PAB
Pesquisa Agropecuária Brasileira
60 years

Genetics/ Original Article

ISSN 1678-3921

Journal homepage: www.embrapa.br/pab

For manuscript submission and journal contents, access: www.scielo.br/pab

Extreme learning machine for genomic prediction of rust disease resistance in Arabica coffee

Abstract – The objective of this work was to investigate the use of Extreme Learning Machines (ELM) for the genomic prediction of rust resistance in *Coffea arabica*. With the objective of identifying an effective predictive model for the selection of resistant genotypes, ELM was compared to Artificial Neural Networks (ANN) and Bayesian Generalized Linear Regression (GBLR) in terms of accuracy measures and computational time. To this end, an F2 population of 245 *C. arabica* plants genotyped with 137 markers was used to evaluate the application of ELM for the genomic prediction of coffee rust resistance. The results indicate that ELM and ANN show a higher accuracy – on average 15% greater than that of GBLR – in predicting rust resistance. Additionally, ELM proves to be computationally more efficient, with a processing speed 5.5 and 19.45 times slower than that of ANN and BGLR, respectively, making it promising for large-scale analyses.

Index terms: *Coffea arabica*, computational intelligence, ELM, genomic selection, plant breeding, statistical learning.

Jackson Tavela da Silva
Universidade Federal de Viçosa, Viçosa, MG, Brazil. E-mail: jackson.tavela@ufv.br

Cynthia Aparecida Valiati Barreto
Universidade Federal de Viçosa, Viçosa, MG, Brazil. E-mail: cynthiavaliatibarreto@gmail.com

Ana Carolina Campana Nascimento
Universidade Federal de Viçosa, Viçosa, MG, Brazil. E-mail: ana.campana@ufv.br

Camila Ferreira Azevedo
Universidade Federal de Viçosa, Viçosa, MG, Brazil. E-mail: camila.azevedo@ufv.br

Dênia Pires de Almeida
Universidade Federal de Viçosa, Viçosa, MG, Brazil, E-mail: denia_pires@hotmail.com

Eveline Teixeira Caixeta
Embrapa Café, Brasília, DF, Brazil. E-mail: eveline.caixeta@embrapa.br

Filipe Ribeiro Formiga Teixeira
Universidade Federal do Piauí, Teresina, PI, Brazil, E-mail: filiperfteixeira@ufpi.edu.br

Moysés Nascimento ✉
Universidade Federal de Viçosa, Viçosa, MG, Brazil. E-mail: moysesnascim@ufv.br

✉ Corresponding author

Received
December 4, 2024

Accepted
June 9, 2025

How to cite
SILVA, J.T. da; BARRETO, C.A.V.; NASCIMENTO, A.C.C.; AZEVEDO, C.F.; ALMEIDA, D.P. de; CAIXETA, E.T.; TEIXEIRA, F.R.F.; NASCIMENTO, M. Extreme learning machine for genomic prediction of rust disease resistance in Arabica coffee. **Pesquisa Agropecuária Brasileira**, v.61, e03985, 2026. DOI: https://doi.org/10.1590/S1678-3921.pab2026.v61.03985.

Máquina de aprendizado extremo para predição genômica de resistência à ferrugem em café arábica

Resumo – O objetivo deste trabalho foi investigar o uso de Máquinas de Aprendizagem Extrema (ELM) para a predição genômica da resistência à ferrugem em *Coffea arabica*. Com o objetivo de identificar um modelo preditivo eficaz para a seleção de genótipos resistentes, o ELM foi comparado a Redes Neurais Artificiais (RNA) e Regressão Linear Generalizada Bayesiana (GBLR) em termos de medidas de acurácia e tempo computacional. Para tanto, uma população F2 de 245 plantas de *C. arabica* genotipadas com 137 marcadores foi utilizada de modo a avaliar a aplicação do ELM na predição genômica da resistência à ferrugem-do-café. Os resultados indicam que o ELM e a RNA apresentam maior acurácia – em média 15% superior ao GBLR – na predição da resistência à ferrugem. Adicionalmente, o ELM se mostra computacionalmente mais eficiente, com velocidades de processamento 5,5 e 19,45 vezes menores que a de RNA e o BGLR, respectivamente, tornando-o promissor para análises de larga escala.

Termos para indexação: *Coffea arabica*, inteligência computacional, seleção genômica, melhoramento de plantas, aprendizado estatístico.

Introduction

Artificial neural network (ANN) has been used to predict genomic genetic values in genomic selection studies (Nascimento et al., 2024; Sousa et al., 2024). ANNs are interesting due to their great flexibility (non-linearity) to deal with non-additive effects (e.g. dominance and epistasis) and the no need to make a priori assumptions about the functional relationship between the input (Single Nucleotide Polymorphism - SNP) and output (phenotypic values) variables of the model.

One limitation to the use of ANNs, especially when applied to high-dimensional genomic data, is the computational cost required to train the model. Sousa et al. (2021) used several machine learning methods to predict rust resistance in Arabica coffee (*Coffea arabica* L.) genotypes and discovered that the time spent fitting an ANN is 9.20 times longer than the time spent fitting a Bayesian LASSO (BLASSO). Ehret et al. (2015), aiming to reduce the computational cost inherent to the use of ANNs in genomic selection studies, used a subset of markers for prediction of genomic genetic values in animal breeding.

