Electricity Energy Demand Prediction Using Computational Intelligence Techniques

ARTICLE HISTORY

Received 06 February 2024
Accepted 19 April 2024

Bruno da S. Macêdo
Federal University of Lavras
Lavras, Brazil
bruno.macedo2@estudante.ufla.br
ORCID: 0009-0009-4375-8464

Camila Martins Saporetti
Polytechnic Institute, Rio de Janeiro State University
Nova Friburgo, Brazil
camila.saporetti@iprj.uerj.br
ORCID: 0000-0002-8145-7074

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2024

B. da S. Macêdo, C. M. Saporetti, "Electricity Energy Demand Prediction Using Computational Intelligence Techniques", Latin-American Journal of Computing (LAJC), vol. 11, no. 2, 2024.

Electricity Energy Demand Prediction Using Computational Intelligence Techniques

Bruno da S. Macêdo
Systems and Automation Engineering Graduate Program
Federal University of Lavras
Lavras, Brazil
bruno.macedo2@estudante.ufla.br
ORCID: 0009-0009-4375-8464

Camila Martins Saporetti
Department of Computational Modeling
Polytechnic Institute, Rio de Janeiro State University
Nova Friburgo, Brazil
camila.saporetti@iprj.uerj.br
ORCID: 0000-0002-8145-7074

***Abstract*—** Energy is an important pillar for the economic development of a country. The demand for electricity is something that continues to grow, one of the contributing factors is the emergence of various technological equipment and the consequent use by the population. There are several resources that can be exploited to generate electricity, with hydroelectric power stations being one of the most used resources. As electrical energy cannot be stored, there is a need to estimate its consumption, looking for a way to meet this energy demand. In this context, this study seeks to apply machine learning techniques, using the Grey Wolf Optimization (GWO) meta-heuristic to optimize regression models, to predict the demand for electricity in Brazil, and it aims to estimate how much energy should be produced. For the predictions, the period between the years 2017 to 2022 was used, totaling around 2,190 samples. The methodology involves pre-processing, crossvalidation, parameters optimization and regression. The results show that Random Forest performed well in the experiments carried out, presenting a coefficient of determination (R^2) of 0.8751, Root Mean Squared Error (RMSE) of 0.0554 and Mean Absolute Error (MAE) of 0.0348 in the best model.

Keywords*—Electric Energy, Machine Learning, Meta-Heuristic, Grey Wolf Optimization*

I. INTRODUCTION

Energy is a fundamental input in the current economy. The economic and social development of countries is deeply related to the growth and increase in the supply of electrical energy [1]. It is estimated that global electrical energy generation will increase by approximately 77% between 2006 and 2030 [2].

The demand for electrical energy is something that continues to grow. One of the contributing factors is the emergence of various technological equipment and the consequent use by the population. Brazil has several resources that can be explored to generate electrical energy, and one of the most used resources is hydroelectric plants, which are renewable energy sources [1].

The contribution of energy from hydroelectric plants is around 63% in Brazil, being responsible for generating approximately 70% of all energy used in the country. Despite incentives for the use of other energy sources, it is estimated that in the following years at least 50% of the energy consumed will still come from hydroelectric plants [3].

Although Brazil has large sources of renewable energy, the country has problems in terms of electricity supply, and issues to be observed regarding energy-related investment [4]. There is a segment of the Brazilian population that has difficulty accessing electricity, with the majority of problems being in the way energy is distributed.

As electrical energy cannot be stored, its production and consumption must be accurately idealized in order to avoid circumstances of energy insufficiency as well as overproduction. Simultaneously, load and demand projections serve as the foundation for several decisions made in the energy markets, enabling the planning and operation in a way that is secure, clear, effective, and meets industry demands. Thus, one can observe the need to estimate energy consumption, looking for some way to meet this energy demand. In this context, computational tools can assist in the prediction process and when it comes to the use of data, machine learning techniques appear as an alternative. Given the above, this study seeks to apply machine learning methods to predict electricity demand in Brazil. Thus, it will be possible to assist in decision-making for the distribution of electrical energy.

When using machine learning techniques, a very important factor is to define attributes of the methods to maximize performance. To overcome this situation, metaheuristics can be applied to optimize the models, seeking the best parameters to obtain estimates with the lowest error.

