forests

MDPI

Article

Cork Oak Regeneration Prediction Through Multilayer Perceptron Architectures

Angelo Fierravanti [1,2,*], Lorena Balducci [3] and Teresa Fonseca [2]

[1] Universidade de Trás-os-Montes e Alto Douro, UTAD, Quinta de Prados, 5000-801 Vila Real, Portugal
[2] Centre for the Research and Technology of Agroenvironmental and Biological Sciences, CITAB, Inov4Agro, Universidade de Trás-os-Montes e Alto Douro, UTAD, Quinta de Prados, 5000-801 Vila Real, Portugal; tfonseca@utad.pt
[3] Département des Sciences Fondamentales, Université du Québec à Chicoutimi, Chicoutimi, QC G7H 2B1, Canada; lbalducc@uqac.ca
* Correspondence: angelofierravanti@utad.pt

Abstract: In Mediterranean ecosystems, a thorough understanding of seedling regeneration dynamics as well as a good predictive ability of the process is essential for sustainable forest management. Leveraging the predictive capacity of the multilayer perceptron (MLP) as recognized as artificial intelligence methodology, the authors analyzed a real case study with a dataset encompassing environmental, ecological, and forestry variables. The study focused on the cork oak (*Quercus suber*, L.) seedling regeneration dynamic, which is a critical process for maintaining ecosystem resilience. A set of 10 MLP with a block from 5 to 50 neurons with hyperbolic tangent (TanH), linear (LIN), and Gaussian (GAUS) activation function were tested and their performance for predictive purposes was compared with traditional quantitative approaches. The MLP configured with 40–50 neurons per activation function (TanH, LIN, GAUS) demonstrated outstanding predictive performance, achieving an area under the curve (*AUC*) of the receiver operating characteristic and precision-recall scores above 0.80. These models made few prediction errors, effectively explaining the majority of the data variance, as indicated by a high generalized R^2 and a low mislearning ratio. This approach outperformed traditional statistical models in predicting seedling regeneration. Tree density, stand density index, and acorn number played an important role, influencing the cork oak seedling prediction. In conclusion, the results of this research determined the importance of an AI classification modeling technique in the prediction of cork oak regeneration, providing practical references for future forest management strategy decisions.

Keywords: acorns; boosting; neural network; artificial intelligence; machine learning; *Quercus suber*; seedling; survival; regeneration; Mediterranean forest

check for updates

Academic Editor: Nikolay S. Strigul

Received: 15 February 2025
Revised: 27 March 2025
Accepted: 31 March 2025
Published: 8 April 2025

Citation: Fierravanti, A.; Balducci, L.; Fonseca, T. Cork Oak Regeneration Prediction Through Multilayer Perceptron Architectures. *Forests* **2025**, *16*, 645. https://doi.org/10.3390/f16040645

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

For decades, the use of artificial intelligence (AI) in various fields (from healthcare to environmental sciences) was limited by computational constraints and skepticism about its applications [1–3]. Nowadays, there is a huge interest in AI given its enhanced ability to process big data, identify patterns, and make predictions. Over time, the skepticism turned into the recognition of work on AI methodologies. Notably, in 2024, John Hopfield from Princeton University and Geoffrey Hinton from the University of Toronto, who pioneered tools for understanding neural networks (NN) as far back as 1982, were awarded the Nobel Prize in Physics [4–7]. At that time, AI models were constrained by the rudimentary

hardware available including basic graphics support provided by cards like the Graphics Controller Adapter CGA [8]. Notwithstanding these constraints, their pioneering efforts established the groundwork for contemporary developments in AI. Consequently, the exponential enhancement in computer component capabilities and the corresponding advancements in AI technologies have facilitated the resolution of complex issues that are challenging to evaluate using non-AI models. In contrast to AI models, non-AI models demonstrate less accuracy and precise recall due to their inherent rigidity, as rule-based systems rely on predetermined rules and logic, making them inflexible in processing unstructured data [9]. Similarly, models as well as linear regression have set decision boundaries, which can pose difficulties for nonlinearly separable data [9]. Moreover, the non-AI model frequently employs static datasets, constraining its capacity to adapt to emerging patterns when these datasets are updated with new information. Consequently, within the existing scientific framework, the incorporation of AI models across diverse domains (e.g., engineering, medicine, informatics, physics, and environmental science) has markedly enhanced adaptive and resilient methodologies for addressing complex problem-solving, especially in contexts necessitating nonlinear and dynamic data processing. Recent breakthroughs in AI have shown the ability to enhance predictive accuracy and optimize decision-making processes, enabling innovative solutions to persistent scientific challenges. The trend of development in AI-related research is undergoing exponential expansion [10]. Scopus.com reports that during the period 2024–2025, there were 33,118 original scientific articles in English exclusively related to the term "artificial intelligence", predominantly within the fields of medicine, engineering, social sciences, biochemistry, and genetics, with a lesser representation in environmental sciences and agriculture.

In particular, forestry research with AI is an emerging domain, analyzing ecological and management forest issues such as weather forecasting, forest monitoring, and wildfire detection [11,12] and has become extremely useful for database classification problems [13,14]. However, AI-based approaches in reforestation and regeneration research are still limited in employment. While the ecological and reproductive limitations of cork oak have been extensively studied using conventional statistical methods [15–18], the utilization of AI techniques, such as MLP, for predicting cork oak regeneration is still a relatively unexamined area of study. However, a focus on the functional verification of the prediction accuracy of forest regeneration modeling is still lacking.

Nowadays, the capabilities of NN in both environmental and forestry studies have been widely increased [19–22]. The NN provides a mechanism for forecasting nonlinear and categorical data, and it is intriguing to assess how NN can address challenges with enhanced accuracy and precision-recall. Furthermore, given that NN is adept at addressing multicollinearity issues among variables (input–output), this versatility applies to the analysis of ecological data, hence facilitating the prediction of seedling density based on readily obtainable data. In many forest studies with NN, the methods most used are the single hidden layer of multilayer perceptron (MLP). Already used in biology, ecology, and environmental studies, the MLP represents a robust analysis for resolving classification problems [23,24]. However, due to the complexity of the analysis, the limitations of computer calculation, and the difficulty of compression, MLP (like other type of NN) had not found a place in forestry topics until recently. Currently, MLP is being used more frequently, and due to the time lag of knowledge accumulated, further studies are needed to verify its performance in forestry disciplines. In particular, there is a significant lack of knowledge about the use on NN in modeling the natural regeneration dynamics of forest species. Selected as a case study, the authors focused on the regeneration of cork oak (*Quercus suber* L.) species in a Mediterranean environment, elucidating the utility of utilizing MLP to model natural regeneration during a defined period.

Cork oak is an evergreen oak that grows in the western Mediterranean Basin is and more abundant in the Iberian Peninsula [25,26], where it plays a crucial role in sustaining both environmental regulation and local economy [25,27]. Portugal has the largest cork oak cover, spanning 719.9 thousand hectares (22.3% ± 1% of mainland Portugal's forests [28]) with a major bioeconomic impact [25,29]. The species is well-adapted to the Mediterranean climate, coping with hot and dry summers. Evolution has led the cork oak to have a high thermotolerance and ability to acclimatize to drought [30]. Cork oak's natural regeneration consists of different reproductive stages (e.g., flowering, seed dispersal, seedling emergence, seedling survival, establishment and sapling growth). Each stage is constrained by various biotic and abiotic factors [31,32], including the vapor pressure deficit (*VPD*), minimum absolute temperature (*Tn*), canopy density, tree age, and diameter, are incorporated as factors influencing acorn development and seedling establishment. Additionally, variability across measurements and locations is introduced by factors such as predation, tree health, seed dispersal by jays, soil mineral availability, and insect infestations, significantly impacting the model predictions and reliability of ecological evaluations [17]. All of these effects may be evaluated as random effects or latent effects if the values are not available. Cork oak has unisexual reproduction with monoecious flowers. The key challenge of the cork oak reproductive system is the high degree of self-incompatibility due to the lack of synchronism between male pollen release and female flower receptiveness, which affects fertilization [33,34]. Trees begin to flower between March and June, where the wind acts mainly in pollination, and the seed dispersal of cork oak acorns between October and February is limited around individual trees [35]. An adequate seed supply starts in trees aged 15–20 years old [36].

Environmental factors, particularly low temperatures and aridity, substantially influence acorn production, leading to annual and biennial variations in yield [37]. Furthermore, acorn production varies among individual trees, and constraints in dispersion, together with post-dispersal mortality due to predation, further constrain spontaneous regeneration [35]. Then, the negative effects of climate anomalies, such as extreme heat, prolonged drought, and precipitation events (severe storms with high rainfall) lead to a general deterioration in the seedling establishment conditions and increase the susceptibility to pests and diseases, weakening the regeneration system [38,39]. In this context, the development from acorn to seedling (seedling growth) is becoming increasingly challenging for forest managers, requiring appropriate measures to ensure successful regeneration. The early prediction of the endogenous variables of oak regeneration could improve the success rate of natural or artificial plantations, although, the definitions of seedling size and survival at the beginning of germination should also be better considered.

Together with the reproductive and ecological challenges linked to cork oak regeneration, there is a deficiency in predictive understanding regarding the initial stages of this process, particularly in recognizing the factors that influence seedling survival and establishment. Currently, only a few research studies have employed AI models in cork oak forests. Studies in Portugal used data mining to predict oak habitats [40] and applied machine learning to identify pastures using remote sensing [41]. To our knowledge, there have been no studies that used neural networks or other AI models to predict how many cork oak seedlings will grow, especially when the prediction is based on forestry measurement data (e.g., using stand density indicators). While the ecological and reproductive limitations of cork oak have been extensively established using conventional statistical methods, the utilization of AI techniques, such as MLP, for predicting cork oak regeneration is still a relatively unexamined area of study. Utilizing AI to understand the limits has the potential to improve natural regeneration strategies and forest management. Specifically, the AI research on development from acorn to seedling should provide challenges for forest man-

agers, necessitating suitable strategies to guarantee effective regeneration. Certainly, the initial predictions derived from a precise AI methodology regarding endogenous variables of oak regeneration could enhance the success rate of both natural and artificial plantations.

Our study aimed to determine which factors have the most substantial impact on the natural regeneration densities by using MLP models. The second goal was to determine the capability of MLP models to classify the cork oak seedling density within subplots with the maximum accuracy and precision-recall for a comparison with traditional quantitative approaches. We hypothesized: (i) environmental factors strongly drive the natural regeneration establishment, and (ii) MLP models perform properly, predicting natural regeneration with increased unbiasedness and accuracy.

2. Materials and Methods

2.1. Study Area and Experimental Design

The study was conducted in a permanent forest cork oak site linked to the Agenda TransForm project, which focuses on the sustainable management and the regeneration of cork oak [42]. The study site is located in Mogadouro (Bragança region), in northeastern Portugal (Lat 41.3525 N; Long −6.7617 W). The region is situated at an elevation of 570 m above sea level, is characterized by dystric and umbric leptosols, and exhibits a typical Mediterranean climate with hot, dry summers and cold, damp winters. The mean annual precipitation is 558 mm, with the rainy season spanning eight months from October to May, peaking in November and December [43].

The measurements presented in this study were carried out in 2022 and 2023. The experimental design consisted of two circular plots (A1 and A2) with an area of 500 m^2 (radius = 12.62 m). In each circular plot, we used two types of inventory sampling designs: (1) linear transect and (2) two-step radial cluster. (1) The linear transect design had 12 square subplots, yielding a total of 24 subplots when accounting for both plots. Each subplot encompassed an area of 1 m^2 (1 m × 1 m), arranged sequentially in a rectangular structure measuring 12 m in length and 1 m in breadth. The center of the linear transect was situated at the midway of each plot (Figure 1). (2) The radial cluster design consisted of five clusters of subplots, each covering a square area of 4 m^2. Each cluster consisted of four-square subplots measuring 1 m^2 (1 m × 1 m), resulting in a total of 20 subplots (4 subplots × 5 clusters) in each plot, and an overall total of 40 subplots for plots A1 and A2 combined. One cluster was located at the center of the plot, while the other four clusters were placed radially, equidistantly positioned from the center (6 m), the plot boundary, and from each other [44]. Due to the presence of two types of inventory sampling designs, we evaluated the cumulative area of overlap. This overlap was 8 m^2 out of 64 m^2 (12.5%), inclusive of all plots. We did not eliminate the overlapping observations, since recent statistical studies conducted on the same plot on the overlap issues by Fierravanti and Fonseca (submitted, [44]) indicated that the inclusion or exclusion of these measurements, representing a 12.5% design overlap and a 10% sample measurement overlap relative to the total, yielded non-statistically significant differences in the findings. The overlap area, mostly located in the plot center and encompassing half of the central cluster (refer to Figure 1), did not create statistical bias.