An alternative to ANNs, which allows the use of all information in the adjustment and has lower computational cost when compared to ANNs, is the use of extreme learning machine (ELM). ELM tends to achieve lower training errors compared to traditional learning algorithms, good generalization performance, and it performs analysis extremely fast (Huang et al., 2006). In the ELM algorithm, neurons are randomly started and then corrected without iterative tuning. The only free parameters that need to be learned are the connections (or weights) between the hidden layer and the output layer. Compared to traditional machine learning methods, ELM is very efficient and tends to achieve a global optimum (Huang et al., 2015). Advantages of ELM in efficiency as well as in performances have already been demonstrated in a wide range of problems from different fields (Taherei Ghazvinei et al., 2018; Silva et al., 2017; Qi & Majda, 2020).

In genomic selection, ELMs have been used to predict the genetic merit of black tea quality and drought tolerance (Koech et al., 2020). Despite presenting good results when compared to those coming from traditional machine learning methodologies, the use of ELMs is still not a common practice in genomic selection studies.

Coffea arabica represents about 60% of the world's coffee production, and Brazil is one of the largest world producers of this tropical plant (Consórcio Pesquisa Café, 2022). However, as in any culture with economic value, there are limiting factors in the production of Arabica coffee, such as economically important diseases like coffee leaf rust. This disease, caused by the fungus *Hemileia vastatrix*, can produce early leaf fall, dry branches, compromise the following harvest, and cause great damage to the producer as well as loss of yield (Silva et al., 2006).

Despite the advances in research, after its introduction in Brazil, rust disease still poses a great threat, increasing the cost of coffee production in the country. Resistant varieties, currently, are the most effective control measures against rust and have been explored in breeding programs (Zambolim, 2016). It is also important to complement that the inclusion of molecular markers allied to these programs can reduce the time that a new cultivar is launched on the market (Sousa et al., 2019).

Considering the high costs of phenotyping for several agronomic traits, especially in perennial plants, the use of models which are able to accurately and quickly predict traits of economic importance, such as the resistance to coffee leaf rust disease, are important tools to accelerate selection in breeding programs.

Several genomic prediction techniques have been used for this purpose, including the Genomic Best Linear Unbiased Predictor (G-BLUP), neural networks, classification trees and their refinements such as *random forest* (Sousa et al., 2021). However, there is still no known study using the ELM model for genomic prediction of rust resistance in *C. arabica*, although its application in other areas revealed to be efficient, being in some instances superior to that of neural network models (Abdullah et al., 2015; Silva et al., 2017).

This study aimed to investigate the use of ELM for the genomic prediction of rust resistance in *Coffea arabica*. With the objective of identifying an effective predictive model for the selection of resistant genotypes, ELM was compared to Artificial Neural Networks (ANN) and Bayesian Generalized Linear Regression (GBLR) in terms of accuracy measures and computational time.

Materials and Methods

The evaluated population consisted of 245 *Coffea arabica* F_2 plants from two parents: Catuaí amarelo IAC 64 (UFV 2148-57), which is susceptible to the rust disease; and Timor Hybrid UFV 443-03, resistant to the rust disease, which is an important source of resistance used in breeding programs.

The experiments were carried out in the municipality of Viçosa, in the state of Minas Gerais (MG), Brazil (20°45'37"S, 42°52'4"W, 648 m altitude). The 245 F2 plants were inoculated with uredospore of pathotype 001 of *H. vastatrix* (Berk & Br), originated in East Africa (Talhinhas et al., 2017).

The symptom evaluation (phenotyping) was carried out in May, June, and August 2009, comprising the first, second, and third repetitions, respectively. Phenotypes were scored on a scale from 1 to 6, in which the highest score obtained in the three repetitions was used to represent the symptom of each plant. Those with no symptoms (score 1) or small chlorotic lesions (score 2) were considered resistant. Phenotypes corresponding to plants that contained large lesions without sporulation (score 3), large chlorotic lesions with small sporulation occupying less than 25% of the area (score 4), lesions with sporulation occupying 25 to 50% of the area (score 5) as well as lesions with sporulation occupying more than 50% of the area (score 6) were considered susceptible to the rust disease (Pestana et al., 2015).

This study used 137 markers – 74 AFLP, 58 SSR, 4 RAPD, and one specific primer – previously described in Pestana et al. (2015). Marker data for each individual were coded for genomic selection analyses. For dominant markers linked in the mating phase to a resistant allele of the Timor Hybrid UFV 443-03 parent, codes -1 and 1 were assigned to the presence and absence of the band, respectively.

Concerning dominant markers in repulsion – allele from the susceptible parent Catuaí Amarelo IAC 64 –, 1 and -1 were also assigned to the presence and absence of the band, respectively. Codominant markers were coded 0 for heterozygote, -1 for bands from the resistant parent and 1 for bands from the susceptible parent (Silva et al., 2017). Quality control of the genotype data used is described in Pestana et al. (2015).

According to genetic mapping studies, the distance between pairs of loci can be divided into four classes, that is to say strongly linked (<1 cM), moderately linked (1-10 cM), weakly linked (11-20 cM) and unlinked (> 20 cM) (Jun et al., 2008). Based on Pestana et al. (2015) map, we considered the markers with a distance shorter than 10 cM – strongly and moderately linked – as being linked to the QTL region.

