Investigating the effects of hyperparameter sensitivity on machine learning algorithms for PV forecasting

Muhammad Ehtsham[1,*], *Marianna Rotilio*[1], Federica Cucchiella[2], *Gianni Di Giovanni*[1], and Domenico Schettini[2]

[1]Department of Civil, Construction-Architectural and Environmental Engineering, University of L'Aquila, Italy
[2]Department of Industrial and Information Engineering and Economics, University of L'Aquila, Italy

Abstract. Machine Learning (ML) models have been introduced in the past, and users have debated whether to tune the hyperparameters of the models. This study investigates the effects of tuning the hyperparameters of the ML models and summarizes the models that are most sensitive to hyperparameter tuning. This study leveraged the historic energy production data of two already operational PV plants. Four state-of-the-art ML models, namely Decision Trees (DT), Random Forest (RF), K-Nearest Neighbors (KNN), and Support Vector Regression (SVR) were investigated. All the ML models were trained with the same training features (meteorological estimates) obtained from the National Aeronautics and Space Administration's (NASA) Power project, with the daily PV energy production selected as the target variable. Models were developed and executed with default and tuned hyperparameters using an 85-15% train-test split. The results revealed that all the models showed improved performance with the tuned hyperparameters. However, the DT and SVR models depicted significantly improved RMSE after tuning of the hyperparameters. The RMSE of DT improved from 111 kWh/d to 75 kWh/d for one plant and from 442 kWh/d to 270 kWh/d for the second plant after tuning the hyperparameters. Similarly, the RMSE of SVR improved from 59 kWh/d to 50 kWh/d in the first case, and in the second case, the improvement of RMSE from 536 kWh/d to 294 kWh/d was observed. The efficiency of the RF and KNN models also improved to some extent after tuning, but the RMSE closely agreed with the default hyperparameters in one case study, making the RF and KNN less prone to hyperparameter sensitivity. This study concluded with the finding that it is necessary to tune the hyperparameters of the DT and SVR models, specifically for energy forecasting. Moreover, the results of this study also highlight the significance of meteorological estimates from NASA's Power project, as models successfully discerned the complex energy forecast patterns. The dataset is deemed suitable for energy forecasting for areas with sparse ground-based observatories and may serve as a baseline dataset for training the ML models.

[*] Corresponding author: muhammad.ehtsham@graduate.univaq.it

1 Introduction

The need for a transition towards more eco-friendly sources of energy has been highlighted by strong evidence of climate change in recent years. Although scientists and researchers have highlighted the scenario in recent decades, it was mostly limited to scientific communities. However, in recent years the climate change scenario has imposed strong effects on weather patterns and significant events have been reported across the globe [1]. These extreme events have given mankind a final warning and even the laymen are now concerned about the climate change scenario. The main factor triggering climate change is the huge reliance on fossil fuels [2], but now people around the globe are developing a sense of ownership towards climate sustainability and adopting more clean and green sources of energy.

The most attractive source of renewable energy is PV energy, the main attraction of PV energy is the ease of installation, operation, and flexibility in the investment [3, 4]. One can install a very small PV plant to meet the specific energy demands of a household. On the other hand, industries and public sector agencies can install PV plants on a large industrial scale. This can fulfill the needs of these industries and the excess amount of energy can be sold to the national grid which results in less stress on other sources of energy.

However, the main shortcoming of the PV energy is its unpredictability. The amount of PV energy that may be produced in the next few weeks or a few months is highly unpredictable [5]. Although there are many models available for calculating rough estimates of PV energy forecasts, these estimates are crude as they utilize the basic metrological data, and the data from the energy produced by the plant is not utilized during the forecasting process. However, the historic energy production can play a vital role in the forecast of the energy production by a plant, because it is highlighted by many researchers that the actual energy production by a plant is different from the energy calculations done in a controlled environment or by the computer models.