The aim of this study is to use energy load data made available by ONS in order to predict demand based on consumption from the previous seven days. To this end, we propose the use of machine learning techniques commonly used in other applications, such as MultiLayer Perceptron (MLP), Random Forest (RF) and Support Vector Machine (SVM), and the metaheuristic Grey Wolf Optimizer (GWO). The results show that the computational methodology developed performs well in prediction and can assist specialists, providing direction for strategic management and it will anticipate future needs, that will serve as a roadmap for the development and execution of strategies. For businesses in

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2024
DOI: 10.5281/zenodo.12192271

the energy sector, selecting the forecasting model and approach for demand prediction is a crucial decision. Companies operating in the energy sector can set their strategic goals and have the opportunity to improve performance based on this study.

The present study is divided into five sections: section 2 presents the research related to this study. Section 3 discusses the study area, as well as the methodology used. In section 4, there is a discussion of the results obtained and, finally, in section 5, the conclusions of this study are presented.

II. RELATED RESEARCHES

Forecasting energy demand in Brazil is a topic that has been studied by many researchers. In these studies, analysis is carried out and it is proposed tools in the context of Data Mining to understand the problem and seek solutions to predict the results.

Ruas *et al*. [5] carried out a study on predicting short-term energy demand in the state of Paraná between 2004 and 2006. Artificial Intelligence methods were used to predict the results, such as Recurrent Artificial Neural Networks (RNNs) and Support Vector Machine (SVM). The SVM algorithm, with 84 days of input, with sub-bands for the forecast, was the one that obtained the best result.

Alves [6] conducted a study on short-term electrical load forecasting, with historical data from periods of 24 and 48 hours forward, from a company in the electrical sector. Multiple Linear Regression (MLR) and Multilayer Perceptron (MLP) algorithms were used. The MLP was the one that achieved the optimal results.

In the research by Drebes [7], the energy demand for a given day was forecast for the Certel Cooperativa Operations Center Company, responsible for the operation of distribution systems, operation of substations and responsible for controlling active demand. The algorithms used were the MLP, Linear Regression (LR) and Random Forest (RF). The LR algorithm was the one that presented really good results.

Schreiber *et al*. [8] made a prediction of the performance of transformers at the State Electricity Distribution Company in the city of Porto Alegre, Rio Grande do Sul. The MLR algorithm was used. The best results showed an average relative error of 0.050 of the real and estimated yield.

In Marcos and Júnior's work [1], machine learning techniques were used to predict electricity consumption in the Northeast region of Brazil, between the years 2004 and 2019. MLP and Convolutional Neural Networks (CNN) were those that obtained the best outcomes.

Oliveira [9] used the GWO meta-heuristic to minimize the objective function total cost of a shell and tube heat exchanger project, which are used to heating and cooling in various applications such as petroleum refineries, chemical processing, among other applications.

In Pizzolato *et al*. [10], the GWO meta-heuristic was used to obtain the optimal configuration of relay actuation and optimize relay time, which allows faults to be identified, locate and alert the operation of an electrical system so that circuit breakers are open, isolating a given defect. Using GWO, it was possible to coordinate the relays, maintaining the adjustments to the protection system.

The papers found do not forecast energy demand for Brazil as a whole, but rather for specific regions, in addition to not using approaches to find the optimal model. The application of machine learning algorithms is very promising and employing meta-heuristics will help to find the best model, making it possible to predict demand with less error.

III. METHODOLOGY

A. Database

The National Electric System Operator (ONS) has diverse information about energy in Brazil. In this study, the variable Energy Load (EL) was used, which indicates the population's demand, that is, how much energy is used.

The database has daily records of the energy load across the country, where this information is separated by regions. As the objective is to analyze the entire country, a sum of information from all regions was carried out to obtain the demand of the Brazilian population as a whole. The period used for predictions is between the years 2017 and 2022, around 8,764 samples.

B. Pre-Processing

There are four attributes available: id_subsistema, nom_subsistema, din_instante and val_cargaenergiamwmed. The id_subsistema attribute contains the initial letter of each region of Brazil. For example, for the North region the representation is N. The nom_subsistema attribute represents the name of the regions of Brazil, being North, South, Southeast and Northeast. Information for the Central-West region is not available on the base. Furthermore, the din_instante attribute indicates a respective date, in the format (YYYY-MM-DD). Finally, the val_cargaenergiamwmed attribute presents the load value in milliwatts (MW).

