Evolutionary Machine Learning-Based Energy Consumption Prediction for the industry

Mouad Bahij, [1, *], *Moussa* Labbadi[2], Chakib Chatri[1], Cheikhne Cheikh Ahmed[1]*, and* Mohamed Cherkaoui[1]

[1] Engineering for Smart and Sustainable Systems Research, Mohammadia School of Engineers, Mohammed V University in Rabat, Morocco
[2] LAMIH, CNRS, UMR-8201, INSA HdF UPHF, Valenciennes 59313, France

Abstract. In the digitalization of industry and the industry 4.0 environment, it is important to master the accurate forecasting of energy demand in order to guarantee the continuity of production service as well as to improve the reliability of the electrical system while promoting energy efficiency strategies in the industrial sector. This paper proposes machine learning models to predict the energy consumption demand in an industrial plant, which takes into account the at-tributes that directly the consumption. The proposed models in this work include Multiple Linear Regression (MLR), Decision Tree (DT), Recurrent Neural Networks (RNN) and Gated Recurrent United (GRU), which are compared according to their performances criteria which help to find the best forecasting models. Basing on simulation results, it is proven that the MLR approach is the best forecasting method.

1 Introduction

The industrial revolution and the rapid development of the economy play an important part in energy consumption for the industry sector where this expansion of progress contributes to the global warming [1]. In Morocco, the industrial sector ac-counts for 22% of the nation's energy consumption. Here, the good management of electricity consumption in the industry can help us to reduce negative environmental risks such as greenhouse gas emissions.

Industry 4.0 has an intelligent operating space where all its elements are connected in an industrial network which leads to an improvement of the products management and manufacturing as well as reduction of the energy consumption, this flexibility of control is provided by the integration of artificial intelligence in the management and optimization of manufacturing process [2].

Machine Learning techniques (ML) create intensive opportunities to reduce the high footprint of energy billing to industry. Hence, load forecasting is used to build future energy estimation models based on the historical data (data of weather, production or electricity consumption) collected by using intelligent meters in plants. Here, the interest, behind using ML application for energy consumption prediction in the industry, is to improve the planning and scheduling of energy production, thus, to create savings benefits [3].

The forecasting models offer the promising solutions to achieve energy efficiency in the industry due to their accuracy and effectiveness. Based on some databases, the ML techniques perform a relationship and identification of the different characteristics of data when training model to predict the desired energy consumption.

In the ML layer, several techniques have been elaborated to realize prediction models such as the DT. This one is widely used to solve prediction problems thanks to their ability to paraphrase the training data into groups when trying to formulate the model [4]. Referring to [5], the author proposed DT as a prediction tool. This technique is used to form a prediction model (for enterprises working in the oil and electricity sector) with a MAPE value of around 4.556% for a 50% training database.

The RNN is another technique established to forecast energy consumption load, thus, to improve efficiency of the desirable model [6]. Effectively, RNN-based learning models show advantages when processing the time series data [7]. In [5], the RNN model is built in order to estimate the short-term residential power consumption load that presents an error of around 7.03% at 20 neurons, this study exploits the Gated Recurrent United (GRU) approach for im-proving the results. Effectively, the GRU presents an error of 6.37% which proves its good forecasting performance [8]. The study, realized in [9], compares the RNN and GRU techniques applied in the traffic flow prediction, in this case of study, it is proven that the GRU approach is more efficient than the RNN with an error of approximatively 9.26%.

For statistical analysis, the linear regression technique is presented to estimate future regression coefficients. In [10], the MLR approach is established to predict energy consumption in a New Zealand food industry. At present, in the paper [11], the MLR approach as well as some other techniques are used to express the primary energy prediction model in China. Here, the results show that the MLR expresses a training error of approximatively 5.55%.

[*] Corresponding author: mouad.t18@gmail.com

The contribution of this paper is to develop four approaches for predicting energy consumption in the industry, namely: MLR, GRU, ANN, and RNN. These approaches are based on attributes collected in the industrial setting, and compared in terms of the accuracy when predicting energy consumption. Effectively, the more the prediction accuracy is improved, the more the planning, of the reasonable energy plan of the production plant, is improved.