Extreme Learning Machine is an algorithm proposed by Huang et al. (2015) and aims to train an artificial neural network with only one hidden layer and without iterative processes. In the algorithm, the weights of the hidden layer are chosen randomly, and those of the output layer are chosen analytically.

The general formula of the algorithm is as follows: $o_{ik} = \psi(c_k + \sum_{j=1}^{l} \varphi(b_j + \sum_{d=1}^{m} x_{id}\, w_{dj})u_{jk})$, where ψ and φ are the activation functions; c_k and b_j are the biases; w_{dj} and u_{jk} are the layer weights, and x_{id} are the inputs, and o_{ik} is the output of the network.

In matrix notation $O_{N\times q} = (o_{ik}) = \psi(1c^T + \mathrm{HU}) = \psi(1c^T + \varphi(1b^T + XW)U)$, where $X = [x_1, x_2, \ldots, x_m, 1]$ represents the input vector, and $Y = [y_1, \ldots, y_s]$ is the vector of outputs.

Matrix W is the matrix with the connection weights of the input layer, and matrix β is the weights of the hidden layer. To facilitate the calculations, the biases of the neurons are inserted in the last row of the matrix W.

$$W = \begin{bmatrix} w_{11} & \cdots & w_{1d} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{md} \\ b_1 & \cdots & b_d \end{bmatrix} \qquad \beta = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1s} \\ \vdots & \ddots & \vdots \\ \beta_{d1} & \cdots & \beta_{ds} \end{bmatrix}$$

Determining the matrix H:

$$H^i = \left[x_1^i, \ldots, x_m^i, 1\right] \times \begin{bmatrix} w_{11} & \cdots & w_{1d} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{md} \\ b_1 & \cdots & b_d \end{bmatrix} \Rightarrow H = \begin{bmatrix} f(H^1) \\ f(H^2) \\ \vdots \\ f(H^N) \end{bmatrix}$$

The activation function is denoted by $f(.)$. The matrix H stores the result of all neurons in the hidden layer, obtained from the multiplication between X and W.

By determining the matrix H, we obtain, by solving the linear system below, the matrix of weights β: $H\beta = Y \Rightarrow \beta = H^{+}$.

Matrix β stores information obtained from the network. The generalized Moore-Penrose operator is used to solve the inverse problem, if the matrix is not square, and minimizes the mean square error between the signal arriving from the intermediate layer and

the desired one. In the present work, the number of neurons in the hidden layer varied from 25 to 100 thousand – 25, 50, 100, 200, 500, 1,000, 2,000, 5,000, 7,000, 10,000, 15,000, 20,000, 25,000, 50,000, 75,000, and 100,000.

The learning algorithm of the multilayer neural network is called *backpropagation* or generalized Delta rule and can be divided into two phases: 1) *Forward* phase: the input vector, $x = \{x1,x2,...,xn\}$ is propagated to the output, with the weights and thresholds remaining constant. In the output of each neuron, the respective deviations (errors) are obtained between the desired and obtained outputs. 2) *Backward* phase: in this step, adjustments are made in all neurons at each iteration, resulting in a gradual adjustment to the desired output (Cruz & Nascimento, 2018).

A neural network with only one hidden layer was used, the number of neurons varying from 1 to 100 (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 25, 50, and 100), and as a learning algorithm, the *backpropagation* algorithm.

Proposed by Pérez & Los Campos (2014), the generalized linear models under the Bayesian approach applied to genomic selection can be used for characters of both quantitative and qualitative nature. When continuous characters y_i; $i = 1, 2, \ldots, n$ are present, the equation is given by $y_i = n_i + \varepsilon_i$, where n_i corresponds to the linear predictor and ε_i expresses the independent residuals and present distribution $N(0,w_i^2\sigma_\varepsilon^2)$, in which the weights are represented by w_i. The linear predictor (n) can be represented as follows: $n = 1\mu + \sum_{j=1}^{J} X_j\beta_j + \sum_{l=1}^{L} \mu_l$, where μ is the intercept; X_j represents the matrix of predictors; β_j are the effect vectors that are associated with the columns of X_j, and u_l are the random effect vectors.

The conditional distribution is given by: $p(y|\theta) = \prod_{i=1}^{n} N(y_i \vee \mu+\sum_{j=1}^{J} \sum_{k=1}^{Kj} x_{ijk}\beta_{jk} + \sum_{l=1}^{L} \mu_{li}, \sigma_\varepsilon^2 w_i^2)$.

In the model presented, θ represents the collection of unknown values such as the random effects, the regression coefficients, the intercept, and the residual variance. The a priori distribution assumes the following factorization: $p(\mu)p(\sigma_\varepsilon^2)\prod_{j=1}^{j}p(\beta_j)\prod_{l=1}^{L}p(u_l)$.

The regression coefficients (β_{jk}) can assume informative or uninformative priors. In the present work, the prior is assumed to be a double exponential – used in the Bayesian Lasso.

The selection accuracy was obtained by the estimator $r = r_{gy}/\sqrt{h^2}$, where r_{gy} is the predictive capacity of the genomic selection – correlation between the predicted genomic value and the associated phenotypic value –, and h^2 is the heritability. The heritability used in this work for resistance to *H. vastatrix* pathotype 001 was 0.50, considering phenotypic data (Pestana et al., 2015).