To predict energy demand, only the variables din_instante and val_cargaenergiamwmed were considered. They were renamed to DATE and EL, respectively. The DATE variable represents a single date and the EL variable represents the sum of energy loads between the North, South, Southeast/Central-West and Northeast regions. After summing up the energy load of the regions, the database had 2,191 samples. Furthermore, a normalization of the EL variable was performed, resulting in values from 0.10 to 0.90. The attributes that refer to the energy load value of each region, as well as those that identify a specific region, were excluded, as the DATE and EL attributes, which contain the sum of the loads between the regions, will be taken into consideration for the analysis. A lag was also created in the database, creating 7 variables: EL1, EL2, EL3, EL4, EL5, EL6 and EL7. EL1 has a charge from the second day of EL and as it contains one less piece of information, this remains as NaN. EL2, from the third day onwards, contains two NaN information and so on. These samples containing NaN were excluded, 7 in total. After exclusion, 2,184 samples were obtained. This way, it will be possible to predict the energy load for the eighth day considering the previous seven. With the creation of these variables, a new base was created, having only the following

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2024
DOI: 10.5281/zenodo.12192271

attributes: DATE, EL, EL1, EL2, EL3, EL4, EL5, EL6 and EL7 and this data was used in this study.

McNemar's statistical test [11] was used to verify the dependence between variables. When applying the test, some factors must be considered, such as the variables are of the same nature; are identical; have the same values; each variable was entered only once in the sample. If the result is less than (0.05) which is the significance level, the null hypothesis (HO) is rejected, that is, there is an association between variables.

C. Cross Validation

Cross Validation (CV) is used to analyze how much a method can generalize across a set of data. CV is widely used in problems whose aim is to make predictions. Using this approach, it is possible to divide the database into training and test sets. The training set is applied to define the parameters that will be used in the model and the test set is to evaluate the model after training the method.

To evaluate the performance of the regression methods, the K-Fold (KF) technique was used [12]. When using KF, k subsets are divided from N samples, where K>1. After separating the subsets, the k-1 subsets are used to train the methods, and the rest of the sets is used to perform them. Thus, at the end of the process, the validation error is calculated. This procedure is repeated K times, using a different test set for each iteration. In order for regression methods to be able to predict future inputs, tests are repeated several times to best train the models.

D. Methods

Random Forest (RF) [13] is a learning algorithm that works as an ensemble. Builds k decision trees on a training data set in k iterations. During each iteration, a set of samples is randomly selected first. To construct a decision tree from this subset, attributes are randomly chosen by the RF. In this case, as the variable used is the EL and the lags, the decision trees are created based on the randomly chosen lags. The McNemar statistical test was applied where the null hypothesis was not rejected, indicating that there is no statistical evidence that there is an association between pairs of variables in a way that allows the use of RF. Each decision tree is constructed by considering independent random subsets of features and samples. The prediction of a new sample is made using an average or median of the individual tree predictions.

SVR is the Support Vector Machine method for regression [14]. The SVR can be linear or non-linear according to the kernel functions employed. Given the data set of points $\{(x_1, y_1), \ldots (x_l, y_l)\}$, where $x_i \in \Re^n$ is a vector of features and $y_i \in \Re^1$ is the vector of target values. It has the parameters > 0 and $C > 0$ and the SVR is formulated as ε the optimization issue in (1).

$$min \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i + C \sum_{i=1}^{l} \xi_i' \qquad (1)$$

subject to

$$w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i,$$

$$y_i - w^T \phi(x_i) - b \leq \varepsilon + \xi_i',$$

$$\xi_i, \xi_i' \geq 0, i=1,\ldots,l.$$

The dual form of the optimization problem can be written as (2).

$$min \frac{1}{2} (\alpha - \alpha')^T K(x_i, x_j)(\alpha - \alpha') + \text{¿} \qquad (2)$$

$$+ \sum_{i=1}^{l} (y_i + \varepsilon)(\alpha_i - \alpha_i')$$

subject to

$$e^T (\alpha - \alpha') = 0,$$

$$0 \leq \alpha_i, \alpha_i' \leq C, i=1,\ldots,l.$$

where $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)^T$, and $\phi(.)$ is the kernel function.

Solving (2) allows determining the parameters to build the SVR approximation.

Then, SVR estimates are given by

$$\hat{y}_i = \varepsilon \sum_{i=1}^{l} (\alpha_i - \alpha_i') K(x_i, x) + b.$$

Multilayer Perceptron (MLP) neural networks were developed to solve problems that are non-linearly separable, in which they cannot be separated by a hyperplane. The structure of a MLP is composed of layers of neurons, where the first layer is for data input, followed by one or more hidden layers to process the information, which uses activation functions, and a last output layer to return the result. In the MLP training phase, the neuron weights are updated to minimize errors using the Backpropagation algorithm [15]. Fig. 1 illustrates an MLP neural network, with a tgh activation function, with six and two neurons in the first and second hidden layers, respectively.

