WIND ENERGY STORAGE PREDICTION USING MACHINE LEARNING

S A K Jainulabudeen[1], M.Sushil [2] ,Mohammed Afrith A [3],Rohith S[4]

[1]Assistant Professor[+], Students [2,3,4]

Department of Computer Science and Engineering,

Panimalar Engineering College, Chennai, Tamilnadu, India

---***---

Abstract - *Traditional (conventional) sources of power generation are Thermal (Coal), Hydro, Gas and Nuclear, but they are depleting and causing carbon emission. Many countries are taking harsh decision to close down thermal and nuclear. An alternative source, catching the attention of everyone is renewable energy from Solar and wind, which require no fuel and abundantly available according to geographical location. However, it has its own characteristic and throw upon challenges to integrate into transmission and distribution grid. Though available throughout the year. Wind potential varies location to location (that's why installed in specific areas mostly in remote locations) and season to season Wind is variable, intermittent and unpredictable during 24 hrs of the day. Location wise wind potential makes the task of transmission and distribution utility/grid operation more difficult in absence of local consumption as well as adequate network. Wind Energy is at peak during monsoon. This is the season when power demand is low. Grid operation has a challenge of handling the excess renewal energy. Grid operation is planned day ahead by taking supply (generator) and demand (utility) commitment. Grid operator is bound to control supply-demand balance to maintain frequency, but it becomes challenging when wind energy is accounted as part of supply in the day ahead planning due to variable and intermittent nature of wind. Grid operator is compelled to back-down conventional sources of generation to minimum level (inefficient operation) for load balancing.*

Key Words **Wind Energy ,Machine Learning ,Grids, renewable energy, storage**

1.INTRODUCTION

To solve climate change, society needs to rapidly decarbonize Incorporation of greater amounts of renewable wind power will be one of many essential strategies for decarbonization But the wind does not always blow.

To incorporate high amounts of wind power, the future smart grid will require accurate forecasts to operate effectively and efficiently.

Vertically-integrated electric utility company seeking to incorporate higher amounts of wind power generation into its electricity portfolio .employ machine learning to accurately predict hourly wind power generation at 7 wind farms, based on historical wind speeds and wind directions.

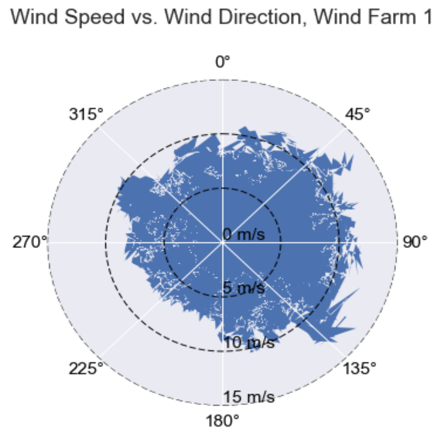

Fig1. Wind Speed vs Wind Direction

Evaluation:

Root Mean Square Error (RMSE): a value of 0 indicates a perfect fit. The lower the RMSE value, the more accurate the predictive capability of the model

Solar energy refers to capturing the energy from the Sun and subsequently converting it into electricity. We can then use that electricity to light up our homes.

Data provided by Institute of Electrical and Electronics Engineers (IEEE) .Retrieved from Kaggle database.Time series dataset with wind speed, wind direction and wind power production data.Data for 7 separate wind farms.Unitless for anonymity.

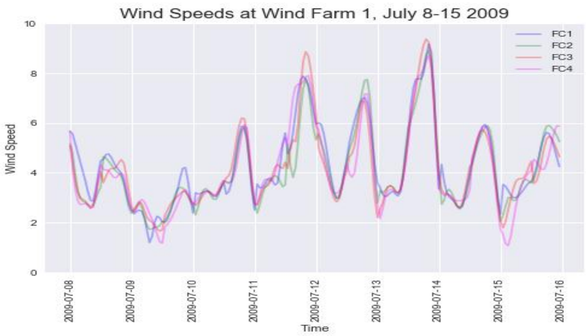

Fig 2. Wind Speed vs Wind Farm

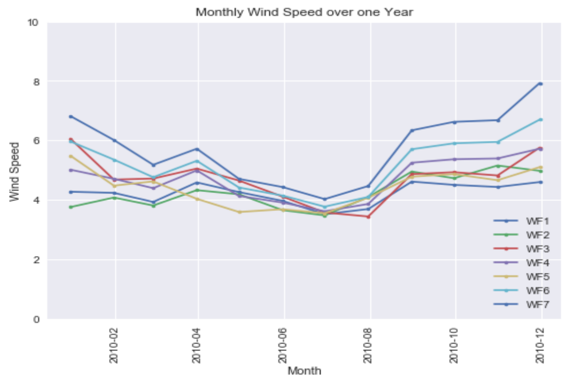

Fig 3. Wind Speed Over the year

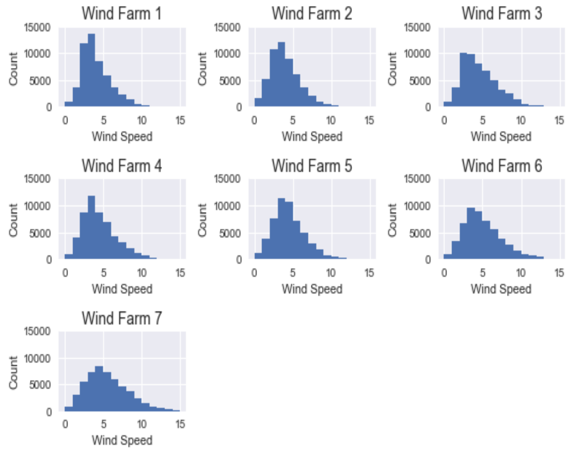

Fig 4. Wind farm

How is solar energy produced

- **Solar energy** refers to capturing the energy from the Sun and subsequently converting it into electricity.
- We can then use that electricity to light up our homes.

2.RELATED WORKS

In their survey, Costa et al. [2] present an expansive diagram of different techniques and scientific, measurable and physical models utilized over the most recent 30 years for shortterm expectation. Soman et al. [19] give a broad overview of the conceivable methods for various conjecture skylines. Past outcomes have demonstrated that techniques from factual learning are incredible methodologies for transient vitality forecast. For instance, Juban et al. [8] displayed a bit thickness estimation approach for a probabilistic determining for various breeze parks. Foresti et al. [5] utilized different part learning relapse as an all-inclusive help vector demonstrate that self-rulingly recognizes the significant highlights for wind speed expectations.

Additionally neural systems have been connected to wind control forecast before, e.g., by Mohandes et al. [13], who contrasted an autoregressive model and a traditional backpropagation arrange. In this line of research, Catalao et al. [1] prepared a three-layered feedforward connect with the Levenberg-Marquardt calculation for momentary breeze control anticipating, which beat the constancy model and ARIMA approaches. Further, Han et al. [7] concentrated on an outfit strategy for neural systems for wind control forecast. As to total of wind turbines,

Focken et al. [4] considered the decline of the forecast mistake of a totaled power yield brought about by spatial smoothing impacts. From the point of view of electrical specialists, Pöller and Achilles [16] investigated how unique breeze turbines can be totaled to a solitary generator. The spatio-transient breeze control forecast methodology that is premise of our line of research has been presented in [10] with an increasingly broad portrayal in [11]. In [20], we exhibited a methodology for preselection of turbines for kNN-based expectation. As the enhancement issue is hard to explain, we proposed a transformative Wind Power Prediction with Machine Learning 15 blackbox technique for a proficient element choice, which compares to a determination of suitable turbines. In [6], we proposed a gathering approach for SVR, where little subsets of preparing information are arbitrarily tested and the expectations of various SVRs are consolidated to a solid classifier. As wind control inclines are troublesome occasions for the joining into the matrix, we considered this issue in a different work [12].

We treat incline expectation as arrangement issue, which we tackle with SVMs. Recursive component determination delineates how the quantity of neighbored turbines influences this methodology. The issue of imbalanced preparing and test sets is broke down concerning the quantity of no-incline occasions. Practically speaking, sensors may fall flat for different reasons and the expectation models can't be connected. In [17], we looked at different missing information strategies for the attribution issue. Another commitment of this work is a kNN-based relapse technique, which is utilized as geo-attribution preprocessing venture by considering the time arrangement of the neighbored turbines. Last, in [22] we broadened the collection of forecast techniques with a cross-connection weighted k-closest neighbor relapse (x-kNN) variation.

The kNN-based comparability measure utilizes loads that depend on the cross-connection of the time arrangement of the neighboring turbines and the objective. In the event that the cross-relationship coefficient is high, the turbine gets a noteworthy impact for the forecast by growing the comparing measurement in the relapse display.

3.PROPOSED SYSTEM

1.Dataset Description

The data used for the project is provided by the Institute of Electrical and Electronics Engineers (IEEE), Power & Energy Society, and retrieved through the Kaggle database. The dataset is a time series dataset with historical power generation, wind speeds and wind directions, for the time period from July 2009 to December 2010.

Data Cleaning Process and Reading the files

The raw dataset is comprised of several csv files. The pandas library is imported as pd, and used to import the csv files using pd.read_csv. Converting dates and times to DateTime objects

The dataset features a 'date' column, featuring dates and times in integer (int64) format as YYYYMMDDHH, where YYYY = year, MM = month, DD = day, and HH = hour. To be more useful, these values are converted from int64 into Date Time objects using pd.to_datetime. Furthermore, these Date Time objects are converted into a standardized ISO 8601 format for convenience. A function called convert_to_iso is defined, to convert a date time in int64 format to a Date Time object in ISO 8601 format. The function convert_to_iso is applied to the 'date' columns in the dataset.

Creating a column for modified dates, 'mod_date' in addition to the 'date' column, the dataset features an 'hors' column, featuring hour values in int64 format. The values in 'hors'ranges from 1 to 48, representing the number of hours-ahead being forecasted. For example, let's say it is July 1st, 2009 at 12am. At this time and date, there is forecast data associated with hors=1. This forecast data thus refers to the forecast for the following time:

(July 1st, 2009, 12am) + (1-hour-ahead forecast) = July 1st, 2009, 1am

Thus, a new column, 'mod_date', is created to feature the DateTimes that include both the original datetimes ('date'), and the hour-ahead forecasts ('hors').

'date' + 'hors' = 'mod_date'

In the equation above, values for 'date' are in DateTime format. Values for 'hors' are originally int64 but are converted to timedelta format using pd.to_timedelta.

Creating a column for forecast categories, 'forecast_cat'

Forecasting data is split into the following four categories:

Category 1: 1-hour to 12-hour ahead data

Category 2: 13-hour to 24-hour ahead data

Category 3: 25-hour to 36-hour ahead data

Category 4: 37-hour to 48-hour ahead data

Thus, a new column 'forecast_cat' is created, featuring a value ranging from 1 to 4 for each data row, corresponding to the appropriate forecast category shown above. Boolean selection is achieved using .loc.

Merging wind data with training data

Then, the wind data is merged with the wind power generation training data using pd.DataFrame.merge. The specific merge method is 'left', specified by the 'how' argument, and is analogous to a LEFT OUTER JOIN SQL Join. Thus, the left outer join in this situation returns all of the rows for which there is both wind speed, direction data and wind power generation data available. Rows that have wind power generation data but no wind speed, direction data available are not included.

3.Machine learning Models

Predicting wind power production based on wind speed and wind direction is a regression problem. We model a dependent variable (wind power) based on two independent variables (wind speed and wind direction). In total, five machine learning models were trained and tested on the data: Linear Regression, Ridge Regression, Lasso Regression, Decision Tree Regression and Neural Network Regression. Root Mean Square Error (RMSE) was the metric used to evaluate prediction accuracy; an RMSE closer to a value of 0 is indicative of higher predictive accuracy. Generally, Decision Tree Regressors performed best, producing the lowest RMSE values across all wind farms. Neural Network Regressors performed worse. Also, Lasso Regression and Ridge Regression produced virtually the same results as standard Linear Regression. This makes sense given that there are only two independent variables at play; Lasso and Ridge are useful when there are many more feature variables involved, and it is necessary to incorporate variable selection. Linear/Lasso/Ridge Regression generally did not perform as well as Decision Tree Regression but always outperformed Neural Networks.

Experimental results:

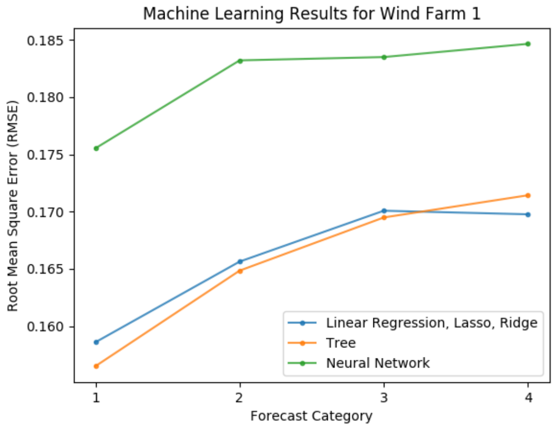

Fig 5 Wind Farm 1

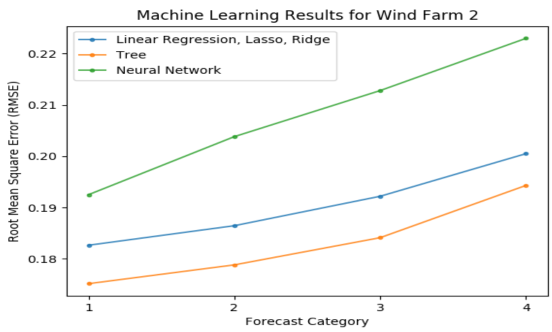

Fig 6 Wind Farm 2

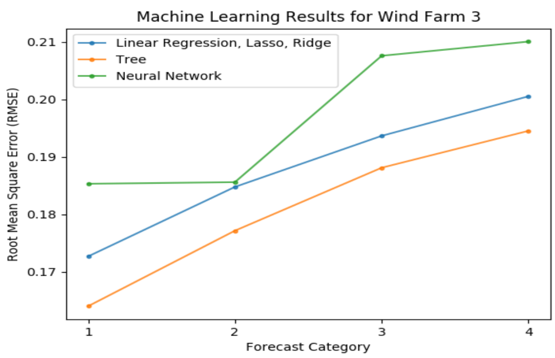

Fig 7 Wind Farm 3

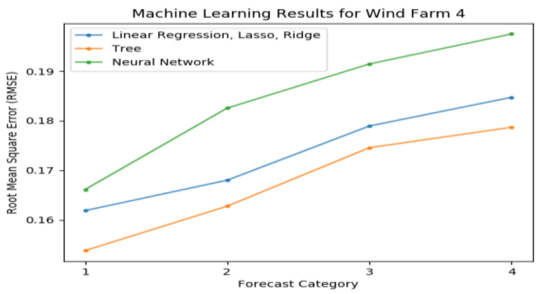

Fig 8 Wind Farm 4

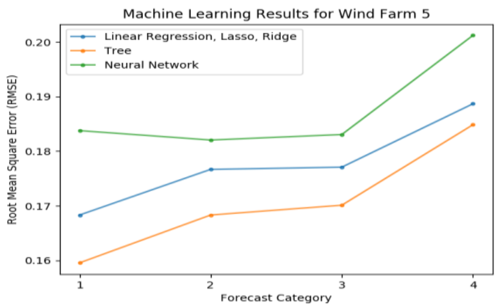

Fig 9 Wind Farm 5

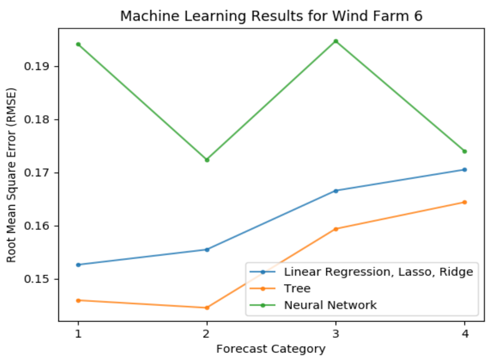

Fig 10 Wind Farm 6

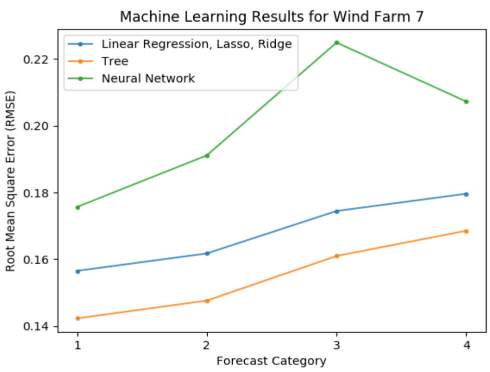

Fig 11 Wind Farm 7

4. CONCLUSIONS

A variety of Machine Learning Models were trained and tested on wind speed, wind direction and wind power production data. Prediction accuracy was evaluated using Root Mean Square Error (RMSE). Decision Tree Regressors performed best as compared to Standard Linear Regression, Lasso Regression, Ridge Regression and Neural Networks. Therefore, Decision Tree Regression is recommended for future predictions of wind speed, wind direction and wind power production data at the various wind farms.

REFERENCES

[1] Treiber, N.A et.al O.: Evolutionary turbine selection for wind power predictions. In: 37th Annual German Conference on AI, pp. 267–272 (2019).

[2] Soman, S.S., et.al.:A reviewof wind power and wind speed forecasting methods with different time horizons. In: North American Power Symposium (NAPS),pp. 1–8 (2018).

[3] Lew, D., Milligan, et.al, R.: How do wind and solar power affect grid operations: the western wind and solar integration study. In:8th International Workshop on Large Scale Integration of Wind Power and on Transmission Networks for Offshore Wind Farms (2018)

[4] Catalao, J.P.S., et.al V.M.F.: An artificial neural network approach for short-term wind power forecasting in Portugal. In: 15th International Conference on Intelligent System Applications to Power Systems (2009)

[5] Costa, A., et.al A review on the young history of thewind power short-term prediction. Renew. Sustain. Energy Rev. 12(6), 1725–1744 (2008)

[6] Ernst, B., et.al Predicting the wind. Power Energy Mag. 5(6), 78–89 (2007)

[7] Focken, U., et.al A.: Short-term prediction of the aggregated power output of wind farms—a statistical analysis of the reduction of theprediction

error by spatial smoothing effects. J.Wind Eng. Ind. Aerodyn. 90(3), 231–246 (2002)

[8] Foresti, L., et.al A.: Learning wind fields with multiple kernels. Stoch. Env. Res. Risk Assess. 25(1), 51–66 (2011)

[9] Heinermann, et.al Precise wind power prediction with SVM ensemble regression.In: Artificial Neural Networks and Machine Learning—ICANN 2014, pp. 797–804. Springer,Switzerland (2014)

[10] Han, S., et.al Neural network ensemble method study for wind power prediction.In: Asia Pacific Power and Energy Engineering Conference (APPEEC) (2011)

[11] Juban, J., et.al Probabilistic short-term wind power forecasting based on kernel. In: Density Estimators. European Wind Energy Conference, pp. 683–688. IEEE (2007)

[12] Kramer, O et.al A framework for data mining in wind power time series. In: Proceedings of ECML Workshop DARE (2014)

[13] Kramer, O et.al Short-term wind energy forecasting using support vector regression.In: 6th International Conference on Soft Computing Models in Industrial and Environmental Applications (2011)

[14] Kramer, O., et.al Analysis of wind energy time series with kernel methods and neural networks. In: 7th International Conference on Natural Computation (2011) .

[15] Kramer, O., et.al Wind power ramp event prediction with support vector machines. In: 9th International Conference on Hybrid Artificial Intelligence Systems (2014)

[16] Mohandes, et.al A neural networks approach for wind speed prediction. Renew. Energy 13(3), 345–354 (1998) .

[17] Pedregosa, F., et.al Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830 (2011)

[18] Pöller, M., Achilles, S.: Aggregated wind park models for analyzing power system dynamics.In: 4th International Workshop on Large-scale Integration of Wind Power and Transmission Networks for Offshore Wind Farms, Billund (2003)

[19] Poloczek, J., et.al KNN regression as geo-imputation method for spatiotemporal wind data. In: 9th International Conference on Soft Computing Models in Industrial and Environmental Applications (2014) .

[20] Robusto, C.C.: The Cosine-Haversine formula. Am. Math. Mon. 64(1), 38–40 (2018)

[21] . Treiber, N.A., et.al Aggregation of features for wind energy prediction with support vector regression and nearest neighbors. In: European Conference on Machine Learning, DARE Workshop (2017)

[22] Treiber, N.A., Kramer, O.: Wind power prediction with cross-correlation weighted nearest neighbors. In: 28th International Conference on Informatics for Environmental Protection (2017).

[23] Wegley,H., et.al Subhourly wind forecasting techniques for wind turbine operations. echnical report, Pacific Northwest Lab., Richland, WA (USA) .