The apparent error rate (AER) is used to calculate the proportion of observations that are misclassified and is given by $AER = 1 - máx_k\ (\hat{p}_{mk})$, where $\hat{p}_{mk}$ represents the proportion of observations in the m-th region belonging to the k-th class.

Data analysis was performed on a computer with a 2.20 GHz Core i5 processor and 8192 MB of RAM, using R software 4.1.1 (R Core Team, 2021). The genomic prediction of the artificial neural networks was performed with the help of the nnet function from the nnet package (Ripley et al., 2016), in which several network topologies were used, varying the number of neurons in the hidden layer. For network convergence, an apparent error rate (AER) was accepted at a maximum of 15% or a maximum of 5,000 iterations on the validation set. The BGLR function, belonging to the BGLR package (Pérez & Los Campos, 2014), was used to estimate the Bayesian generalized models starting at 100,000 iterations, with the first 20,000 observations being discarded (burnin) and the values being saved every 10 observations (thin). The ELM, ANN, and GBLR analyses were repeated 100 times to estimate an average of the processing time of the analyses.

Results and Discussion

The results for the best network topologies for ANN and ELM, as well as the results obtained in GBLR, are presented in Table 1. When compared to the machine learning methodologies, GBLR underperformed, showing predictive capacity about 15% lower than the others (0.345±0.013). Values found for the predictive ability of ELM (0.473±0.010) and ANN (0.468±0.011) were similar.

The average AER obtained for the three analyses performed ranged from 22.8% to 23.7%. Specifically, the lowest average was obtained by GBLR and ELM. When comparing the computational cost between these three techniques for their fitted models, GBLR required the longest analysis time (101.568s±0.366s), followed by ANN (28.719s±0.631s) and ELM (5.229s±0.007s). ELM stood out by being almost 20 times faster than GBLR and about five times faster than ANN.

The way of obtaining parameters for analyses involving machine learning varies according to the methodology used. In this work, different numbers of neurons were used to evaluate the behavior of ELM and ANN in order to choose the best topology for each technique. To evaluate ANN, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 25, 50 and 100 neurons were used, whereas for ELM the variation was 25, 50, 100, 200, 500, 1,000, 2,000, 5,000, 7,000, 10,000, 15,000, 20,000, 25,000, 50,000, 75,000, and 100,000 neurons. This difference can be explained by the fact that the number of neurons in the ELM needs to be much larger than the ANN in order to compensate for random initialization (Huang et al., 2015).

The behavior of the accuracy for ELM and ANN with respect to the number of neurons can be observed in Figure 1. As the number of neurons increased, the accuracy of ELM increased until it practically stabilized at 5000 neurons. ANN presented the same pattern of accuracy values up to 100 neurons. The highest accuracy values were 0.473±0.010 and 0.468±0.011 for ELM with 25,000 neurons, and ANN with 15 neurons, respectively. The lowest accuracy values were 0.138±0.012 and 0.439±0.010 for ELM with 200 neurons, and ANN with four neurons, respectively. The network topologies that allowed a lower error rate in the classification of individuals resistant or susceptible to the rust disease in the validation stage were the same as those observed for accuracy (Figure 2).

As shown in Figure 3, the average processing time of the ELM and ANN analyses is in seconds. The comparative results show that with a much larger number of neurons in its hidden layer, ELM demands much less computational time than ANN. With regard to both analyses, there is a positive correlation between increasing number of neurons and increasing time. If we evaluate only the computational time of the analyses, we can see that using 50 neurons in the hidden layer, the ELM performed the 100 analyses with an average of 0.007s±0.001s, while the ANN had an average of 563.859s±13.086s, being the ELM more than 80,000 times faster than the ANN, in this case.

Results of this work demonstrate the potential of ELM for the genomic prediction of rust resistance in *C. arabica*, a method still with limited application in our specific context. In this work, machine learning algorithms based on ELM and ANN were used and compared among themselves and with those derived from Bayesian generalized linear regression (GBLR) to test the efficiency of genomic prediction for this trait in Arabica coffee.

The machine learning algorithms were more efficient in performing the genomic prediction for rust resistance, presenting a higher accuracy when compared to GBLR. Similar results were observed by Sousa et al. (2021), who used decision trees and their refinements, ANN and GBLR so that to predict rust resistance. The highest accuracy values were observed for ANN, while the lowest were for GBLR. Sousa et al. (2019), in analyses with the G-BLUP method for rust resistance in C. *arabica*, found 46% selective accuracy, a value close to that observed by the ELM and ANN methodologies in this study. As for TAE, ELM, and GBLR stood out for presenting a lower error rate when compared to ANN. Sousa et al. (2021) observed in their comparative work between machine learning and a Bayesian method that the largest errors were found for decision trees and GBLR.

The following algorithms and their respective accuracy and time were compared: ELM (extreme learning machine) -99.03%; ML-ELM (multilayer extreme learning machine) -97.79%; ELM with Gaussian kernel -98.75%; DBN (deep belief network) -98.87%; DBM (deep Boltzmann machine) -99.05%; SAE (stacked auto-encoder) -98.6%; and SDAE (stacked denoising auto-encoder) -98.72%. On the other hand, Ahmad et al. (2018) found that extreme

Table 1. Accuracy (ACC), mean apparent error rate (AER), mean time, and their respective standard errors and total time for the fitted Extreme Learning Machine (ELM), Artificial Neural Network (ANN) and Bayesian Generalized Linear Regression (GBLASSO) models.