This paper is structured as follows: Section 2 presents some prediction methods. The methodology of the energy forecasting is described in section 3. The simulation results are discussed in section 4. Finally, the last section is dedicated to the conclusion.

2 Theoretical Background

This section is reserved to present the ML techniques used in the prediction of the energy consumption. These forecasting approaches have been identified as a major actor in improving energy efficiency in industry.

2.1. Multiple Linear Regression

Multiple Linear regression is an important method for estimating regression coefficients in multivariate statistical analysis, the MLR is applied in different disciplines to forecast the output variables by a linear combination of the independent attributes. This method is introduced to estimate the energy consumption load in the building [11].

The multiple linear regression model is described as follows:

$$S = \beta_0 + \beta_1 a_1 + \beta_2 a_2 + \beta_3 a_3 \ldots\ldots + \beta_P a_p + \varepsilon \quad (1)$$

where, (a_i) expresses the values of the systems input, (S) is the expected output, (β_0) and (ε) are the regression coefficients and the random error term, receptively.

The creation of the MLR model requires a long time for the identification of the parameters, as well as the quantity of data can affect the efficiency of the accuracy. The fluidity afforded by this approach will make MLR more usable for linear problems.

2.2. Decision Tree

The DT is a tree structure based on the composition of axis nodes and branch nodes to make the decision. This approach works through the formation of the first node of the tree to the lowest node, passing through intermediate root nodes. Each node shows the value of the attribute giving the result of the decision to the branches [12].

This repetitive splitting process supports the DT in the build the forecasting model. In prediction, DT has a major advantage of producing a model by logically interpret-able regels comparable to other modeling approaches.

The process of gain splitting for DT is expressed as follows:

$$S_r(T) = \sum_{k=1}^{k} p_{k\tau}\ (1 - p_{k\tau}) \quad (2)$$

where, $(p_{k\tau})$ presents the proportion of points in the class (k).

2.3. Recurrent Neural Networks

During the learning process, the RNN has the ability to catch temporal dependences using a nodes training pattern oriented on a graph along a sequence. The RNN architecture contains three layers, the connection between the nodes is based on the recurrent information to the same neurons in the previous state for the construction of the current neuron [13].

The RNN makes use of an internal memory concept to process the input variables where key is the capturing of computed in-formation on previous states up to the present state in order to generate the prediction model.
The computing of the output as a result of the previous output state and the current input state is defines as follows:

$$S_t = \sigma\,(W_1\, x^t + W_2\, S^{(t-1)} + b\,) \quad (3)$$

where, S_t is the output, σ means the activation function of each element, b presents the vector bias of the hidden layer. W_1 and W_2 denote the weight parameters. $S^{(t-1)}$ is the output from a prior time step and x^t is the input at the current time step.

2.4. Gated recurrent unit

The GRU technique has a simple structure that reduces the internal cache state and gives a fast modulation of the information flow.

The GRU architecture contains two prediction structures, the update gate for the measurement and control of the previous state information is maintained in the current state and the reset gate presents the combination operation to be performed by the previous state information [14].

The mathematical formulation of GRU is described as follows:

$$r_t = \sigma(w_r h_{t-1} + V_r\, x_t + b_r) \quad (4)$$

$$z_t = \sigma\,(w_z h_{t-1} + V_Z\, x_t + b_z) \quad (5)$$

Where, x_t is the input values, h_{t-1}presents the activation to the last state, the matrix parameters are specified by the values V and $w.\, z_t$ represented the update gate determining the amount of information provided to make the new output, the reset gate is indicated by r_t to find out what information to ignore.

The memory content necessary to store the reset gate information related to the past and the tangent function is expressed as follows:

$$\tilde{h}_t = tanh\,(w_r(h_{t-1} * R) + V_M x_t) \quad (6)$$

$$h_t = (h_{t-1}) * (1 - z) + z * \tilde{h}_t) \quad (7)$$

where, $\tilde{h}_t$ expresses the candidate activation at time t, V_M and w_r are parameters of the matrix in the learning process $\tilde{h}_t$, "$*$" signifies the matrix multiplication.