ML is a subdivision of Artificial intelligence, which is undoubtedly the most debated topic of the past few years [6]. ML models are utilized to make predictions based on the historical data. In contrast to traditional methods, in which forecasts were made upon specific set rules, the ML has been showing promising results as it makes decisions based on the already recorded data points. The application of ML models whether they are supervised or unsupervised, does not require top-notch knowledge of the coding. The simplicity of these models is the main attraction among scientists from various fields to adopt them and validate their application in their respective fields. Researchers have been utilizing ML for forecasting PV energy plants at various time horizons utilizing different datasets and models [7–9].

However, there is a need to address one specific topic, which is the tuning of hyperparameters in ML models [10–12]. Almost every ML model has a specific set of hyperparameters controlling the application of the model. Some researchers have concluded that the optimization of the ML model is necessary and results in improved performance of the models [13–15]. Others highlighted that the optimization of the models did not affect the efficiency of the models and even in some cases resulted in decreased efficiency [16, 17].

This research work is conducted to focus on the hyperparameter tuning of the ML models. The historic PV energy production data obtained from already operational PV plants is utilized to train and test the ML models. The effectivity of the models is assessed on the default and the tuned hyperparameters and the conclusions are drawn upon the effectivity of both techniques. The results may serve as a baseline for the students and researchers associated with PV energy forecasting and create an opportunity to visualize the effect and need of tuning the hyperparameters.

2 Methodology

Different researchers have adopted different techniques for training and testing the ML models for forecasting variables. In the following paragraphs, the methodology of this research work is explained. Moreover, Figure 1 shows the methodological flowchart of the process adopted.

The research works started with data acquisition the daily energy production data of already operational PV plants was accessed from the database, specifically tailored Python queries combined with the energy production data of inverters connected with a single PV plant. The operational data from two operational PV plants was collected and analyzed for any anomalies and missing data points in the dataset. The analysis revealed the missing data points were less than 3 percent of the total data. The missing data was replaced with values using interpolation techniques. Moreover, data was analyzed for outliers and the rows were deleted where anomalies were detected.

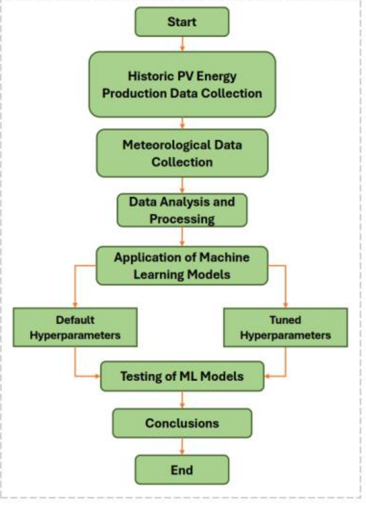

Fig. 1. Methodological flowchart of research work

In the second phase, the metrological data from the web sources was collected for training the ML models as it is necessary to train the models on the satellite data, because the ground data is limited and availability of the data is not so easy, especially the students and researchers associated with the PV energy forecasting. Moreover, some remote areas make it more challenging, as the recording stations are few and availability takes time and consumes resources. Hence, in this research work we relied on the satellite metrological resources and assessed the efficiency of the models using notable satellite sources, the data acquired which was later used for training the models has been discussed in detail in the next sections.

In the third phase, the Machine learning models that depicted attractive results in various scientific fields were shortlisted and the scripts of the codes were developed using Python in PyCharm. Various libraries needed for the implementation, results visualization, graph plotting, and performance metrics calculation were imported. Some of the libraries which were imported were “pandas as pd”, “numpy as np”, “matplotlib.pyplot as plt”. Moreover, the models were imported from the scikit-learn library using commands:

- “from sklearn.svm import SVR”
- “from sklearn.tree import DecisionTreeRegressor”
- “from sklearn.ensemble import RandomForestRegressor”
- “from sklearn.neighbors import KNeighborsRegressor”