Methodology	ACC	AER	Time (s)	Total time (min)
ELM	0.473±0.010	0.228±0.003	5.229±0.007	8.715
ANN	0.468±0.011	0.237±0.004	28.719±0.007	47.864
GBLASSO	0.345±0.013	0.228±0.006	101.568±0.007	169.280

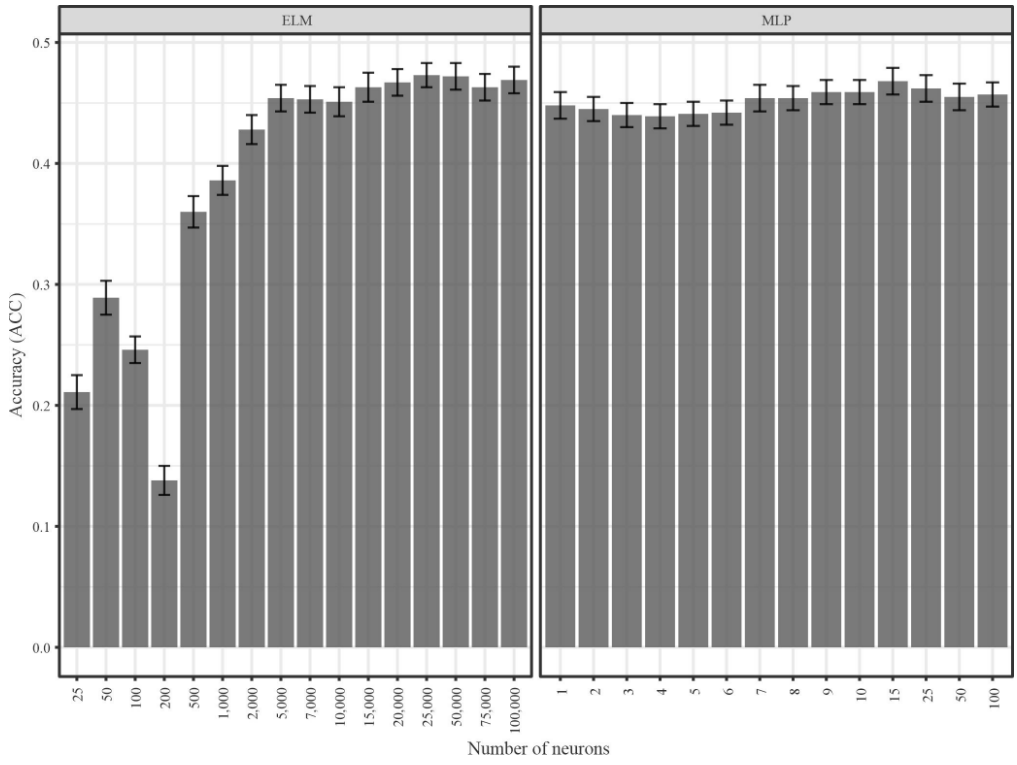

Figure 1. Average accuracy and standard errors for different Extreme Learning Machine (ELM) and Artificial Neural Network (ANN) topologies.

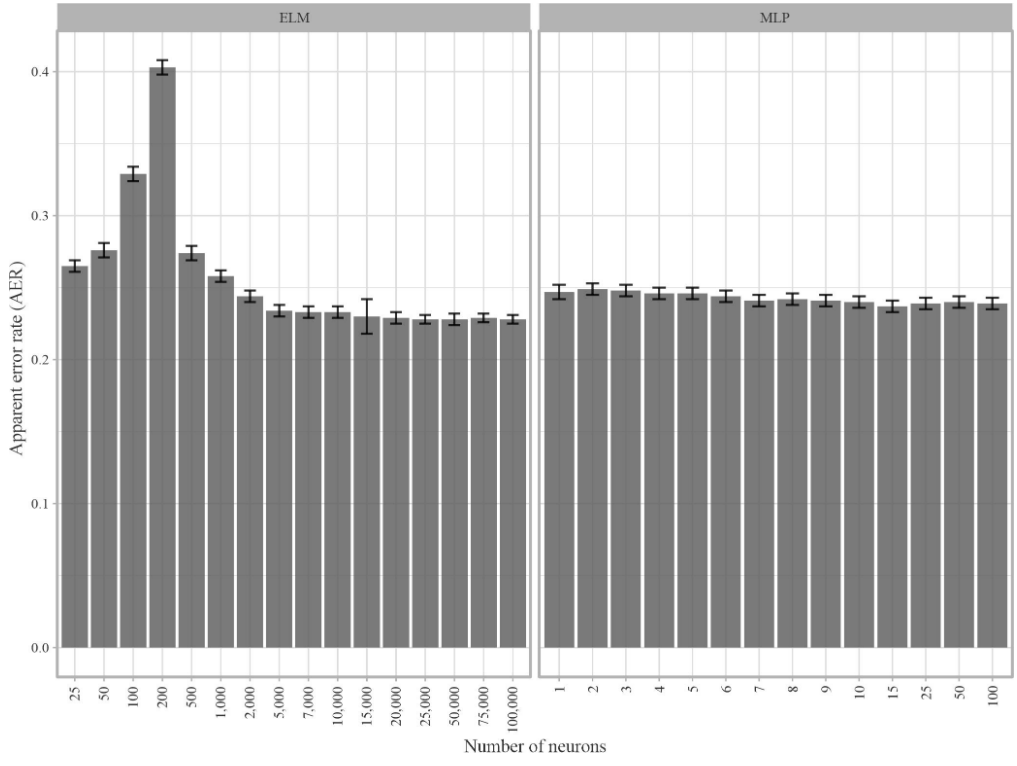

Figure 2. Mean Apparent Error Rate (AER) and standard errors for the different Extreme Learning Machine (ELM) and Artificial Neural Network (ANN) topologies.