3 Methodology

In the industrial sector, the energy consumption forecasting is an important systematic operation to know the future energy demands. The four models employed in this paper are DT, MRL, RNN, and GRU run in Python. The phases of the generation and validation of the ML model for industrial energy forecasting are illustrated in Fig. 1.

The prediction of the energy load is based on ML approaches, where the formation of the models depends on the use of the database collected from the indicators of the factory production line. The prediction is based on production and climate variables, which affects the change in the energy demands.

The chosen attributes of the database are temperature, humidity, lighting, emissions and consumption. This database consists of two sections, one for training and the second for the testing of the model. To build the model, a normalization step is taken from a database, described in equation 8 in order to minimize the loss function.

$$\acute{x_i} = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (8)$$

where, $\acute{x_i}$ indicates the normalization value, x_i is the original value, $min(x)$ and $max(x)$ express the minimum and maximum values in the vector x_i. This data processing methodology eliminates outliers, improves model efficiency and accelerates the process of forecasting [7].

To measure the energy precision and the efficiency of the approaches proposed, performance indices are presented to assess the quality of each prediction model, including the percentage score in training and also in testing plus over the mean absolute percentage error MAPE (%) criterion expressed as follows:

$$MAPE = \frac{100\%}{n} \sum_{I=0}^{n} \left| \frac{y_i - \hat{y_i}}{y_i} \right| \quad (9)$$

Where, n is the number of samples, y_i is the actual value, and $\hat{y_i}$ the predicted value. The calculation of MAPE (%) was carried out on the basis of the non-normalized values [15].

The building of the model starts by measuring and collecting data from the manufacturing site. The second step is concerned with the preparation of the database for training and the other one for testing, both types of data followed a process of normalization in order to filter out the impurities. Subsequently, the implementation of the chosen algorithms and the validation of the models deduced are included in the last part. The prediction models are tested using the test database in each case in sequence to calculate the MAPE error and to check the efficiency in precision.

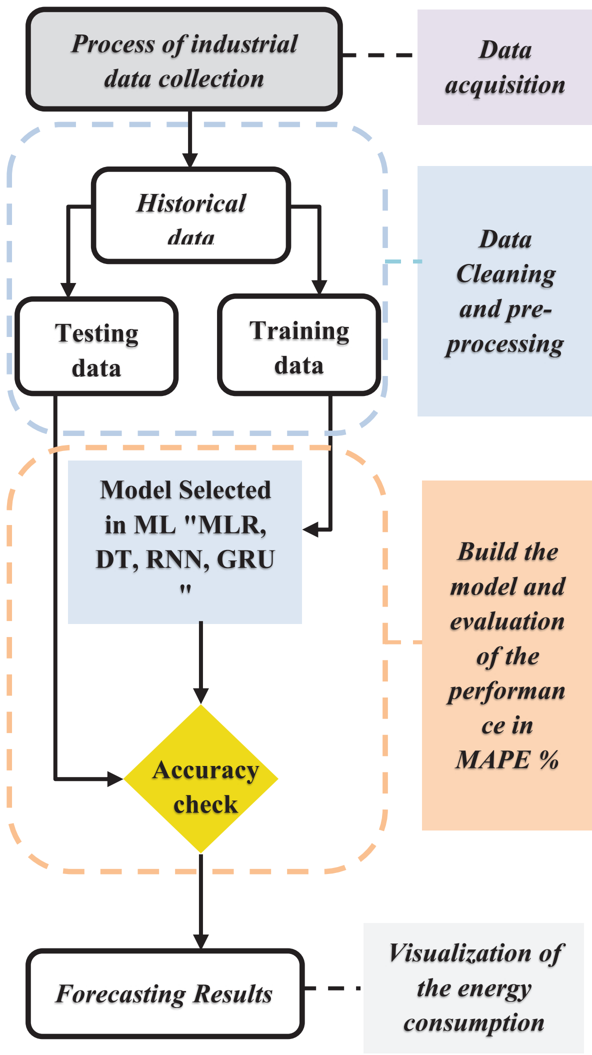

Fig. 1. Methodology structure for the energy forecasting.