In the fourth phase of the research, the ML models were executed to forecast the energy using default hyperparameters. After analyzing the models for energy forecasting using default hyperparameters, the authors trained the models on tuned hyperparameters. Different techniques have been introduced for tuning the ML models, some of the commonly used techniques are:

- Grid search
- Random search
- Gradient-based optimization
- Bayesian optimization
- Babysitting or 'Trial and Error'

The optimization of hyperparameters was done using Babysitting Hyperparameter Optimization (BHO). This technique is simple and is also called 'Trial and Error', its simplicity makes it very famous among students and researchers [1]. The method is implemented fully manual basis and various combinations of hyperparameters are used for the best possible output results the user has to always check the results of previous hyperparameters with the new inputs and at the end he comes with an optimum set of hyperparameters. This method is time-consuming and requires sufficient prior experience with the models [2].

In the last phase of the research, the conclusions were drawn, performance metrics of the default and tuned hyperparameters were reported for the various models and the effectivity of each model was discussed and concluded.

3 Application of ML Models

All the selected ML models were executed in PyCharm using the Python programming language as previously reported in the methodology section. The input features were mainly the metrological estimates from the satellite resources. The input features were:

- Wind speed
- Precipitation
- Specific Humidity
- All Sky Surface Shortwave Downward Irradiance
- Earth Skin Temperature
- Standard Global Solar Radiation
- Nominal power

Daily PV energy production in kWh was selected as the target variable. All the models were trained using the same set of features to compare the efficiency of each model on the same basis. Data was analyzed for the consistency of the data and outliers and the rows with missing data points were removed before loading the dataset for training the ML models. First, 85 percent of the data rows were used for training the models, and the last 15 percent of the rows were used to test the efficiency of the models, this separation of the dataset overruled the possibility of overfitting and created the opportunity for fair comparison in light of PV energy forecasts.

Table 1 depicts the details of the PV plants, plant specification; number of inverters, date of commencement of plant, nominal power of pant, and data availability period.

Table 1. Details of the PV Plants

Code of	**Plant Specification**

Plant	*Nominal power*	*Number of Inverters*	*Date of Commencement*	*Location*
COA	405 kWp	14	2012-06-29	Mosciano Sant'Angelo (TE)
CRB	51.6 kWp	14	2010-10-15	L'Aquila (AQ)

The hyperparameters control the models' complexity, generalization, and performance. For DT and RF models, depth, sample splitting, and leaf sizes were tuned to control overfitting and enhance decision-making. SVR was tuned with a linear kernel and regularization to balance error minimization with model complexity. In KNN, the number of neighbors was adjusted for smooth decision boundaries and improved generalization. Table 2 depicts the details of all the models.

Table 2. Details of the ML Models CRB

Model	Training and Testing Dataset		
	Train-Test Split	*Training Period*	*Testing Period*
Tuned Hyperparameters (max_depth=25, min_samples_split=75, min_samples_leaf=20)			
DT	85-15%	01/01/18 to 20/03/22	20/03/22 to 30/11/22
Tuned Hyperparameters (n_estimators=70, random_state=30, max_depth=20, min_samples_split=12, min_samples_leaf=67)			
RF	85-15%	01/01/18 to 20/03/22	20/03/22 to 30/11/22
Tuned Hyperparameters (kernel='linear', C=2.0, gamma='scale', degree=3)			
SVR	85-15%	01/01/18 to 20/03/22	20/03/22 to 30/11/22
Tuned Hyperparameters (n_neighbors=55)			
KNN	85-15%	01/01/18 to 20/03/22	20/03/22 to 30/11/22

Table 3. Details of the ml Models COA

Model	Training and Testing Dataset		
	Train-Test Split	*Training Period*	*Testing Period*
Tuned Hyperparameters (max_depth=35, min_samples_split=5, min_samples_leaf=20)			
DT	85-15%	01/01/2018 to 04/11/2021	05/11/2021 to 31/08/2022
Tuned Hyperparameters (n_estimators=20, random_state=30, max_depth=40, min_samples_split=52, min_samples_leaf=7)			
RF	85-15%	01/01/2018 to 04/11/2021	05/11/2021 to 31/08/2022
Tuned Hyperparameters (kernel='linear', C=0.2, gamma='scale', degree=3)			
SVR	85-15%	01/01/2018 to 04/11/2021	05/11/2021 to 31/08/2022

Model	Training and Testing Dataset		
	Train-Test Split	***Training Period***	***Testing Period***
Tuned Hyperparameters (n_neighbors=50)			
KNN	85-15%	01/01/2018 to 04/11/2021	05/11/2021 to 31/08/2022

4 Results and Discussion

Evaluation of ML models is important before selection of the best suitable Model. The effectiveness of the ML models is greatly hindered by the hyperparameters, this study leveraged the historic energy production data obtained from two already-operated PV plants and applied four state-of-the-art ML models for the energy forecast. Apart from exploiting the sensitivity of the ML models concerning hyperparameters, this study also offers the applicability potential of the dataset provided by "The NASA Langley Research Centre POWER Project" [3]. The applicability of the satellite data for ML models is specifically important in terms of PV energy, as the installations are spatially diverse and installed at various places starting from city centers to the most remote areas, making the availability of the ground-based data difficult. As ground-based observatories are often far-flung, and the acquisition of the data is always time-consuming due to the documentation required by the officials involved in the handling of the data. These issues are eliminated in satellite-based resources, as they provide the datasets from a user-friendly interface and the acquisition of the data is done within minutes. The results of this study revealed insights into the impact of tuning the hyperparameters on the performance of the models. All the models were trained with 85 percent of the initial data rows and were tested on the last 15 percent of the data rows.

RMSE is one of the best performance metrics to check the efficiency of the ML models. Figures 2 and 3 depict the RMSE of both plants for the testing dataset and table 4 depicts the RMSE of all the selected models. It can be interpreted from the figures that the RMSE has decreased in almost all the cases with tuned hyperparameters. Specifically for the DT and SVR the change of RMSE is significant as compared to RF and KNN models. DT has shown significantly improved performance for both plants. However, the abnormal trend was observed with default hyperparameters in the case of SVR, where we observed the worst RMSE among all the investigated models and RMSE significantly improved after the tuning of the hyperparameters. This trend highlighted the sensitivity of the SVR models with respect to hyperparameter tuning, making the tuning of the hyperparameters unavoidable when employing the SVR model for PV energy forecasting.

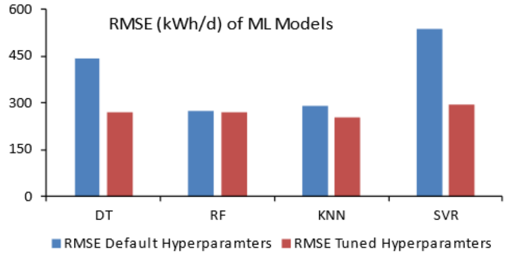

Fig. 2. Comparison of RMSE of ML models using Default and Tuned Hyperparameters, Plant COA

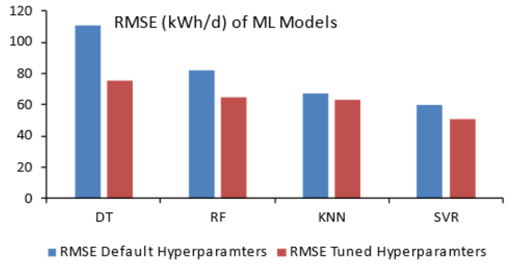

Fig. 3. Comparison of RMSE of ML models using Default and Tuned Hyperparameters, Plant CRB

Lastly, the comparison of energy forecasts on the testing data has been depicted in Figures 4 to Figure 11. All the figures depict the energy forecast of the models using the default and tuned hyperparameters, creating the opportunity for the readers to visualize how the tuning of the hyperparameters changed the forecast pattern of the models. Specifically, the forecast generated by the SVR in the case of plant A depicts huge anomalies and the margin of improvement which can be achieved by tuning the hyperparameters of the SVR. Overall, it can be interpreted from the figures that RF and KNN models were able to capture the intrusive patterns of energy forecasts.

Table 4. RMSE comparison of both Plants

ML Model	**RMSE kWh/d**			
	Plant A		**Plant B**	
	Default Hyperparameters	***Tuned Hyperparameters***	***Default Hyperparameters***	***Tuned Hyperparameters***
DT	442	271	111	75
RF	276	269	82	65
KNN	292	253	67	63
SVR	536	295	60	51

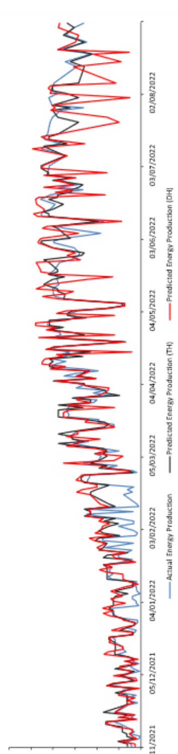

Fig. 4. Actual vs Predicted Energy, DT Model, Plant COA

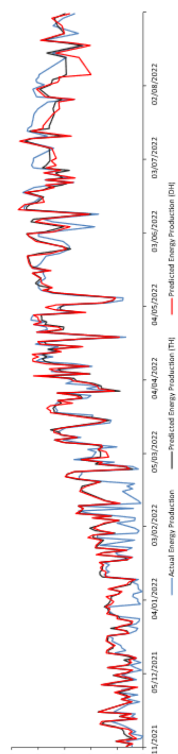

Fig. 5. Actual vs Predicted Energy, RF Model, Plant COA

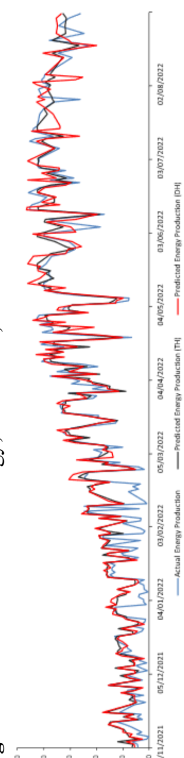

Fig. 6. Actual vs Predicted Energy, KNN Model, Plant COA

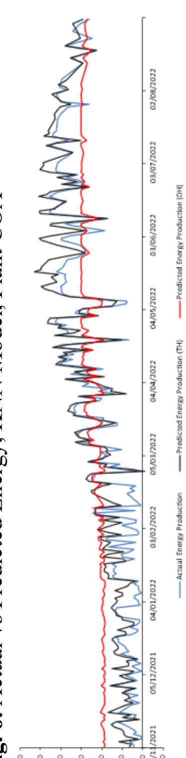

Fig. 7. Actual vs Predicted Energy, SVR Model, Plant COA

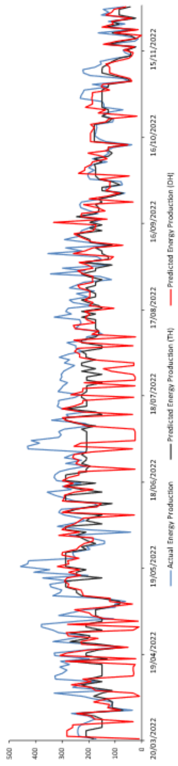

Fig. 8. Actual vs Predicted Energy, DT Model, Plant CRB

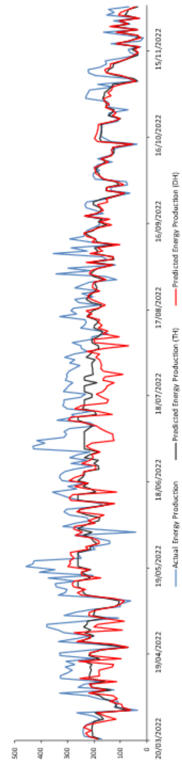

Fig. 9. Actual vs Predicted Energy, KNN Model, Plant CRB

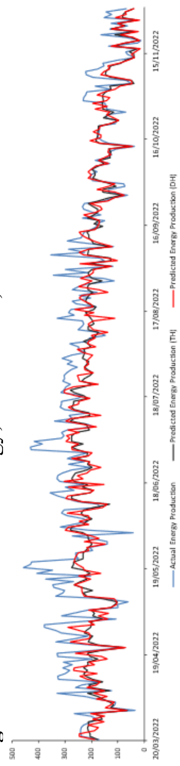

Fig. 10. Actual vs Predicted Energy, KNN Model, Plant CRB

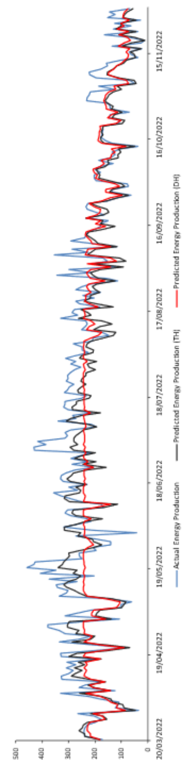

Fig. 11. Actual vs Predicted Energy, SVR Model, Plant CRB

5 Conclusions

This study was conducted to summarize the effect of tuning the hyperparameters on the ML models. Specifically focusing on the use of ML models for the PV energy forecasting. The study utilized the historical energy production data of two already operating PV plants. One is a small-scale system of 51.6 kWp, and the other is a large-scale system of 405 kWp. The study used the metrological estimates from NASA's POWER project. Specific parameters affecting the generation of PV energy were acquired from the web database. This dataset is an important source of metrological estimates and has proved its significance among engineering and scientific communities. Training features (comprising of variables obtained from NASA's power project were employed to train the ML models to generate a forecast of the Daily PV energy production, which was selected as the target variable.

The results of the study depicted important insights into the use of the ML models with and without tuning the hyperparameters. Patterns revealed that it is always important to adjust the hyperparameters of the models and compare them with the default hyperparameters, as they change with the dataset which one is dealing with. In some cases, the default hyperparameters showed satisfactory results; on the other hand, the results were significantly improved after tuning the hyperparameters.

Specifically, users employing DT and SVR should carefully analyze the hyperparameters as both models are found prone to high errors with default hyperparameters. The magnitude of error is related to the type and nature of the training parameters. However, as far as this study is concerned, both models depicted significantly improved performance after tuning the hyperparameters. RF and KNN models depicted slightly better results with tuned hyperparameters in both case studies, however, they are found less sensitive to hyperparameter tuning as compared to the aforementioned two models. The comparison of actual vs predicted energy production of the testing dataset depicted in Figures 4 to 11, highlights the importance of dataset of the NASA's power project, as the models were able to capture the intrusive patterns of the energy forecast.

Future research directions include the real-time verification of machine learning and deep learning models using databases that provide historical data and forecasts of meteorological conditions. This approach will undoubtedly enhance the effective management of PV energy sources, as decision-makers will have more reliable data to make informed interventions and decisions in advance.

This research is carried out as part of the XXXVIII Cycle of the PhD programme in Civil, Building-Architecture and Environmental Engineering at the University of L'Aquila. The research topic is "Monitoring System for the Production of Renewable Energy and Data Acquisition System" co-financed by the company Solis Spa. The authors express their gratitude to Solis SpA for the economic support and for their invaluable cooperation in this study. This encompasses providing technical and production data for BIPVs and BAPVs, supplying the necessary computer processing system and required software, as well as the contributions of the monitoring team and the company's technical head in validating the research findings. The authors would like to express their gratitude to NASA Langley Research Center for providing the metrological data, which were obtained from the NASA Langley Research Center (LaRC) POWER Project funded through the NASA Earth Science/Applied Science Program.

References

[1] Raihan, A.: A review of the global climate change impacts, adaptation strategies, and mitigation options in the socio-economic and environmental sectors. Journal of Environmental Science and Economics. 2, 36–58 (2023).

[2] Abbasi, K.R., Shahbaz, M., Zhang, J., Irfan, M., Alvarado, R.: Analyze the environmental sustainability factors of China: The role of fossil fuel energy and renewable energy. Renew Energy. 187, 390–402 (2022).
[3] Cucchiella, F., Rotilio, M., Capannolo, L., De Berardinis, P.: Technical, economic and environmental assessment towards the sustainable goals of photovoltaic systems. Renewable and Sustainable Energy Reviews. 188, 113879 (2023).
[4] Cucchiella, F., Rotilio, M., Barile, G., De Berardinis, P., Leoni, A., Ragnoli, M., Scarsella, M., Stornelli, V.: Renovation Wave: a bioeconomy panel produced with waste. J Clean Prod. 142868 (2024).
[5] Krein, P.T., Galtieri, J.A.: Active management of photovoltaic system variability with power electronics. IEEE J Emerg Sel Top Power Electron. 9, 6507–6523 (2020).
[6] Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R.: Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information fusion. 58, 82–115 (2020).
[7] Qu, Y., Xu, J., Sun, Y., Liu, D.: A temporal distributed hybrid deep learning model for day-ahead distributed PV power forecasting. Appl Energy. 304, 117704 (2021).
[8] ÖNAL, Y.: Gaussian kernel based SVR model for short-term photovoltaic MPP power prediction. Computer Systems Science and Engineering. 41, (2022).
[9] Di Giovanni, G., Rotilio, M., Giusti, L., Ehtsham, M.: Exploiting building information modeling and machine learning for optimizing rooftop photovoltaic systems. Energy Build. 313, 114250 (2024).
[10] Probst, P., Boulesteix, A.-L., Bischl, B.: Tunability: Importance of hyperparameters of machine learning algorithms. Journal of Machine Learning Research. 20, 1–32 (2019).
[11] Weerts, H.J.P., Mueller, A.C., Vanschoren, J.: Importance of tuning hyperparameters of machine learning algorithms. arXiv preprint arXiv:2007.07588. (2020).
[12] Hossain, M.R., Timmer, D.: Machine learning model optimization with hyper parameter tuning approach. Global Journal of Computer Science and Technology. 21, 7–13 (2021).
[13] Karl, F., Pielok, T., Moosbauer, J., Pfisterer, F., Coors, S., Binder, M., Schneider, L., Thomas, J., Richter, J., Lang, M.: Multi-Objective Hyperparameter Optimization in Machine Learning—An Overview. ACM Transactions on Evolutionary Learning and Optimization. 3, 1–50 (2023).
[14] Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A.: Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. Wiley Interdiscip Rev Data Min Knowl Discov. 13, e1484 (2023).
[15] Luo, G.: A review of automatic selection methods for machine learning algorithms and hyper-parameter values. Network Modeling Analysis in Health Informatics and Bioinformatics. 5, 1–16 (2016).
[16] Ridd, P., Giraud-Carrier, C.G.: Using Metalearning to Predict When Parameter Optimization Is Likely to Improve Classification Accuracy. In: MetaSel@ ECAI. pp. 18–23 (2014).
[17] Hoque, K.E., Aljamaan, H.: Impact of hyperparameter tuning on machine learning models in stock price forecasting. IEEE Access. 9, 163815–163830 (2021).
[18] Abreu, S.: Automated architecture design for deep neural networks. arXiv preprint arXiv:1908.10714. (2019).
[19] Yang, L., Shami, A.: On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing. 415, 295–316 (2020).
[20] NASA, "Power Data Access Viewer: Prediction of Worldwide Energy Resource." [Online]. Available: https://power.larc.nasa.gov/data-access-viewer/. [Accessed: Aug. 02, 2024].