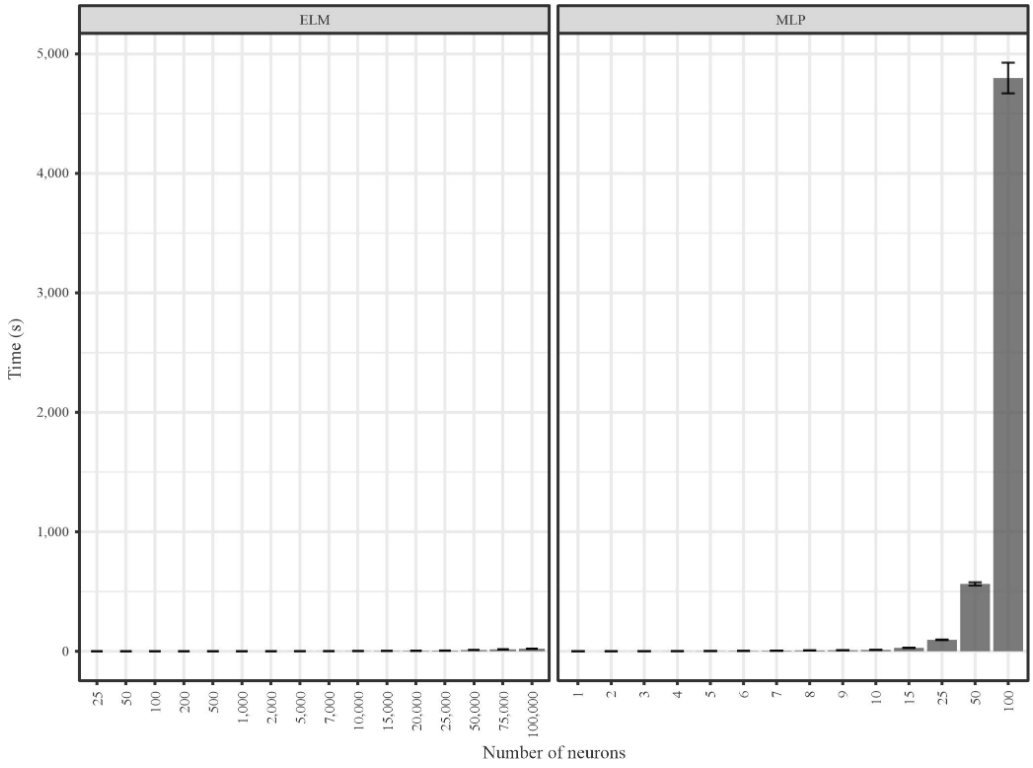

Figure 3. Average time in seconds and standard errors for the different Extreme Learning Machine (ELM) and Artificial Neural Network (ANN) topologies.

learning algorithm obtained higher accuracy (reaching values close to 99%) when compared to the random forest and support vector machine methods.

Ghoneim et al. (2020), using convolutional neural networks and ELM for cervical cancer classification, obtained better predictive results with the ELM algorithm, which presented more than 99% accuracy for problem detection. According to this present work, the ELM algorithm has great generalization ability, as demonstrated by Chen et al. (2019), Huang et al. (2006) and Wang et al. (2022).

The ELM algorithm was the method that showed the shortest processing time, i.e., the one that used the least computational capacity. The results obtained in this study revealed that the algorithm required only 5.22 seconds to process the genomic data, whereas the multi-layer network required 28.79 seconds and the BGLR method, 101.56 seconds. This means that the extreme learning method demanded 5.5 times less time than the multilayer network – which used the backpropagation algorithm – and 19.45 times less time than BGLR. These results are similar to those of Huang et al. (2006), who found that the ELM algorithm was 170 times faster than the backpropagation algorithm. Bai et al. (2014), comparing support vector machines (SVM) with ELM, obtained results in which the network using the ELM algorithm was about 500 times faster than the SVM.

Regarding all the scenarios considered, the extreme learning machine algorithm presented the shortest processing time in relation to the other methods. This behavior is observed due to the fact that this algorithm quickly operates the least squares solution by means of the generalized Moore-Penrose inverse.

Conclusions

1. Extreme learning machines (ELM) and artificial neural networks (ANN) models demonstrate superior accuracy in predicting *Coffea arabica* rust resistance, exhibiting, on average, a 15% greater performance compared to the Bayesian generalized linear regression (GBLR) model.

2. The ELM model stands out for its computational efficiency, revealing processing times 5.5 times shorter than the observed for the ANN model, and 19.45 times shorter in relation to the GBLR model.

3. There is a potential of the ELM model for applications in large-scale analyses, establishing it as an interesting alternative for the genomic prediction of rust resistance in breeding programs.