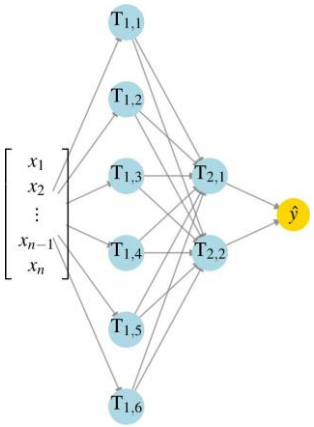

Fig. 1. Architecture of a MLP

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2024
DOI: 10.5281/zenodo.12192271

One of the more actual meta-heuristic swarm intelligence techniques is the Grey Wolf Optimizer (GWO). Because of its remarkable advantages over other swarm intelligence techniques, namely, that it requires no derivation information during the initial search and has very few parameters, it has been extensively used for a wide range of optimization problems. Additionally, it is straightforward, easy to apply, adaptable, scalable, and has the unique ability to balance exploration and exploitation in a way that promotes favorable convergence during the search.

The GWO meta-heuristic is based on the social behavior of grey wolves, seeking to simulate the social hierarchy of wolves in a pack [16]. In the GWO, the inhabitants are separated in alpha (α), beta (β), delta (δ) and omega (ω). Theα wolves that have more capability to survive environment are the first denominated, β and δ that lead other wolves ωα towards promising locations in the search space. During the process, the wolves advance and modify their, β or δα positions as follows:

$$\vec{M}=|\vec{L}\,\vec{X}_p(t)-\vec{X}(t)| \quad (3)$$

$$\vec{X}_{(t+1)}=\vec{X}_p(t)-\vec{J}\,\vec{M} \quad (4)$$

where t is the most recent epoch, $\vec{J}=2a\,\vec{r}_1 a$, $\vec{L}=2\vec{r}_2$, $\vec{X}_p$ is the prey position vector, $\vec{X}$ is the position vector of a grey wolf, $\vec{r}_1$, $\vec{r}_2$ are random vectors in [0,1] and the parameter a decreases linearly from two to zero. The GWO supposes that, β and δ are the assumed (optimal) preyα position. The three best solutions obtained so far are considerate, β and δ respectively. Other wolves are thenα indicated as ω and with the capacity to reposition themselves in relation to, β and δ. The proposed mathematical modelα that reports the location to be readjusted of the ω wolves:

$$\vec{M}_\alpha=\left|\vec{L}_1\vec{X}_\alpha-\vec{X}\right| \quad (5)$$

$$\vec{M}_\beta=\left|\vec{L}_2\vec{X}_\beta-\vec{X}\right| \quad (6)$$

$$\vec{M}_\delta=\left|\vec{L}_3\vec{X}_\delta-\vec{X}\right| \quad (7)$$

where $\vec{X}_\alpha$, $\vec{X}_\beta$, $\vec{X}_\delta$ present the position of, β and δα respectively, $\vec{L}_1$, $\vec{L}_2$, $\vec{L}_3$ are random vectors and the $\vec{X}$ is the position of the current solution.

Equations (5), (6) and (7) calculate the distance between the current solution and, β and δ. The final position of the α actual solution is calculated as follows:

$$\vec{X}_1=\vec{X}_\alpha-\vec{J}_1\vec{M}_\alpha \quad (8)$$

$$\vec{X}_2=\vec{X}_\beta-\vec{J}_2\vec{M}_\beta \quad (9)$$

$$\vec{X}_3=\vec{X}_\delta-\vec{J}_3\vec{M}_\delta \quad (10)$$

$$\vec{X}_{(t+1)}=\frac{\vec{X}_1+\vec{X}_2+\vec{X}_3}{3} \quad (11)$$

where $\vec{X}_\alpha$, $\vec{X}_\beta$, $\vec{X}_\delta$ show the current location of the wolves, β and δ, α $\vec{J}_1$, $\vec{J}_2$, $\vec{J}_3$ are randomly generated vectors and *t is* the number of epochs.

As shown above, (5), (6) and (7) represent the step of the wolf ω toward, β and δ respectively. Equations (8), (9), α (10) and (11) are the final location of the ω wolves. There are two vectors too, as can be seen: $\vec{J}$ and $\vec{L}$.

E. Metrics

The following metrics were used to evaluate performance: Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Coefficient of Determination (R^2).

RMSE calculates the square root of the mean of the difference between the true value and the estimated value for the data set. The calculation of the difference is squared. The higher the RMSE value in the calculation, the worse the model will be [17]. In (12), y_i the true value, $\hat{y}_i$ the estimated value and n the data set number.

$$RMSE=\sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2\right)} \quad (12)$$

In MAE, the average difference between the true value and the estimated value for the data set is calculated, but as there may be negative values in the difference, the value in the module is considered. The lower the value obtained in the MAE calculation, the better the predicted results will be [17]. In (13), y_i the true value, $\hat{y}_i$ the estimated value and n the data set number.

$$MAE=\frac{1}{n}\sum_{i=1}^{n}\left|y_i-\hat{y}_i\right| \quad (13)$$

R^2 measures the variance of a model data. The variability value is between 0 and 1. The value obtained in the R^2 calculation indicates that the higher the value, the prediction is closer to what is expected according to the original data [17]. In (14), y_i represents the true value, $\hat{y}_i$ the value to be predicted and $\bar{y}$ the average value for y.

$$R^2=\frac{\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2} \quad (14)$$

IV. RESULTS AND DISCUSSIONS

The computational methodology was implemented using the programming language Python, which is a language used to perform data analysis and which has several libraries with optimized functions [18]. All experiments were run on a computer with the following specifications: Intel(R) Core(TM) i5-1135G7, 8 GB RAM, and Windows 10 operating system. The Scikit-Learn and PyGMO libraries were used. Scikit-Learn is a library that allows you to work with machine learning, it has a set of resources, such as algorithms to perform data analysis, metrics for prediction,

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2024
DOI: 10.5281/zenodo.12192271

among others [19]. PyGMO is a library used to work with optimization problems, it has several optimization algorithms to be used in conjunction with machine learning algorithms for a better performance [20]. As for the library, 30 independent iterations were carried out to evaluate the methodology. A KF with a value of K equal to 5 was used.

Table 1 presents the description of the models used in the GWO meta-heuristic, indicating the method, the parameters for the model to perform well, the description of these parameters and their configuration, indicating the values used during the executions. In MLP, the activation function settings to be used in GWO were represented as follows: 0: Identity, 1: Logistic, 2: Tanh and 3: ReLU, the configuration being [0, 3]. In this configuration, for example, it means that the values will be generated randomly in the range from 0 to 3. The máximum number of GWO generations was 30.

TABLE I. DESCRIPTION OF THE MODELS

Methods	Parameters	Description	Settings
MLP	hidden_layer_sizes	Number of hidden layers and Number of neurons in each layer	[1, 4] and [1, 50]
	activation	Activation Function	0: Identity; 1: Logistic; 2: Tanh; 3: ReLU
RF	n_estimators	Number of trees in the forest	[100, 200]
	max_depth	Maximum depth of the tree	[2, 10]
SVM	C	Adjusts the penalty for regression errors	[20, 200]
	gamma	Defines how far the influence of a single training example extends	[0.001, 0.1]
	epsilon	Sets a limit on insensitivity to errors in predictions	[0.001, 0.1]

Table 2 presents the results obtained using the average and standard deviation, the R^2 metrics of the iterations, RMSE and MAE. Using the GWO meta-heuristic, it was possible to find the ideal parameters to maximize the performance of the regression models. The best model is highlighted in bold.

The model that achieved the best performance was RF, with a R^2 of 0.8704, MAE of 0.0352 and RMSE of 0.0564, thus indicating that it was the method that obtained the best predictions in relation to the EL variable. The SVM also showed good results, with a R^2 of 0.8618, MAE of 0.0353 and RMSE of 0.0583. The closer the value of R^2 is to 1 and the lower the values of MAE and RMSE, the better the model performance.

TABLE II. MODELS PERFORMANCE

Methods	R^2	RMSE	MAE
MLP	0.7982 ± 0.0288	0.0703 ± 0.0047	0.0496 ± 0.0042
RF	**0.8704 ± 0.0016**	**0.0564 ± 0.0003**	**0.0352 ± 0.0003**
SVM	0.8618 ± 0.0005	0.0583 ± 0.0001	0.0353 ± 0.0003

Table 3 shows the optimal MLP, RF and SVM parameters to maximize performance. The best result obtained by RF has a number of trees equal to 130, a maximum tree depth equal to 10, R^2 of 0.8751, RMSE of 0.0554 and MAE of 0.0348. The results show that using ensemble methods generates good results, as the method combines several models.