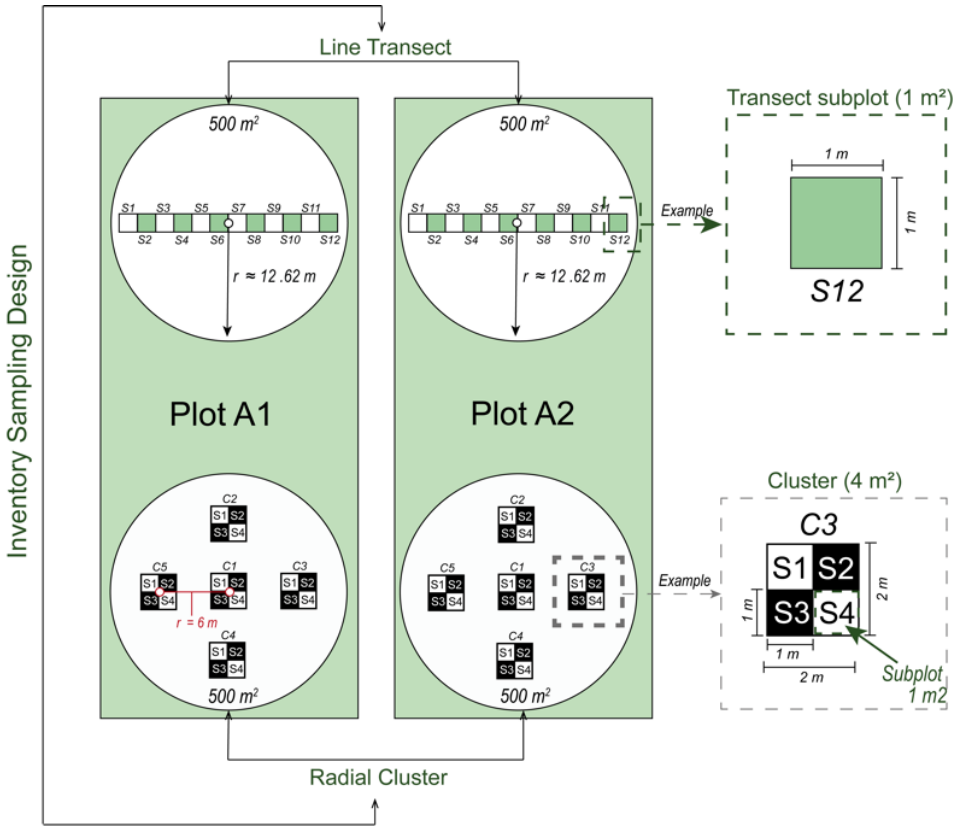


Figure 1. Inventory sampling design (ISD) for the sampling seedlings and acorns, featuring two linear horizontal transect plots and two radial cluster plots. Subplots were labeled with the prefix "S" and numbered 1 to 12 for the LT plots and 1 to 4 for the RC plots. Clusters in the RC plots are denoted with the prefix "C" and numbered 1 to 5. Note that due to graphical considerations, the sizes of the ISD in the plots are not represented to scale; in actuality, the ISD are smaller. Additionally, the plots are consistently displayed separately for visual clarity, but in actuality, they overlap.

2.2. Regeneration and Tree Samplings

The number of acorns (*Acorns*) and total number of living and dead seedlings per m^2 (*TS* and *DS*, respectively) were recorded in two circular plots (A1 and A2). In particular, we sampled the study plots on 46 *DOY* (February 2022), 157 *DOY* (June 2022), and 263 *DOY* (September 2022). In 2023, there were 4 *DOY* (January), 181 *DOY* (June), 304 *DOY* (October exclusively for plot A2), and 312 *DOY* (November). The dates were chosen to cover the flowering and seed dispersal periods of the cork oak acorns, which are between February and June and between October and February, respectively [35]. The living seedlings were discriminated on the basis of stem height measured from the base to the top (H), following a threshold of $H \leq 10$ cm (TS1) and $H > 10$ cm (TS2). The acorns (Acorns) and Seedlings were classified in three classes: Class 0 (C0), Class 1 (C1), and Class 2 (C2), where C0 represents the absence of elements (acorns or seedlings), C1 encompasses instances with a count of acorns and seedlings ranging from 1 to 4, and C2 for both the acorns and seedlings encompasses instances with counts exceeding 4. In the study, the presence of dead seedling (*PDS*) with stem height (*H*) under 50 cm was classified in two classes (YES and NO) (Table A1). The dead seedlings during each sampling data were between 0 and 2 (98.5% of cases), and 1.5% was between 3 to 4, therefore, due to the reduced number of values up to 2, we decided to only use two classes.

To describe the stand structure, a circular plot with an area of 500 m^2 was established, within which all trees with a diameter equal to or greater than 7.5 cm were recorded, along with their spatial location (distance and azimuth relative to the center). Each plot was

then subdivided in four quadrants (Q1, Q2, Q3, and Q4) of 125 m^2 each, arranged in a clockwise direction. The number of trees in each quadrant was assessed, and the density was extrapolated to a per-hectare unit, allowing for variability assessment within the plot (Figure 2).

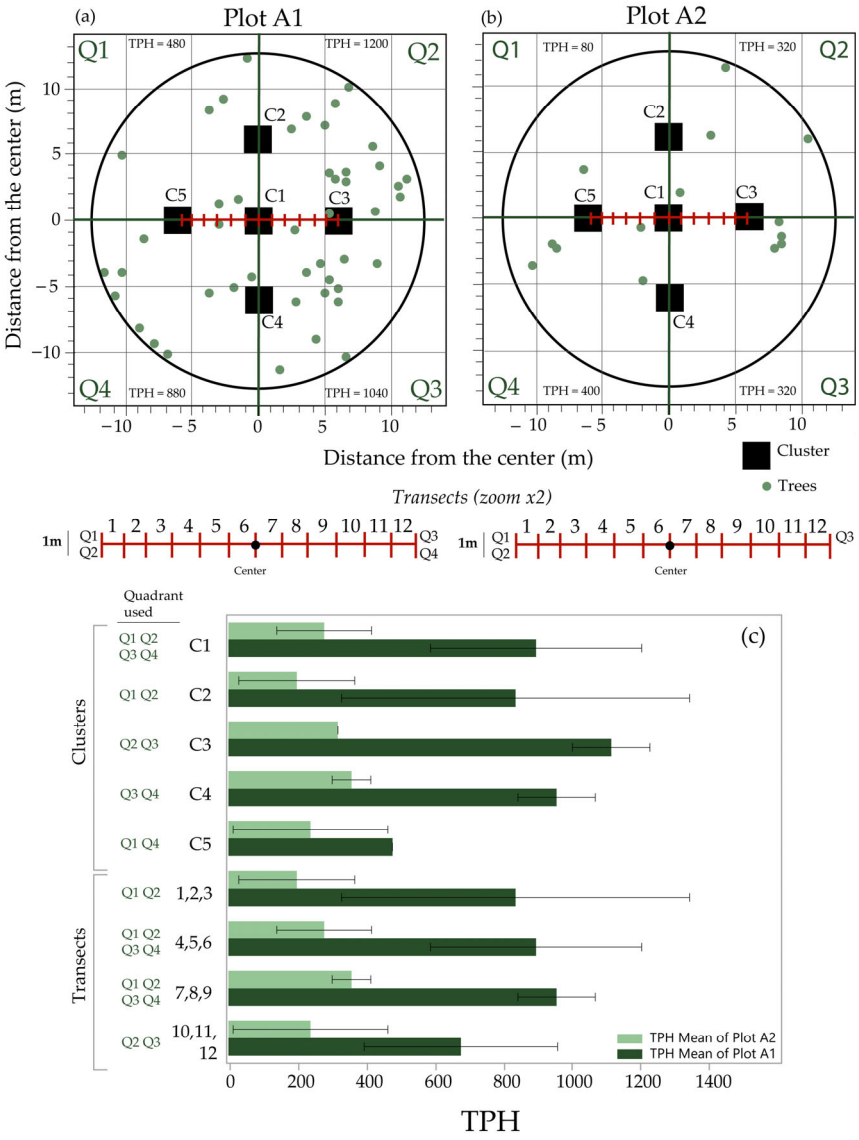


Figure 2. Demonstration of how the tree per hectare (*TPH*) values were derived for each quadrant (Q1, Q2, Q3, and Q4). (**a**,**b**) Represents plots A1 and A2, respectively, containing the tree used for the calculation of *TPH*. (**c**) Panoramic representation indicating the quadrant with which the *TPH* was related for each cluster and set of three transect subplots.

The variable stand density index (*SDI*) was calculated using the expression:

$$SDI = TPH \left(\frac{dg}{25} \right)^{b}$$

where *TPH* is the number of trees per hectare, *dg* is the quadratic mean diameter of the trees, and *b* is the allometric coefficient whose value for cork oak is defined as −1.806 [17,45]. In this application, the calculation of the index was carried out for the subset of mature trees, identified here as those with a minimum diameter of 25 cm.

We also characterized the surrounding environment, measuring the slope of each plot, classified as the absence or presence of 5° (≈8.74%) inclination, then further classified into

two classes (YES and NO) (Table A1). The measurements were made at the same time of the acorn and seedling samplings.

The *RND* was derived through a multi-step process: (1) Exploratory factor analysis (EFA) using the maximum likelihood estimation with Varimax rotation; (2) assessment of the sampling adequacy (Kaiser–Meyer–Olkin ≥ 0.60) and sphericity (Bartlett's Test, $p < 0.05$); and (3) validation of the factor structure via confirmatory factor analysis (CFA) using the *CFI* (>0.95), *TLI* (>0.95), *NFI* (>0.90), *RMSEA* (<0.08), *SRMR*, and χ^2 ($p \geq 0.05$). The *RND* values for each plot, along with the normalized variables contributing to it, are detailed in Table A2. Factor selection was based on eigenvalues, with Factor 1 (eigenvalue = 2.52) explaining 62.01% of the variance. Factor 2 (eigenvalue = 0.8) was excluded. Variables retained from Factor 1 for CFA included *TS1* (0.94), *TS2* (0.57), *Acorns* (0.67), and dead seedlings (0.66) (Table A2).

2.3. Multilayer Perceptron Analysis

The multilayer perceptrons (MLPs) used in this study were constructed with a block of five neurons each, employing the hyperbolic tangent (TanH), linear (LIN) and Gaussian (GAUS) activation functions across two consecutives hidden layers (H1 and H2), from input to output (Figure 3).

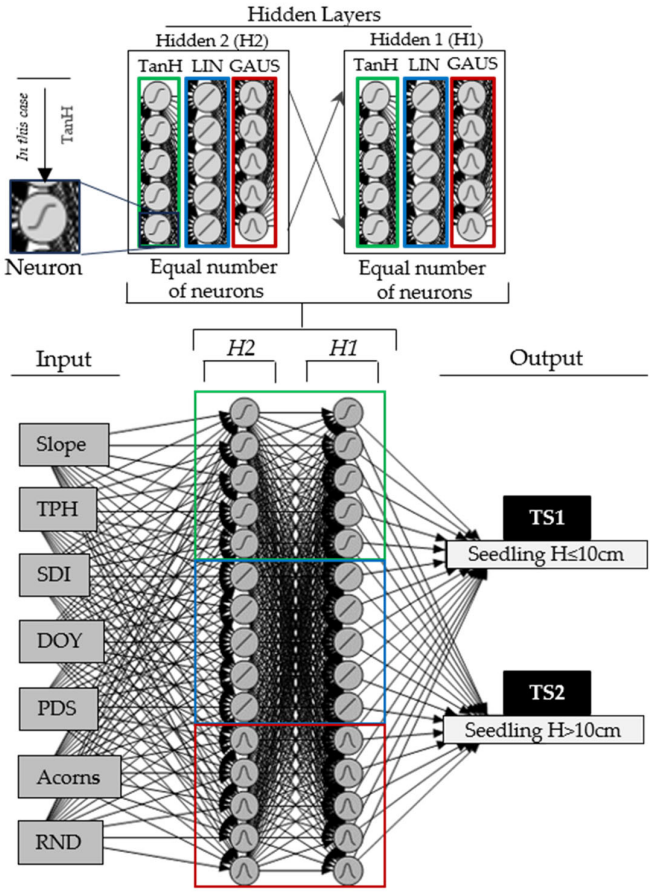


Figure 3. Schematic description of the ecological process, highlighting the variables and AI methodologies employed. Descriptive variables (input): the slope position, the number of trees per hectare (*TPH*), stand density index (*SDI*), sampling day of the year (*DOY*), the presence of dead seedlings (*PDS*), the random normal distribution (*RND*). For the dependent variable (output): seedlings height, *TS1* and *TS2* correspond to $H \leq 10$ cm and $H > 10$ cm, respectively. H2 and H1 represent the neurons' hidden layers connected to the outputs and neurons and inputs and neurons, respectively. The green, blue, and red rectangles signify the cluster of neurons ($\times 5$) in the hidden layer, utilizing the hyperbolic tangent (TanH), linear (LIN), and Gaussian (GAUS) activation functions, respectively.

Consistently, the number of neurons remained uniform across both layers. A total of 10 MLPs models were tested (Figure 4), ranging from a minimum of 30 neurons (e.g., MLP-5 = (5 TanH + 5 LIN + 5 GAUS neurons) × 2 hidden layers) and a maximum of 300 neurons (e.g., MLP-50 = (50 TanH + 50 LIN + 50 GAUS) × 2). The use of these three activation functions (TanH, LIN, and GAUS), dispersed across hidden layers with a comparable number of neurons, improved the model's capacity to generalize and forecast regeneration patterns. The TanH function translates inputs to a range of −1 to 1, encapsulating intricate relationships. The LIN preserves linearity, which is advantageous when input–output interactions are proportional. GAUS, characterized by its bell-shaped curve, represents localized responses and gradual fluctuations. This balanced amalgamation utilizes nonlinearity (TanH and GAUS) and linearity (LIN), guaranteeing consistency and stability while averting bias toward any singular activation function. In the context of classification, the integrated use of TanH, LIN, and GAUS enables the model to adjust to various data structures and enhances generalization [46,47]. To assess the predictive performance of the model, each MLP model underwent initial validation using a portion of the dataset generated through the *k*-fold method [48–50]. In our study, we employed a *k*-fold value of 6, resulting in a cross-validation dataset comprising 16.6% (83/17 split) of the total data, as it balanced the accuracy between $k = 5$, which performed poorly on validation and training (GR^2), and $k = 7$, which affected the prediction quality. The MLPs' tested inputs were slope (surrounding environment, categorical variable), number of trees per hectare (*TPH*), stand density index (*SDI*) (stand tree variables), *DOY* (time, categorical variable), *PDS*, and number of acorns classified (regeneration, categorical variable), and *RND* (random numerical variable), while the output of the models was the seedling categorized by height (*H*) in *TS1* ($H \leq 10$) and *TS2* ($H > 10$) (regeneration categorical variable) (Figure 4).

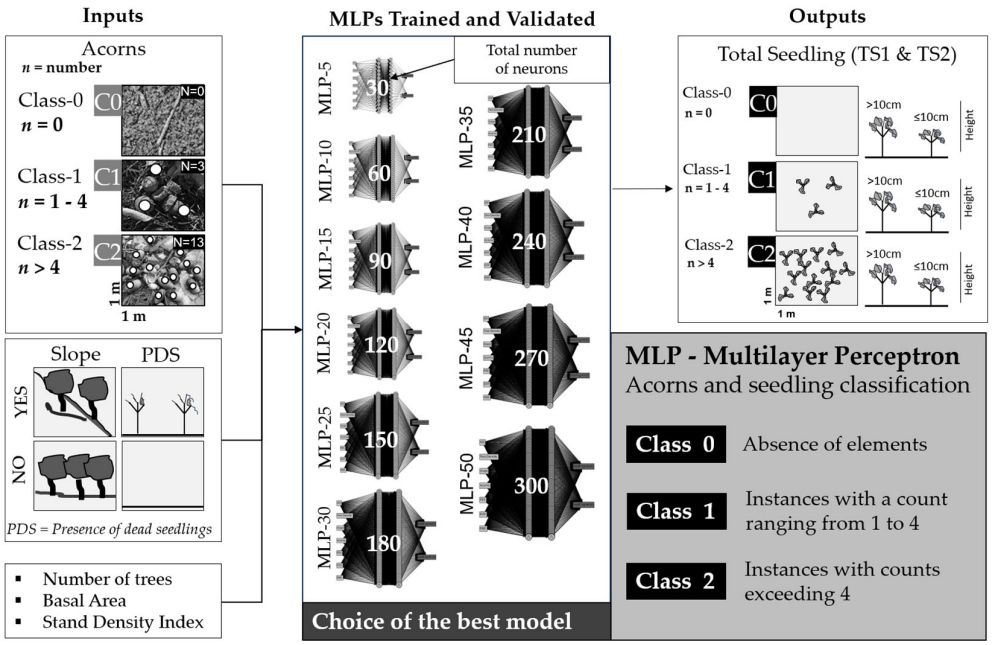


Figure 4. MLP models and the inputs and output classification of the variables used in our study.

2.4. Performance Analysis of MLP Models

The performance of the MLP models was assessed systematically in five steps (Figure 2): (1) measuring the accuracy of the MLPs; (2) measuring the precision; (3) measuring the recall (sensitivity); (4) identifying whether there were inputs that were problematic in the prediction of output; and (5) comparisons with other AI and non-AI statistical analyses. In general, one method to obtain an initial understanding is the confusion matrix,

which was used to gain insights into the MLPs' accuracy, precision, and recall. The confusion matrix was employed to assess the misclassification in both the training and validation predictions [51,52]. A comprehensive metric analysis was conducted to thoroughly examine these indicators of model prediction accuracy. Eight metric parameters were utilized as indicators to evaluate the model's fit regarding the accuracy, precision, and recall of the MLP models for both the validation and training datasets (please refer to Table 1 for further information).

Table 1. List of metric parameters used to evaluate the accuracy and precision of the MLP model, along with their interpretations and relevant references (Ref.).

Metric Parameter	Abbr.	Definition	Interpretation	Ref.
Generalized R^2	GR^2	Measures model performance; higher values close to 1 indicate better performance, while negative values imply worse performance compared with a model predicting the data's average.	Higher GR^2 values indicate superior model fit, while negative values reflect a poor performance.	[53]
Entropy R^2	ER^2	A variant of GR^2 based on entropy theory; higher values close to 1 represent better performance.	Similar to GR^2, higher values signify improved accuracy and model fit.	
Negative log-likelihood	*NLL*	Evaluates the likelihood of the model predictions given the data; lower *NLL* values denote a better fit.	Smaller *NLL* values reflect a superior model performance.	[54]
Root average squared error	*RASE*	Calculates the square root of the mean squared differences between the predictions and actual data values.	Smaller *RASE* values indicate a better predictive performance.	
Mean absolute deviation	*MAD*	Computes the average absolute difference between the predictions and actual data values.	Lower *MAD* values suggest a higher model accuracy.	[55,56]
Misclassification rate	*MR*	Proportion of misclassifications to total classifications; ranges from 0 (perfect classification) to 1 (all data misclassified).	Smaller *MR* values denote better a classification accuracy.	
Receiver operating characteristic (ROC) area under curve (*AUC*)	ROC_{AUC}	The area under the receiver operating characteristic curve; it evaluates a binary classification model's performance by plotting the true positive rate against the false positive rate at various thresholds.	*AUC* ranges from 0.5 (random classification) to 1 (perfect classification); higher values indicate better discrimination between classes.	[51,57]
Precision-recall (PR) area under curve	PR_{AUC}	Evaluates the precision and recall for imbalanced datasets, focusing on positive class performance; ranges from 0 to 1.	A PR_{AUC} value of 1 indicates perfect precision and recall, highlighting strong performance, particularly in imbalanced data scenarios.	[51,52,58]

2.5. Comparison Between MLP Models and Other Prediction Statistical Models

The effectiveness of MLP in terms of modeling and performance was assessed by comparing the same accuracy, precision, and recall indicators used in the comparisons among the MLPs. For the comparison, we chose to compare the MLP with the most widely used AI and non-AI methodology as well as studies in various sectors of forestry ecology. Therefore, as AI modeling methods, we chose to use bootstrap forest (BT), decision

trees (DT), k-nearest neighbor (kNN), Naive Bayes (NB), support vector machine (SVM), and neural network boosted (NNB). Notably, the NNB models consisted of single-layer architectures with three neurons utilizing the TanH activation function. Specific parameters for boosting were defined including a maximum of 100 interactions and a learning rate to 0.1. In the case of non-AI models, we chose the nominal logistic (NL) and generalized regressions with Lasso (GRL), Elastic Net (GRE), and Bridge (GRB) estimation methods. To identify a more suitable traditional model for comparison with the MLPs, the NL and generalized regressions models were further evaluated against each other using Akaike information criterion correct (*AICc*) and Bayesian information criterion (*BIC*) indicators. Moreover, the fit of the NLs and their source of variation was further evaluated in detail using the square tests (χ^2). Similar to the MLPs, a *k*-fold of 6 was employed for these statistical analyses.

A detailed analysis based on the G^2 (the likelihood ratio chi-square statistic) was carried out for the BF. The G^2 is a statistic used to evaluate how well a split (or division) works within a decision tree. The G^2 measures the difference between the expected and the observed data distribution after the split. A higher G^2 value indicates a better separation between the target classes. When constructing BF trees, the G^2 is calculated for each possible split, and the one that maximizes this statistic is selected. The best MLP, AI, and traditional statistical (non-AI) prediction models were also visually compared with a mosaic plot to check the accuracy.

2.6. Assessing the Effects of Inputs on Outputs

An independent resampled input (IRI) technique was used to identify the role of the surrounding environment, stand characteristics, regeneration variables, time, and random inputs on the prediction of the seedling density output after the best MLP was used. This methodology permitted us to assess the effects of different inputs while maintaining their independence through resampling methods on the *TS1* and *TS2* [59]. From the IRI, we considered the total effect, which is the comprehensive impact of a predictor variable (e.g., *Acorns*) on the input (*TS1* and *TS2*) including all direct and indirect pathways and the main effect, which is the direct impact of a predictor variable on the response, without considering the interactions or the effects of other predictors [59]. All analyses performed in this research were conducted using JMP®, Version 18 Pro, Student Edition (SAS Institute Inc., Cary, NC, USA, 1989–2024).

3. Results

3.1. Performance Analysis of the MLP Models

Our study showed that the MLP-40 and MLP-50 multilayer perceptron models best showed regeneration, as their GR^2 and *NLL* values were similar. The advanced GR^2 in MLP-40 showcased the model's capacity to clarify data discrepancies. MLP-40 demonstrated more accuracy than MLP-50 (which has more neurons) in data classification (Figure 5). The GR^2 of MLP-40 varied between 0.47 and 0.57, depending on the training or validation dataset. GR^2 had a variation of ± 0.06 between the training and validation datasets (Table A3). The model accurately predicted the growth of short, high-density seedlings in *TS1* better than that of tall, low-density seedlings in *TS2*. We achieved this by examining the seedling density and various other metrics. There were the same number of observations in both the training dataset and validation dataset. However, the validation dataset had 80% less mean negative log-likelihood (*NLL*) than the training dataset. The *NLL* differences in the training datasets for *TS1* and *TS2* were about 5%–8% on average. Predictions for short seedlings were lower, which showed that TS1 had better MLP-40 accuracy for seedling density classifications per square meter (C1, C2, and C3). The MLP-40 demonstrated accu-

racy, as evidenced by the training and validation confusion matrix (Table 2). In addition to the GR^2 and *NLL*, the results showed that MLP-40 was better than all other MLP models in every way (Figure 5). These results included model accuracy, prediction error, variance explanation, misclassification, and deviation. As shown by ER^2, MLP-40 was better at explaining the variance in the data. It also produced lower *RASE* and *MAD* values, which means that it was more accurate at making predictions and had fewer errors. Furthermore, MLP-40 had a reduced *MR*, signifying enhanced classification efficacy.

The MLP models with the highest and lowest neuron counts (MLP-50 and MLP-5) exhibited suboptimal performance. The MLP-50 model's increased NLL fit with more neurons suggests difficulties in data modeling. The performance of the MLP-50 network might indicate overfitting due to the increased number of neurons, but this was not found, as the GR^2 of the training set was equal to or slightly lower than that of the validation set, indicating that the model generalized well, rather than memorizing the training data. Despite MLP-50 possessing 60 more neurons compared with MLP-40, the GR^2 and *NLL* fit–accuracy trade-off indicates an imbalance. Analogous to the GR^2–*NNL* connections, MLP-5 forecast seedling quantities with the least precision and fit. MLP-5 was less complex to compute than MLP-50 due to having 10 times fewer neurons. The GR^2–*NLL* relationship indicates that MLP-40 optimally balanced model accuracy and efficiency (Figure 6).

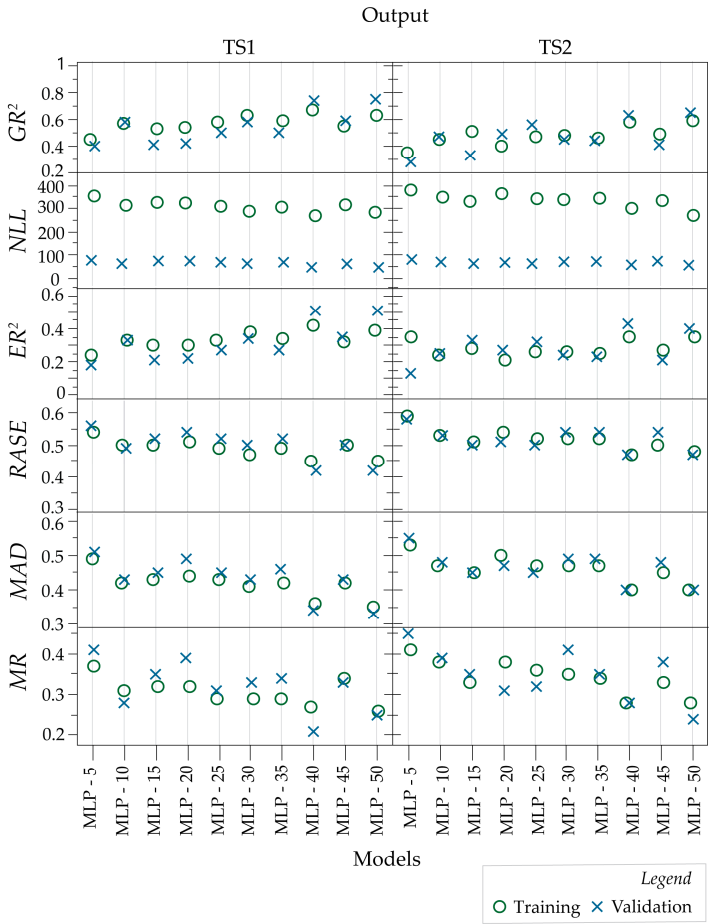


Figure 5. Metric indicators of the prediction capacity of *TS1* and *TS2* (output) made by the MLP-40 model on the training and validation dataset. Generalized R-square (GR^2), negative log-likelihood (*NLL*), entropy R-square (ER^2), root average square error (*RASE*), mean absolute deviation (*MAD*), and misclassification rate (*MR*).

Table 2. Confusion ratio of the MLP-40 model. Accuracy of the MLP-40 model in classifying the three seedling classes (C1, C2, and C3) based on the output (*TS1* and *TS2*); actual values were obtained during the field sampling and predicted by MLP-40. The numbers written in bold are related to the correct prediction in terms of the actual-predicted ratio.

	Output	Actual	Predicted Rate					
			Training			Validation		
			C0	C1	C2	C0	C1	C2
MLP-40 model	*TS1*	C0	**0.82**	0.13	0.04	**0.94**	0.03	0.03
		C1	0.17	**0.77**	0.07	0.16	**0.79**	0.05
		C2	0.14	0.38	**0.48**	0.17	0.33	**0.50**
	TS2	C0	**0.75**	0.20	0.05	**0.67**	0.29	0.04
		C1	0.15	**0.78**	0.07	0.07	**0.86**	0.07
		C2	0.11	0.38	**0.51**	0.24	0.29	**0.48**

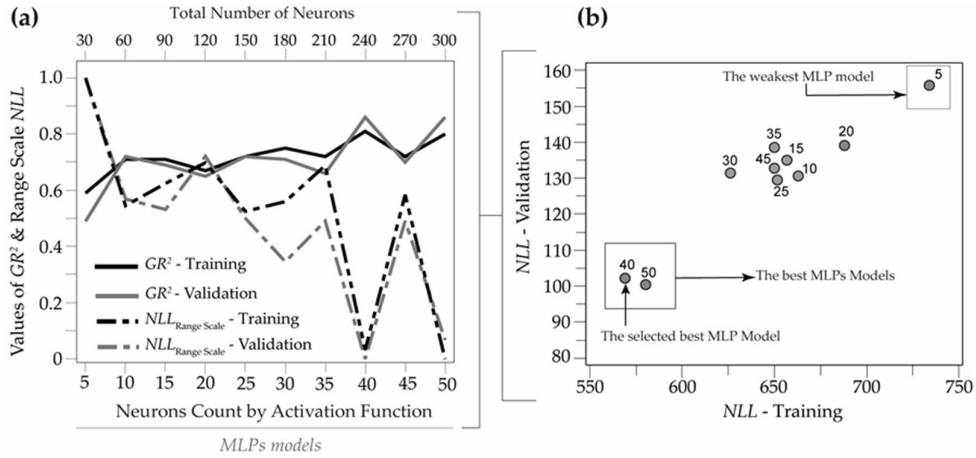


Figure 6. (**a**) Representation of the generalized R-square (GR^2) and the scale log-likelihood range (*NLL*) for all of the MLPs tested; (**b**) relationship between the *NLL* of the training dataset and *NLL* validation dataset using the actual *NLL* values.

3.2. Comparison of the Best MLPs with Other AI and Non-AI Models

The assessment of the model's effectiveness in distinguishing the seedling counts at different choice thresholds validated the compromise between sensitivity and specificity for the area under the ROC curve (ROC_{AUC}). MLP-5 was recognized as the least effective model, and MLP-40 and MLP-50 were designated as the most accurate AI models in this study. Moreover, MLP-40 exhibited an enhanced performance on the training datasets, attaining the best overall accuracy, which was attributed to the increased ROC_{AUC} values (Figure 7), whereas MLP-50 displayed inferior ROC_{AUC} values, indicating a 10% difference in the ROC curve and true positive sensitivity relative to the less complicated MLP-40.

The ROC_{AUC} value aligns with the other metrics used for comparing AI and non-AI models, reinforcing their applicability in resource-efficient scenarios (Figure 8).

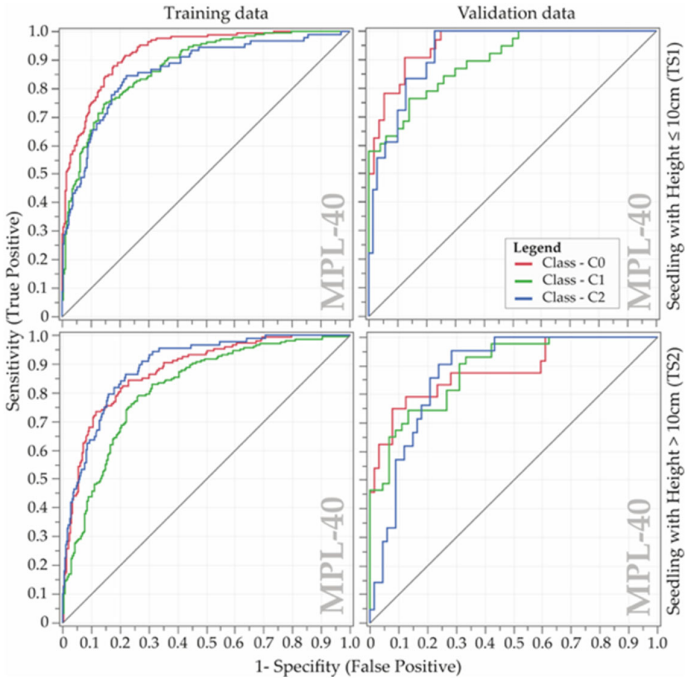


Figure 7. Training and validation dataset. Receiver operating characteristic (ROC) curve for the MLP-40 model (the best MLP model) considering the seedling density (C0, C1, and C2), distinguishing between those with height (*H*): $H \leq 10$ (*TS1*) and $H > 10$ (*TS2*).

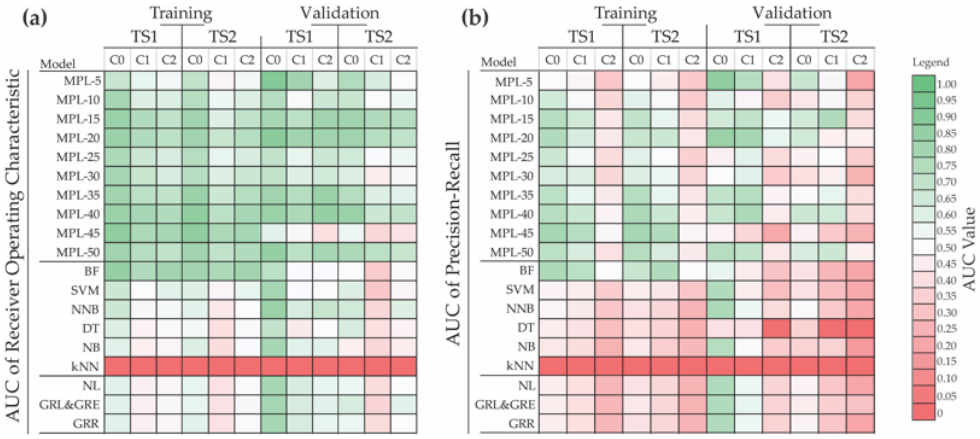


Figure 8. The area under the curve (*AUC*) values represent: (**a**) the receiver operating characteristic (ROC) and (**b**) the precision-recall (PR) curves. Both (**a**) and (**b**) were calculated for both the training and validation datasets. *AUC* values were assessed for the output variables TS1 and TS2, categorized into three classes (C0, C1, and C2).

Furthermore, MLP-40 showed significant effectiveness versus the most challenging AIs such as NB and kNN. The Naive Bayes and k-nearest neighbor algorithms exhibited suboptimal performance due to elevated entropy (ER^2). Within the metric parameters, the precision of the AI, bootstrap forest (BF), and the other neural network category, NNB, was comparable to that of the MLPs. Nonetheless, the BF model often generated erroneous positive predictions, particularly with the prediction of taller seedlings (*TS2*) in count class C2 or the representative class, which denotes more than four seedlings per square meter. Misunderstandings on the false positive predictions of BF persisted across the validation dataset. Nonetheless, it was established that MLP-40 surpassed BF, while confusion was observed in MLP-40 (Figure 9).

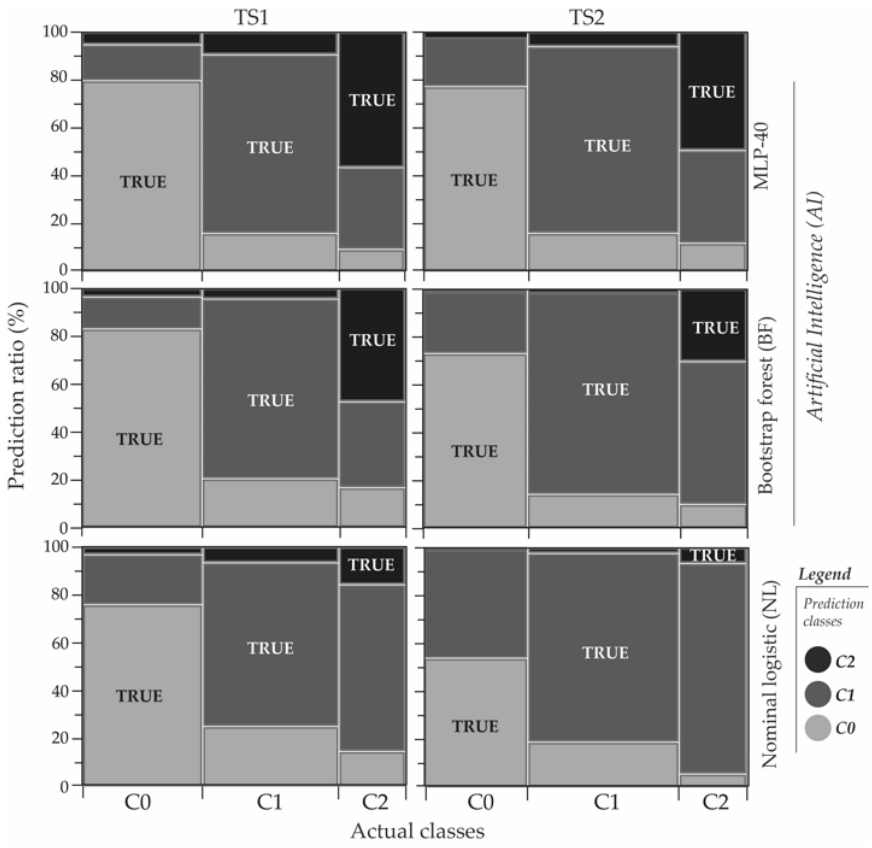


Figure 9. Mosaic plots of the best models. The mosaic plots show the predicted data compared with the actual data, for the best MLP model (MLP-40), the best other AI model (bootstrap forest, BF), and the best traditional statistical predictive method (nominal logistic, NL).

3.2.1. Comparison Between MLP Models and Other Dedicated AI Models

As above-mentioned, the MLP models demonstrated a superior performance across various metrics, and their high accuracy and precision-recall were well-documented by the high values of the *AUC*. This was followed by BF, the other AI, which was competitive when focusing on training datasets, and certain models, such as NNB and SVM, showed moderate predictive performance. However, there were more negative differences in the classification of C2 seedlings, particularly between the MLPs and the NB, or DT or kNN models. The BF model also achieved a robust level of precision, as shown by PR_{AUC} values in the training phase (up to 0.88), although it slightly declined during validation. The study of other several AI models showed that improvements to the neural network (e.g., NNB models) did not make them as accurate as the more complex MLP models. The GR^2 for the NNB remained largely unchanged at 0.27 during training and 0.20 during validation. While NNB is a faster one-layer neural network model for prediction than MLP-40 and is somewhat good at boosting, it struggled with classification accuracy, especially in the C2 category (seedling number above four), which is a problem that BF also had (Table 3). The MLP-40 and MLP-50 models did well on both the training and validation datasets for *TS1* and *TS2*. This was shown by their ROC_{AUC} scores, which measured how accurate they were. Neglecting the MLP models, the BF model was the best AI model, obtaining the same ROC_{AUC} value of 0.92 as MLP-50 when using the training dataset. Nonetheless, the NNB and support vector machine (SVM) models exhibited modest ROC_{AUC} values of 0.82 and exhibited modest PR_{AUC} values of 0.69 and 0.70, respectively, in terms of precision-recall. The study ultimately confirmed that the C2 group faced the greatest challenge in predicting seedling outcomes. In comparison to BF's accuracy and precision, the NB, DT,

and KNN models performed poorly in categorizing the C2 category across all seedling types, particularly in terms of precision and recall (Figure 9).

Table 3. Confusion ratio of the NNB model. Accuracy of the NNB model in classifying the three seedling classes (C1, C2, and C3) based on the output (*TS1* and *TS2*); actual values were obtained during the field sampling and predicted by NNB. The numbers written in bold are related to the correct prediction in terms of the actual-predicted ratio.

	Output	Actual	Predicted Rate					
			Training			Validation		
			C0	C1	C2	C0	C1	C2
NNB model	*TS1*	C0	**0.80**	0.20	0.01	**0.82**	0.16	0.02
		C1	0.28	**0.69**	0.03	0.20	**0.78**	0.03
		C2	0.18	0.71	**0.11**	0.25	0.70	**0.05**
	TS2	C0	**0.60**	0.40	0.00	**0.60**	0.41	0.00
		C1	0.26	**0.75**	0.00	0.34	**0.66**	0.00
		C2	0.09	0.91	**0.00**	0.11	0.90	**0.00**

3.2.2. Comparison Between MLP Models and Non-AI Models

The non-AI models did not surpass the MLP-40 in accuracy and precision-recall during the predictive tasks. The ROC_{AUC} and PR_{AUC} values, which were inferior to those of MLP-40, suggest that the non-AI model was less effective in categorizing seedlings. The least powerful MLP-5 model outperformed the non-AI statistical methods, demonstrating superior capability in managing imbalanced class distributions. The goodness-of-fit uncertainty measurements (entropy) in classification via GR^2 and ER^2 indicated values that were 22%–33% lower in the non-AI models compared with MLP-40, hence affirming the enhanced predictive capability of AI models (Table A3). The confusion matrices indicated that the nominal logistic (NL) models consistently misclassified C2 as C1 across all seedling dimension categories (Figure 8). The MLP models, in contrast, had superior accuracy in classifying the items properly.

3.3. Main Effect of the Input Factors

The independent resampling input (IRI) of the MLP-40 model indicated that the stand variables associated with tree density and the stand density index (*SDI*) for mature trees over 25 cm in diameter had the most pronounced influence on the seedling density prediction (main effect 25%, total effect 51%) (Figure 10). The second variable that defined the response of the number of trees per hectare (*TPH*) (main effect 16%, total effect 40%) (Figure 10).

The quantity of acorns in terms of absence, scarce, or moderate presence was the third most significant factor for the prediction of seedlings. The total effects of time and slope were also considerable, ranging from 22% to 17%, respectively, for the seedling density, while them main effects were 2% and 3%, respectively. Furthermore, among the variables influencing the seedling density over time, the presence of perished seedlings demonstrated a relatively minor effect (main effect 12%, total effect 2%). The *RND* may accommodate random effects absent from the major stated variables. In comparison to the other non-AI models, the NL model demonstrated the optimal statistical fit for both shorter and taller seedlings (*TS1* and *TS2*), as indicated by *Δ (AICc)* and *Δ (BIC)* values of zero; hence it was classified as NL (Table 4). Furthermore, the overall model test of NL demonstrated a significant $\chi^2_{(Wald)}$ with a $p < 0.0001$ for both models with seedlings in *TS1* and *TS2* as the responses, indicating that the predictors of *TS1* and *TS2* are associated with the actual

responses of *TS1* and *TS2*. Conversely, the lack of fit test results indicated a $\chi^2_{(Wald)}$ that was not significant ($p > 0.05$) (Table A4), suggesting that the model adequately fit the data.

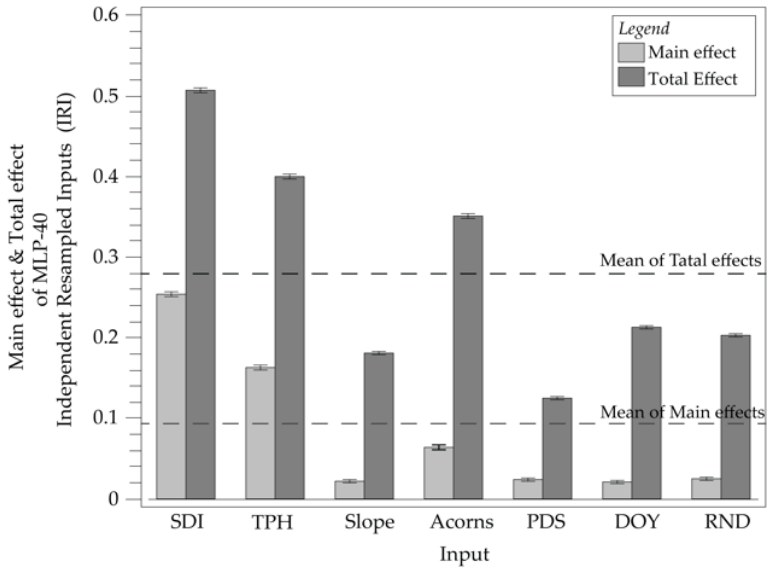


Figure 10. Total (in dark gray) and main (in light gray) effect values obtained by the independent resampled input (IRI) of the MLP-40 model (the model with best metric values compared with other models).

Table 4. Akaike information criterion correct (*AICc*) and Bayesian information criterion (BIC) indicators concerning the nominal logistic (NL) and generalized regression models (Lasso, Ridge, Elastic Net). The *Δ (AICc)* and *Δ (BIC)* were the difference between the *AICc* and *BIC* of the best model and other models for each response (*TS1* and *TS2*). The values of 0.00 in *Δ (AICc)* and *Δ (BIC)* represent the best models.

Response	Models	*AICc*	*Δ (AICc)*	*BIC*	*Δ (BIC)*
TS1	NL	790.36	0	869.06	0
	GRL	808.91	−18.55	880.07	−11.01
	GRE	811.13	−20.77	886.14	−17.08
	GRR	813.72	−23.36	892.58	−23.52
TS2	NL	800.02	0	878.88	−38.52
	GRL	806.14	−6.12	861.77	−21.41
	GRE	820.26	−20.24	840.36	0
	GRR	814.58	−14.56	893.44	−53.08

The comprehensive model test validated that NL may elucidate phenomena (χ^2_{Wald}, $p < 0.0001$), and the findings offer partial endorsement for MLP-40's IRI. The NL demonstrated that *Acorns* were the primary variable for *TS1* in all models (χ^2_{Wald}: 73.03–78.37, $p < 0.0001$). *SDI* and *TPH* significantly influenced *TS2*, particularly in the generalized regression Lasso (GRL), Elastic Net (GRE), and Ridge (GRR) models. *SDI* had more significance than *Acorns* (χ^2_{Wald}: 57.80–57.95, $p < 0.0001$) (Table 5).

The NL and regression models had similar GR^2 and ER^2 values, differing by around 4%; however, they remained inferior to MLP-40. The BF results, closely resembling those of MLP-40, indicated that, similar to MLP-40, the presence of acorns significantly influenced seedling selection. However, in contrast to MLP-40, timing (or, in this instance, the *DOY*) significantly influenced the model predictions. The acorns significantly influenced tiny seedlings (*TS1*), yielding a G^2 score of 71.94. Nonetheless, they still influenced bigger seedlings (*TS2*), yielding a G^2 score 53% less than *TS1*. Moreover, the temporal variable

DOY strongly affected all groups of seedlings, exhibiting similar G^2 values in *TS1* and *TS2* (Table 6).

Table 5. Effect tests of nominal logistic (NL), generalized regressions (GRs), Lasso (GRL), Elastic Net (GRE), and Ridge (GRR). The $\chi^2_{(Wald)}$ is Wald's chi-squared significance test with its *p*-value: $p \geq 0.05$ (ns), $p < 0.05$ (*); $p < 0.01$ (**), $p < 0.0001$ (***).

Response	Source of Variation	$\chi^2_{(Wald)}$			
		NL	GRL	GRE	GRR
TS1	*SDI*	0.00 ^{ns}	28.62 ***	21.51 ***	9.35 **
	TPH	0.00 ^{ns}	77.45 ***	71.01 ***	57.58 ***
	Slope	6.41 *	6.39 **	6.39 *	6.35 *
	Acorns	78.37 **	73.2 ***	73.19 ***	73.03 ***
	PDS	5.37 ^{ns}	5.08 ^{ns}	5.08 ^{ns}	5.09 ^{ns}
	DOY	0.84 ^{ns}	0.82 ^{ns}	0.82 ^{ns}	0.85 ^{ns}
	RND	0.22 ^{ns}	0.21 ^{ns}	0.21 ^{ns}	0.22 ^{ns}
TS2	*SDI*	0.00 ^{ns}	57.95 ***	57.80 ***	12.59 ^{ns}
	TPH	0.00 ^{ns}	6.78 *	7.03 *	57.50 ***
	Slope	2.33 ^{ns}	1.08 ^{ns}	1.16 ^{ns}	2.24 ^{ns}
	Acorns	24.27 ***	19.48 ***	19.66 ***	25.57 ***
	PDS	2.99 ^{ns}	1.84 ^{ns}	1.95 ^{ns}	3.33 ^{ns}
	DOY	8.49 **	6.03 **	6.17 *	8.15 *
	RND	1.65 ^{ns}	0.31 ^{ns}	0.39 ^{ns}	1.66 ^{ns}

Table 6. Bootstrap forest (BF) model splits, G^2 is the likelihood ratio chi-square statistic, and portion results.

	Output	Input	Splits	G^2	Portion
BF Model	*TS1*	*SDI*	51	12.77	0.05
		TPH	61	9.73	0.04
		Slope	94	17.99	0.06
		Acorns	84	71.94	0.26
		PDS	40	8.41	0.03
		DOY	349	53.13	0.19
		RND	513	103.39	0.37
	TS2	*SDI*	8	19.62	0.07
		TPH	11	25.21	0.09
		Slope	141	20.75	0.08
		Acorns	132	33.46	0.13
		PDS	43	8.94	0.03
		DOY	302	53.60	0.2
		RND	493	104.51	0.39

4. Discussion

The results of the study confirm both hypotheses, showing that the environmental factors effectively and strongly drive the establishment of natural regeneration over time. In detail, the findings corroborate the notion that stand characteristics associated with the stand density index (*SDI*) significantly affect the seedling density, alongside the critical impact of acorn quantity and seasonal variations. Moreover, the findings prove that the MLP outperformed earlier prediction models, including both AI and non-AI systems, in predicting natural regeneration with enhanced accuracy and precision-recall.

4.1. The Relative Importance of Main Explicative Factors on Seedling Regeneration

The study indicates that the MLP-40 IRI data demonstrated that *SDI* and *TPH* are significant variables influencing the seedling density of cork oak. These forests, marked by recurrent drought conditions, experience intensified competition for water, light, and nutrients, negatively impacting forest health and resilience. The MLP-40 model showed that the *SDI* and tree density of the stand had a large influence on the prediction of the classes of the number of seedlings. Even without modeling, differences in the *SDI* and tree density were evident across the stands. The model effectively incorporated and predicted these distinctions. Anticipating differences in the *SDI* and tree density are very important as they impact germination, and as a whole process, regeneration from acorn to seedling. These regeneration factors influence the availability of essential resources for growth and the survival of seedlings such as light, water, and nutrients [17,60]

The *SDI*, as a vital dendrometric parameter, is an essential metric for forest managers, allowing them to evaluate when a stand nears its biological carrying capacity. In this setting, management actions like thinning are essential for reducing competition, fostering tree development, and improving cork quality.

Our findings corroborate previous studies, indicating that sustaining ideal *TPH* or *SDI* values benefits both the environment and forestry [17,45]. Particularly for *TPH* in the previous two decades, studies indicate that *TPH* levels between 50% and 75% are recommended to promote natural regeneration [17]. These values ensure that sufficient light penetrates to facilitate the growth of understory vegetation and the regeneration of oaks while simultaneously reducing the fuel loads and enhancing the resilience of the stand. By adding the *SDI* or *TPH* to predictive models, forest managers can better predict and deal with the problems caused by environmental stresses, which will make the cork oak ecosystems in northern Portugal healthier and more resilient. Indeed, recent studies show that high tree density can intensify competition for resources, potentially impeding the growth and survival of seedlings [17,61,62]. Conversely, a lower tree density may afford the seedlings with ample resources, promoting their growth and survival [63]. Recent studies have highlighted the importance of these factors in oak regeneration. However, we contend that this axiom may not hold true in regions with Mediterranean climates, where rainfall and transpiration levels differ [64]. Instead, in warm areas within cork oak habitats, the density of trees and the *SDI* in a cork oak population can influence microclimates, resource availability, and competition, all of which are vital for seedling survival and growth [65]. An increase in tree density per hectare could promote regeneration in cork oak forests in Portugal [66,67], and higher tree densities could create a more favorable microclimate for seedling establishment and growth [68,69]. This is due to several factors; for instance, higher tree densities can provide more shade, thereby reducing the soil temperature and evaporation, which in turn conserves soil moisture [70]. One of the most significant factors contributing to this favorable microclimate is humidity.

The transpiration from a larger number of trees can increase the local humidity, which can be beneficial for seedling growth [71,72]. Increased tree density can enhance plant nutrients in soil, primarily because more trees result in greater leaf litter accumulation. As this litter decomposes, it enriches the soil with organic matter and nutrients [15,73]. Indeed, while higher tree densities can facilitate regeneration, it is essential to recognize that there is a threshold beyond which competition for resources (such as light, water, and nutrients) can become detrimental to seedling growth [74,75]. Therefore, finding the optimal tree density is crucial for promoting successful regeneration [76,77]. A model considering the impact of density on natural regeneration dynamics offers valuable insights for decision-making support in forest management.

In terms of seedling survival, several factors can contribute to acorn and seedling survival in cork oak woodlands. An interesting environmental factor identified was topography [78,79], specifically the slope of the terrain where acorns initiate the germination process, ultimately giving rise to seedlings. The MLP-40 model confirmed that the slope of the terrain plays a role in the regeneration of oaks, especially during the transition from acorn to seedling. The terrain slope can impact the water runoff, soil erosion, and solar radiation levels, all of which are pivotal factors affecting germination and growth [80]. Water is essential for acorn germination and seedling growth. However, steep slopes can exacerbate water runoff, potentially drying out the soil and rendering it unsuitable for germination. Conversely, gentler slopes have a greater capacity to retain water, fostering successful germination and growth. Soil erosion poses another challenge on sloping terrain [70]. Erosion can displace acorns and young seedlings or deplete essential nutrients from the soil [78]. Therefore, ensuring soil stability on the slopes is decisive for the successful establishment of cork oak seedlings. Finally, the orientation of the slope can influence solar radiation, which is a key factor for photosynthesis. For example, north-facing slopes in the Northern Hemisphere receive less direct sunlight than south-facing slopes, potentially resulting in cooler and wetter conditions [81,82].

4.2. The Ability of MLP to Prediction of Seedling Regeneration Compared with Other Models

The research identified superior performance in some MLPs (e.g., MLP-40 and MLP-50) for the prediction and categorization of seedlings. MLP-40, with a total of 240 neurons, had the maximum accuracy and carried out an exceptional fit to the data, as seen by the metric indicators. This outcome corroborates extensive research across several fields indicating that the utilization of MLP yields significantly high accuracy, precision, and recall [83–88]. Furthermore, our use of various activation functions, evenly distributed throughout the two hidden layers, demonstrated an enhancement in the capacity to discern complicated, linear, and nonlinear connections, along with category categorization of the data related to regeneration. The MLP results affirm that a two-hidden-layer configuration is accurate in predicting and classifying regeneration concerning seedlings, surpassing boosted neural networks and other artificial intelligence methods [89,90]. In cork oak regeneration situations, the integration of TanH, LIN, and GAUS activation functions demonstrated efficacy in predicting outputs, even in the presence of negative values, gradient-based optimization, and nonlinear input values [91–93]. Moreover, the study validates the significant efficacy of MLPs, including MLP-40, in comparison to conventional quantitative methods. Our work confirms the premise that MLP models effectively forecast natural regeneration, demonstrating that MLPs can model cork oak regeneration with a high degree of impartiality and accuracy, surpassing both AI and non-AI models.

In particular, MLPs have shown more flexibility and more accuracy and precision-recall than models that can only capture linear relationships, such as the nominal logistic, which is a statistical model classically used for binary classification problems. Non-AI models are simple, fast, and provide good performance when the relationship between the input and output is linear or nearly linear. The most closely related method to MLP is NNB. However, our findings revealed that NNB yields inferior metric parametric results, translating in lower accuracy. Despite NNB's capability of forming large neural network models through boosting, with a single hidden layer of neurons, it exhibited lower accuracy compared with MLP in our specific case. The confusion matrix underscores the high classification performance of MLP-40 and MLP-50, aligning with findings from environmental research that validate their utility in assessing outcomes with high precision. In addition, the MLP results outperformed other ML statistical analysis. Both BF and SVM demonstrated weaker performance compared with MLP. BF exhibited robustness against overfitting

and could handle complex nonlinear relationships [91,94,95]. Similar to MLP, it can be computationally intensive and less interpretable than simpler models. However, unlike MLP with a high number of neurons, it had less accuracy in the validation.

The MLP-40 also surpassed SVM in accuracy. SVM is a type of machine learning model that can perform both linear and nonlinear classifications as well as MLP, and it is known for its robustness and ability to handle high-dimensional data (see [96,97]). However, in our case, the SVM was less accurate than MLP.

We found that SVM mistook the prediction of classes during the training and validation. We speculate that this mistake of typology was due to the fact that SVM can be sensitive to the choice of kernel and hyperparameters and may not perform well with many overlapping classes [98]. This was the case in our instance, where SVM misclassified the number of seedlings in C2.

5. Conclusions

The research was conducted over two years in a mature Mediterranean cork oak forest in Portugal, focusing on natural regeneration and using MLP-models. Our approach highlighted the key role of stand tree density in the regeneration success of cork oak. The slope was also identified as a vital factor for seedling growth. In our study, the MLP models provided elevated accuracy and precision in recall, with a discrete number of neurons and combination of activation functions. MLP models surpassed non-AI models in recognizing the significance of acorn quantity for seedling density. The MLP-40 model demonstrated strong performance in predicting the seedling density in pure cork oak stands in northern Portugal, highlighting the stand density index (*SDI*) and trees per hectare (*TPH*) as key predictors. Finally, this study offers practical guidance for forthcoming decisions regarding forest management strategies. The determination of the optimal density values should integrate insights from this study alongside recommendations from previous research [17,45] to support the more effective management of cork oak natural regeneration.

The results emphasize the importance of managing the mature tree density in pure cork oak stands to enhance the establishment and recruitment of natural seedlings. Thinning practices should be designed to optimize tree spacing while preserving ecosystem functionality and cork productivity. Further research is recommended to validate these management strategies across different environmental conditions and investigate additional ecological factors that may influence seedling dynamics and stand structure. By integrating AI-driven models into management practices, forest practitioners can better anticipate the outcomes of different interventions, ultimately promoting sustainable cork oak regeneration in Mediterranean ecosystems.

Author Contributions: Conceptualization, A.F. and T.F.; Methodology, A.F.; Software, A.F.; Validation, A.F., T.F. and L.B.; Formal Analysis, A.F.; Investigation, A.F.; Resources, T.F.; Data Curation, A.F.; Writing—Original Draft Preparation, A.F. and L.B.; Writing—Review and Editing, T.F. and L.B.; Visualization, A.F.; Supervision, T.F.; Project Administration, T.F.; Funding Acquisition, T.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work obtained partial funding from Agenda Transform, project no. C644865735-00000007, following the Mobilization Agendas for Business Innovation (Notice No. 02/C05-i01/2021), supported by the Recovery and Resilience Plan (PRR) and European Funds NextGeneration EU. The author AF is grateful for his grant (BI/UTAD/56/2023) within the scope of the project Agenda Transform. Part of the research was partially funded by the INTERREG-SUDOE Program through the European Regional Development Fund (ERDF) [project "ForManRisk-Forest Management and natural Risks"], operation number SOE3/P4/F0898, coordinated at UTAD by TF, and supported by National Funds by FCT—Portuguese Foundation for Science and Technology, under projects UIDB/04033/2020 (https://doi.org/10.54499/UIDB/04033/2020), and UID/04033:

Centro de Investigação e de Tecnologias Agro-Ambienteis e Biológicas and LA/P/0126/2020 (https://doi.org/10.54499/LA/P/0126/2020).

Data Availability Statement: The raw data utilized in this study are not publicly accessible due to their protected status. Summarized data and analytical results can be obtained upon reasonable request by contacting the corresponding author.

Acknowledgments: The authors express their gratitude to Carlos Fernandes for his significant assistance in data collecting and acknowledge Stéphanie Ribeiro for her involvement during the initial phase of this process. Gratitude is also expressed to Maria Emília Silva, the researcher overseeing the Agenda Transform project at UTAD. Moreover, the authors wish to express their gratitude toward the anonymous reviewers for their precious advice.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
AICc	Corrected Akaike information criterion
AUC	Area under curve
BIC	Bayesian information criterion
BT	Bootstrap forest
C0	Classes with no acorns or seedlings
C1	Classes with quantity of acorns and seedlings between one to four
C2	Classes with a quantity of acorns and seedlings above four
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
DOY	Day of year
DT	Decision tree
EFA	Exploratory Factor Analysis
ER^2	Entropy R-square
GAUS	Gaussian activation function
GR^2	Generalized R-square
GRB	Generalized regressions with Bridge
GRE	Generalized regressions with Elastic Net
GRL	Generalized regressions with Lasso
H	Height
H1	Hidden layer link with output variable
H2	Hidden layer ling with input variable
IRI	Independent resampled input
kNN	k-nearest neighbor
LIN	Linear activation function
MAD	Mean absolute deviation
MLP	Multilayer perceptron
MR	Misclassification rate
NB	Naive Bayes
NFI	Normed Fit Index
NL	Nominal logistic
NLL	Negative log-likelihood
NN	Neural network
NNB	Neural network boosted
PDS	Presence of dead seedling
PR	Precision-recall
PR_{AUC}	Area under curve of precision-recall

Q1	First quadrant
Q2	Second quadrant
Q3	Third quadrant
Q4	Fourth quadrant
RASE	Root average squared error
RMSEA	Root Mean Square Error of Approximation
RND	Random normal distribution
ROC	Receiver operating characteristic
ROC_{AUC}	Area Under Curve of Receiver Operating Characteristic
SDI	Stand density index
SRMR	Standardized Root Mean Square Residual
SVM	Support vector machine
TanH	Hyperbolic tangent activation function
TLI	Tucker-Lewis Index
TPH	Tree per hectare
TS1	Total seedling (Living) with height $\leq$ 10 cm
TS2	Total seedling (Living) with height > 10 cm

Appendix A

Table A1. Detailed information and percentages (total and by plots) of the categorical variables in the input and output, which were used in the multilayer perceptrons (MLPs).

Layer	Variables	Classes	Details	Percentages (%)		
				Total	Plot A1	Plot A2
Input	*Acorn*	C0	Absence of acorns	35.64	41.67	29.69
		C1	Presence between 1 and 4 acorns per m^2	30.94	26.04	35.42
		C2	Presence of more than 4 acorns per m^2	33.43	32.29	34.90
	PDS	No	Absence of dead seedling	90.88	89.58	89.58
		Yes	Presence of dead seedling (max 4)	9.12	10.42	10.42
	Slope	No	Flat terrain (<$\pm$5° or ~8.75%)	37.29	37.50	37.50
		Yes	Sloping terrain measured between 5° and 20° (36.4%)	62.71	62.50	62.50
Output	*TS1*	C0	Absence of seedling H $\leq$ 10 per m^2	39.78	37.50	43.23
		C1	Presence between 1 and 4 seedling H $\leq$ 10 per m^2	39.78	42.19	36.46
		C2	Presence of more than 4 seedling H $\leq$ 10 per m^2	20.44	20.31	20.31
	TS2	C0	Absence of seedling with H > 10 per m^2	37.85	20.31	56.77
		C1	Presence between 1 and 4 seedlings with H > 10 per m^2	48.07	56.25	38.54
		C2	Presence of more than 4 seedlings with H > 10 per m^2	14.09	23.44	4.69

Table A2. Distribution of each parameter used to determine the random normal distribution (*RND*) obtained by the CFA. *Acorns*, *TS1*, *TS2*, and dead seedlings in this case were normalized continuous numeric variables. *SD* is the standard deviation, *SE* is the standard error, *U 95%* is the upper 95% mean and *L 95%* is the lower 95% mean. The values of the variables were normal in the case of *RND* and normalized in the case of components of *RND*.

Dataset	Variables		Statistics				
			Mean	*SD*	*SE*	*U 95%*	*L 95%*
Total		*RND*	0.00	0.99	0.11	0.24	−0.24
	Components of *RND*	*Acorns*	8.00	4.90	0.61	9.22	6.78
		TS1	20.47	20.17	2.52	25.51	15.43
		TS2	20.47	20.17	2.52	25.51	15.43
		Dead seedling	1.75	2.48	0.31	2.37	1.13

Table A2. *Cont.*

Dataset	Variables		Statistics				
			Mean	***SD***	***SE***	***U 95%***	***L 95%***
Plot A1		*RND*	−0.01	0.93	0.16	0.32	−0.35
	Components of *RND*	*Acorns*	7.25	4.17	0.74	8.75	5.75
		TS1	20.61	19.5	3.45	27.64	13.58
		TS2	26.46	18.02	3.19	32.95	19.96
		Dead seedling	1.27	2.72	0.48	2.25	0.29
Plot A2		*RND*	0.01	1.04	0.18	0.39	−0.36
	Components of *RND*	*Acorns*	8.75	5.51	0.97	10.74	6.77
		TS1	20.33	21.12	3.73	27.95	12.72
		TS2	18.54	20.4	3.61	25.90	11.19
		Dead seedling	2.23	2.14	0.38	3.00	1.46

Table A3. Performance of MLPs in predicting the seedling height (*H*) categorized as $H \leq 10$ (*TS1*) and $H > 10$ (*TS2*). Metric parameters GR^2 and ER^2 were generalized R-squared and the entropy and *NLL* was the log-likelihood. *RASE* represents the root mean square error, while *MAD* and *MR* denote the mean absolute deviation and the misclassification rate, respectively. (*) represents the least effective MLP model, while (**) denotes the most accurate MLP model. 'NA' indicates that the result is not available.

Response	Model	Training					Validation				
		GR^2	***ER^2***	***RASE***	***MAD***	***MR***	***GR^2***	***ER^2***	***RASE***	***MAD***	***MR***
TS1	MLP-40 **	0.67	0.42	0.45	0.36	0.27	0.74	0.51	0.42	0.34	0.21
	MLP-5 *	0.45	0.24	0.54	0.49	0.37	0.4	0.18	0.56	0.51	0.41
	BF	0.63	0.38	0.48	0.45	0.26	0.06	0.03	0.68	0.56	0.45
	SVM	0.42	0.22	0.54	0.49	0.37	0.24	0.12	0.56	0.52	0.38
	NNB	0.36	0.18	0.56	0.52	0.4	0.27	0.13	0.56	0.52	0.4
	DT	0.35	0.17	0.56	0.53	0.4	0.25	0.12	0.56	0.52	0.36
	NB	0.32	0.16	0.57	0.53	0.4	0.16	0.07	0.58	0.54	0.38
	kNN	NA	0	NA	NA	0.51	NA	−0.21	NA	NA	0.48
	NL	0.34	0.17	0.57	0.53	0.38	0.21	0.1	0.57	0.53	0.42
	GRL&GRN	0.3	0.15	0.58	0.55	0.43	0.21	0.1	0.58	0.58	0.36
	GRR	0.34	0.17	0.57	0.53	0.39	0.21	0.1	0.57	0.53	0.41
TS2	MLP-40 **	0.58	0.35	0.47	0.4	0.28	0.63	0.43	0.47	0.4	0.28
	MLP-5 *	0.35	0.35	0.59	0.53	0.41	0.28	0.13	0.58	0.55	0.45
	BF	0.57	0.33	0.5	0.48	0.28	0.12	0.05	0.6	0.57	0.48
	SVM	0.34	0.17	0.56	0.52	0.39	0.17	0.07	0.6	0.56	0.47
	NNB	0.28	0.13	0.58	0.56	0.44	0.21	0.1	0.59	0.57	0.43
	DT	0.16	0.07	0.6	0.59	0.49	0.17	0.07	0.6	0.58	0.45
	NB	0.07	0.03	0.58	0.55	0.52	−0.02	−0.01	0.58	0.55	0.5
	kNN	NA	−0.05	NA	NA	0.54	NA	0.05	NA	NA	0.45
	NL	0.28	0.13	0.58	0.55	0.44	0.18	0.08	0.59	0.57	0.43
	GRL&GRN	0.27	0.13	0.58	0.56	0.45	0.2	0.09	0.59	0.57	0.44
	GRR	0.26	0.12	0.59	0.57	0.44	0.19	0.08	0.6	0.57	0.42

Table A4. Presentation of the nominal logistic (NL) "whole model test" and "lack of fit" results. The χ^2 is the chi-square of the log-likelihood with its *p*-value (*p*).

Response	Metric Indicators	Whole Model Test	Lack of Fit
TS1	χ^2	121.85	768.43
	p	<0.001	0.91
TS2	χ^2	110.39	769.32
	p	<0.001	0.91

References

1. Lindner, A.; Berges, M. *Can You Explain AI to Me? Teachers' Pre-Concepts about Artificial Intelligence*; IEEE Xplore: Piscataway, NJ, USA, 2020; pp. 1–9.
2. Ngie, H.M.; Nderu, L.; Mutanu, L.; Gicuku, D.M. *Mitigating Preconception in Machine Learning Classifiers*; IEEE Xplore: Piscataway, NJ, USA, 2021; pp. 1–6.
3. Bewersdorff, A.; Zhai, X.; Roberts, J.; Nerdel, C. Myths, Mis- and Preconceptions of Artificial Intelligence: A Review of the Literature. *Comput. Educ. Artif. Intell.* **2023**, *4*, 100143. [CrossRef]
4. Haumann, J.R.; Daly, R.T.; Worlton, T.G.; Crawford, R.K. *IPNS Distributed Processing Data Acquisition System*; IEEE Xplore: Piscataway, NJ, USA, 1982; Volume 29, pp. 62–66.
5. Toong, H.D.; Gupta, A. Personal Computers. *Sci. Am.* **1982**, *247*, 86–107.
6. Liu, Y.-C.; Gibson, G.A. *Microcomputer Systems: The 8086/8088 Family: Architecture, Programming, and Design*; Prentice-Hall, Inc.: Saddle River, NJ, USA, 2000; ISBN 978-0-13-580944-0.
7. Gibney, E.; Castelvecchi, D. Physics Nobel Scooped by Machine-Learning Pioneers. *Nature* **2024**, *634*, 523–524. [CrossRef]
8. Pellequer, J.L.; Westhof, E. PREDITOP: A Program for Antigenicity Prediction. *J. Mol. Graph.* **1993**, *11*, 204–210. [CrossRef]
9. Tian, M.; Xing, Q.; Wang, X.; Yuan, X.; Cheng, X.; Ming, Y.; Yin, K.; Li, Z.; Wang, P. Prediction of Junior High School Students' Problematic Internet Use: The Comparison of Neural Network Models and Linear Mixed Models in Longitudinal Study. *Psychol. Res. Behav. Manag.* **2024**, *17*, 1191–1203. [CrossRef] [PubMed]
10. Scopus.Com. Available online: https://www.scopus.com/results/results.uri?sort=plf-f&src=s&sid=d69c7e537319a89bae27b729bcdd6da2&sot=a&sdt=cl&sl=23&s=Artificial+intelligence&origin=resultslist&editSaveSearch=&txGid=1532825316 9b12993ac2e9b2ba37dff3&sessionSearchId=d69c7e537319a89bae27b729bcdd6da2&limit=10&yearFrom=2024&yearTo=2025 &cluster=scoexactkeywords,%22Artificial+Intelligence%22,t+scolang,%22English%22,t+scosubtype,%22ar%22,t (accessed on 21 March 2025).
11. Lang, N.; Jetz, W.; Schindler, K.; Wegner, J.D. A High-Resolution Canopy Height Model of the Earth. *Nat. Ecol. Evol.* **2023**, *7*, 1778–1789. [CrossRef]
12. Kiani Shahvandi, M.; Adhikari, S.; Dumberry, M.; Modiri, S.; Heinkelmann, R.; Schuh, H.; Mishra, S.; Soja, B. Contributions of Core, Mantle and Climatological Processes to Earth's Polar Motion. *Nat. Geosci.* **2024**, *17*, 705–710. [CrossRef]
13. Jevšenak, J.; Levanič, T. Should Artificial Neural Networks Replace Linear Models in Tree Ring Based Climate Reconstructions? *Dendrochronologia* **2016**, *40*, 102–109. [CrossRef]
14. Kumaraswamy, B. Neural Networks for Data Classification. In *Artificial Intelligence in Data Mining*; Binu, D., Rajakumar, B.R., Eds.; Academic Press: Cambridge, MA, USA, 2021; pp. 109–131, ISBN 978-0-12-820601-0.
15. Arosa, M.L.; Ceia, R.S.; Costa, S.R.; Freitas, H. Factors Affecting Cork Oak (*Quercus suber*) Regeneration: Acorn Sowing Success and Seedling Survival under Field Conditions. *Plant Ecol. Divers.* **2015**, *8*, 519–528. [CrossRef]
16. Caldeira, M.C.; Ibáñez, I.; Nogueira, C.; Bugalho, M.N.; Lecomte, X.; Moreira, A.; Pereira, J.S. Direct and Indirect Effects of Tree Canopy Facilitation in the Recruitment of Mediterranean Oaks. *J. Appl. Ecol.* **2014**, *51*, 349–358. [CrossRef]
17. Ribeiro, S.; Cerveira, A.; Soares, P.; Ribeiro, N.A.; Camilo-Alves, C.; Fonseca, T.F. Natural Regeneration of Cork Oak Forests under Climate Change: A Case Study in Portugal. *Front. For. Glob. Change* **2024**, *7*, 1332708. [CrossRef]
18. Mechergui, T.; Pardos, M.; Jacobs, D.F. Effect of Acorn Size on Survival and Growth of Quercus Suber L. Seedlings under Water Stress. *Eur. J. Forest Res.* **2021**, *140*, 175–186. [CrossRef]
19. Peng, C.; Zuezhi, W. Recent Applications of Artificial Neural Networks in Forest Resource Management: An Overview. *AAAI* **1999**, *1*, W1.
20. Hilbert, D.W.; Ostendorf, B. The Utility of Artificial Neural Networks for Modelling the Distribution of Vegetation in Past, Present and Future Climates. *Ecol. Model.* **2001**, *146*, 311–327. [CrossRef]

21. Özbay, B.; Keskin, G.A.; Doğruparmak, Ş.Ç.; Ayberk, S. Predicting Tropospheric Ozone Concentrations in Different Temporal Scales by Using Multilayer Perceptron Models. *Ecol. Inform.* **2011**, *6*, 242–247. [CrossRef]
22. Gürsoy, M.İ.; Orhan, O.; Tekin, S. Creation of Wildfire Susceptibility Maps in the Mediterranean Region (Turkey) Using Convolutional Neural Networks and Multilayer Perceptron Techniques. *For. Ecol. Manag.* **2023**, *538*, 121006. [CrossRef]
23. Muñoz-Mas, R.; Martínez-Capel, F.; Alcaraz-Hernández, J.D.; Mouton, A.M. Can Multilayer Perceptron Ensembles Model the Ecological Niche of Freshwater Fish Species? *Ecol. Model.* **2015**, *309–310*, 72–81. [CrossRef]
24. Wang, Y.; Fang, Z.; Hong, H.; Costache, R.; Tang, X. Flood Susceptibility Mapping by Integrating Frequency Ratio and Index of Entropy with Multilayer Perceptron and Classification and Regression Tree. *J. Environ. Manag.* **2021**, *289*, 112449. [CrossRef]
25. Silva, J.; Araújo, S.D.S.; Sales, H.; Pontes, R.; Nunes, J. *Quercus suber* L. Genetic Resources: Variability and Strategies for Its Conservation. *Forests* **2023**, *14*, 1925. [CrossRef]
26. Camarero, J.J.; Sánchez-Miranda, Á.; Colangelo, M.; Matías, L. Climatic Drivers of Cork Growth Depend on Site Aridity. *Sci. Total Environ.* **2024**, *912*, 169574. [CrossRef]
27. Ramos, A.M.; Usié, A.; Barbosa, P.; Barros, P.M.; Capote, T.; Chaves, I.; Simões, F.; Abreu, I.; Carrasquinho, I.; Faro, C.; et al. The Draft Genome Sequence of Cork Oak. *Sci. Data* **2018**, *5*, 180069. [CrossRef] [PubMed]
28. Uva, J.S.; Faias, S.P. *6º Inventário Florestal Nacional—IFN*; Instituto da Conservação da Natureza e das Florestas: Lisboa, Portugal, 2019.
29. Reis, F.; Pereira, A.J.; Tavares, R.M.; Baptista, P.; Lino-Neto, T. Cork Oak Forests Soil Bacteria: Potential for Sustainable Agroforest Production. *Microorganisms* **2021**, *9*, 1973. [CrossRef] [PubMed]
30. Aronson, J.; Pereira, J.S.; Pausas, J.G. *Cork Oak Woodlands on the Edge: Ecology, Adaptive Management, and Restoration*; Island Press: Washington, DC, USA, 2012; ISBN 978-1-61091-130-6.
31. Holmgren, M.; Scheffer, M.; Huston, M.A. The Interplay of Facilitation and Competition in Plant Communities. *Ecology* **1997**, *78*, 1966–1975.
32. Pulido, F.J.; Díaz, M. Regeneration of a Mediterranean Oak: A Whole-Cycle Approach. *Écoscience* **2005**, *12*, 92–102. [CrossRef]
33. Sobral, R.; Costa, M.M.R. Role of Floral Organ Identity Genes in the Development of Unisexual Flowers of *Quercus suber* L. *Sci. Rep.* **2017**, *7*, 10368. [CrossRef]
34. Sobral, R.; Silva, H.G.; Laranjeira, S.; Magalhães, J.; Andrade, L.; Alhinho, A.T.; Costa, M.M.R. Unisexual Flower Initiation in the Monoecious *Quercus suber* L.: A Molecular Approach. *Tree Physiol.* **2020**, *40*, 1260–1276. [CrossRef]
35. Pérez-Ramos, I.M.; Rodríguez Urbieta, I.; Zavala, M.A.; Marañón, T. *Regeneration Ecology of Quercus Suber (Cork Oak) in Southern Spain*; Universidad de Huelva: Huelva, Spain, 2008; ISBN 978-84-96826-47-2.
36. Arosa González, M.L. The Decline of Cork Oak Woodlands: Biotic and Abiotic Interactions in Portuguese Montados. Ph.D. Thesis, Universide Coimbra, Coimbra, Portugal, 2015; p. 137. Available online: https://hdl.handle.net/10316/29554 (accessed on 1 April 2025).
37. Díaz-Fernández, P.M.; Climent, J.; Gil, L. Biennial Acorn Maturation and Its Relationship with Flowering Phenology in Iberian Populations of *Quercus suber*. *Trees* **2004**, *18*, 615–621. [CrossRef]
38. Diffenbaugh, N.S.; Pal, J.S.; Giorgi, F.; Gao, X. Heat Stress Intensification in the Mediterranean Climate Change Hotspot. *Geophys. Res. Lett.* **2007**, *34*. [CrossRef]
39. Azul, A.M.; Mendes, S.M.; Sousa, J.P.; Freitas, H. Fungal Fruitbodies and Soil Macrofauna as Indicators of Land Use Practices on Soil Biodiversity in Montado. *Agrofor. Syst.* **2011**, *82*, 121–138. [CrossRef]
40. Pino-Mejías, R.; Cubiles-de-la-Vega, M.D.; Anaya-Romero, M.; Pascual-Acosta, A.; Jordán-López, A.; Bellinfante-Crocci, N. Predicting the Potential Habitat of Oaks with Data Mining Models and the R System. *Environ. Model. Softw.* **2010**, *25*, 826–836. [CrossRef]
41. Morais, T.G.; Domingos, T.; Falcão, J.; Camacho, M.; Marques, A.; Neves, I.; Lopes, H.; Teixeira, R.F.M. Permanent Pastures Identification in Portugal Using Remote Sensing and Multi-Level Machine Learning. *Front. Remote Sens.* **2024**, *5*, 1459000. [CrossRef]
42. Ribero, S.; Fonseca, T.F. Bonnes pratiques pour favoriser la régénération du chêne-liège (*Quercus suber* L.) dans le nord-est de Trás-os-montes (Por-tugal)—Fiche 4. In *Garantir a Regeneração e Reduzir o Risco de Incêndio: Um Desafio para o Futuro das Florestas no Sudoeste da Europa*; Lehoucq, A., Araújom, D., Beltrán, M., Dalgé, B., Destribat, B., Ferreira, P., Fonseca, T.F., A López, A., Magalhães, M., Maugard, F., Eds.; Interreg Sudoe; ForManRisk, UTAD: Vila Real, Portugal, 2023; p. 136, ISBN 978-989-704-538-7.
43. Weatherspark.com. Climate and Average Weather Year Round in Mogadouro, Portugal, 25-01-2025. Available online: https://weatherspark.com/y/33558/Average-Weather-in-Mogadouro-Portugal-Year-Round#Figures-Rainfall (accessed on 21 March 2025).
44. Fierravanti, A. Teresa, Fidalgo Fonseca Regeneration Patterns in Cork Oak Stands: Insights from Transect and Cluster Sampling Inventory Designs. *Forests* 2025, *Submitted*.

45. Fonseca, T.F.; Monteiro, L.; Enes, T.D.; Cerveira, A. Self-Thinning Dynamics in Cork Oak Woodlands: Providing a Baseline for Managing Density. *For. Syst.* **2017**, *26*, e006. [CrossRef]
46. Liu, P.; Nie, X.; Liang, J.; Cao, J. Multiple Mittag-Leffler Stability of Fractional-Order Competitive Neural Networks with Gaussian Activation Functions. *Neural Netw.* **2018**, *108*, 452–465. [CrossRef] [PubMed]
47. Lederer, J. Activation Functions in Artificial Neural Networks: A Systematic Overview. *arXiv* **2021**, arXiv:2101.09957. [CrossRef]
48. Aran, O.; Yildiz, O.T.; Alpaydin, E. An Incremental Framework Based on Cross-Validation for Estimating the Architecture of a Multilayer Perceptron. *Int. J. Pattern Recognit. Artif. Intell.* **2009**, *23*, 159–190. [CrossRef]
49. Jung, K.; Bae, D.-H.; Um, M.-J.; Kim, S.; Jeon, S.; Park, D. Evaluation of Nitrate Load Estimations Using Neural Networks and Canonical Correlation Analysis with K-Fold Cross-Validation. *Sustainability* **2020**, *12*, 400. [CrossRef]
50. Vu, H.L.; Ng, K.T.W.; Richter, A.; An, C. Analysis of Input Set Characteristics and Variances on K-Fold Cross Validation for a Recurrent Neural Network Model on Waste Disposal Rate Estimation. *J. Environ. Manag.* **2022**, *311*, 114869. [CrossRef]
51. Lever, J.; Krzywinski, M.; Altman, N. Classification Evaluation. *Nat. Methods* **2016**, *13*, 603–604. [CrossRef]
52. Altman, N.; Krzywinski, M. Graphical Assessment of Tests and Classifiers. *Nature Met.* **2021**, *18*, 840–842. [CrossRef]
53. Lukočienė, O.; Varriale, R.; Vermunt, J.K. The Simultaneous Decision(s) about the Number of Lower- and Higher-Level Classes in Multilevel Latent Class Analysis. *Sociol. Methodol.* **2010**, *40*, 247–283. [CrossRef]
54. Hyndman, R.J.; Koehler, A.B. Another Look at Measures of Forecast Accuracy. *Int. J. Forecast.* **2006**, *22*, 679–688. [CrossRef]
55. Favre, J.; Hayoz, M.; Erhart-Hledik, J.C.; Andriacchi, T.P. A Neural Network Model to Predict Knee Adduction Moment during Walking Based on Ground Reaction Force and Anthropometric Measurements. *J. Biomech.* **2012**, *45*, 692–698. [CrossRef] [PubMed]
56. Rossi, L.; Bagheri, M.; Zhang, W.; Chen, Z.; Burken, J.G.; Ma, X. Using Artificial Neural Network to Investigate Physiological Changes and Cerium Oxide Nanoparticles and Cadmium Uptake by *Brassica Napus* Plants. *Environ. Pollut.* **2019**, *246*, 381–389. [CrossRef]
57. Gonçalves, L.; Subtil, A.; Oliveira, M.R.; Bermudez, P. de Z. ROC Curve Estimation: An Overview. *REVSTAT-Stat. J.* **2014**, *12*, 1–20. [CrossRef]
58. Coppock, H.; Nicholson, G.; Kiskin, I.; Koutra, V.; Baker, K.; Budd, J.; Payne, R.; Karoune, E.; Hurley, D.; Titcomb, A.; et al. Audio-Based AI Classifiers Show No Evidence of Improved COVID-19 Screening over Simple Symptoms Checkers. *Nat. Mach. Intell.* **2024**, *6*, 229–242. [CrossRef]
59. Saltelli, A. Making Best Use of Model Evaluations to Compute Sensitivity Indices. *Comput. Phys. Commun.* **2002**, *145*, 280–297. [CrossRef]
60. Iacona, G.D.; Kirkman, L.K.; Bruna, E.M. Effects of Resource Availability on Seedling Recruitment in a Fire-Maintained Savanna. *Oecologia* **2010**, *163*, 171–180.
61. Fonseca, T.F.; Duarte, J.C. A Silvicultural Stand Density Model to Control Understory in Maritime Pine Stands. *iForest* **2017**, *10*, 829. [CrossRef]
62. Pu, X.; Weemstra, M.; Jin, G.; Umaña, M.N. Tree Mycorrhizal Type Mediates Conspecific Negative Density Dependence Effects on Seedling Herbivory, Growth, and Survival. *Oecologia* **2022**, *199*, 907–918. [CrossRef]
63. Honda, E.A.; Pilon, N.A.L.; Durigan, G. The Relationship between Plant Density and Survival to Water Stress in Seedlings of a Legume Tree. *Acta Bot. Bras.* **2019**, *33*, 602–606. [CrossRef]
64. Camarasa-Belmonte, A.M.; Rubio, M.; Salas, J. Rainfall Events and Climate Change in Mediterranean Environments: An Alarming Shift from Resource to Risk in Eastern Spain. *Nat. Hazards* **2020**, *103*, 423–445. [CrossRef]
65. Faias, S.P.; Paulo, J.A.; Tomé, M. Inter-Tree Competition Analysis in Undebarked Cork Oak Plantations as a Support Tool for Management in Portugal. *New For.* **2020**, *51*, 489–505. [CrossRef]
66. Catry, F.X.; Moreira, F.; Duarte, I.; Acácio, V. Factors Affecting Post-Fire Crown Regeneration in Cork Oak (*Quercus suber* L.) Trees. *Eur. J. For. Res.* **2009**, *128*, 231–240. [CrossRef]
67. Dias, A.C.; Boschmonart-Rives, J.; González-García, S.; Demertzi, M.; Gabarrell, X.; Arroja, L. Analysis of Raw Cork Production in Portugal and Catalonia Using Life Cycle Assessment. *Int. J. Life Cycle Assess.* **2014**, *19*, 1985–2000. [CrossRef]
68. Yu, F.; Wang, D.; Shi, X.; Yi, X.; Huang, Q.; Hu, Y. Effects of Environmental Factors on Tree Seedling Regeneration in a Pine-Oak Mixed Forest in the Qinling Mountains, China. *J. Mt. Sci.* **2013**, *10*, 845–853. [CrossRef]
69. Dinis, C.; Surový, P.; Ribeiro, N.; Oliveira, M.R.G. The Effect of Soil Compaction at Different Depths on Cork Oak Seedling Growth. *New Forests* **2015**, *46*, 235–246. [CrossRef]
70. Costa, D.; Tavares, R.M.; Baptista, P.; Lino-Neto, T. The Influence of Bioclimate on Soil Microbial Communities of Cork Oak. *BMC Microbiol.* **2022**, *22*, 163. [CrossRef]
71. Bréda, N.; Granier, A.; Aussenac, G. Effects of Thinning on Soil and Tree Water Relations, Transpiration and Growth in an Oak Forest (*Quercus petraea* (Matt.) Liebl.). *Tree Physiol.* **1995**, *15*, 295–306. [CrossRef]

72. Besson, C.K.; Lobo-do-Vale, R.; Rodrigues, M.L.; Almeida, P.; Herd, A.; Grant, O.M.; David, T.S.; Schmidt, M.; Otieno, D.; Keenan, T.F.; et al. Cork Oak Physiological Responses to Manipulated Water Availability in a Mediterranean Woodland. *Agric. For. Meteorol.* **2014**, *184*, 230–242. [CrossRef]
73. Robert, B.; Caritat, A.; Bertoni, G.; Vilar, L.; Molinas, M. Nutrient Content and Seasonal Fluctuations in the Leaf Component of Coark-Oak (*Quercus suber* L.) Litterfall. *Vegetatio* **1996**, *122*, 29–35. [CrossRef]
74. Paulo, J.A.; Tomé, M. An Individual Tree Growth Model for Juvenile Cork Oak Stands in Southern Portugal. *Silva Lusitana* **2009**, *17*, 27–38.
75. Pasalodos-Tato, M.; Pukkala, T.; Cañellas, I.; Sánchez-González, M. Optimizing the Debarking and Cutting Schedule of Cork Oak Stands. *Ann. For. Sci.* **2018**, *75*, 1–11. [CrossRef]
76. Acácio, V.; Holmgren, M. Pathways for Resilience in Mediterranean Cork Oak Land Use Systems. *Ann. For. Sci.* **2014**, *71*, 5–13. [CrossRef]
77. Dey, D.C. Oak Regeneration Ecology and Dynamics. In *Proceedings, Wildland Fire in the Appalachians: Discussions Among Fire Managers and Scientists*; General Technical Report SRS-199; Waldrop, T.A., Ed.; USDA Forest Service, Southern Research Station: Asheville, NC, USA, 2014; Volume 199, pp. 3–11, 208p.
78. Costa, A.; Madeira, M.; Oliveira, Â.C. The Relationship between Cork Oak Growth Patterns and Soil, Slope and Drainage in a Cork Oak Woodland in Southern Portugal. *For. Ecol. Manag.* **2008**, *255*, 1525–1535. [CrossRef]
79. Boehm, A.R.; Hardegree, S.P.; Glenn, N.F.; Reeves, P.A.; Moffet, C.A.; Flerchinger, G.N. Slope and Aspect Effects on Seedbed Microclimate and Germination Timing of Fall-Planted Seeds. *Rangel. Ecol. Manag.* **2021**, *75*, 58–67. [CrossRef]
80. Boussaidi, N.; Ncibi, R.; Hasnaoui, I.; Ghrabi Gammar, Z. Impacts Des Facteurs Orographiques et Anthropiques Sur La Régénération Naturelle Du Chêne-Liège (*Quercus suber*) Dans La Région de Kroumirie, Tunisie. *Rev. D'écologie* **2010**, *65*, 235–242. [CrossRef]
81. Jasińska, J.; Sewerniak, P.; Markiewicz, M. Links between Slope Aspect and Rate of Litter Decomposition on Inland Dunes. *CATENA* **2019**, *172*, 501–508. [CrossRef]
82. Mechergui, T.; Pardos, M.; Boussaidi, N.; Jacobs, D.F.; Catry, F.X. Problems and Solutions to Cork Oak (*Quercus suber* L.) Regeneration: A Review. *iForest* **2023**, *16*, 10–22. [CrossRef]
83. Christin, S.; Hervet, É.; Lecomte, N. Applications for Deep Learning in Ecology. *Methods Ecol. Evol.* **2019**, *10*, 1632–1644. [CrossRef]
84. Rammer, W.; Seidl, R. Harnessing Deep Learning in Ecology: An Example Predicting Bark Beetle Outbreaks. *Front. Plant Sci.* **2019**, *10*, 451705. [CrossRef]
85. César de Lima Araújo, H.; Silva Martins, F.; Tucunduva Philippi Cortese, T.; Locosselli, G.M. Artificial Intelligence in Urban Forestry—A Systematic Review. *Urban For. Urban Green.* **2021**, *66*, 127410. [CrossRef]
86. Tarek, Z.; Elhoseny, M.; Alghamdi, M.I.; EL-Hasnony, I.M. Leveraging Three-Tier Deep Learning Model for Environmental Cleaner Plants Production. *Sci. Rep.* **2023**, *13*, 19499. [CrossRef]
87. Mfetoum, I.M.; Ngoh, S.K.; Molu, R.J.J.; Nde Kenfack, B.F.; Onguene, R.; Naoussi, S.R.D.; Tamba, J.G.; Bajaj, M.; Berhanu, M. A Multilayer Perceptron Neural Network Approach for Optimizing Solar Irradiance Forecasting in Central Africa with Meteorological Insights. *Sci. Rep.* **2024**, *14*, 3572. [CrossRef]
88. Wang, W.; Zhang, J.; Su, Q.; Chai, X.; Lu, J.; Ni, W.; Duan, B.; Ren, K. Accurate Initial Field Estimation for Weather Forecasting with a Variational Constrained Neural Network. *npj Clim. Atmos. Sci.* **2024**, *7*, 223. [CrossRef]
89. Sebastian, A.; Pannone, A.; Subbulakshmi Radhakrishnan, S.; Das, S. Gaussian Synapses for Probabilistic Neural Networks. *Nat. Commun.* **2019**, *10*, 4199. [CrossRef]
90. Nguyen, A.; Pham, K.; Ngo, D.; Ngo, T.; Pham, L. *An Analysis of State-of-the-Art Activation Functions for Supervised Deep Neural Network*; IEEE: Piscataway, NJ, USA, 2021; pp. 215–220.
91. Gundogdu, O.; Egrioglu, E.; Aladag, C.H.; Yolcu, U. Multiplicative Neuron Model Artificial Neural Network Based on Gaussian Activation Function. *Neural Comput. Appl.* **2016**, *27*, 927–935. [CrossRef]
92. Lau, M.M.; Hann Lim, K. Review of Adaptive Activation Function in Deep Neural Network. In Proceedings of the 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Sarawak, Malaysia, 3–6 December 2018; pp. 686–690.
93. Dubey, S.R.; Singh, S.K.; Chaudhuri, B.B. Activation Functions in Deep Learning: A Comprehensive Survey and Benchmark. *Neurocomputing* **2022**, *503*, 92–108. [CrossRef]
94. Gu, H.; Qiao, Y.; Xi, Z.; Rossi, S.; Smith, N.G.; Liu, J.; Chen, L. Warming-Induced Increase in Carbon Uptake Is Linked to Earlier Spring Phenology in Temperate and Boreal Forests. *Nat. Commun.* **2022**, *13*, 3698. [CrossRef]
95. Simon, S.M.; Glaum, P.; Valdovinos, F.S. Interpreting Random Forest Analysis of Ecological Models to Move from Prediction to Explanation. *Sci. Rep.* **2023**, *13*, 3881. [CrossRef]
96. Bisgin, H.; Bera, T.; Ding, H.; Semey, H.G.; Wu, L.; Liu, Z.; Barnes, A.E.; Langley, D.A.; Pava-Ripoll, M.; Vyas, H.J.; et al. Comparing SVM and ANN Based Machine Learning Methods for Species Identification of Food Contaminating Beetles. *Sci. Rep.* **2018**, *8*, 6532. [CrossRef]

97. Siemers, F.M.; Bajorath, J. Differences in Learning Characteristics between Support Vector Machine and Random Forest Models for Compound Classification Revealed by Shapley Value Analysis. *Sci. Rep.* **2023**, *13*, 5983. [CrossRef] [PubMed]
98. Mantovani, R.G.; Rossi, A.L.D.; Vanschoren, J.; Bischl, B.; de Carvalho, A.C.P.L.F. Effectiveness of Random Search in SVM Hyper-Parameter Tuning. In Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 12–17 July 2015; pp. 1–8.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.