References

ABDULLAH, S.S.; MALEK, M.A.; ABDULLAH, N.S.; KISI, O.; YAP, K.S. Extreme learning machines: a new approach for prediction of reference evapotranspiration. **Journal of Hydrology**, v.527, p.184-195, 2015. DOI: https://doi.org/10.1016/J.JHYDROL.2015.04.073.

AHMAD, I.; BASHERI, M.; IQBAL, M.J.; RAHIM, A. Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection. **IEEE Access**, v.6, p.33789-33795, 2018. DOI: https://doi.org/10.1109/ACCESS.2018.2841987.

BAI, Z.; HUANG, G.-B.; WANG, D.; WANG, H.; WESTOVER, M.B. Sparse extreme learning machine for classification. **IEEE Transactions on Cybernetics**, v.44, p.1858-1870, 2014. DOI: https://doi.org/10.1109/TCYB.2014.2298235.

CHEN, Z.; GRYLLIAS, K.; LI, W. Mechanical fault diagnosis using convolutional neural networks and extreme learning machine. **Mechanical Systems and Signal Processing**, v.133, art.106272, 2019. DOI: https://doi.org/10.1016/j.ymssp.2019.106272.

CONSÓRIO PESQUISA CAFÉ. **Produção dos cafés do Brasil ocupa 1,82 milhão de hectares em 2022**. Avaiable at: <http://www.consorciopesquisacafe.com.br/index.php/imprensa/noticias/1102-2022-02-18-19-46-19>. Accessed on: Aug. 8 2022.

CRUZ, C.D.; NASCIMENTO, M. **Inteligência computacional aplicada ao melhoramento genético**. Viçosa: Ed. da UFV, 2018.

EHRET, A.; HOCHSTUHL, D.; GIANOLA, D.; THALLER, G. Application of neural networks with back-propagation to genome-enabled prediction of complex traits in Holstein-Friesian and German Fleckvieh cattle. **Genetics Selection Evolution**, v.47, p.1-9, 2015. DOI: https://doi.org/10.1186/S12711-015-0097-5.

GHONEIM, A.; MUHAMMAD, G.; HOSSAIN, M.S. Cervical cancer classification using convolutional neural networks and extreme learning machines. **Future Generation Computer Systems**, v.102, p.643-649, 2020. DOI: https://doi.org/10.1016/j.future.2019.09.015.

HUANG, G.-B.; ZHU, Q.-Y.; SIEW, C.-K. Extreme learning machine: theory and applications. **Neurocomputing**, v.70, p.489-501, 2006. https://doi.org/10.1016/j.neucom.2005.12.126.

HUANG, G.; HUANG, G.-B.; SONG, S.; YOU, K. Trends in extreme learning machines: a review. **Neural Networks**, v.61, p.32-48, 2015. DOI: https://doi.org/10.1016/j.neunet.2014.10.001.

JUN, T.-H.; VAN, K.; KIM, M.Y.; LEE, S.-H.; WALKER, D.R. Association analysis using SSR markers to find QTL for seed protein content in soybean. **Euphytica**, v.162, p.179-191, 2008. DOI: https://doi.org/10.1007/s10681-007-9491-6.

KOECH, R.K.; MALEBE, P.M.; NYARUKOWA, C.; MOSE, R.; KAMUNYA, S.M.; LOOTS, T.; APOSTOLIDES, Z. Genome□enabled prediction models for black tea (*Camellia sinensis*) quality and drought tolerance traits. **Plant Breeding**, v.139, p.1003-1015, 2020. DOI: https://doi.org/10.1111/PBR.12813.

NASCIMENTO, M.; NASCIMENTO, A.C.C.; AZEVEDO, C.F.; OLIVEIRA, A.C.B. de; CAIXETA, E.T.; JARQUIN, D. Enhancing genomic prediction with stacking ensemble learning in Arabica coffee. **Frontiers in Plant Science**, v.15, art.1373318, 2024. DOI: https://doi.org/10.3389/fpls.2024.1373318.

PÉREZ, P.; LOS CAMPOS, G. de. Genome-wide regression and prediction with the BGLR statistical package. **Genetics**, v.198, p.483-495, 2014. DOI: https://doi.org/10.1534/genetics.114.164442.

PESTANA, K.N.; CAPUCHO, A.S.; CAIXETA, E.T.; ALMEIDA, D.P. de; ZAMBOLIM, E.M.; CRUZ, C.D.; ZAMBOLIM, L.; PEREIRA, A.A.; OLIVEIRA, A.C.B. de; SAKIYAMA, N.S. Inheritance study and linkage mapping of resistance loci to *Hemileia vastatrix* in Híbrido de Timor UFV 443-03. **Tree Genetics & Genomes**, v.11, art.72, 2015. DOI: https://doi.org/10.1007/s11295-015-0903-9.

QI, D.; MAJDA, A.J. Using machine learning to predict extreme events in complex systems. **Proceedings of the National Academy of Sciences**, v.117, p.52-59, 2020. DOI: https://doi.org/10.1073/pnas.191728511.

RIPLEY, B.; VENABLES, W.; RIPLEY, M.B. Package 'nnet'. **R package version**, v.7, p.700, 2016.

SILVA, P.P. dos S.; SOUSA, M.B. e; OLIVEIRA, E.J. de. Prediction models and selection of agronomic and physiological traits for tolerance to water deficit in cassava. **Euphytica**, v.215, art.73, 2019.