TABLE III. BEST MODELS

Methods	Parameters	R^2	RMSE	MAE
MLP	hidden_layer_sizes = (47, 49, 48, 44), activation = ReLU	0.8244	0.0656	0.0450
RF	n_estimators: 130, max_depth: 10	0.8751	0.0554	0.0348
SVM	C = 100, gamma = 0.1, epsilon = 0.0229	0.8626	0.0580	0.0353

Fig. 2 presents boxplots for the R^2, RMSE and MAE metrics, showing the performance of the RF, SVM and MLP regression models in predicting the load during the 30 iterations. It can be seen that RF was the model that presented the best results, with the lowest values for RMSE and MAE and the highest for R^2. The MLP, considering the network topology adopted, presented some variations in the prediction of the load variable, resulting in higher values for MAE. For the MLP, some MAE values were not concentrated like the other predictions, thus indicating that for some data the predictions were not correct. Such analysis is appropriate considering the search range for the parameters used and the meta-heuristic employed.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2024
DOI: 10.5281/zenodo.12192271

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XI, Issue 2, July 2024

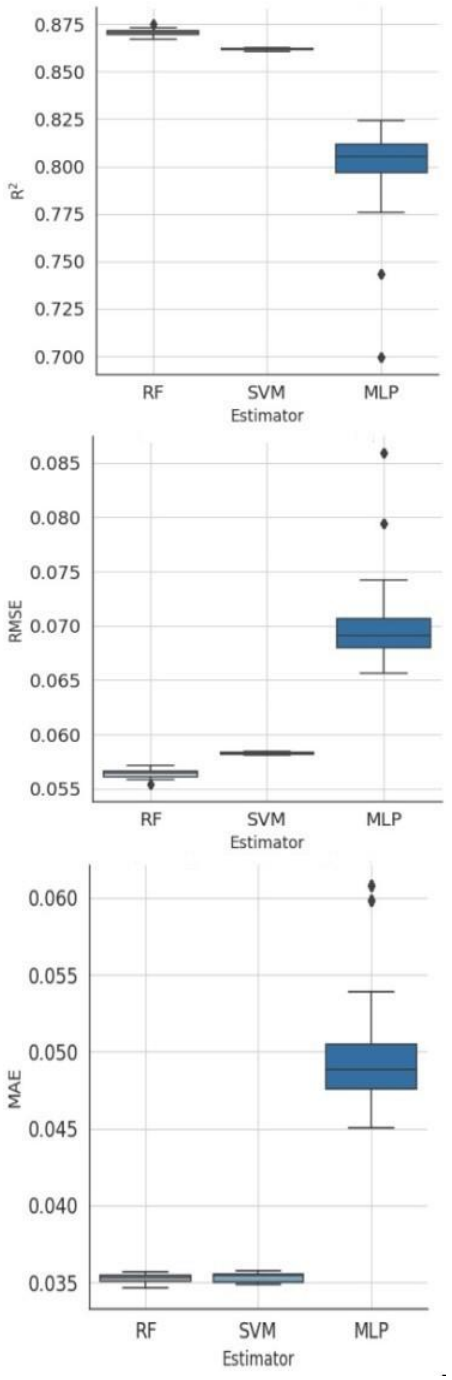

Fig. 2. Boxplot for R^2, RMSE and MAE metrics

Fig. 3 shows a comparison of the true energy load already normalized with that predicted by the best MLP, RF and SVM models in the analyzed period. It can be observed that the error between the predicted and the true is less for the RF comparing to other models. The time series formed by these values have similar behavior. This shows that the RF model has the capacity to assist in the energy load prediction process, which helps to verify population demand.

