



Article

Optimizing the In Vitro Propagation of Tea Plants: A Comparative Analysis of Machine Learning Models

Taner Bozkurt ^{1,*} , Sezen İnan ¹, İjlal Dündar ¹, Musab A. Isak ² and Özhan Şimşek ³

¹ Tekfen Agricultural Research Production and Marketing Inc., 01355 Adana, Türkiye; sezen.inan@tekfen.com.tr (S.İ.); ijilal.dundar@tekfen.com.tr (İ.D.)

² Agricultural Sciences and Technology Department, Graduate School of Natural and Applied Sciences, Erciyes University, 38030 Kayseri, Türkiye; musabisak11@gmail.com

³ Horticulture Department, Agriculture Faculty, Erciyes University, 38030 Kayseri, Türkiye; ozhansimsek@erciyes.edu.tr

* Correspondence: taner.bozkurt@tekfen.com.tr or 01tanerbozkurt@gmail.com

Abstract: In this study, we refine in vitro propagation techniques for *Camellia sinensis* using a machine learning approach to ascertain the influence of different shooting and rooting conditions on key growth metrics. This was achieved by applying random forest (RF), XGBoost, and multilayer perceptron (MLP) models to dissect the complexities of micropropagation and rooting processes. The research unveiled significant disparities in growth metrics under varying media conditions, underscoring the profound impact of media composition on plant development. The meticulous statistical analysis, employing ANOVA, highlighted statistically significant differences in growth metrics, indicating the critical role of media composition in optimizing growth conditions. Methodologically, the study utilized explants from 2–3-year-old tea plants, which underwent sterilization before being introduced to two distinct culture media for their micropropagation and rooting phases. Statistical analyses were conducted to evaluate the differences in growth outcomes between media, while machine learning models were employed to predict the efficacy of micropropagation and rooting based on various growth regulators. This approach allowed for a comprehensive evaluation of the model's performance in simulating plant growth under different conditions, leveraging metrics like R^2 , RMSE, and MAE. The findings from this study significantly advance the understanding of tea plant micropropagation, highlighting the utility of machine learning models in agricultural optimization. This research contributes to enhancing micropropagation strategies for the tea plant and exemplifies the transformative potential of integrating machine learning into plant science, paving the way for improved agricultural and horticultural practices. This interdisciplinary approach offers a novel perspective on optimizing in vitro propagation processes, contributing substantially to plant tissue culture and biotechnology.

Keywords: tea plant micropropagation; machine learning in horticulture; in vitro propagation optimization; growth parameter prediction; random forest and MLP models



Citation: Bozkurt, T.; İnan, S.; Dündar, İ.; Isak, M.A.; Şimşek, Ö. Optimizing the In Vitro Propagation of Tea Plants: A Comparative Analysis of Machine Learning Models. *Horticulturae* **2024**, *10*, 721. <https://doi.org/10.3390/horticulturae10070721>

Academic Editor: Michailidis Michail

Received: 23 May 2024

Revised: 25 June 2024

Accepted: 26 June 2024

Published: 9 July 2024



Copyright: © 2024 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

The tea plant (*Camellia sinensis* (L.) O. Kuntze) is one of the world's most popular and well-known non-alcoholic beverages. This plant, grown in more than 30 countries, including China, India, Kenya, Sri Lanka, Vietnam, Japan, and Türkiye, is native to Southeast Asia and is a member of the Theaceae family of the order Ericales [1]. The natural height of *C. sinensis*, an evergreen perennial plant, can reach up to 15 m. It produces white flowers and is characterized by cross-pollination [2]. This plant's immature shoot tips, rich in chemical compounds, are used in tea processing. Tea has been documented and provenanced as early as the 6th and 7th centuries BCE. It was reportedly first used medicinally, especially to cure stomach problems. However, insufficient conclusive information exists in the literature about turning tea into a drink [3]. Based on the oxidation and fermentation processes,

consumers are offered various teas, including oolong, black, green, and white. Oolong is semi-fermented, black tea is fermented, and white and green teas are unfermented. While oolong and green tea are the most popular teas in Asia, black tea is more commonly drunk in India and other European nations [4,5].

Around 4000 bioactive compounds have been found in tea, with polyphenols accounting for one-third of the total. Aside from polyphenols, tea also has small quantities of minerals, amino acids, carbohydrates, and caffeine. Remarkably, polyphenols can differ according to the degree of fermentation used [6,7]. Furthermore, tea has been reported to include mg/g levels of the elements Na^+ , Mg^{2+} , Ca^{2+} , Mn^{2+} , and K^+ and $\mu\text{g/g}$ levels of Zn^{2+} , Ni^{2+} , Cr^{3+} , Cd^{2+} , Fe^{2+} , Cu^{2+} , and Co^{2+} [8]. In addition to its antioxidant and anti-cancer qualities, tea has been shown to have numerous other health advantages, including the ability to repair DNA damage. Tea can lower the incidence of the mouth, stomach, esophagus, bladder, breast, small intestine, pancreas, and lung cancers [9,10]. There is now more interest in this plant because of its many health benefits and its antifungal and antibacterial activity [4,11]. The production of tea plants has been rising annually in parallel with consumption, as they are both commercially valuable and health promoting. China leads the world in tea leaf production overall, accounting for 48.87% of global production in 2022 according to the FAO data for 2023. India (20%), Kenya (7.8%), Sri Lanka (4.7%), and Turkey (4.3%) produce the highest percentages after China [12].

Traditionally, cuttings and seeds are used to propagate the tea plant. Until the early 1800s, the only propagation mode was through seeds. On the other hand, genetic changes arising during seed propagation have led to notable plant variances. As a result, cuttings are employed in different ways. However, there have also been documented issues with cutting propagation. The most remarkable is how long a single plant takes (approximately 12 months) to grow from a cutting and be prepared for planting [13,14]. As a result, increasing the number of plants for the commercial market and propagating tea clonally have become complex tasks [15]. In vitro tissue culture technologies promise to address these issues and facilitate tea plants' rapid and high-quality commercial propagation.

Machine learning (ML) modeling has gained popularity in plant biotechnology. However, there is a paucity of supportive data, particularly on in vitro plant tissue culture. Various machine learning algorithms have been applied to optimize diverse aspects such as in vitro sterilization, in vitro germination, in vitro cell culture, in vitro micropropagation and rooting, in vitro drought stress, and in vitro somatic embryogenesis [16–22].

Artificial neural networks (ANNs), like the human brain, are constructed from interconnected processing neurons that can decipher information through weighted communication links [23]. These networks are utilized to statistically establish relationships between independent (input) and dependent (output) variables, facilitating the analysis of information exchange between them.