4 Results and Discussion

In this section, the simulation results are presented, obtained from real production data collected in the industry for energy estimation. The proposed GRU, DT, MLR and RNN approaches are compared in order to extract the most precise model for the energy forecast. The training of the model uses a 12-month data set of an industrial manufacturing plant in Morocco that are recorded each shift of production monitoring. The evaluation criteria for each model are described in section 3. Table 1 shows the values of the performance indicators calculated for the generated models.

In Fig. 2 the performance of DT is presented, where the perturbations from the actual consumption profile to the prediction profile produced by the model are low. The MAPE % value attained for this application is of order 5.75 %. On the other hand, the efficiency of the MLR is given in Fig.3, the MAPE % value does not exceed 3.45 10 $^{-26}$ % suggesting the MLR forecasting scenario has a superior training capability.

The scores % obtained in table 1 by the MLR and DT methods are respectively 100% and 99.99 % in the training phase and 100% and 99.99 % in the test phase, which means that the MLR has a considerably better prediction power compared to the DT.

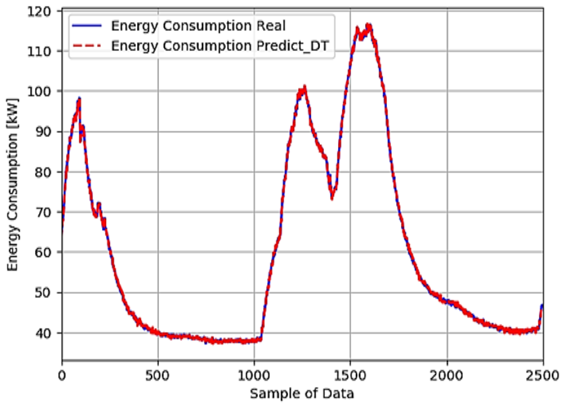

Fig. 2. Energy profile estimated by DT performance.

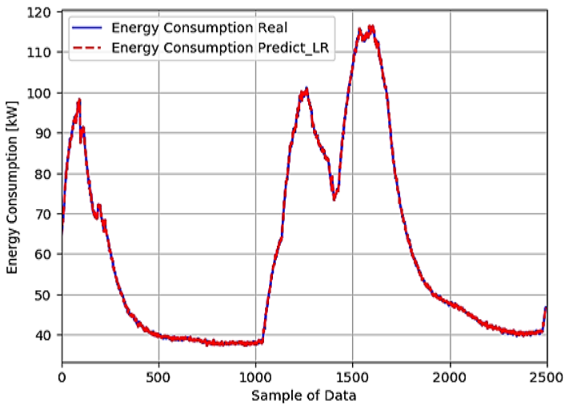

Fig. 3. Energy profile estimated by MLR performance.

Table 1. The performance criteria for the compared ML approaches.

Approaches	Training Score (%)	Predict Score (%)	MAPE (%)
MRL	100.00	100.00	$3.45\ 10^{-26}$
GRU	99.947	99.990	2.5800
RNN	99.970	99.970	1.5600
DT	99.999	99.990	5.7500

For the RNN, the Fig.4 expresses the correlation between the real consumption curve and the consumption estimated by the model, so this approach expresses a little bit of vibration at the energy peaks. According to table 1 the performance criteria expressed for this approach is 1.56% for the MAPE % as in 99.97% training score and 99.97% for the test.

In the following Fig.5, we propose a visualization of the GRU prediction, here the curve built by the model estimation followed the path of consumption with the presents of some perturbation at the level of the zones of high or low energy demands.

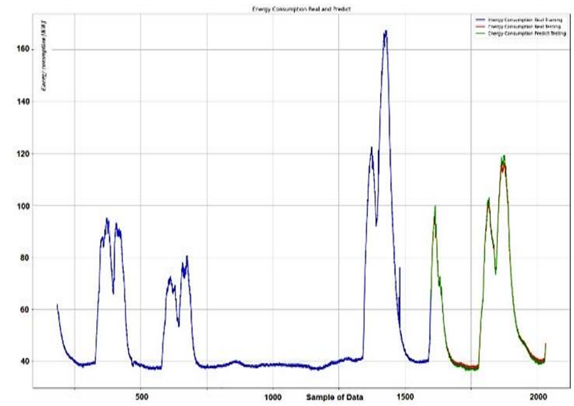

Fig. 4. Energy profile estimated by RNN performance.