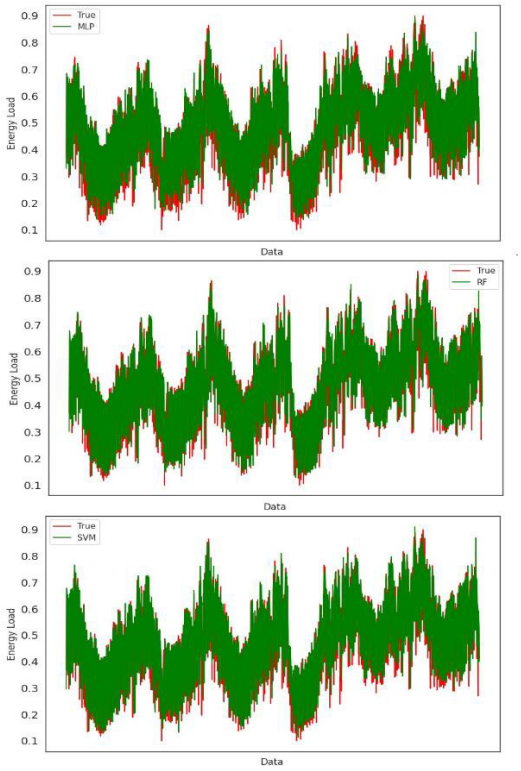

Fig. 3. Energy Load True x Energy Load Predict – MLP, RF and Svm respectively

Fig. 4 illustrates the parameter distribution of the MLP, RF and SVM models. One can see that for MLP the ReLU was the activation func chosen in all runs and 4 hidden layers in the most runs. For RF, the max depth was 10 in all runs and No. estimators around 130 in 8 runs followed by 200 in 5 runs. In the SVM case, C equal 100 in all runs, γ equal 0.1 and ε values were well distributed between 0.001 and 0.0275.

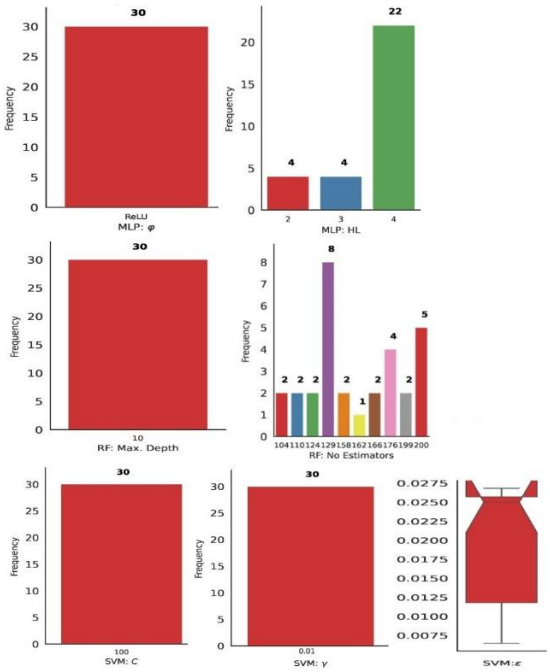

Fig. 4. MLP, RF and SVM parameters distribution

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2024
DOI: 10.5281/zenodo.12192271

V. CONCLUSION

In this paper, the performance of three regression methods (MLP, RF, SVM) for forecasting electricity demand in Brazil, it was analyzed for the algorithms to have good results, the GWO meta-heuristic was used to improve their performance. The RF and SVM methods showed the best results. However, the RF was the one that presented the best result, with R^2 = 0.8751, RMSE = 0.0554 and MAE = 0.0348 and, which shows that using ensemble methods generates good results, by combining a set of models. For the MLP, some predictions were not correct, due to the fact that some metric values were not concentrated. This analysis was carried out considering the search space for the parameters adopted and the meta-heuristic used.

Future research includes applying deep learning techniques, such as Long Short-Term Memory, to evaluate whether methods considered more robust will perform better in predicting energy demand. In addition, analyzing the insertion of other variables that affect daily energy consumption and that can assist in prediction.

Furthermore, the proposed approach proved to be effective, as it used cross-validation techniques to enable the model ability to generalize data from the tests carried out. Moreover, by using the GWO meta-heuristic, it was possible to search for the best parameters to maximize the performance of the regression models, as well as by the use of an ensemble algorithm that combines multiple models.

REFERENCES

[1] I. P. Marcos and A. P. P. Júnior," Forecast of Electricity Consumption in the Northeast Region of Brazil". Journal of Engineering and Applied Research. v. 6, n. 3, p. 21-30, 2021.

[2] L. C. Morais, Study on the panorama of electrical energy in Brazil and future trends. Dissertation - Electrical Engineering. UNESP. 2015.

[3] B., Stearns, F. Rangel, F. Firmino F, Rangel and J. Oliveira, Predicting performance of enem candidates through socioeconomic data. In: Proceedings of the XXXVI SBC
Scientific Initiation Paper Competition. SBC, 2017.

[4] Energy research company, Dea Technical Note 22/16, Projection of electricity demand for the next 10 years (2016 – 2026). 2016.

[5] G. I. S. Ruas, T. A. C. Bragatto, M. V. Lamar, A. R. Aoki and S. M. Rocco," Forecasting electrical energy demand using artificial neural networks and support vector regression". VI National Artificial Intelligence Meeting. p. 1262-1271, 2007.

[6] M. F. Alves, Forecasting electrical load demand by stepwise variable selection and artificial neural networks. Dissertation Electrical Engineering. UNESP. 2013.

[7] F. Drebes, Electricity demand forecast using artificial intelligence. 2020. Monograph (Graduation in Electrical Engineering) – University of Vale do Taquari - Univates, Lajeado, 03 Dec. 2020.

[8] J. F. Schreiber, I. E. M. Kühne, L. A. Destefani, A. T. Z. R. Sausen, M. Campos and P. S. Sausen, "Intelligent Networks: Data Mining as a Support Tool for the Analysis of Large Volumes of Data in Underground Energy Substations". In: Brazilian Automatic Congress-CBA. 2020.

[9] C. T. Oliveira, Optimization of a shell and tube heat exchanger using the grey wolf algorithm. Dissertation - Mechanical Engineering. Unisinos. 2015.2018.

[10] G. P. Pizzolato, E. M. dos Santos, A. R. Fagundes, J. O. dosSantos and H. Hasselein, Optimizing the Operating Time of Overcurrent Relays Using the Grey Wolf Algorithm. Brazilian Symposium on Electrical Systems, 1(1). 2020.

[11] P. A. Lachenbrunch, McNemar test, Wiley StatsRef: StatisticsReference Online. 2014.

[12] R., Kohavi, "A study of cross-validation and bootstrap foraccuracy estimation and model selection", Ijcai. Vol. 14. No. 2. 1995.

[13] L. Breiman, Random forests. Machine learning, v. 45, p. 5-32,2001.

[14] A. J. Smola, Learning with Kernels, PhD Thesis. TechnicalUniversity of Berlin, 1998.

[15] S. Haykin, *Neural networks: a comprehensive foundation*. Prentice Hall PTR, 1998.

[16] S. Mirjalili, S. M. Mirjalili and A. Lewis, "Grey wolf optimizer".Advances in engineering software, v. 69, p. 46-61, 2014.

[17] B., Stearns, F. Rangel, F. Firmino F, Rangel and J. Oliveira,Predicting performance of enem candidates through socioeconomic data. In: Proceedings of the XXXVI SBC
Scientific Initiation Paper Competition. SBC, 2017.

[18] A. Boschetti and L. Massaron, *Python data science essentials*. Packt Publishing Ltd, 2016.

[19] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B.Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrotand and É. Duchesnay, "Scikit-learn: Machine learning in Python". the Journal of machine Learning research, v. 12, p. 2825-2830, 2011.

[20] F. Biscani and D, Izzo, "A parallel global multiobjectiveframework for optimization: pagmo". Journal of Open Source Software, v. 5, n. 53, p. 2338, 2020.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2024
DOI: 10.5281/zenodo.12192271

AUTHORS

Bruno da S. Macêdo

Mestrando em Engenharia de Sistemas e Automação na Universidade Federal de Lavras. Possui graduação em Engenharia da Computação pela Universidade do Estado de Minas Gerais e curso técnico profissionalizante em Informática para Internet pelo Centro Federal de Educação Tecnológica de Minas Gerais (2018). Tem desenvolvidos trabalhos na área de Inteligência Computacional.

Camila Martins Saporetti

Doutora em Modelagem Computacional pela Universidade Federal de Juiz de Fora. Mestre em Modelagem Computacional pela Universidade Federal de Juiz de Fora. Graduada em Bacharelado em Ciências Exatas pelo Instituto de Ciências Exatas da Universidade Federal de Juiz de Fora, em Engenharia Computacional pela Faculdade de Engenharia da Universidade Federal de Juiz de Fora e em Ciência da Computação pelo Instituto de Ciências Exatas da Universidade Federal de Juiz de Fora. Atualmente é professora adjunta do Instituto Politécnico da Universidade do Estado do Rio de Janeiro e membro permanente do Programa de Pós-Graduação em Modelagem Computacional do Instituto Politécnico. Tem experiência na área de Inteligência Computacional, atuando principalmente nos seguintes temas: aprendizado de máquina e metaheurísticas.