The application of conventional statistical approaches for decrypting the vast amount of encrypted data within extensive datasets of biological interactions, such as micropropagation, poses significant challenges attributed to intricate, noisy, and misleading datasets entailing multifactorial processes [24]. Nevertheless, recent advancements have achieved accuracy, prediction, and optimization in plant tissue culture procedures by utilizing various machine learning models. These approaches employ neural networks [25,26] and machine learning models based on decision trees [27].

ML has emerged as a pivotal asset in plant science research, notably within in vitro plant studies, enhancing the breadth and depth of investigations in plant breeding, genomics, and biotechnology [28]. These innovative techniques have facilitated the analysis of bioelectrical patterns in plants, revealing their acute sensitivity to environmental changes [29]. Furthermore, ML has significantly propelled plant molecular research by capitalizing on the vast sequencing data [30]. The application of ML algorithms in plant omics and the enhancement of agronomic traits has shown promise in amplifying research outcomes [31]. Additionally, modeling plant behavior through data from plant sensors has yielded a more nuanced comprehension of plant performance over time [32]. Deep learning,

a subset of machine learning, has been instrumental in plant disease detection, phenotyping, and image analysis, marking a significant advancement in the field [33]. A focused area of ML application includes plant disease detection, classification, and prediction, underscoring its transformative potential in plant health management strategies [34–37]. Employing ML algorithms to differentiate between various plant stress phenotypes and identifying disease indicators through hyperspectral imaging has showcased the technology’s nuanced capability [38,39].

Furthermore, ML’s role in predicting the effects of electromagnetic radiation on plants and developing models to assess plant sensitivity to radiofrequency electromagnetic fields has been pivotal [40]. The synergy of machine learning with cutting-edge technologies like hyperspectral imaging, LIDAR sensors, and uncrewed aerial vehicles (UAVs) has unveiled new avenues for plant species classification, canopy density mapping, and health assessment [41]. Moreover, ML’s application in automating weed identification and enhancing precision agriculture showcases its utility in generating novel insights [42,43]. In vitro plant research, particularly when integrated with machine learning techniques, enables the detailed analysis and modeling of plant responses under controlled conditions that simulate various stress factors, including pathogenic and environmental stresses. We can analyze large datasets from in vitro experiments using machine learning algorithms to identify patterns and biomarkers indicative of disease susceptibility or resilience to environmental stressors. This can be extrapolated to real-world applications by developing predictive models that aid in early disease detection and assessing the impact of environmental stresses on plant health. For instance, machine learning models can be trained using data from in vitro studies where plants are exposed to specific pathogens or stress conditions. These models can then predict disease progression or stress responses based on early indicators detected in plant tissues, which are often subtle and complex [44]. Researchers can unlock profound insights into plant physiology, behavior, and health by leveraging ML algorithms, significantly advancing agricultural and plant science endeavors.

The cornerstone of this investigation is to pioneer the micropropagation of the commercially paramount tea plant, *C. sinensis*, a crop witnessing a global surge in production. This research ventures into uncharted territories by integrating machine learning (ML) and artificial neural networks (ANNs) into plant tissue culture, marking a novel convergence of biotechnology and computational intelligence aimed at deciphering and enhancing the intricacies of plant growth under in vitro conditions. Unique in its approach, the study not only harnesses the predictive power of ML and ANNs to navigate the complex dynamics of micropropagation but also signifies the inaugural application of these advanced computational models within tea plant research. Through meticulous biotechnological scrutiny and the adept application of ML models, we aim to transcend traditional analytical methods, offering a pathway to demystify and foresee the intricate interactions that govern plant tissue culture. This innovative integration of biotechnology and machine learning enhances the precision and efficiency of tea plant micropropagation and significantly impacts commercial cultivation practices. By producing genetically uniform and vigorous plantlets faster, this approach promises to elevate commercial tea cultivation, ensuring high-quality crop production and meeting the increasing demands of global markets.

2. Materials and Methods

2.1. Plant Material

Fresh shoots from 2–3-year-old tea plants (*Camellia sinensis*) were sourced from a commercial tea producer in Ardeşen, Rize, Turkey. These shoots served as explants for subsequent in vitro culture.

2.2. Sterilization and Culture Conditions

Explants were rinsed with tap water for 3–5 min to remove surface debris. Subsequently, they were sterilized by immersion in a 70% ethanol solution for two minutes within a sterile laminar airflow cabinet. Three washes with sterile distilled water followed this.

Explants were further treated with a 3% sodium hypochlorite solution for three minutes and washed three times with sterile distilled water to ensure complete sterilization. Following this, explants were transferred to culture media.

2.3. *In Vitro* Micropropagation

The micropropagation of plantlets was performed using two different culture media with different PGRs and a PGR-free medium as the control. The first medium (SM1) was comprised of MS medium (Duchefa Biochemie, Haarlem, The Netherlands) supplemented with 30 g/L sucrose, 0.1 g/L myo-inositol, 1.0 mg/L BAP (6-Benzylaminopurine) (Duchefa Biochemie, Haarlem, The Netherlands), 0.5 mg/L GA₃ (Gibberellic Acid) (Duchefa Biochemie, Haarlem, The Netherlands), 0.1 mg/L Thiamine HCl, and 2 mg/L Glycine. The second medium (SM2) was similarly composed of MS medium supplemented with 30 g/L sucrose, 0.1 g/L myo-inositol, 2.0 mg/L BAP, 0.5 mg/L GA₃, 0.1 mg/L Thiamine HCl, and 2 mg/L Glycine. Agar (7.5 g/L) was added to solidify both media, and their pH was adjusted to 5.8 using 1 N HCl and 1 N NaOH. The media were then sterilized by autoclaving at 121 °C and 103.4 kPa for 20 min before being dispensed into sterilized disposable plastic containers and glass tubes. The sterilized plantlets were subsequently transferred to these media for micropropagation. Every 4 weeks, the plantlets were subcultured into fresh media three times. The growth chambers were used to carry out the incubation under a 16/8 h photoperiod at 25 ± 2 °C. The photosynthetically active radiation (PAR) was set to an optimal level of 100 µmol/m²/s to provide adequate light for photosynthesis while preventing photoinhibition. The relative humidity within the culture environment was maintained at an optimal 75% to facilitate efficient physiological processes without risking tissue desiccation. The multiplication coefficient, which measures the propagation efficiency over the subculture period, was calculated by subtracting the initial count of plantlets at the beginning of the subculture cycle from the total number of plantlets obtained at the end and then dividing the result by the initial count. This calculation provides a quantitative measure of the proliferative capacity of the plantlets under the given culture conditions.

2.4. *In Vitro* Rooting

Root induction was performed using modified MS media. Media 1 (RM1) contained 30 g/L sucrose, 0.1 g/L myo-inositol, 0.5 mg/L indole-3-butyric acid (IBA) (Duchefa Biochemie, Haarlem, The Netherlands), 0.5 mg/L GA₃, 0.1 mg/L thiamine HCl, and 2 mg/L glycine. Media 2 (RM2) had a higher concentration of IBA at 2 mg/L and media without PGRs was used as a control. Media were prepared and sterilized as described for micropropagation. Rooting was initiated in growth chambers under conditions identical to the micropropagation conditions.

2.5. Statistical Analysis

In this study, the analysis of variance (ANOVA) was employed using R-programming software (version 4.3.1) to conduct a comprehensive statistical evaluation, aiming to explore the differences across various shoot and root mediums. ANOVA, a robust statistical method, was utilized to analyze the mean values of specific traits in the shoot and root mediums. This approach allowed for an in-depth examination of the variations attributable to the experimental conditions. Exhibiting significant changes were subjected to the least significant difference (LSD) test. The analysis incorporated three replications, each containing 30 plantlets.

2.6. Modeling Procedure

This study used two decision-tree-based machine learning (ML) algorithms, random forest (RF) and extreme gradient boosting (XGBoost), along with an artificial neural network (ANN) based multilayer perceptron (MLP), to model and predict the efficiency of tea micropropagation and rooting through various plant growth regulators. Using a 10-fold cross-validation method, the dataset was split into training and testing subsets to

thoroughly evaluate the prediction capabilities of ANN and decision-tree-based machine learning algorithms.

The two distinct PGR mediums were included in the input variables. Conversely, the micropropagation rate, plant height (cm), leaf number, number of roots, and length of roots (cm) were the target (output) variables. Coding was implemented using R-programming (version 4.3.1), with the usage of the Caret and Kernlab packages.

These metrics thoroughly evaluate how well the ANN-based MLP and decision-tree-based RF and XGBoost performed in simulating and forecasting the micropropagation and rooting effectiveness of tea using plant growth regulators. Metrics like the coefficient of determination (R^2), which evaluates the correlation between the model and dependent variable, root mean square error (RMSE), which shows the alignment of the regression line with the observed data points, and mean absolute error (MAE), which determines the average error between observed and predicted values, were employed for the evaluation and comparison of the effectiveness and accuracy of the models. Equations (1)–(3).

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (3)$$

Y_i = actual value, $\hat{Y}_i$ = Predicted value, $\bar{Y}$ = mean of the actual values, and n = sample count.

MLP is organized into several levels, including an input layer, an output layer, and one or more hidden layers. Using both input and output variables as part of the training set, MLP uses a supervised training technique. The training procedure continues until the minimization of Equation (4). [20]

$$E = \frac{1}{n} \sum_{s=1}^n (y_s - \hat{y}_s) \quad (4)$$

As a gradient-boosting decision trees class member, XGBoost is well-known for its exceptional speed and performance. Within a gradient-boosting architecture, XGBoost excels at learning from errors and progressively reducing the error rate over multiple iterations.

$$y_i = F(x_i) = \sum_{(d=1)}^D f_d(x_i), f_d \in F, i = 1, \dots, n \quad (5)$$

$$L_j = \sum_{i=1}^n l(y_i, \hat{y}^{(j-1)}) + f_j(x_i) + \Omega(f_j) \quad (6)$$

Equation (6) shows the XGBoost iterative model, and Equation (5) shows the XGBoost objective function.

In Equation (5), F is the function space of the tree model and f_d represents an independent tree structure that classifies each individual i into one leaf. In Equation (6), l is the convex loss function; it measures the difference between the observed y and predicted $\hat{y}$.

Breiman [45] created the RF approach, effectively an ensemble of unpruned trees. The RF approach is well known for its effectiveness and simplicity of usage, and it has proven successful in both regression and classification problems. The trained tree makes the final

decision in the RF technique, which uses bagging, also referred to as bootstrap aggregation. The basic idea of the RF model is shown in Equation (7).

$$y = \sum_{i=1}^n \left(a_i - a_i^*\right)k(x, x_i) + b$$

(7)

3. Results

In this study, we investigate how different shooting and rooting conditions affect important growth parameters during the micropropagation of tea. The multiplication coefficient, number of leaves, plant height (cm), number of roots, and root length (cm) are important metrics that provide information on the effectiveness of various media compositions. Using a careful examination, we reveal the complex relationship between growth results and media compositions, giving insight into the elements that propel productive tea plant propagation.

Table 1 presents the findings from a study examining the impact of various plant growth regulators (PGRs) on micropropagation traits. The experiment was structured into two segments, shooting traits and rooting traits, assessing the efficacy of different concentrations and combinations of PGRs.

Table 1. The effect of plant growth regulators (PGRs) on micropropagation traits.

Micropropagation			
	PGR Free (Control)	1 mg/L BAP + 0.5 mg/L GA ₃	2 mg/LBAP + 0.5 mg/L GA ₃
Multiplication Coefficient	0.61 ± 0.21 c	1.15 ± 0.28 b	1.74 ± 0.36 a
Number of Leaves	3.52 ± 0.75 c	6.53 ± 1.17 b	12.68 ± 1.43 a
Plant Height	3.34 ± 0.74 c	4.23 ± 0.83 b	9.96 ± 1.80 a
Rooting			
	PGR free (Control)	1 mg/L IBA + 0.5 mg/L GA ₃	2 mg/L IBA + 0.5 mg/L GA ₃
Number of Roots	0.18 ± 0.39 c	0.68 ± 0.57 b	1.64 ± 0.48 a
Root Length	0.13 ± 0.31 c	0.69 ± 0.61 b	1.56 ± 0.30 a

LSD: 0.1, small letters show statistical differences between PGRs.

In the examination of shooting traits, the study evaluated three distinct treatments: a control group without PGRs, a treatment group receiving 1 mg/L BAP combined with 0.5 mg/L GA₃, and another group treated with a higher concentration of 2 mg/L BAP and 0.5 mg/L GA₃. The results demonstrated a clear dose–response relationship, with increasing concentrations of BAP significantly enhancing both the multiplication coefficient and the number of leaves. Specifically, the highest concentration of BAP (2 mg/L) combined with 0.5 mg/L GA₃ yielded the greatest multiplication coefficient of 1.74 and the highest number of leaves, recorded at 12.68. A similar trend was evident for plant height, with the tallest plants, measuring 9.96 in height, observed in the treatment with the highest concentration of BAP and GA₃. The statistical significance of these findings was robustly supported by an LSD value of 0.1, with a *p*-value less than 0.0001 confirming significant differences across the treatment groups. These differences were denoted by alphabetical notations, with ‘a’ representing the highest means and ‘c’ the lowest, illustrating the effectiveness of the PGR treatments in enhancing plant growth metrics.

In the rooting traits section, the focus shifted to the effects of IBA combined with GA₃. Like the shooting traits, increasing the concentration of IBA from 0 mg/L in the control up to 2 mg/L significantly improved the number and length of roots from averages of 0.18–1.64 and 0.13–1.56, respectively. This positive correlation suggests that IBA is particularly effective in promoting root development, a critical aspect of plant establishment in tissue culture.

Machine Learning Analysis

Table 2 thoroughly analyzes the predictive performance of three machine learning models, RF, XGBoost, and MLP, regarding micropropagation and rooting parameters. The utilized evaluation measures, specifically R^2 , RMSE, and MAE, are essential for determining how well each model captures the nuances of the examined plant growth processes.

Table 2. Assessment metrics for ANN-based MLP and decision-tree-based RF and XGBoost.

ML Models	Assessment Matrices	Micropropagation Parameters			Rooting Parameters	
		Multiplication Rate	Number of Leaves	Plant Height (cm)	Number of Roots	Root Length (cm)
MLP	R^2	0.74	0.93	0.86	0.93	0.91
	RMSE	0.16	0.08	0.12	0.07	0.12
	MAE	0.13	0.10	0.10	0.12	0.08
RF	R^2	0.73	0.91	0.85	0.92	0.91
	RMSE	0.17	0.09	0.11	0.10	0.11
	MAE	0.14	0.10	0.09	0.05	0.06
XGboost	R^2	0.72	0.90	0.87	0.93	0.92
	RMSE	0.16	0.10	0.12	0.10	0.10
	MAE	0.14	0.09	0.10	0.06	0.08

MLP: Multilayer perceptron; ML: Machine learning; XGBoost: Extreme gradient boosting; RF: Random forest; R^2 : Coefficient of determination; RMSE: Root mean square; MAE: Mean absolute error.

The MLP model demonstrates excellent predictive accuracy, as indicated by high R^2 values, particularly for the number of leaves and roots, which are both at 0.93. This suggests that MLP is particularly effective in capturing the underlying patterns for these traits. The parameters' low RMSE and MAE values further affirm the model's precision, indicating minimal deviations from the actual values.

While the RF model shows slightly lower R^2 values compared to MLP, it still maintains robust predictive capabilities, especially for rooting parameters, where it matches the high R^2 values noted in XGBoost. The RF model exhibits slightly higher RMSE values, yet it achieves very low MAE values for rooting parameters, highlighting its efficiency in average error minimization.

XGBoost shows a strong performance, particularly in rooting parameters, with R^2 values peaking at 0.93 for the number of roots and 0.92 for root length. These high values reflect XGBoost's capability to handle complex data interactions, likely prevalent in root development. The RMSE and MAE values are competitive and consistent, suggesting that XGBoost effectively balances accuracy and error minimization.

Overall, all three models exhibit high consistency in predictive accuracy across the measured parameters, with particular strengths noted in specific areas. The results underscore the potential of these models in enhancing predictive modeling applications in plant science, providing valuable insights that could help optimize micropropagation practices and improve understanding of plant growth dynamics under controlled conditions. Each model has shown specific strengths that make them suitable for different aspects of predictive modeling in plant science, highlighting the importance of model selection based on the nature of the data and the specific traits being analyzed. However, overall, the XGBoost model showed the best performance for several reasons: consistently high R^2 values, balanced error metrics, and performance in complex data patterns. While the MLP and RF models also performed well, particularly MLP with very high R^2 values in some parameters, XGBoost's consistent performance across all metrics and parameters slightly edges out the others, making it the best model in this analysis for predicting micropropagation and rooting parameters. This makes XGBoost particularly valuable for complex and varied datasets commonly encountered in plant science research.

Figures 1–5 illustrate how the estimated and actual values are distributed among different models when analyzing metrics connected to plants.

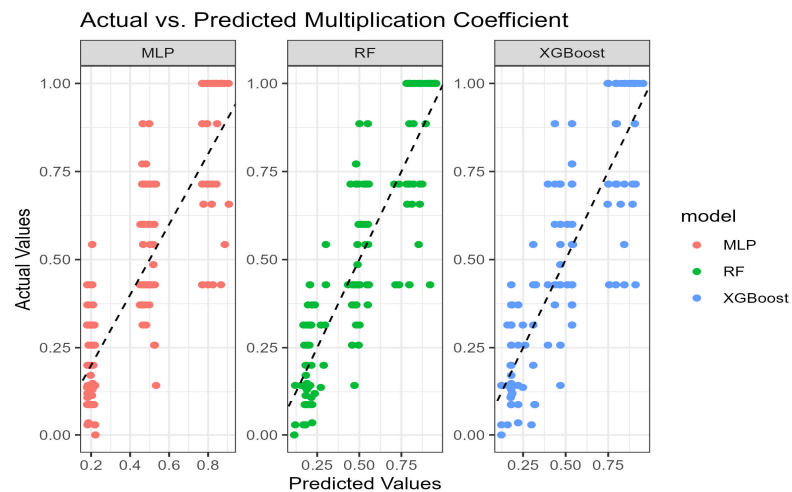


Figure 1. Actual vs. predicted values of the three models (MLP, RF, and XGBoost) for the multiplication coefficient.

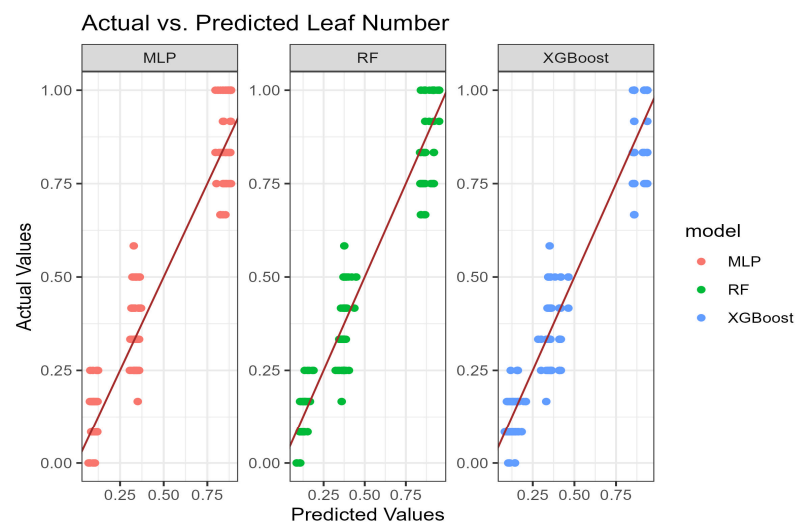


Figure 2. Actual vs. predicted values of the three models (MLP, RF, and XGBoost) for leaf number.

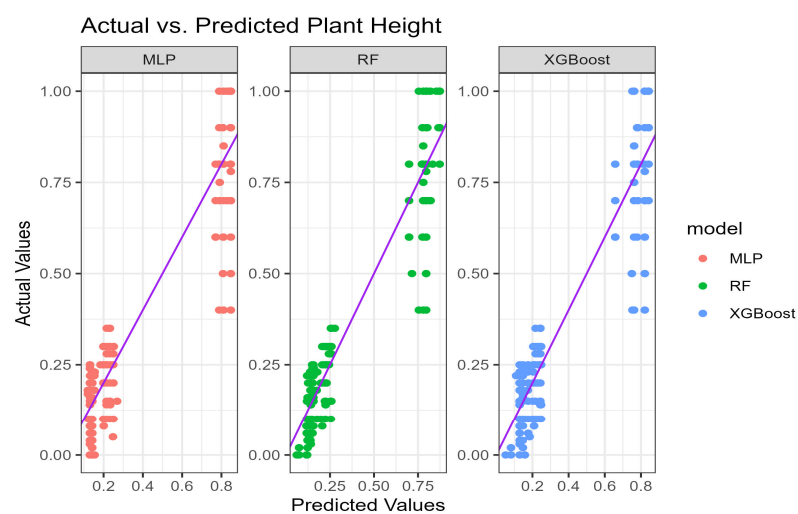


Figure 3. Actual vs. predicted values of the three models (MLP, RF, and XGBoost) for plant height.

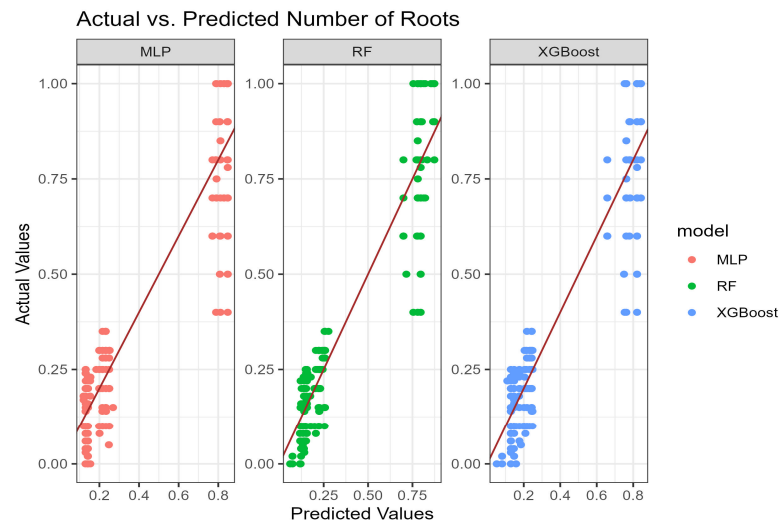


Figure 4. Actual vs. predicted values of the three models (MLP, RF, and XGBoost) for the number of roots.

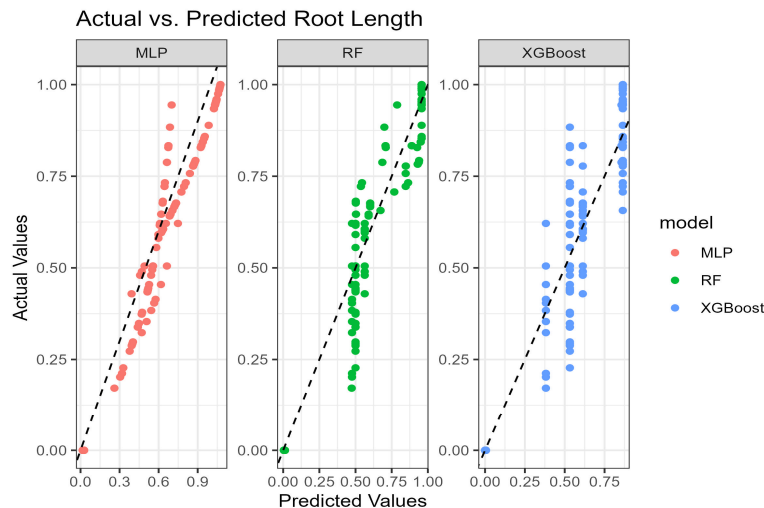


Figure 5. Actual vs. predicted values of the three models (MLP, RF, and XGBoost) for root length.

4. Discussion

In our study, we explore the impact of different shooting and rooting conditions on crucial growth parameters during the micropropagation of tea. This examination delves into how media compositions influence the multiplication coefficient, number of leaves, plant height, number of roots, and root length—vital metrics that shed light on the effectiveness of various media compositions. We have uncovered the nuanced relationship between growth outcomes and media compositions through meticulous analysis, providing valuable insights into the factors that drive successful tea plant propagation. Our results demonstrate significant differentiation in growth metrics across two distinct media conditions, SM1&RM1 and SM2&RM2. This differentiation was evidenced by the mean values and corresponding *p*-values derived from an ANOVA, followed by post hoc comparisons using the LSD test. These findings indicate the pronounced influence of media environments on various growth metrics, with the composition of the media playing a crucial role in determining plant growth and development. Like many other plants, the tissue culture of tea has a rich history that spans several years. This research aims to identify a more practical process for the commercial production of tea and establish a protocol that facilitates the production of many plants per unit of area and time. Despite potential variations due to the plant material studied, the source of the explant, or laboratory conditions, reporting on micropropagation is invaluable for such a significant plant. Our study preferred

chemicals that are easily accessible and commonly found in many countries. Growth regulators such as BAP for micropropagation and IBA for in vitro rooting were selected based on their widespread availability and effectiveness. Similar studies corroborate our findings, highlighting the responsiveness of tea plant nodal explants to BAP and IAA in specific media compositions, leading to an increase in multiple shoots per explant. For instance, a study reported that nodal explants of a tea clone exhibited better responses and produced 12.6 ± 1.9 multiple shoots per explant when supported with BAP (4 mg/L) and IAA (0.2 mg/L) in $\frac{1}{2}$ MS medium [46]. Bag et al. [47] conducted practices to enhance the acclimatization success of tea, noting the effectiveness of waiting for 12 weeks for the root induction of micropropagated plantlets and the positive influence of using larger bottles (500 mL) instead of 250 mL containers on growth, stem thickness, leaf number, and leaf area.

Further, Sun et al. [48] utilized plant growth regulators for callus induction and subsequent plant regeneration in different explants of an elite tea (*C. sinensis*) clone. The combination of Tidiazuron (TDZ), BAP, and IBA was used, with varying degrees of success depending on the explant and growth regulator concentrations. Similarly, in another study, TDZ was found to induce in vitro shoot proliferation in tea seedlings at different and lower concentrations and to induce callus formation at higher concentrations [49]. Sarathchandra et al. [50] investigated somatic embryogenesis in callus culture studies using tea plants' young leaves and nodal explants, observing callus formation and embryo-like structures under specific conditions. Samarina et al. [51] showed genetic stability and variability among micropropagated tea plantlets over a long-term period. The morphological variability and instances of aneuploidy in their findings complement our results, where different media compositions led to variable growth outcomes. These results underline the importance of selecting and stabilizing media conditions to minimize genetic and morphological variability in micropropagation. Additionally, the impact of plant growth regulators on embryogenesis in tea reported by Wachira and Ogada [52] correlates with our observations on how these regulators influence different growth metrics in micropropagation. Wachira's findings, where embryogenic competence was reduced by increasing concentrations of growth regulators, support our recommendations for optimizing regulator concentrations to enhance growth outcomes. Furthermore, Molina et al. [53]'s work on the regeneration effectiveness of different media for tea clones provides a practical foundation that aligns with our findings. Their identification of the optimal media conditions for specific clones complements our study's emphasis on media optimization, illustrating the broader applicability of these findings in tailoring micropropagation practices to different genetic backgrounds. Also, the study by Gonbad et al. [54] on protocol optimization using nodal segments highlights the critical role of precise growth regulator combinations and concentrations, as we also noted in our study. The comparison between Gonbad's findings and ours further emphasizes the significance of detailed, condition-specific studies for improving micropropagation efficiency and success.

In summary, the comparative discussion of our results within the context of the existing literature, including studies [46,55], underscores the significance of media composition in the micropropagation of tea. The substantial differences observed between Medium 1 and Medium 2 across all metrics underline the crucial role of media environments in optimizing growth conditions. These insights contribute to the broader body of knowledge, enabling researchers and practitioners to refine agricultural and horticultural practices for improved plant propagation.

In our study on tea micropropagation, we investigated the performance of three machine learning models: RF, XGBoost, and MLP, across various growth parameters, including the multiplication coefficient, leaf count, plant height, number of roots, and root length. Our analysis, encapsulated in Table 2, demonstrates that the MLP model, with its highest R^2 scores across multiple metrics, is superior in capturing the nuances of tea plant growth under varied conditions, outperforming both RF and XGBoost in consistency across evaluated parameters.

Comparing our findings with those from similar studies reveals a broader landscape of machine learning applications in plant science. For instance, the study by Şimşek [20] on strawberry cultivars under drought stress highlighted the RF model's exceptional accuracy in predicting water stress effects, demonstrating the model's adaptability across strawberries with different genetic backgrounds. This echoes our findings, where RF also showed strong performance, especially in predicting the number of roots with an R^2 of 0.90, underscoring the effectiveness of RF in plant sciences when dealing with specific growth traits. Further, the application of machine learning in the lavender study by Şimşek et al. [20] revealed varying degrees of effectiveness across different models, including MLP, RBF, XGBoost, and GP, in predicting plant characteristics. The study noted XGBoost's impressive performance in predicting root length, which is similar to our observations, where XGBoost displayed intermediate performance across the micropropagation parameters of tea. This similarity underscores the potential of XGBoost in handling complex datasets across different plant species.

On the other hand, Pepe et al.'s [24] exploration into *Cannabis sativa* L. in vitro culture optimization using MLP, GRNN, and ANFIS models indicated GRNN's superior predictive performance over the others. This contrasts with our findings, where MLP outshined RF and XGBoost, suggesting that the choice of model may depend on the specific characteristics of the plant species and the growth parameters under investigation. The study by Aasim et al. [55] on the in vitro germination of hemp further elaborates on the potential of machine learning, particularly emphasizing the standout performance of the random forest model. This aligns with our results, highlighting RF's strong predictive capability, especially in rooting parameters. It underscores the utility of RF in scenarios requiring high precision in predictions, such as in vitro germination and seedling growth traits.

These comparative insights across studies illuminate the machine learning models' diverse utility and performance in plant science research. While MLP emerged as the most consistent performer in our tea micropropagation study, the superior performance of RF in predicting specific growth traits, as seen in both our study and others, showcases the model's robustness. XGBoost's intermediate performance, alongside the varying effectiveness of models like GRNN in cannabis research, highlights the importance of model selection based on the specific dataset and research objectives. The comparison validates the significant role of machine learning in enhancing agricultural and horticultural practices and emphasizes the necessity for tailored approaches in choosing the most appropriate model for specific plant studies.

5. Conclusions

Applying machine learning algorithms to studying tea plant propagation in vitro has dramatically improved our comprehension of ideal growth environments. A comparative study of the RF, XGBoost, and MLP models revealed that the MLP model accurately and reliably predicts important growth parameters under various media circumstances. This result emphasizes how important media composition is in determining how plants grow. Machine learning is being incorporated into plant science research to help develop more sustainable and productive agriculture and horticulture techniques and improve our understanding of intricate biological processes. This study offers a practical route toward these improvements by highlighting the potential of machine learning to enhance plant propagation methods.

Author Contributions: Conceptualization, T.B. and Ö.Ş.; data curation, M.A.I.; methodology, T.B., M.A.I., S.İ., İ.D. and Ö.Ş.; supervision, Ö.Ş. and T.B.; visualization, M.A.I., S.İ. and İ.D.; writing—original draft, T.B., M.A.I., İ.D., S.İ. and Ö.Ş.; writing—review and editing, S.İ., İ.D., T.B. and Ö.Ş. All authors have read and agreed to the published version of the manuscript.

Funding: This research and APC was funded by Tekfen Agricultural Research Production and Marketing Inc.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Authors Taner Bozkurt, Sezen İnan and İjlal Dündar were employed by the company Tekfen Agricultural Research Production and Marketing Inc. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Bedrood, Z.; Rameshrad, M.; Hosseinzadeh, H. Toxicological effects of *Camellia sinensis* (green tea): A review. *Phytother. Res.* **2018**, *32*, 1163–1180. [CrossRef] [PubMed]
- Mondal, T.K.; Bhattacharya, A.; Laxmikumaran, M.; Singh Ahuja, P. Recent advances of tea (*Camellia sinensis*) biotechnology. *Plant Cell Tissue Organ Cult.* **2004**, *76*, 195–254. [CrossRef]
- Karwowska, K.; Skotnicka, M.; Smiechowska, M. Tea production and its forecasts, and the possibility of tea cultivation in the context of environmental requirements in China. *Zesz. Nauk. Szkoły Głównej Gospod. Wiej. W Warszawie* **2019**, *19*, 180–191. [CrossRef]
- Camargo, L.E.A.; Pedroso, L.S.; Vendrame, S.C.; Mainardes, R.M.; Khalil, N.M. Antioxidant and antifungal activities of *Camellia sinensis* (L.) Kuntze leaves obtained by different forms of production. *Braz. J. Biol.* **2016**, *76*, 428–434. [CrossRef] [PubMed]
- Bancirova, M. Comparison of the antioxidant capacity and the antimicrobial activity of black and green tea. *Food Res. Int.* **2010**, *43*, 1379–1382. [CrossRef]
- Tariq, M.; Naveed, A.; Barkat, A.K. The morphology, characteristics, and medicinal properties of *Camellia sinensis* tea. *J. Med. Plants Res.* **2010**, *4*, 2028–2033. [CrossRef]
- Prasanth, M.I.; Sivamaruthi, B.S.; Chaiyasut, C.; Tencomnao, T. A review of the role of green tea (*Camellia sinensis*) in antiphotaging, stress resistance, neuroprotection, and autophagy. *Nutrients* **2019**, *11*, 474. [CrossRef]
- Kumar, A.; Nair, A.G.C.; Reddy, A.V.R.; Garg, A.N. Availability of essential elements in Indian and US tea brands. *Food Chem.* **2005**, *89*, 441–448. [CrossRef]
- Vasishth, K.; Sharma, P.D.; Karan, M.; Rakesh, D.; Vyas, S.; Sethi, S.; Manktala, R. *Study to Promote the Industrial Exploitation of Green Tea Poly-Phenols in India*; ICS-UNIDO: Rome, Italy, 2003; pp. 15–22.
- Mukhopadhyay, M.; Mondal, T.K.; Chand, P.K. Biotechnological advances in tea (*Camellia sinensis* [L.] O. Kuntze): A review. *Plant Cell Rep.* **2016**, *35*, 255–287. [CrossRef]
- Evensen, N.A.; Braun, P.C. The effects of tea polyphenols on *Candida albicans*: Inhibition of biofilm formation and proteasome inactivation. *Can. J. Microbiol.* **2009**, *55*, 1033–1039. [CrossRef]
- FAOSTAT. 2024. Available online: <https://www.fao.org/faostat/en/#data> (accessed on 4 May 2024).
- Mondal, T.K. Tea breeding. In *Breeding Plantation Tree Crops: Tropical Species*; Springer: New York, NY, USA, 2009; pp. 545–587.
- Mondal, T.K. *Camellia*. In *Wild Crop Relatives: Genomic and Breeding Resources*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 15–39. [CrossRef]
- Widhianata, H.; Taryono. Organogenesis responses of tea (*Camellia sinensis* (L.) O. Kuntze) var. *assamica* and *sinensis*. *AIP Conf. Proc.* **2019**, *2099*, 020026. [CrossRef]
- Hesami, M.; Jones, A.M.P. Application of artificial intelligence models and optimization algorithms in plant cell and tissue culture. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 9449–9485. [CrossRef] [PubMed]
- Hesami, M.; Pepe, M.; Monthony, A.S.; Baiton, A.; Jones, A.M.P. Modeling and optimizing in vitro seed germination of industrial hemp (*Cannabis sativa* L.). *Ind. Crop. Prod.* **2021**, *170*, 113753. [CrossRef]
- Hesami, M.; Naderi, R.; Tohidfar, M.; Yoosefzadeh-Najafabadi, M. Application of adaptive neuro-fuzzy inference system-non-dominated sorting genetic Algorithm-II (ANFIS-NSGAI) for modeling and optimizing somatic embryogenesis of *Chrysanthemum*. *Front. Plant Sci.* **2019**, *10*, 455523. [CrossRef] [PubMed]
- Kirtis, A.; Aasim, M.; Katurci, R. Application of artificial neural network and machine learning algorithms for modeling the in vitro regeneration of chickpea (*Cicer arietinum* L.). *Plant Cell Tissue Organ Cult.* **2022**, *150*, 141–152. [CrossRef]
- Şimşek, Ö. Machine Learning Offers Insights into the Impact of In Vitro Drought Stress on Strawberry Cultivars. *Agriculture* **2024**, *14*, 294. [CrossRef]
- Demirel, F.; Uğur, R.; Popescu, G.C.; Demirel, S.; Popescu, M. Usage of Machine learning algorithms for establishing an effective protocol for the in vitro micropropagation ability of black chokeberry (*Aronia melanocarpa* (Michx.) Elliott). *Horticulturae* **2023**, *9*, 1112. [CrossRef]
- Şimşek, Ö.; Dalda Şekerci, A.; Isak, M.A.; Bulut, F.; İzgü, T.; Tütüncü, M.; Dönmez, D. Optimizing Micropropagation and Rooting Protocols for Diverse Lavender Genotypes: A Synergistic Approach Integrating Machine Learning Techniques. *Horticulturae* **2024**, *10*, 52. [CrossRef]
- Emamgholizadeh, S.; Parsaeian, M.; Baradaran, M. Seed yield prediction of sesame using artificial neural network. *Eur. J. Agron.* **2015**, *68*, 89–96. [CrossRef]
- Pepe, M.; Hesami, M.; Jones, A.M.P. Machine learning-mediated development and optimization of disinfection protocol and scarification method for improved in vitro germination of cannabis seeds. *Plants* **2021**, *10*, 2397. [CrossRef]
- García-Pérez, P.; Lozano-Milo, E.; Landin, M.; Gallego, P.P. Machine Learning unmasked nutritional imbalances on the medicinal plant *Bryophyllum* sp. cultured in vitro. *Front. Plant Sci.* **2020**, *11*, 576177. [CrossRef] [PubMed]

26. Izquierdo, P.F.; Coral, M.P.; Benavides, F.O. Application of an artificial neural network coupled to a genetic algorithm for the production of polyphenols in *Parachlorella kessleri* grown under mixotrophic conditions. *Algal Res.* **2024**, *77*, 103331. [\[CrossRef\]](#)
27. Özcan, E.; Atar, H.H.; Ali, S.A.; Aasim, M. Artificial neural network and decision tree-based models for prediction and validation of in vitro organogenesis of two hydrophytes—*Hemianthus callitrichoides* and *Riccia fluitans*. *In Vitro Cell. Dev. Biol.-Plant* **2023**, *59*, 547–562. [\[CrossRef\]](#)
28. Niazian, M.; Niedbala, G. Machine learning for plant breeding and biotechnology. *Agriculture* **2020**, *10*, 436. [\[CrossRef\]](#)
29. Gimenez, V.; Pauleti, P.; Silva, A.; Costa, E. Bioelectrical pattern discrimination of miconia plants by spectral analysis and machine learning. *Theor. Exp. Plant Physiol.* **2021**, *33*, 329–342. [\[CrossRef\]](#)
30. Sun, S.; Wang, C.; Ding, H.; Zou, Q. Machine learning and its applications in plant molecular studies. *Brief. Funct. Genom.* **2019**, *19*, 40–48. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Isewon, I.; Apata, O.; Oluwamuyiwa, F.; Aromolaran, O.; Oyelade, J. Machine learning algorithms: Their applications in plant omics and agronomic traits' improvement. *F1000research* **2022**, *11*, 1256. [\[CrossRef\]](#)
32. Naik, K.; Pandey, M.; Panda, A.; Albasir, A.; Taneja, K. Data driven modelling of nuclear power plant performance data as finite state machines. *Modelling* **2021**, *2*, 43–62. [\[CrossRef\]](#)
33. Mohanty, S.; Hughes, D.; Salathé, M. Using deep learning for image-based plant disease detection. *Front. Plant Sci.* **2016**, *7*, 215232. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Munjal, D.; Singh, L.; Pandey, M.; Lakra, S. A systematic review on the detection and classification of plant diseases using machine learning. *Int. J. Softw. Innov.* **2022**, *11*, 1–25. [\[CrossRef\]](#)
35. Wei, X.; Johnson, M.; Langston, D.; Mehl, H.; Song, L. Identifying optimal wavelengths as disease signatures using hyperspectral sensor and machine learning. *Remote Sens.* **2021**, *13*, 2833. [\[CrossRef\]](#)
36. Sinshaw, N.; Assefa, B.; Mohapatra, S.; Beyene, A. Applications of computer vision on automatic potato plant disease detection: A systematic literature review. *Comput. Intell. Neurosci.* **2022**, *2022*, 7186687. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Sujatha, R.; Chatterjee, J.; Jhanjhi, N.; Brohi, S. Performance of deep learning vs machine learning in plant leaf disease detection. *Microprocess. Microsyst.* **2021**, *80*, 103615. [\[CrossRef\]](#)
38. Guo, D.; Juan, J.; Chang, L.; Zhang, J.; Huang, D. Discrimination of plant root zone water status in greenhouse production based on phenotyping and machine learning techniques. *Sci. Rep.* **2017**, *7*, 8303. [\[CrossRef\]](#)
39. Huang, Y.; Xin, Z.; Zhang, P.; Reddy, K.; Zhang, J. Hyperspectral plant sensing for differentiating glyphosate-resistant and glyphosate-susceptible johnsongrass through machine learning algorithms. *Pest Manag. Sci.* **2022**, *78*, 2370–2377. [\[CrossRef\]](#)
40. Halgamuge, M. Machine learning for bioelectromagnetics: Prediction model using data of weak radiofrequency radiation effect on plants. *Int. J. Adv. Comput. Sci. Appl.* **2017**, *8*, 223–235. [\[CrossRef\]](#)
41. Ayu, F.; Manessa, M.D.M.; Stevanus, C.T.; Efriana, A.F. Machine learning approach to assess rubber plant health through canopy density mapping using very high-resolution aerial photographs. In *Eighth Geoinformation Science Symposium 2023: Geoinformation Science for Sustainable Planet*; SPIE: Bellingham, WA, USA, 2024; Volume 12977, pp. 667–676. [\[CrossRef\]](#)
42. Espejo-Garcia, B.; Malounas, I.; Vali, E.; Fountas, S. Testing the suitability of automated machine learning for weeds identification. *AI* **2021**, *2*, 34–47. [\[CrossRef\]](#)
43. Dimitriadis, S.; Goumopoulos, C. Applying machine learning to extract new knowledge in precision agriculture applications. In *Proceedings of the 2008 Panhellenic Conference on Informatics*, Samos, Greece, 28–30 August 2008; pp. 100–104. [\[CrossRef\]](#)
44. Liu, Z.; Bashir, R.N.; Iqbal, S.; Shahid, M.M.A.; Tausif, M.; Umer, Q. Internet of Things (IoT) and machine learning model of plant disease prediction—blister blight for tea plant. *IEEE Access* **2022**, *10*, 44934–44944. [\[CrossRef\]](#)
45. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
46. Borchetia, S.; Das, S.C.; Handique, P.J.; Das, S. High multiplication frequency and genetic stability for commercialization of the three varieties of micropropagated tea plants (*Camellia* spp.). *Sci. Hortic.* **2009**, *120*, 544–550. [\[CrossRef\]](#)
47. Bag, N.; Palni, L.M.S.; Nandi, S.K. An efficient method for acclimatization: In vitro hardening of tissue culture-raised tea plants (*Camellia sinensis* (L.) O. Kuntze). *Curr. Sci.* **2019**, *117*, 288–293. [\[CrossRef\]](#)
48. Sun, J.; Lei, P.; Zhang, Z.Z.; Shi, G.; Tang, Z.; Zhu, S.; Jiang, C.J.; Wan, X.C. Shoot basal ends as novel explants for in vitro plantlet regeneration in an elite clone of tea. *J. Hortic. Sci. Biotechnol.* **2012**, *87*, 71–76. [\[CrossRef\]](#)
49. Mondal, T.K.; Amita, B.; Anil, S.; Paramvir, S.A. Micropropagation of tea (*Camellia sinensis* (L.) O. Kuntze) using thidiazuron. *Plant Growth Regul.* **1998**, *26*, 57–61. [\[CrossRef\]](#)
50. Sarathchandra, T.M.; Upali, P.D.; Wijewardena, R.G.A. Studies on the tissue culture of tea (*Camellia sinensis* (L.) O. Kuntze) 4 somatic embryogenesis in stem and leaf callus cultures. *Sri Lanka J. Tea Sci.* **1988**, *57*, 50–54.
51. Samarina, L.; Gvasaliya, M.; Koninskaya, N.; Rakhmangulov, R.; Efremov, A.; Kiselyova, N.; Ryndin, A.; Hanke, M.-V. A comparison of genetic stability in tea [*Camellia sinensis* (L.) Kuntze] plantlets derived from callus with plantlets from long-term in vitro propagation. *Plant Cell Tissue Organ Cult.* **2019**, *138*, 467–474. [\[CrossRef\]](#)
52. Wachira, F.; Ogada, J. In vitro regeneration of *Camellia sinensis* (L.) O. Kuntze by somatic embryogenesis. *Plant Cell Rep.* **1995**, *14*, 463–466. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Molina, S.P.; Rey, H.Y.; Pérez, M.L.; Mroginski, L.A. Plant regeneration of tea (*Camellia sinensis*) by in vitro culture of meristems, axillary buds and uninodal segments. *J. Fac. Agric. Sci. Natl. Univ. Cuyo* **2013**, *45*, 127–134.

54. Gonbad, R.A.; Sinniah, U.R.; Abdul Aziz, M.; Mohamad, R. Influence of cytokinins in combination with GA 3 on shoot multiplication and elongation of tea clone Iran 100 (*Camellia sinensis* (L.) O. Kuntze). *Sci World J.* **2014**, *2014*, 943054. [[CrossRef](#)] [[PubMed](#)]
55. Aasim, M.; Katırcı, R.; Akgur, O.; Yildirim, B.; Mustafa, Z.; Nadeem, M.A.; Baloch, F.S.; Karakoy, T.; Yılmaz, G. Machine learning (ML) algorithms and artificial neural network for optimizing in vitro germination and growth indices of industrial hemp (*Cannabis sativa* L.). *Ind. Crop. Prod.* **2022**, *181*, 114801. [[CrossRef](#)]

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.