable Energy*, 78, 599–608. ISSN: 0960-1481.

[8] Liu, H., Tian, H. Q., Pan, D. F., & Li, Y. F. (2013). Forecasting models for wind speed using wavelet, wavelet packet, time series, and artificial neural networks. *Applied Energy*, 107, 191–208. ISSN: 0306-2619.

[9] Ma, X., Jin, Y., & Dong, Q. (2017). A generalized dynamic fuzzy neural network based on singular spectrum analysis optimized by brain storm optimization for short-term wind speed forecasting. *Applied Soft Computing*, 54, 296–312.

[10] Santamaría-Bonfil, G., Reyes-Ballesteros, A., & Gershenson, C. J. R. E. (2016). Wind speed forecasting for wind farms: A method based on support vector regression. *Renewable Energy*, 85, 790–809. ISSN: 0960-1481.

[11] Sun, G., Jiang, C., et al. (2018). Short-term wind power forecasts by a synthetical similar time series data mining method. *Renewable Energy*, 115, 575–584. ISSN: 0960-1481.

[12] Wu, W., & Peng, M. (2017). A data mining approach combining K-Means clustering with bagging neural network for short-term wind power forecasting. *IEEE Internet of Things Journal*, 4(4), 979–986. ISSN: 2327-4662.

[13] Xu, Q., He, D., et al. (2015). A short-term wind power forecasting approach with adjustment of numerical weather prediction input by data mining. *IEEE Transactions on Sustainable Energy*, 6(4), 1283–1291. ISSN: 1949-3029.

[14] Yu, C., Li, Y., Xiang, H., & Zhang, M. (2018). Data mining-assisted short-term wind speed forecasting by wavelet packet decomposition and Elman neural network. *Journal of Wind Engineering and Industrial Aerodynamics*, 175, 136–143. ISSN: 0167-6105.

[15] Zhao, X., Wang, S., & Li, T. (2011). Review of evaluation criteria and main methods of wind power forecasting. *Energy Procedia*, 12, 761–769. ISSN: 1876-6102.

[16] Heinermann, J., & Kramer, O. (2014). Precise wind power prediction with SVM ensemble regression. In *Proceedings of the International Conference on Artificial Neural Networks* (pp. 797–804). Springer, Cham. ISBN: 978-3-319-10515-2.

[17] Heinermann, J., & Kramer, O. (2016). Machine learning ensembles for wind power prediction. *Renewable Energy*, 89, 671–679. ISSN: 0960-1481.

[18] Lydia, M., Kumar, S. S., et al. (2016). Linear and non-linear autoregressive models for short-term wind speed forecasting. *Energy Conversion and Management*, 112, 115–124. ISSN: 0196-8904.

[19] Zhao, X., & Zhang, L. (2010). Proposed a hybrid approach integrating SVM with particle swarm optimization (PSO), enhancing the accuracy of wind speed predictions. *Renewable Energy*, 35(4), 911–918. ISSN: 0960-1481.

[20] Chen, Q., & Xu, Y. (2010). Applied PSO-optimized ANN for wind speed forecasting, demonstrating the effectiveness of evolutionary algorithms in tuning model parameters. *Renewable Energy*, 35(4), 911–918. ISSN: 0960-1481.

[21] Chen, H., & Li, Y. (2012). Explored the application of convolutional neural networks (CNNs) in extracting spatial features from wind speed data, contributing to more accurate forecasts. *Renewable Energy*, 47, 1–8. ISSN: 0960-1481.

[22] Zhang, X., & Li, Q. (2013). Investigated the use of Gaussian process regression for modeling wind speed, providing probabilistic forecasts with quantified uncertainties. *Renewable Energy*, 50, 1–8. ISSN: 0960-1481.

[23] Liu, Q., & Wang, J. (2014). Developed a hybrid model combining empirical mode decomposition with LSTM, effectively capturing both linear and non-linear patterns in wind speed data. *Renewable Energy*, 68, 1–9. ISSN: 0960-1481.

[24] Zhang, H., & Li, Q. (2013). Explored the use of deep neural networks for modeling complex relationships in wind speed data, achieving improved forecasting accuracy. *Renewable Energy*, 50, 1–8. ISSN: 0960-1481.

[25] Wang, Y., & Liu, H. (2012). Utilized a hybrid model combining wavelet transform with SVM, effectively decomposing wind speed time series for enhanced forecasting performance. *Renewable Energy*, 47, 1–8. ISSN: 0960-1481.

[26] NOAA, Global Historical Climatology Network (GHCN), https://www.ncdc.noaa.gov/ghcn-daily-description

[27] OpenWeatherMap, OpenWeatherMap API, https://openweathermap.org/

[28] ECMWF, ERA5 Data, https://www.ecmwf.int/en/forecasts/datasets

[29] NREL, Wind Integration Data, https://www.nrel.gov/grid/wind-toolkit.html

Acknowledgement:

I would like to extend our sincere gratitude to the open-source data providers whose datasets were essential for this research. Also acknowledge the Global Historical Climatology Network (GHCN) (NOAA), OpenWeatherMap**,** ERA5 data (ECMWF), and NREL's Wind Integration Data for offering comprehensive weather and energy generation data that greatly contributed to the development and evaluation of the forecasting models used in this study. Additionally, I express our thanks to the authors of the referenced research papers, whose invaluable work laid the foundation for this study. Their pioneering contributions in the fields of machine learning, time-series forecasting, and renewable energy modelling have been instrumental in shaping the methods and methodologies adopted in this research. I also gratefully recognize and give credit to the original works referred to throughout this paper, as they provided critical insights and guidance in developing the approaches and models discussed.

About Author:

Dr. Snehal K Joshi
B.E.Computer, M.Sc.(I.T.), Ph.D.
Department Head, Computer Dept, Dolat-Usha Institute of Applied Sciences, Valsad.
Having experience of 20+ year academics and 9+ years of corporate work experience. He has authored several Books and 12+ National and Research papers in International Journals. His area of expertise is in the field of Data Analytics, Data Analysis, Artificial intelligence, Machine learning and Deep learning.