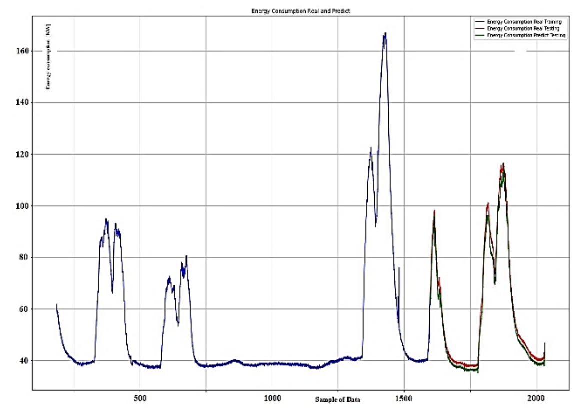

Fig. 5. Energy profile estimated by GRU performance.

5 Conclusion

This paper proposed four techniques used in the ML including Multiple Linear Regression (MLR), Decision Tree (DT), Recurrent Neural Network (RNN), and Gated Recurrent Unit (GRU) to forecast the energy consumption pattern for Industrial Production Factory.

The four solutions proposed in the present work are evaluated by performance criteria to find out their accuracy in predicting. Finally, the simulation results indicated the MLR solution gives the best performance with a minimum MAPE % of $3.45\ 10^{-26}$ compared to the RNN and GRU techniques, the DT method expresses the least module efficient with MAPE value of 5.75%.

References

1. H. Lasi, P. Fettke, H. G. Kemper, T. Feld, M. Hoffmann, Bus. Inf. Syst. Eng **6**, 239 (2014)
2. S. Vaidya, P. Ambad, S. Bhosle, Procedia Manuf **20**, 233 (2018)
3. A. Rahman, V. Srikumar, A. D. Smith, Appl. Energy **212**, 372 (2018)

4. Y. Wei, X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, X. Zhao, Renew. Sust. Energ. Rev **82**, 1027 (2018)
5. C. Zhang, H. Zhang, D. Liu, *IEEE Access* **8**, 11635 (2019)
6. A. Almalaq, G. Edwards, *A review of deep learning methods applied on load forecasting*, in 16th IEEE international conference on machine learning and applications, ICMLA, 18-21 December 2017, Cancun, Mexico (2017)
7. L. Wen, K. Zhou, S. Yang, *Electr.* Power Syst. Res **179**, 106073 (2020)
8. J. Zheng, X. Chen, K. Yu, L. Gan, Y. Wang, K. Wang, *Short-term power load forecasting of residential community based on GRU neural network*, in International Conference on Power System Technology, POWERCON, 6-8 November 2018, Guangzhou, China (2018)
9. Z. Bartlett, L. Han, T. T. Nguyen, P. Johnson, *Prediction of road traffic flow based on deep recurrent neural networks*, in IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI, 19-23 August 2019, Leicester, UK (2019)
10. A. C. Cleland, M. D. Earle, I. F. Boag, Int. J. Food Sci. Technol **16**, 481 (1981)
11. B. Liu, C. Fu, A. Bielefield, Y. Q. Liu, Energies **10**, 1453 (2017)
12. L. Romeo, J. Loncarski, M. Paolanti, G. Bocchini, A. Mancini, E. Frontoni, Expert Syst. Appl **140**, 112869 (2020)
13. J. Wang, J. Yan, C. Li, R. X. Gao, R. Zhao, Comput Ind **111**, 1 (2019)
14. J. Duan, H. Zuo, Y. Bai, J. Duan, M. Chang, B. Chen, Energy **217**, 119397 (2021)
15. T. Ahmad, C. Huanxin, D. Zhang, H. Zhang, Energy **198**, 117283 (2020)