SILVA, G.N.; NASCIMENTO, M.; SANT'ANNA, I. de C.; CRUZ, C.D.; CAIXETA, E.T.; CARNEIRO, P.C.S.; ROSADO, R.D.S.; PESTANA, K.N.; ALMEIDA, D.P. de; OLIVEIRA, M. da S. Artificial neural networks compared with Bayesian generalized linear regression for leaf rust resistance prediction in Arabica coffee. **Pesquisa Agropecuária Brasileira**, v.52, p.186-193, 2017. DOI: https://doi.org/10.1590/s0100-204x2017000300009.

SILVA, M. do C.; VÁRZEA, V.; GUERRA-GUIMARÃES, L.; AZINHEIRA, H.G.; FERNANDEZ, D.; PETITOT, A.-S.; BERTRAND, B.; LASHERMES, P.; NICOLE, M. Coffee resistance to the main diseases: leaf rust and coffee berry disease. **Brazilian Journal of Plant Physiology**, v.18, p.119-147, 2006. DOI: https://doi.org/10.1590/S1677-04202006000100010.

SOUSA, I.C. de; BARRETO, C.A.V.; CAIXETA, E.T.; NASCIMENTO, A.C.C.; AZEVEDO, C.F.; ALKIMIM, E.R.; NASCIMENTO, M. The trade-off between density marker panels size and predictive ability of genomic prediction for agronomic traits in *Coffea canephora*. **Euphytica**, v.220, art.46, 2024. DOI: https://doi.org/10.1007/s10681-024-03303-8.

SOUSA, I.C. de; NASCIMENTO, M.; SILVA, G.N.; NASCIMENTO, A.C.C.; CRUZ, C.D.; ALMEIDA, D.P. de;

PESTANA, K.N.; AZEVEDO, C.F.; ZAMBOLIM, L.; CAIXETA, E.T. Genomic prediction of leaf rust resistance to Arabica coffee using machine learning algorithms. **Scientia Agricola**, v.78, e20200021, 2021. DOI: https://doi.org/10.1590/1678-992X-2020-0021.

SOUSA, T.V.; CAIXETA, E.T.; ALKIMIM, E.R.; OLIVEIRA, A.C.B.; PEREIRA, A.A.; SAKIYAMA, N.S.; ZAMBOLIM, L.; RESENDE, M.D.V. Early selection enabled by the implementation of genomic selection in *Coffea arabica* breeding. **Frontiers in Plant Science**, v.9, art.1934, 2019. DOI: https://doi.org/10.3389/FPLS.2018.01934.

TAHEREI GHAZVINEI, P.; HASSANPOUR DARVISHI, H.; MOSAVI, A.; YUSOF, K. BIN W.; ALIZAMIR, M.; SHAMSHIRBAND, S.; CHAU, K. Sugarcane growth prediction based on meteorological parameters using extreme learning machine and artificial neural network. **Engineering Applications of Computational Fluid Mechanics**, v.12, p.738-749, 2018. DOI: https://doi.org/10.1080/19942060.2018.1526119.

TALHINHAS, P.; BATISTA, D.; DINIZ, I.; VIEIRA, A.; SILVA, D.N.; LOUREIRO, A.; TAVARES, S.; PEREIRA, A.P.; AZINHEIRA, H.G.; GUERRA-GUIMARÃES, L.L.; VÁRZEA, V.; SILVA, M. do C. The coffee leaf rust pathogen *Hemileia vastatrix*: one and a half centuries around the tropics. **Molecular Plant Pathology**, v.18, p.1039-1051, 2017. DOI: https://doi.org/10.1111/mpp.12512.

TEAM, R.D.C. R: A language and environment for statistical computing., 2021.

WANG, J.; LU, S.; WANG, S.-H.; ZHANG, Y.-D. A review on extreme learning machine. **Multimedia Tools and Applications**, v.81, p.41611-41660, 2022. DOI: https://doi.org/10.1007/s11042-021-11007-7.

ZAMBOLIM, L. Current status and management of coffee leaf rust in Brazil. **Tropical Plant Pathology**, v.41, p.1-8, 2016. DOI: https://doi.org/10.1007/S40858-016-0065-9

Author contributions

Jackson Tavela da Silva: formal analysis, writing - original draft; **Cynthia Aparecida Valiati Barreto**: formal analysis, writing - original draft; **Ana Carolina Campana Nascimento**: project administration, supervision, writing - reviewing & editing; **Camila Ferreira Azevedo**: writing - reviewing & editing; **Dênia Pires de Almeida**: data curation; **Eveline Teixeira Caixeta**: data curation; **Filipe Ribeiro Formiga Teixeira**: writing - reviewing & editing; **Moysés Nascimento**: project administration, supervision, writing - reviewing & editing.

Chief editor: Edemar Corazza
Edited by: Célia Tremacoldi

Data availability statement

Data available upon request: research data are only available upon reasonable request to the corresponding author.

Declaration of use of AI technologies

No generative artificial intelligence (AI) was used in this study.

Conflict of interest statement

The authors declare no conflicts of interest.

Acknowledgments

To Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café (CBP&D/Café), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (408833/2023–8), and Instituto Nacional de Ciência e Tecnologia do Café (INCT) for funding this research. Two of the authors (MN – 310755/2023; CFA – 309856/2023-0) are supported by scientific productivity grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Disclaimer/Publisher's note: