UJRET
Bi-annual refereed journal
University Journal of Research Engineering & Technology
September 2024
Volume 1, (1) pp. 63-72

Performance Evaluation of Machine Learning Techniques on the microgrid and weather data

Ronak Solanki
Assistant Professor, Ganpat University, Mehsana-384012, Gujarat, India

Devang Pandya
Assistant Professor, Ganpat University, Mehsana-384012, Gujarat, India

Ketan Sarvakar
Assistant Professor, Ganpat University, Mehsana-384012, Gujarat, India

Abstract

Microgrid uses distributed power generation. Renewable resources have discontinuous pattern of power generation. Hence it is difficult to forecast wind or solar power accurately. Forecasting methods of renewable energy sources and its consumption have been investigated deeply for the prediction of demand in near future using conventional machine learning techniques. We attempt to analyze the performance of few well-known machine learning techniques for the prediction of target variables. This paper explores the overall performance of Support Vector Regression (SVR) with linear and non-linear kernels and Gradient Boosting Regression (GBR) in predictive modeling on microgrid and weather data. SVR with a linear kernel achieves awesome accuracy (99.9%) on a dataset in which a robust linear relationship exists between input functions and the target variable. This high accuracy highlights SVR's effectiveness for troubles characterized via clear linear trends. In assessment, non-linear SVR kernels, which includes polynomial and Radial Basis Function (RBF), carry out less correctly because of the lack of non-linear relationship among the participating variables. Gradient Boosting Regression, even as showing lower accuracy in a few cases, excels in coping with complicated and non-linear interactions, making it mainly nicely-acceptable for tasks related to difficult relationships, which includes climate forecasting. GBR's ensemble studying approach, which integrates multiple vulnerable decision trees, gives stepped up performance through decreasing bias and variance and taking complicated feature interactions. The outcomes imply that even as SVR with a linear kernel is good for easy, linear datasets, GBR is ultimate for situations with non-linear and multifaceted facts relationships. Thus, model selection must align with the dataset's characteristics and the specific necessities of the predictive project.

Keywords: Microgrid, Support Vector Regression (SVR), Gradient Boosting Regression (GBR), Non-linear Relationships, Ensemble Learning

Corresponding author
E-mail address:atrivedi8866@gmail.com

1. Introduction

Solar panels have been accessible enough to power homes for more than ten years. As these renewable energy sources become increasingly widespread, more and more people are using them to power homes. Whether it's an individual home, a neighborhood, or even a business park, the infrastructure to power the local energy needs is called a microgrid. A building or a neighborhood can maintain electricity during blackouts with the help of a microgrid, which is a tiny, localized electric power system. When solar panels are insufficient or there are other problems, these grids can also be connected to the main grid and take power from it. This means they can function while maintaining independence, as a component of a bigger system. Microgrids are capable of being isolated from the utility grid because to their combination of energy efficiency and generating technologies, switches, and controllers. In addition, if required, these control systems enable the microgrid to be related to the main power grid.

revolutionary process that has transformed the landscape of manufacturing and design. Unlike traditional subtractive manufacturing, which involves

1.1. Why Do We Need Microgrids?

These small, localized power networks have several uses. A microgrid is made up of residential buildings that produce power using solar energy. So do neighborhoods powered by renewable energy sources and electric batteries combined with the grid. Businesses that power their offices with solar panels and energy-efficient appliances are also part of a microgrid. There are larger alternative energy sources such as wind farms and industrial-size solar installations. While they are helping the environment, that power must often be transmitted much further than from fossil fuel plants, which can be built closer to demand. Because generated electricity is more dependable and localized, building microgrids helps solve that issue.

Homes, businesses, and even colleges or universities that have installed solar power and energy-efficient appliances can keep their essential services running, even when the main stream power network is down. Microgrids have the ability to stimulate the economy by supplying energy to isolated Tier 2 and Tier 3 areas and facilitating the expansion of small and medium-sized enterprises. For instance, after being installed for nearly two years, solar microgrids in Ladakh's isolated Himalayan areas have increased visitor revenue by roughly $24,000. India can now generate 375GW of power, 70% of which is from coal. A hybrid microgrid and central grid should be investigated to meet the demand as this system, ideally, undergoes plans for an overhaul and more renewable energy becomes available. This investigation involves optimal planning of power generation and its demand. It is pointed out that, using roughly calculated figures, 85,000 2MW microgrids could provide 170GW at a cost of $430 billion. In India, building a 1.6-GW coal plant costs $2 billion, to put this figure into context (Bhave, 2016)(Khan et al., 2019). Thus, generation is less expensive than microgrids; however, it is important to keep in mind that the cost of the latter is rapidly decreasing over time as greater scale is attained. Additionally, coal has greater government subsidies than renewable energy sources, and its high cost is not offset by the significant harm it causes to the local ecology and climate. Community microgrid is a local energy system which can function on its own as well as can be connected to the grid to provide surplus electricity to other users connected on the grid. Community mircogrid offers the superior advantage of locally controlling the production, storage and delivery of energy (Saha et al., 2023).

2. Brief literature review

There are several studies and experiments have been conducted to carry out microgrid optimization. multi-time scale optimal scheduling strategy integrated with robust optimization and model predictive control (MPC) is proposed by Yang and et al in (Yang et al., 2019). The optimal scheduling strategy is divided into two layers, daily optimization layer and intraday management layer. Based on the uncertainty

of PV and load, a two-stage robust optimization model with a time scale of 15 minutes is established in the daily optimization layer while time scale for intraday management layer is set to 5 minutes. Authors claimed that their proposed scheduling method guarantee the normal operation in various scenario and also succeeds to minimize the cost. In (Dumas et al., 2021), Authors proposed a two-layer approach with a value function to propagate information from operational planning to real-time optimization. The approach is tested by accounting for forecasting errors and high-resolution data monitored on-site corresponding to a "real-life" case. Also the proposed method is found compatible with the energy market because it can deal with imbalance, reserve and dynamic selling/purchasing prices. In (Wang et al., 2015), attempt has been done to minimize the cost by scheduling energy generations for microgrid intelligently. Along with unstable and unreliable renewable energy sources, heat and power generators are combined. Prepared model allows the net demand and heat demand distributions to fluctuate around their reference distributions. In (Javaid et al., 2018), Authors proposed generic Home Energy Management Control System (HEMCS) to efficiently schedule the household load and integrate RESs. For energy cost calculation, real-time pricing (RTP) and inclined block rate schemes are combined, because in case of only RTP, there is a possibility of building peaks during off-peak hours that may damage the entire power system. To coordinate various distributed energy resources (DERs) for an isolated power system, an ideal MPC-based control technique is suggested in (Xia et al., 2009). The intended MPC control strategy can achieve all of the performance goals, which include minimizing fuel expenses and variations in diesel generator power output (reducing mechanical wear and tear), minimizing expenses related to energy storage devices' short battery lives, and promoting regular system operation while optimizing wind penetration. According to the simulation tests, at high wind penetration levels, the closed loop MPC method outperforms the open loop look-ahead dispatch in terms of performance index.

3. Problem Identification

Energy has received extra attention since it is an essential component of any economy's social and economic growth. In conjunction with industrialization and globalization, the demand for energy has been rising steadily over the past few decades. It is anticipated that this need will continue to rise, rising by almost 30% between 2015 and 2035 in line with global economic expansion. Thus, the development of energy demand projections is necessary because precise forecasts of energy demand help policymakers optimize the energy supply schedule and offer insightful recommendations for organizing the operations of the energy supply system. Recently, artificial intelligence and machine learning techniques have immensely influenced the predictive analytics. Here in this article, we investigated the performance of machine learning techniques for predicting the recommended retail price on Daily Electricity Price & Demand Data of Victoria state, Australia and predicting surface temperature on Liege microgrid data.

3.1 Why are forecasts important?

Optimize energy use: By predicting generation from renewable sources likes wind and solar, forecasts allow for proactive planning of load control and possible grid imports.

Improve grid stability: Microgrids have the potential to enhance grid stability by anticipating the main grid's energy exchange requirements.

Reduce costs: Precise forecasts aid in minimizing reliance on expensive backup generators and maximizing usage of cost-effective energy sources.

4. Experimental setup and results

The first dataset is of electrical demand and forecasting which is obtained from Kaggle (Das, 2024) (Das & Wang, 2023). This dataset is related to Daily Electricity Price and Demand Data is a dataset which

contains daily price, demand and weather data in Australia's second largest state Victoria. The dataset covers 2016 days between 1 January 2015 and 6 October 2020. During some intraday intervals RRP was negative, so that energy producers were paying consumers rather than wise versa. Brief description of the dataset-1 is shown in Table 1.

Table 1. Daily Electricity Price and Demand Data, Victoria State-Australia (Dataset-1)

Field Name	Description
Date	The date of the recording
Demand	Total daily electricity demand in MWh
RRP	Recommended retail price in AUD / MWh
demand_pos_RRP	Total daily demand at positive RRP in MWh
RRP_positive	Averaged positive RRP, weighted by the corresponding intraday demand in AUD / MWh
demand_neg_RRP	Total daily demand at negative RRP in MWh
RRP_negative	Average negative RRP, weighted by the corresponding intraday demand in AUD / MWh
frac_at_neg_RRP	Fraction of the day when the demand was traded at negative RRP
min_temperature	Minimum temperature during the day in Celsius
max_temperature	Maximum temperature during the day in Celsius
solar_exposure	Total daily sunlight energy in MJ/m^2
Rainfall	Daily rainfall in mm
school_day	If students were at school on that day
holiday	If the day was a state or national holiday

We begin with the estimation of the daily energy price on the Dataset-1 with the inputs of all the features except the school_day and holiday. Regression analysis (Freedman, 2009) is a supervised learning technique which is based on statistical concepts. It provides us sound platform to estimate the relationships between a response variable and one or more independent variables.

Authors in (Freedman, 2009) (*Victoria Electricity Data Forecasting*, n.d.) (Das, 2024) have shown the prediction of daily energy price in Victoria energy using linear regression model without rolling mean and with a rolling window. We have compared and investigated the results of daily energy price prediction with Support vector repressor and Gradient boosting. Support Vector Regression (SVR) extends the principles of Support Vector Machines (SVM) to regression problems, offering a powerful tool for predicting continuous outputs. By leveraging various kernels such as quadratic, radial basis function, and sigmoid, SVR Model can handle complex and non-linear relationships in the data.

We begin our experiments with Support vector regression with linear kernel. Obvious reason for doing this, it is observed more linear relationship form the available features in the dataset. SVR with linear kernel has offered very high accuracy of prediction which is clearly depicted in Figure 1. In Figure 1, we can see that predicted curve is found exactly overlapping with actual values.

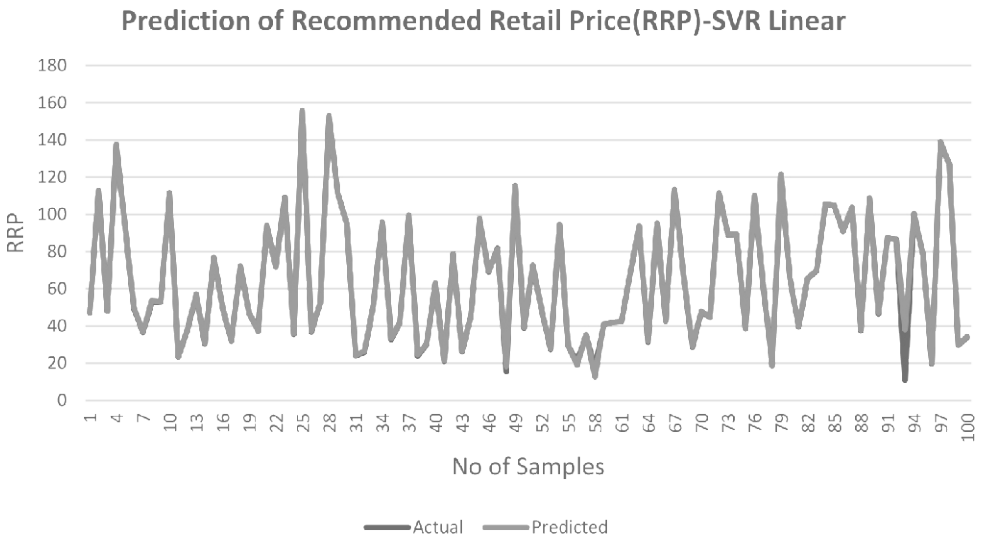


Figure 1. Recommended retail price(RRP) prediction over 100 samples, SVR-Linear

Similarly, to understand the behavior of the dataset, further experiments were conducted with SVR polynomial kernel and RBF. Figure 2 and Figure 3 depict the predicted values over actual value for first 100 samples. During the training phase, we considered first 12 features and ignoring last two features namely; school_day and holiday. These features have minimum impact hence they are avoided. However, it is clearly observed that SVR with polynomial and RBF kernel are found much inaccurate.

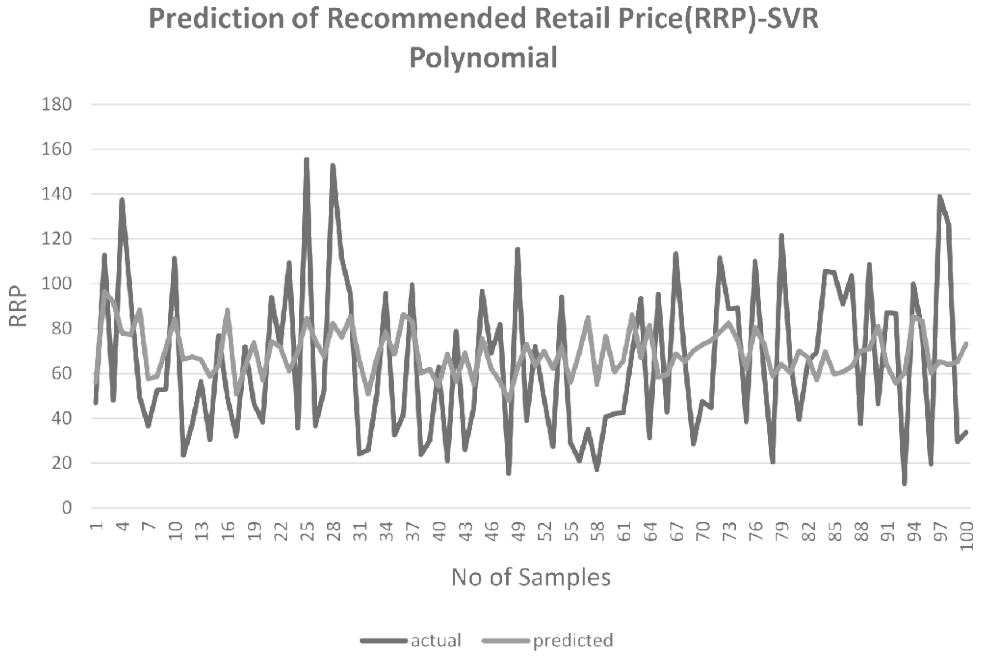


Figure 2. Recommended retail price(RRP) prediction over 100 samples, SVR(Kernel=Polynomial)

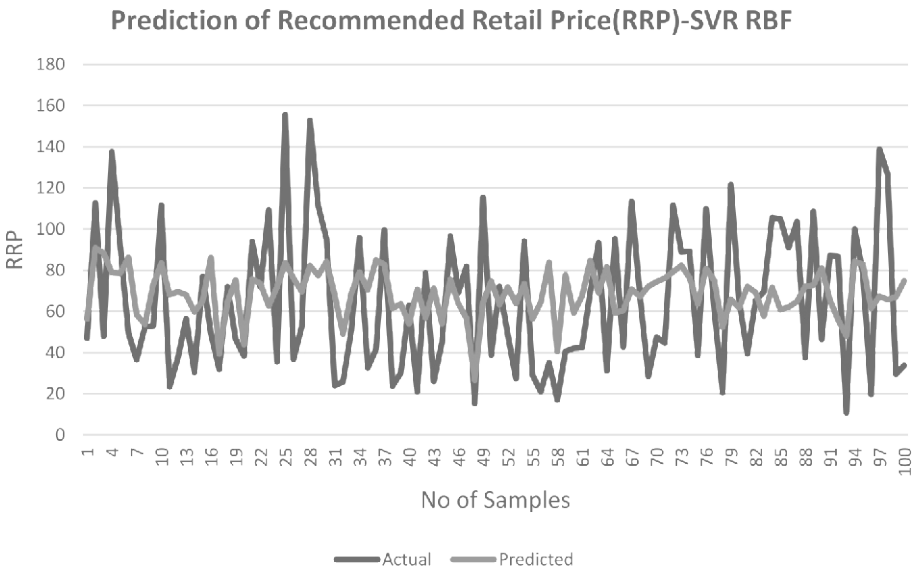


Figure 3. Recommended retail price(RRP) prediction over 100 samples, SVR(Kernel=RBF)

Kernel parameters for the experiments have been fixed as shown in Table 2.

Table 2. Kernel parametric values

Kernels/Values	Values
SVR-Linear	C=1
SVR-Polynomial	C=1, degree=3
SVR-RBF	C=1, gamma='scale'(default)

We reported the following error measures on Dataset-1 with these techniques as shown below in Table 3.

Table 3. Reported error measure (MSE & RMSE) on Dataset-1

Technique/Error Measure	MSE	RMSE
SVR-Linear	3.3871	1.8404
SVR-Polynomial	135.33	18316.2
SVR-RBF	18746	18897117

Gradient boost regression:

Gradient Boosting is a popular boosting technique in machine learning used for classification and regression tasks. It is a one kind of ensemble Learning method which trains the model sequentially and each new model tries to correct the previous model. Purpose is to combine various weak learners and gradually evolving strong learner. We have set the learning rate of 0.1 and depth is set to 3. Figure 4 depicts the prediction of RRP with the use of gradient boosting regression. Technique reported an RMSE of 71 and accuracy of 73.14%

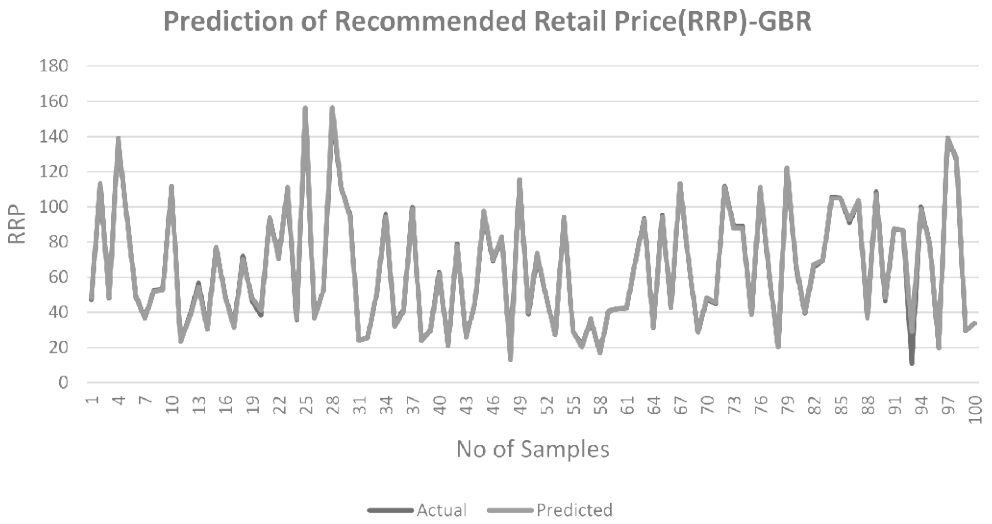


Figure 4. Recommended retail price (RRP) prediction over 100 samples, GBR

Combining all the results on the dataset, it is clearly observed that SVR with linear marginally performs better than linear regression. Hence, SVR emerges powerful than rest techniques which were applied to conduct these experiments on the dataset-1. Comparative results are shown in Figure 5.

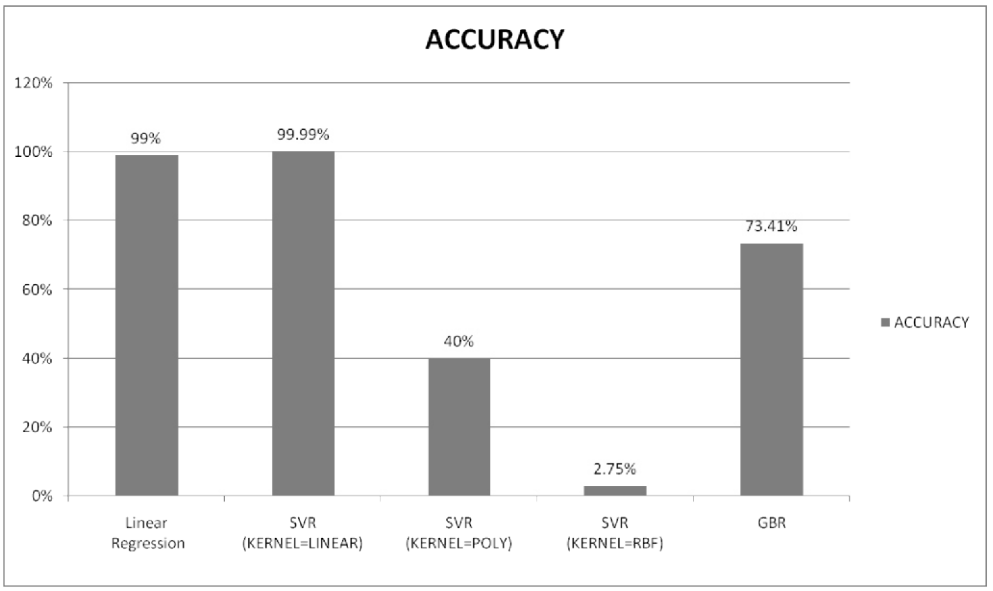


Figure 5. Comparative accuracy reported on Dataset-1

For further validation, we conducted experiments on another dataset "Liege Microgrid Open Data" obtained from Kaggle (*Victoria Electricity Data Forecasting*, n.d.). This dataset contains the weather data. This dataset contains the following field.

Table 4. Dataset Description

Sr.No	Field Name	Description
1	CD	Low clouds (0 to 1)
2	CM	Medium clouds (0 to 1)
3	CU	High clouds (0 to 1)
4	PREC	Precipitation (mm / 15 min)
5	RH2m	relative humidity (%)
6	SNOW	Snow height (mm)
7	ST	Surface Temperature (°C)
8	SWD	Global Horizontal Irradiance (W/m2)
9	SWDtop	Total Solar Irradiance at the top
10	TT2M	Temperature 2 meters above the ground (°C)
11	WS100m	Wind speed at 100m from the ground (m/s)
12	WS10m	Wind speed at 10m from the ground (m/s)

Weather forecast prediction directly helps us for decisions like minimize the total energy cost and revenues, the peak consumption and injection related costs, and plans for reserve requirements. Thus, this dataset is analyzed to investigate the relationship among the features. Surface temperature was considered as a target variable in against of the other input variables. The rise in surface temperature will surely raise the demand of electricity and will eventually have the impact on Microgrid planning. Again same former techniques have been applied and tested on this dataset. Following error measure have been observed on Dataset-2 and are mentioned in Table 5.

Table 5. Reported error measure (MSE & RMSE) on Dataset-2

Technique/Error Measure	MSE	RMSE
SVR-Linear	1.315	1.147
SVR-Polynomial	42.203	6.496
SVR-RBF	0.794	0.892

Also, RMSE with Gradient Boosting Rregessor(GBR) is observed 0.869 with much accurate prediction. It is clearly observed with the experiments that SVM with linear kernel has outperformed the other techniques. Comparative results of accuracy is shown in Figure 6.

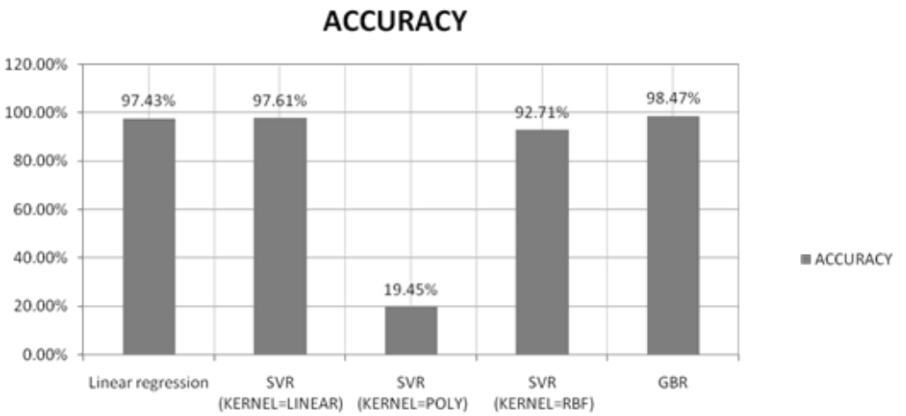


Figure 6. Accuracy of various techniques reported on Dataset-2

5. Conclusion

It is challenging to deal with uncertainty of power generations from the renewable energy sources. Forecasting of the demand can surely help to reduce the impact of uncertainty associated with power generations and load on the microgrid. We investigated the predictive ability of several machine learning techniques on the available microgrid and weather datasets. SVR(with linear kernel) on datsaset-1 establishes strong linear connection between input characteristics and the target variable, SVR with a linear kernel is probably yielding good results (99.99% accuracy). This means that SVR with a linear kernel is a good option because of its high accuracy in capturing linear trends, interpretability, and simplicity. Due to their inability to properly match the underlying data structure, the polynomial and RBF kernels, which introduce non-linearity, performed worse. Though Gradient Boosting Regression has a lower reported accuracy, it may be a better option if the data contains more complicated relationships or non-linear interactions since it can handle non-linear relationships more successfully. The selection of these models ultimately relies on the characteristics of your dataset and the particular objectives of your predictive modeling assignment. Gradient boost regression is a strong learner which is created through assembling of various weak learners. We have observed on dataset-2 that SVR linear offers good accuracy however with a smaller number of parameters its efficiency decreases. Predicting the weather requires understanding of many intricate, non-linear relationships between many variables. More accurately than a single model like SVR with a linear or polynomial kernel, GBR's ensemble technique can capture these relationships. To create a powerful predictor, GBR weaves together a number of weak learners (decision trees). GBR is better at modeling the intricate interactions between factors like wind speed, temperature, and humidity in meteorological data. Gradient Boost Regression performs better in this dataset because, in summary, it can handle complex, non-linear interactions and is well-suited for weather forecasting tasks.

References

[1] Bhave, M. P. (2016). *The microgrid revolution: business strategies for next-generation electricity* (Bloomsbury Publishing USA (Ed.)). Bloomsbury Publishing USA.

[2] Das, S. (2024). *Modeling and Mitigating Power Grid Vulnerabilities: A Comprehensive Analysis of Renewable Energy Integration, Cascading Failures, and Microgrid Resilience*.

[3] Das, S., & Wang, Z. (2023). Optimizing Microgrid Energy Management Systems with Variable Renewable Energy Penetration: Analysis of Data Loss Effects. *Proceedings of the Future Technologies Conference*, 357–374.

[4] Dumas, J., Dakir, S., Liu, C., & Cornélusse, B. (2021). Coordination of operational planning and real-time optimization in microgrids. *Electric Power Systems Research*, *190*, 106634. https://doi.org/https://doi.org/10.1016/j.epsr.2020.106634

[5] Freedman, D. A. (2009). *Statistical models: theory and practice*. cambridge university press.

[6] Javaid, N., Hafeez, G., Iqbal, S., Alrajeh, N., Alabed, M. S., & Guizani, M. (2018). Energy Efficient Integration of Renewable Energy Sources in the Smart Grid for Demand Side Management. *IEEE Access*, *6*, 77077–77096. https://doi.org/10.1109/ACCESS.2018.2866461

[7] Khan, S., Javaid, N., Chand, A., Khan, A. B. M., Rashid, F., & Afridi, I. U. (2019). Electricity load forecasting for each day of week using deep CNN. *Web, Artificial Intelligence and Network Applications: Proceedings of the Workshops of the 33rd International Conference on Advanced Information Networking and Applications (WAINA-2019) 33*, 1107–1119.

[8] Saha, T., Haque, A., Halim, M. A., & Hossain, M. M. (2023). A Review on Energy Management of Community Microgrid with the use of Adaptable Renewable Energy Sources. *International Journal of Robotics and Control Systems; Vol 3, No 4 (2023)DO - 10.31763/Ijrcs.V3i4.1009* . https://pubs2.ascee.org/index.php/IJRCS/article/view/1009%7Cto_array%3A0

[9] *Victoria electricity data forecasting*. https://www.kaggle.com/code/rabiatibrahim/victoria-electricity-data-forecasting/comments

[10] Wang, R., Wang, P., & Xiao, G. (2015). A robust optimization approach for energy generation scheduling in microgrids. *Energy Conversion and Management*, *106*, 597–607. https://doi.org/https://doi.org/10.1016/j.enconman.2015.09.066

[11] Xia, X., Zhang, J., & Elaiw, A. (2009). A Model Predictive Control approach to dynamic economic dispatch problem. *2009 IEEE Bucharest PowerTech*, 1–7. https://doi.org/10.1109/PTC.2009.5282270

[12] Yang, H., Pu, Y., Qiu, Y., Li, Q., & Chen, W. (2019). Multi-Time Scale Integration of Robust Optimization with MPC for Islanded Hydrogen-Based Microgrid. *2019 IEEE Sustainable Power and Energy Conference (ISPEC)*, 1163–1168. https://doi.org/10.1109/iSPEC48194.2019. 8975221