A Method for Estimating Tree Growth Potential with Back Propagation Neural Network

Jianfeng Yao , Cancong Zhao , Xuefan Hu * , Yingshan Jin * , Yanling Li , Liming Cai , Zhuofan Li , Fang Li , Fang Liang

Posted Date: 18 January 2025

doi: 10.20944/preprints202501.1296.v1

Keywords: tree growth potential; BPNN; forest management; classification model

Article

A Method for Estimating Tree Growth Potential with Back Propagation Neural Network

Jianfeng Yao [1,2,3], Cancong Zhao [1], Xuefan Hu [4,*], Yingshan Jin [4,*], Yanling Li [1], Liming Cai [2,3,5], Zhuofan Li [2,3,6], Fang Li [4] and Fang Liang [4]

[1] College of Computer and Information Technology, Xinyang Normal University, Xinyang 464000, China
[2] Henan Dabieshan National Field Observation and Research Station of Forest Ecosystem, Zhengzhou 450046, China
[3] Xinyang Academy of Ecological Research, Xinyang 464000, China
[4] Beijing Key Laboratory of Greening Plants Breeding, Beijing Academy of Forestry and Landscape Architecture, Beijing 100102, China
[5] College of Mathematics and Statistics, Xinyang Normal University, Xinyang 464000, China
[6] College of Tourism, Xinyang Normal University, Xinyang 464000, China
* Correspondence: hufanzi@163.com ;j_ysh@163.com

Abstract: Tree growth potential is crucial for maintaining forest health and sustainable development. Traditional expert-based assessments of growth potential are inherently subjective. To address this subjectivity and improve accuracy, this study proposed a method of using Backpropagation Neural network (BPNN) to classify tree growth potential. 60 *Pinus tabulaeformis* (Carr.) and 60 *Platycladus orientalis* (Linn.) were selected as experimental trees in the Miyun Reservoir Water Conservation Forest Demonstration Zone in Beijing, and 95 *Pinus massoniana* (Lamb.) and 60 *Cunninghamia lanceolate* (Linn.) were selected as experimental trees in the Jigongshan Nature Reserve. The average annual ring width of the outermost 2cm xylem of the experimental trees were measured by discs or increment cores, and the wood volume increment of each experimental trees in recent years were calculated. According to wood volume increment, the growth potential of experimental trees was divided into three levels: strong, medium, and weak. Using tree height, breast height diameter, average crown width as input variables, using growth potential level as output variables, four sub models for each tree species were established; Using tree species, tree height, breast height diameter, average crown width as input variables, using growth potential level as output variables, a generalized model was established for these four tree species. The test results showed that the accuracy of the sub models for *Pinus tabulaeformis*, *Platycladus orientalis*, *Pinus massoniana*, and *Cunninghamia lanceolate* were 68.42%, 77.78%, 86.21%, and 78.95%, respectively, and the accuracy of the generalized model was 71.19%. These findings suggested that employing BPNN is a viable approach for accurately estimating tree growth potential.

Keywords: tree growth potential; BPNN; forest management; classification model

1. Introduction

The concept of tree growth potential encompasses the vigor and vitality of tree growth, serving as an indicator of both growth rate and overall health status [1]. A stronger growth potential signifies a faster growth rate and an enhanced ability to resist diseases and pests [2]. Accurately estimating tree growth potential is crucial for several reasons. Firstly, it aids in maintaining forest health. By accurately assessing growth potential, forestry managers can selectively remove trees with weaker growth potential while preserving those with stronger potential, thereby promoting vitality and reducing the incidence of forest diseases and pests [2]. Secondly, it contributes to improved forest productivity and carbon storage. Accurate estimation allows forest managers to devise appropriate

 doi:10.20944/preprints202501.1296.v1

management plans, such as harvesting trees with declining growth potential to control forest density and foster an environment conducive to the growth of remaining trees [3]. Thirdly, incorporating growth potential factor can significantly enhance the accuracy of tree growth models. Current models often overlook this factor, resulting in lower estimation accuracy, particularly for individual trees. By integrating growth potential as a variable, the predictive accuracy of these models can be substantially improved [4]. Consequently, precise estimation of tree growth potential is vital for sustainable forest development.

The accurate quantification of tree growth potential is a complex challenge influenced by various factors, such as genetic makeup [5], environmental conditions [6], stand density [7,8], tree height [9], crown width [10], leaf characteristics [11], and root system development [12]. Current assessment methods primarily rely on qualitative grading, with limited research on robust quantitative approaches for tree growth potential evaluation. To address this challenge, this study proposes a novel approach that employs Back Propagation Neural Networks (BPNN) to classify tree growth potential. Tree height, DBH, and average crown width are used as input variables to predict the growth potential category of individual trees. The growth potential category is determined by the incremental increase in tree volume, which is calculated from the average annual ring width within the outermost 2 cm xylem of the tree trunk. The classified growth potential was used as the target variable for the BPNN models. Four sub BPNN models were developed for each experimental tree species: *Pinus tabuliformis*, *Platycladus orientalis*, *Pinus massoniana*, and *Cunninghamia lanceolate*. Additionally, in order to establish a universally applicable model with strong generalization ability, a generalized model was constructed using data of these four experimental tree species.

2. Materials and Methods

2.1. Overview of the Study Area

The study areas included Miyun Reservoir Water Conservation Forest Demonstration Zone and Jigong Mountain Nature Reserve.

Miyun Reservoir Water Conservation Forest Demonstration Zone (116°53′ E, 40°25′ N) is located in the north of Beijing, China. The region has a sub-humid, warm temperate, continental monsoon climate, characterized by hot, rainy summers and cold, dry winters. The average annual precipitation is 600 mm, and the average annual temperature is 11.8 °C. There are low mountains at an altitude of approximately 120 m a.s.l., with slopes ranging from 10° to 35°. The south-facing slopes are characterized by thin soil and exposed rocks, whereas the north-facing slopes feature shaded, deep, and loose soil. In 1980, a *Pinus tabuliformis* plantation was established on a north slope. The age of *Pinus tabuliformis* was about 44 years. This plantation exhibits an average tree height of 7.56 m and an average trunk diameter of 14.8 cm, with a stand density of approximately 915 trees per hectare and a canopy density of about 0.8. Conversely, a *Platycladus orientalis* plantation was established on a south slope in 1990. The age of *Platycladus orientalis* was about 34 years. The *Platycladus orientalis* trees have an average height of 6.87 m and an average trunk diameter of 9.1 cm. This plantation has a stand density of about 2550 trees per hectare and a canopy density of approximately 0.95. In 2024, 8 plots of 50m * 50m were established, including 4 plots of *Pinus tabulaeformis* and 4 plots of *Platycladus orientalis*. The distribution map of *Pinus tabulaeformis* plots and *Platycladus orientalis* plots was shown in Figure 1.

 doi:10.20944/preprints202501.1296.v1

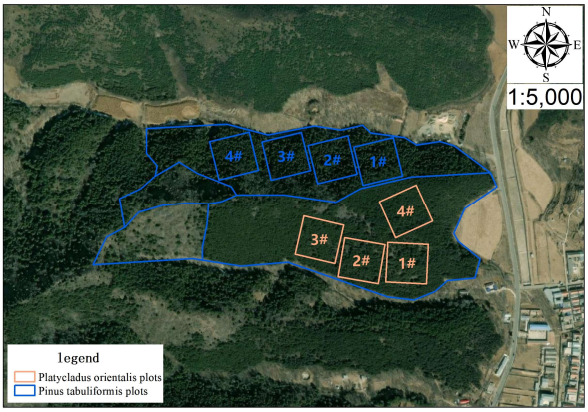
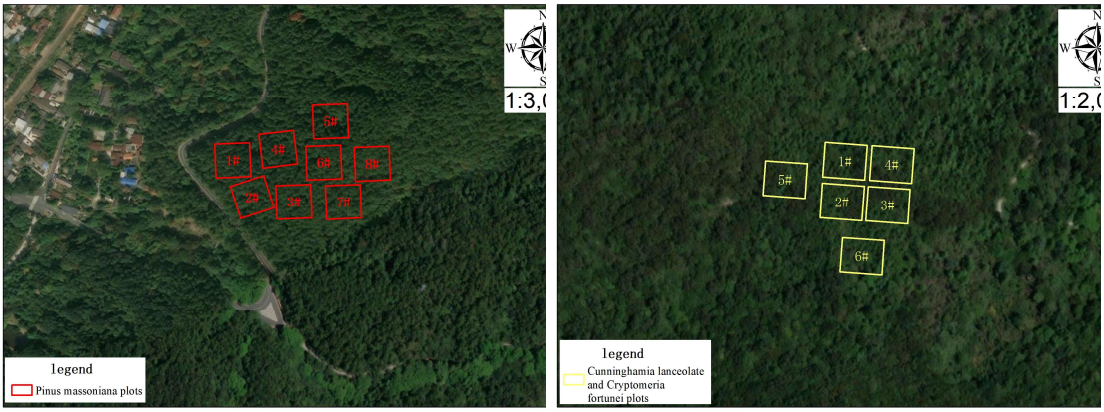


Figure 1. The distribution map of *Pinus tabulaeformis* plots and *Platycladus orientalis* plots in Miyun Reservoir Water Conservation Forest Demonstration Zone.

Jigong Mountain Nature Reserve (114°01′-114°06′ E, 31°46′-31°52′ N) is located in the south of in Xinyang City, Henan Province, China. This region is a typical transition area from the central subtropical to the northern subtropical zone in China, with a unique ecological environment, making it an important natural reserve. It has a subtropical monsoon climate with distinct seasons: hot and humid summers, and mild winters. The average annual precipitation ranges from 1100 mm to 1400 mm, and the average annual temperature is approximately 15.2°C. The vegetation in the area is complex and diverse, mainly consisting of deciduous broadleaf forests, mixed coniferous and broadleaf forests, and evergreen broadleaf forests, with typical vegetation being the central subtropical forest. The soil is mainly yellow-brown soil and mountain brown soil, suitable for the growth of various plants. In 2022, 14 plots of 20m * 20m were established, including 8 *Pinus massoniana* plots and 6 *Cunninghamia lanceolate* and *Cryptomeria fortune* mixed forest plots. The distribution map of *Pinus massoniana plots* and *Cunninghamia lanceolate and Cryptomeria fortune* mixed forest plots was shown in Figure 2.



(**a**) (**b**)

Figure 2. The distribution map of the plots in the Ji gong Mountain Nature Reserve: (**a**) *Pinus massoniana* plots; (**b**) *Cunninghamia lanceolate* and *Cryptomeria fortunei* mixed forests plots.

2.2. Data Collection

In each plot in Miyun Reservoir Water Conservation Forest Demonstration Zone, 15 trees were selected as experimental trees, including 5 superior trees, 5 moderate trees, and 5 suppressed trees. The tree height, DBH, and crown width in the four cardinal directions (east, south, west, and north) were recorded for each experimental tree. Subsequently, a disc approximately 5 cm thick was cut

 doi:10.20944/preprints202501.1296.v1

from each tree at the 1.3m height mark on the trunk. 60 *Pinus tabuliformis* discs and 60 *Platycladus orientalis* discs were sampled.

In each plot in JiGong Mountain Nature Reserve, 12 trees were selected as experimental trees, including 4 superior trees, 4 moderate trees, and 4 suppressed trees. The tree height, DBH, and crown width in the four cardinal directions (east, south, west, and north) were recorded for each experimental tree. An increment core was extracted from each tree using an increment borer, and related tree information was recorded. The number of *Cunninghamia lanceolate* trees was less than 12 in one of the mixed forest plots, and no tree was selected as experimental tree in the plot. Therefore, there were 60 *Cunninghamia lanceolate* trees were selected as experimental trees. One *Pinus massoniana* increment core was decayed, therefore, there were 95 *Pinus massoniana* trees were selected as experimental trees.

In order to accurately classify the growth potential of experimental trees, the average volume increment of experimental trees in recent years was used to classify the types of tree growth potential. The discs and increment cores were polished until the annual ring lines were clear. The width of the annual rings in four directions (discs) or two directions (cores) was measured with the annual ring analysis instrument, Lintab 6.0. The average annual ring width of the outermost 2 cm xylem was then calculated. The volume increment was calculated using Equation (1).

$$V = \mathrm{W} \times D \times H/3, \quad (1)$$

where: W denotes the average annual ring width of the outermost 2 cm, D denotes the DBH, and H denotes the tree height.

The experimental trees of each tree species were sorted in descending order based on volume increment. The top one-third of the trees were classified as strong growth potential type, the middle one-third as medium, and the bottom one-third as weak. The statistical indicators of each growth potential type of each tree species are shown in Table 1.

Table 1. The statistical indicators of each tree species.

Tree species	Number of trees	Tree age range	Growth potential type	Tree height		Average crown width		DBH	
				Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Pinus tabuliformis	**60**	**44~46**	strong	9.645	0.987	2.103	0.443	15.276	2.025
			medium	8.455	1.019	1.972	0.368	13.150	1665
			weak	7.590	1.348	1.787	0.488	12.545	2.056
Platycladus orientalis	60	34~36	strong	8.402	1.025	1.430	0.420	11.675	1.892
			medium	7.929	0.793	1.213	0.372	9.850	1.787
			weak	7.713	1.227	1.115	0.392	8.130	1.425
Pinus massoniana	95	41~53	strong	24.347	4.289	3.165	0.811	39.266	6.755
			medium	21.250	4.196	2.490	0.817	30.208	7.245
			weak	17.197	4.350	2.030	0.848	19.374	6.319
Cunninghamia lanceolate	60	34~49	strong	19.757	2.164	2.706	1.036	33.705	4.824
			medium	17.100	3.212	1.781	0.624	24.255	4.688
			weak	11.185	2.434	1.766	0.788	14.140	5.404

2.3. Establishment of BPNN Model for Predicting Growth Potential

2.3.1. Introduction to the BPNN Algorithm

The artificial neural network algorithm is a computational method designed to emulate the neural systems of the human brain. It comprises numerous artificial neurons and the intricate connections between them, enabling the processing and learning from input data to address classification or regression problems [13]. Among the various types of artificial neural networks, the

 doi:10.20944/preprints202501.1296.v1

BPNN is notable for its excellent fault tolerance in data prediction. By employing the backpropagation of errors, BPNN can continuously optimize neuron weights and thresholds, thereby enhancing the accuracy, fairness, and objectivity of predictions [14]. The structure of the BPNN algorithm consists primarily of input, hidden, and output layers. Each connection between these layers is assigned a specific weight $W(w_{ij}^{(l)})$, forming a linear mapping from input $X(x_1,\ldots,x_m)$ to output $Y(y_1,\ldots,y_n)$ [15]. The underlying principle of this network structure is illustrated in Figure 3.

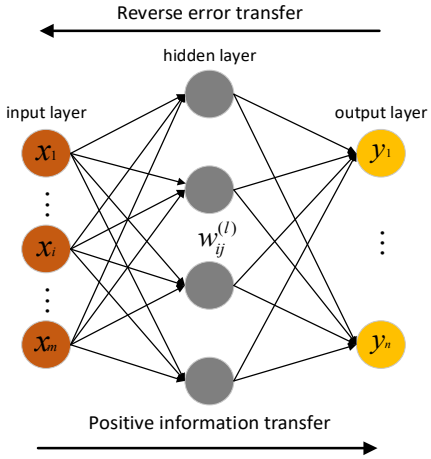


Figure 3. Principle of BPNN.

Figure 3 illustrates the weights w between layers, as shown in equation (1), where: l denotes the current hidden layer, i represents the neuron node in the previous layer, and j indicates the neuron node in the next layer.

$$W = w_{ij}^{(l)}, \tag{2}$$

In the BPNN algorithm, choosing the right activation function can increase data processing flexibility and enhance training efficiency. Common activation functions include Sigmoid, ReLU, and Softmax, with their respective formulas as follows:

$$Sigmoid(z) = \frac{1}{1+e^{-z}}, \tag{3}$$

$$ReLU(z) = \max(0, z), \tag{4}$$

$$Softmax(z) = \frac{e^{z_i}}{\sum_j e^{z_j}}, \tag{5}$$

where: z represents the output value of the previous layer node, i represents the index of the current neuron and j is used to traverse all the neuron nodes.

2.3.2. Evaluation Indicators

In the context of machine learning algorithms for classification, the confusion matrix serves as a crucial tool for summarizing prediction outcomes, particularly in classification problems. It effectively illustrates the degree of alignment between predictions of the model and the actual values in a matrix format [16]. This matrix is widely recognized as a key indicator for assessing the performance of supervised algorithms. Specifically, in a binary classification model, the primary task is to determine whether an instance should be classified as positive or negative. During the training phase, the model systematically compares predicted values with actual values of the samples. Within the confusion matrix, the rows correspond to the actual values, while the columns represent the predicted values. Typically, this matrix comprises four possible outcomes, which provide insights into accuracy of the model and error distribution.

 doi:10.20944/preprints202501.1296.v1

- **True Positive (TP):** The actual class is positive, and the model predicts it as positive;
- **False Positive (FP):** The actual class is negative, but the model predicts it as positive;
- **False Negative (FN):** The actual class is positive, but the model predicts it as negative;
- **True Negative (TN):** The actual class is negative, and the model predicts it as negative.

The structure of the confusion matrix is shown in Figure 4.

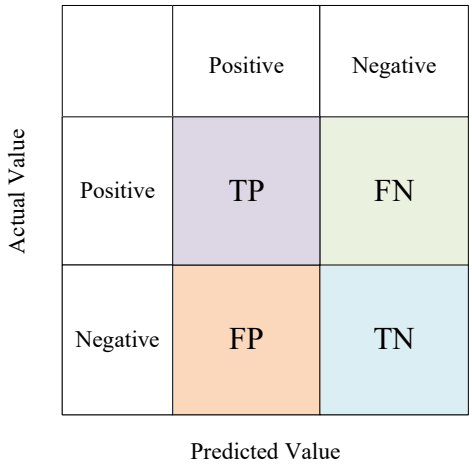


Figure 4. Structure diagram of confusion matrix.

In this study, tree growth potential was categorized into three distinct levels: strong growth potential (denoted as A_1), medium growth potential (denoted as A_2), and weak growth potential (denoted as A_3). Figure 5 illustrates the classification outcomes, where N_{ij} indicates the number of samples for which the actual growth potential is A_i, while the model predicts it as A_j [17].

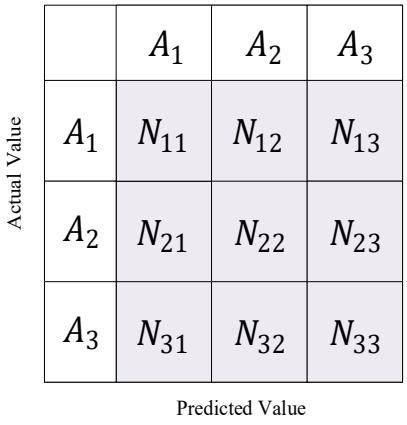


Figure 5. Confusion matrix.

Using a confusion matrix, several key evaluation metrics can be derived to assess the performance of a predictive model:

Precision: The proportion of correctly predicted samples based on the predicted results.

Recall: Based on actual samples, the proportion of correctly predicted positive cases among the total actual positive cases.

F1: The harmonic average of accuracy rate and recall rate, which is a comprehensive index to judge the overall model.

Accuracy: Represents the proportion of all correct predictions of the model. Generally, higher values for these four metrics suggest a more effective model.

 doi:10.20944/preprints202501.1296.v1

The formulas for these metrics are as follows:

$$Precision_i = \frac{N_{ii}}{\sum_{k=1}^{3} N_{ki}}, \tag{6}$$

$$Recall_i = \frac{N_{ii}}{\sum_{k=1}^{3} N_{ik}}, \tag{7}$$

$$F1_i = 2 * \frac{Precision_i * Recall_i}{Precision_i + Recall_i}, \tag{8}$$

$$Accuracy = \frac{\sum_{i=1}^{3} N_{ii}}{\sum_{i=1}^{3} \sum_{j=1}^{3} N_{ij}}, \tag{9}$$

2.3.3. Modeling Training

To enhance the accuracy of the tree growth potential prediction model, this study employed the MLPC classifier from the sklearn library to model datasets of four tree species [18]. The model training process involved several key steps:

Step 1: Data loading. Load the data from the excel file. The input variables selected were tree height, DBH, and average crown width, while the output variable was the tree growth potential classification class.

Step 2: Data preprocessing. Due to the significant differences in the numerical values of tree height, crown width, and DBH, the input variables were normalized to eliminate the scale differences of each input variable and make them equally important to the model. The values of three input variables were converted to the range of [0,1] with MinMaxScaler.

Step 3: Dataset dividing. The dataset was divided into a training set and a test set using the train_test_split function. The training set, comprising 70% of the data, was used to train the model, while the test set, comprising 30%, was used to evaluate t the generalization ability of the model. To ensure consistent dataset splits across runs, the random_state parameter was set to 22, facilitating experimental reproducibility.

Step 4: Instantiating the network model. The MLPClassifier from the sklearn library was imported, and the model was instantiated with specific parameter settings, denoted as model = MLPClassifier(*).

Step 5: Model evaluation. The accuracy of model on the test set was assessed using the accuracy_score function. Additionally, a classification report was generated to document the model's precision, recall, and F1-score.

Step 6: Result Visualization. Import the matplotlib library and the confusion_matrix function from sklearn library. Drew the confusion matrix to display the number of samples correctly and incorrectly classified in each cell of the matrix. Used the One-vs-Rest method to plot the ROC curve for each category and calculated the AUC (Area Under Curve) score to provide an intuitive view of the performance of the model in each tree growth potential category.

The construction of the generalized model requires the integration of datasets from all four tree species, incorporating tree species information as an additional input variable. To maximize the amount of training data for the model while ensuring a uniform distribution of growth potential categories, the dataset is randomly divided into training and test sets at an 8:2 ratio. The remaining steps follow the same procedure as described above.

2.3.4. Hyperparameter Adjustment

In this study, the BPNN model initially focused on determining the number of hidden layers and the number of neurons per hidden layer before proceeding to search for the optimal hyperparameters. Selecting an appropriate number of hidden layers and neurons is crucial to balancing generalization and fitting effectiveness. A single layer neural network, if its limited number of neurons, often results in poor fitting performance. Conversely, a neural network with an excessive number of layers and neurons may suffer from overfitting, thereby reducing its generalization ability

 doi:10.20944/preprints202501.1296.v1

[19]. To address this, the experiment conducted in this paper involved two groups of control experiments: a single hidden layer group (H1) and a multi-hidden layer group (H2). Both groups utilized default hyperparameters. The network structure of the model was first established, after which the Grid Search (GS) algorithm was employed to identify the optimal hyperparameters. Ultimately, the best parameter combination was determined. The default hyperparameters used in the experiment, along with those considered in the GS, are detailed in Table 2.

Table 2. Hyperparameter settings.

Hyperparameter	Default	GS Hyperparameters
Activation	Relu	[identity, logistic, tanh, relu]
Solver	Adam	[sgd, adam]
Alpha	0.0001	[0.01,0.001,0.0001,0.00001]
Learning_rate	Constant	[constant, invscaling, adaptive]
Max_iter	200	-
Random_state	None	-

3. Results

3.1. Results of Network Structure Adjustment

In a single hidden layer network structure, the number of neurons can be estimated using an empirical formula:

$$n = \sqrt{a + b} + m, \tag{10}$$

where n represents the number of neurons in the hidden layer, a is the number of input nodes, b is the number of output nodes, and m was set an adjustable constant ranging from 1 to 60 in this study [20].

In the dataset for the four tree species, the input variables include tree height, DBH, and average crown width, so the input nodes were set to 3. Additionally, in the dataset for the generalized model, tree species information was also included, which led to the input nodes being set to 4. The output variable is the growth potential, divided into three categories, so the output nodes were all set to 3. Consequently, the number of neurons n in the single hidden layer ranged from 3 to 63. Taking into account factors such as the number of training set data, model accuracy, and training time, the number of hidden layers was set to 2 in H2, and the number of neuron nodes in each layer was within the range of n. Due to the extensive results obtained from varying the number of neurons, only the top three models with the highest accuracy in both the H1 and H2 groups are presented. The maximum iteration parameter, max_iter, was set to 1500 to avoid underfitting and improve the accuracy of model. For consistent experimental results, the random_state was fixed at 22, and the default values were kept for the other hyperparameters. Both H1 and H2 used the same hyperparameters settings.

In the H1 group, the models were denoted as BPNN-1i (i=1,2,3), and in the H2 group, they were denoted as BPNN-2i (i=1,2,3). The evaluation results of the H1 and H2 models for *Pinus tabuliformis*, *Platycladus orientalis, Pinus massoniana, Cunninghamia lanceolate* are shown in Table 3, Table 4, Table 5 and Table 6.

Table 3. Model evaluation results of *Pinus tabuliformis*.

Model	Hidden_layer_sizes	Precision (%)	Recall (%)	F1 (%)	Accuracy (%)
BPNN-11	(3)	68.33	64.29	63.26	63.16
BPNN-12	(5)	70.37	63.49	64.96	63.16
BPNN-13	(9)	71.30	69.84	68.18	68.42
BPNN-21	(3,3)	68.89	69.84	67.58	68.42
BPNN-22	(3,9)	65.74	63.49	62.73	63.16

 doi:10.20944/preprints202501.1296.v1

BPNN-23	(13,9)	72.22	69.05	69.86	68.42

Table 4. Model evaluation results of *Platycladus orientalis*.

Model	Hidden_layer_sizes	Precision (%)	Recall (%)	F1 (%)	Accuracy (%)
BPNN-11	(5)	63.89	65.71	63.90	66.67
BPNN-12	(7)	59.37	60.16	59.49	61.11
BPNN-13	(45)	69.44	69.68	69.26	72.22
BPNN-21	(3,5)	71.03	70.48	69.05	72.22
BPNN-22	(3,11)	84.44	74.44	74.28	77.78
BPNN-23	(5,3)	75.79	75.24	74.10	77.78

Table 5. Model evaluation results of *Pinus massoniana*.

Model	Hidden_layer_sizes	Precision (%)	Recall (%)	F1 (%)	Accuracy (%)
BPNN-11	(7)	73.81	71.47	71.21	72.41
BPNN-12	(9)	79.06	75.64	75.02	75.86
BPNN-13	(13)	79.26	79.81	78.66	79.31
BPNN-21	(11,63)	82.32	80.77	80.37	82.76
BPNN-22	(41,3)	81.67	82.37	81.48	82.76
BPNN-23	(61,3)	85.24	84.94	84.59	86.21

Table 6. Model evaluation results of *Cunninghamia lanceolate*.

Model	Hidden_layer_sizes	Precision (%)	Recall (%)	F1 (%)	Accuracy (%)
BPNN-11	(3)	69.31	71.43	65.71	68.41
BPNN-12	(5)	73.02	74.29	73.33	73.68
BPNN-13	(7)	79.05	80.95	78.57	78.95
BPNN-21	(3,5)	68.42	69.31	71.43	65.71
BPNN-22	(3,11)	74.40	76.19	72.39	73.68
BPNN-23	(5,15)	79.05	80.95	78.57	78.95

In Table 3, the number of neurons in the three models of the H1 group is 3, 5, and 9, respectively. From the perspective of accuracy, BPNN-13 performs the best, with higher Precision, Recall, and F1 score compared to BPNN-11 and BPNN-12. In the H2 group, when the number of neurons is set to (3, 3) and (13, 9), the accuracy of the models is the same as that of the BPNN-13 model. Among these, BPNN-23 achieves the highest Precision, reaching 72.22%. Although its Recall is slightly lower, the F1 score is greater than that of both BPNN-13 and BPNN-11. Overall, the model with (13, 9) neurons yields the best performance for *Pinus tabuliformis*.

Similarly, the comparison of the H1 and H2 group experiments for *Platycladus orientalis*, *Pinus massoniana*, *Cunninghamia lanceolate*, and the generalized model reveals the optimal Hidden_layer_sizes, as shown in Table 4, Table 5, Table 6 and Table 7. It can be observed that the accuracy is highest when the hidden_layer_sizes for the *Platycladus orientalis*, *Pinus massoniana*, *Cunninghamia lanceolate*, and Generalized model are set to (5,3), (61, 3), (5, 15) or (7), and (33, 45), respectively.

Table 7. Model evaluation results of Generalized model.

Model	Hidden_layer_sizes	Precision (%)	Recall (%)	F1 (%)	Accuracy (%)
BPNN-11	(9)	63.78	62.74	62.77	62.71
BPNN-12	(15)	63.64	63.83	63.40	64.41
BPNN-13	(57)	65.93	65.41	64.40	66.10
BPNN-21	(9,35)	64.74	65.91	64.16	66.10
BPNN-22	(19,39)	70.17	67.33	66.28	67.80
BPNN-23	(33,45)	71.15	70.68	69.92	71.19

 doi:10.20944/preprints202501.1296.v1

3.2. Results of other Hyperparameters Adjustment

GS is an exhaustive method employed to optimize the parameters of an estimation function through cross-validation, ultimately identifying the most effective algorithm [21]. In this experiment, while the network structure of the BPNN model was kept constant, GS was utilized to fine-tune the remaining hyperparameters. This process typically involves iterative adjustments of hyperparameter values to determine the optimal combination [22].

Through experimentation, it was found that multiple combinations of hyperparameters could achieve the best performance for each model. This could be attributed to the relatively small size of the dataset, where changes in certain hyperparameters do not significantly affect the model's performance. In future studies, increasing the sample size could provide more conclusive insights. A set of final optimal hyperparameter combinations for the four tree species and the generalized model is presented in Table 8.

Table 8. Optimal hyperparameter settings.

Hyperparameter	*P. tabuliformis*	*P. orientalis*	*P. massoniana*	*C. lanceolate*	Generalized model
activate	Relu	Relu	Relu	Relu	Relu
solver	Adam	Adam	Adam	Adam	Adam
alpha	0.0001	0.0001	0.0001	0.0001	0.0001
learning_rate	Constant	Constant	Constant	Constant	Constant

3.3. Analysis of Model Classification Performance

Table 9 presents the final evaluation results of different models. Among them, *Pinus massoniana* demonstrated outstanding performance, achieving the highest accuracy of 86.21%, while also exhibiting strong performance in Precision, Recall, and F1-score, reaching 85.24%, 84.94%, and 84.59%, respectively. *Pinus tabuliformis* achieved an accuracy of 68.42%, with a Precision of 72.22%, Recall of 69.05%, and F1-score of 69.86%, demonstrating a certain degree of classification capability. Both *Platycladus orientalis* and *Cunninghamia lanceolate* achieved accuracies approaching 80%, at 77.78% and 78.95%, respectively. Notably, *Platycladus orientalis* showed excellent performance in Precision, reaching 84.44%, while *Cunninghamia lanceolate* demonstrated superior performance in Recall, achieving 80.95%. The generalized model also achieved a respectable accuracy of 71.19%, with a Precision of 71.15%, Recall of 70.68%, and F1-score of 69.92%, indicating stable performance.

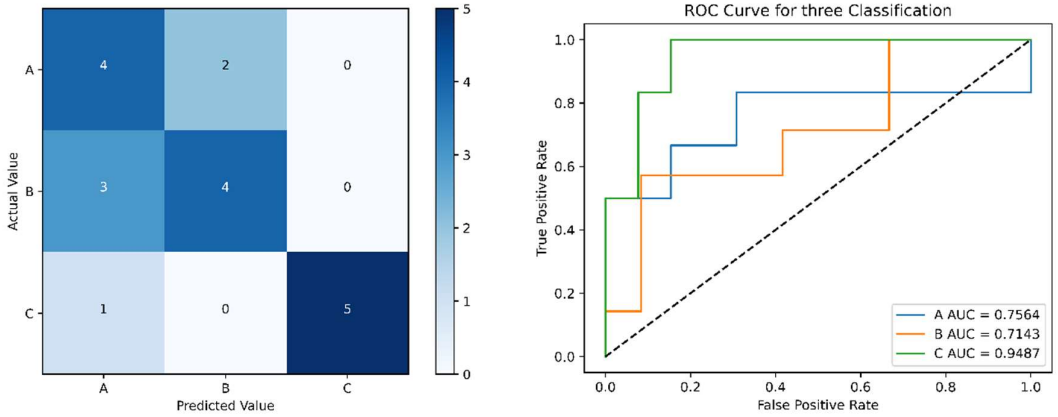
Table 9. Evaluation results.

Model	Precision (%)	Recall (%)	F1 (%)	Accuracy (%)
Pinus tabuliformis	72.22	69.05	69.86	68.42
Platycladus orientalis	84.44	74.44	74.28	77.78
Pinus massoniana	85.24	84.94	84.59	86.21
Cunninghamia lanceolate	79.05	80.95	78.57	78.95
Generalized model	71.15	70.68	69.92	71.19

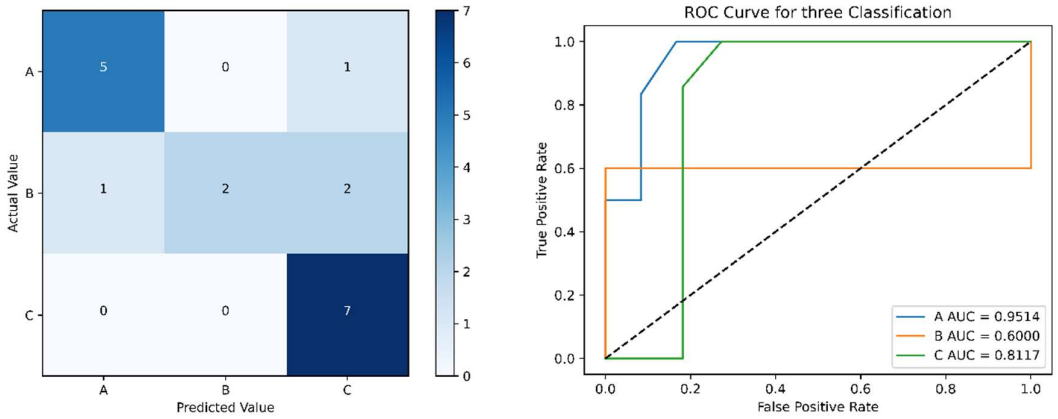
To provide a more intuitive understanding of the classification performance, confusion matrices and ROC curves (Figure 6) were generated for the BPNN models of each species. In Figure 6(a), the model accurately predicted the strong, medium, and weak growth categories for *Pinus tabuliformis* in 4, 4, and 5 cases, respectively. However, there were two instances of strong growth misclassified as medium, three medium misclassified as strong, and one weak case misclassified as strong. The ROC curves in Figures 6(a) illustrate the performance across the three growth categories. The AUC for growth categories A and B is slightly lower than that for category C. This suggests that in the *Pinus tabuliformis* model, trees with weak growth potential are more easily classified, while those with strong or medium growth potential tend to be more easily confused. This discrepancy may be attributed to the growth characteristics of *Pinus tabuliformis*, which has a slower growth rate. At 44

 doi:10.20944/preprints202501.1296.v1

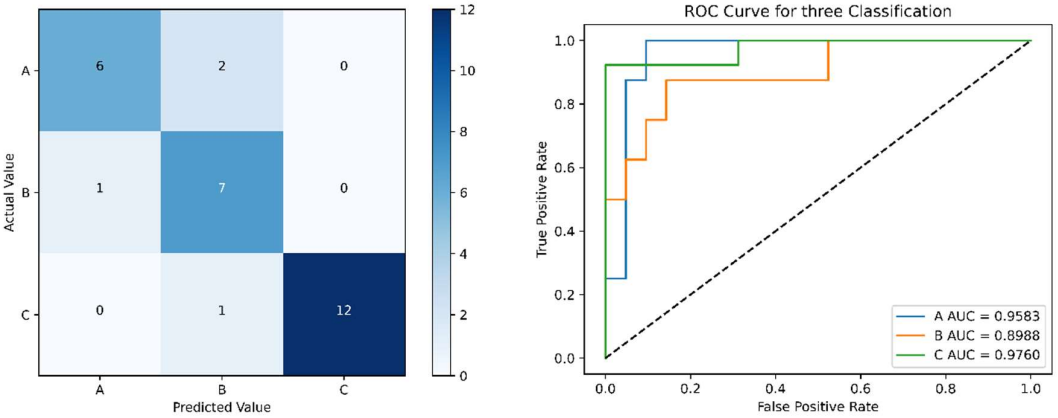
years of age, the tree exhibits relatively uniform annual ring widths in the outermost 2 cm, leading to more difficulty in distinguishing its features.



(**a**)



(b)



(c)

 doi:10.20944/preprints202501.1296.v1

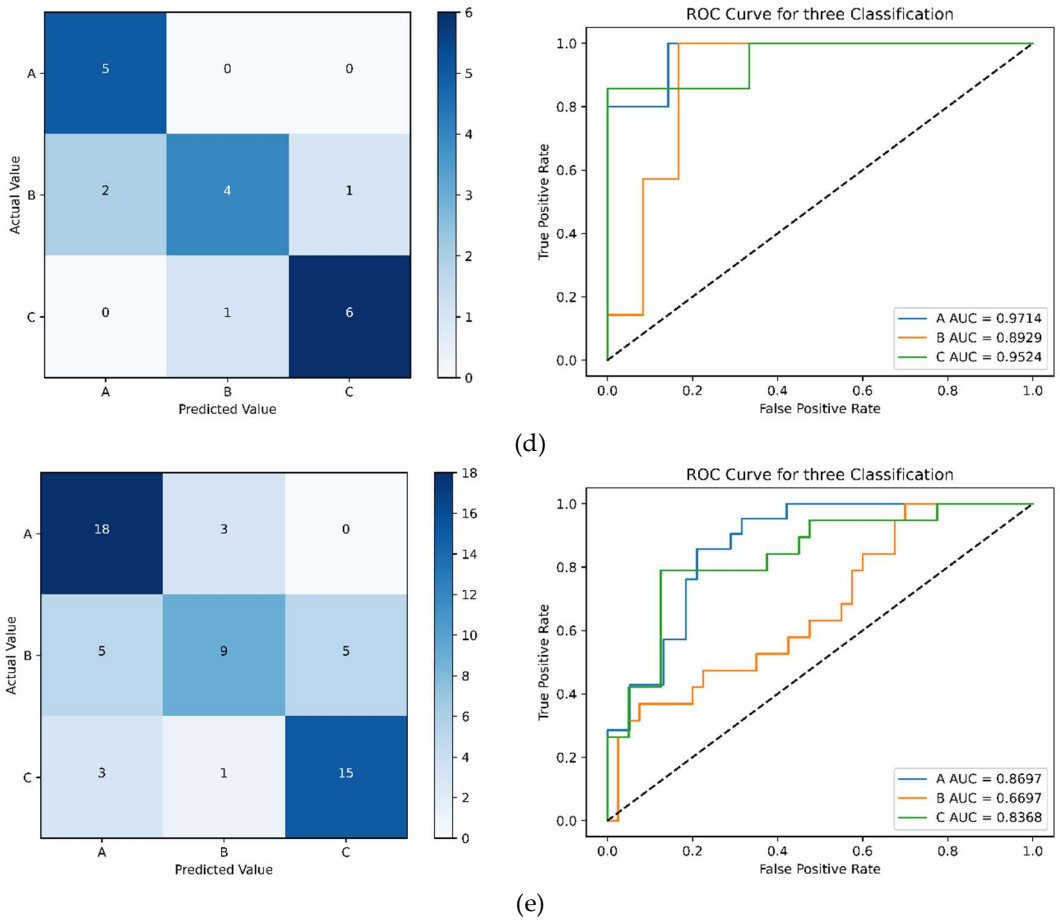


Figure 5. Confusion matrix and ROC curve for each model: (**a**) Confusion matrix and ROC curve of the *Pinus tabuliformis* model; (**b**) Confusion matrix and ROC curve of the *Platycladus orientalis* model; (**c**) Confusion matrix and ROC curve of the *Pinus massoniana* model; (**d**) Confusion matrix and ROC curve of the *Cunninghamia lanceolate* model; (**e**) Confusion matrix and ROC curve of the model for Generalized model.

Similarly, in Figure 6(c), *Pinus massoniana* demonstrated the best classification performance. The model correctly predicted the strong, medium, and weak growth categories for *Pinus massoniana* in 6, 7, and 12 cases, respectively. There were two strong cases misclassified as medium, one medium misclassified as strong, and only one weak case misclassified as medium. Moreover, the AUC for three growth potential is close to 1. These results suggest that while the model is generally effective at identifying tree growth categories, some misclassifications do occur.

It can be seen from the confusion matrices and ROC curves of each model that trees with medium growth potential are easily confused with those of strong and weak growth potential. However, the model can better distinguish between strong and weak growth potential. This indicates that plants with medium growth potential may share similar characteristics with plants of strong or weak growth potential, making it difficult for the model to differentiate them. After all, medium growth potential lies between the two.

4. Discussion

The traditional method of determining growth potential types mainly relied on expert experience and classification results are subjective and inaccurate. To enhance the accuracy of growth potential recognition, researchers have proposed various methods. Waring et al. suggested using the amount of trunk wood produced per square meter of leaf area as an indicator of tree growth potential [23]. However, measuring leaf area is challenging, and mathematical models for estimating it often

 doi:10.20944/preprints202501.1296.v1

yield significant errors, limiting the applicability of model. Hatch et al. proposed assessing growth potential through exposed crown surface area [24], but this approach only considers crown width, resulting in low estimation accuracy. Ding et al. utilized tree volume, incorporating tree height and DBH, as a measure of growth potential [2]. Nevertheless, this method fails to account for variations in site conditions and tree ages, leading to discrepancies in volume for trees with similar growth potential. In forestry management, it is crucial to develop plans tailored to different site conditions and tree ages, selecting trees with high growth potential as targets and harvesting those with lower potential. Consequently, distinguishing growth potential among trees with varying site conditions and ages remains challenging.

Artificial intelligence, which studies human intelligent activities to construct systems capable of performing tasks requiring human intelligence, offers a feasible solution to solving complex problems [25,26]. The BP algorithm consists of two primary processes: forward propagation of signals and error backpropagation. During forward propagation, input signals are transmitted through hidden layers to the output nodes, undergoing nonlinear transformations to generate the predicted output. If the predicted output deviates from the target output, the error is propagated backward through the network [27]. The error is distributed across the units of each layer and is used to adjust the weights of the units. Specifically, the adjustment involves modifying the connection strengths between input and hidden nodes, as well as between hidden and output nodes, along with the thresholds. These adjustments are made to minimize the error through gradient descent. Through iterative training, the optimal network parameters, including weights and thresholds, are determined [28]. Consequently, the BPNN possesses the capability for highly complex pattern classification and excellent multidimensional function mapping [29], making it one of the most widely used neural network models [30–32]. In this study, a growth potential prediction model was established using a BPNN, with tree height, DBH, and average crown width as input variables and growth potential type as the output variable. The test results showed that the accuracy of the sub models for *Pinus tabulaeformis*, *Platycladus orientalis*, *Pinus massoniana*, and *Cunninghamia lanceolate* were 68.42%, 77.78%, 86.21%, and 78.95%, respectively, and the accuracy of the generalized model was 71.19%. These findings suggested that employing BPNN is a viable approach for accurately estimating tree growth potential. However, the discrimination accuracy of some tree species was not high. In future research, the number of experimental trees can be increased to improve the discrimination accuracy of the model.

5. Conclusions

The growth potential of trees can be estimated by tree height, crown width and DBH. The higher the tree height, the larger the crown width and diameter at breast height, and the better the growth potential of the tree. The traditional method of determining growth potential types mainly relied on expert experience. The discrimination results were subjective and had low accuracy. This paper proposed a method for determining tree growth types based on BP neural network. The test results showed that the accuracy of the sub models for *Pinus tabulaeformis*, *Platycladus orientalis*, *Pinus massoniana*, and *Cunninghamia lanceolate* were 68.42%, 77.78%, 86.21%, and 78.95%, respectively, and the accuracy of the generalized model was 71.19%. These findings suggested that employing BPNN is a viable approach for accurately estimating tree growth potential.

Author Contributions: Conceptualization, J.Y. and C.Z.; methodology, X.H.; software, C.Z., Y.L; validation, Y.J., L.C. and Z.L.; formal analysis, Y.J.; investigation, F.L.; resources, F.L..; data curation, C.Z.; writing—original draft preparation, C.Z.; writing—review and editing, J.Y.; visualization, Z.L.; supervision, X.H.; project administration, Y.J.; funding acquisition, X.H. All authors have read and agreed to the published version of the manuscript.” All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Standardization Demonstration of Multi functional Management Technology for Miyun Reservoir Water Conservation Forest in Beijing (JING [2023]TG 06); the Natural Science Foundation of Henan Province (232300421167); Xinyang Academy of Ecological Research Open Foundation (2023XYQN04); Xinyang Academy of Ecological Research Open Foundation (2023XYZD02); Research on the multifunctionality and driving factors of oak forest ecosystems in Beijing (YZQN202405); Research on Key

Techniques for Quercus Planting (YZZD202407); and Postgraduate Education Reform and Quality Improvement Project of Henan Province (YJS2023SZ23).

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

References

1. Li , L.; Zou, S.; Yu, Z.; Wang, B.; Yang, L.; Zhou, L. Evaluation of growth potential of Ginkgo biloba and Viburnum vulgare in road green space in Chongqing[J]. *Hubei Agricultural Science*,**2020**,*59(08)*:111-115. DOI:10.14088/j.cnki.issn0439-8114.2020.08.025.
2. Ding, Y.; Lü, C.; Han, B.; Pu, H.; Wu, M. Relationship between tree growth potential, Monochamus alternatus population density and pine wood nematode disease severity[J]. *Chinese Journal of Applied Ecology*, **2001**, *(03)*: 351-354.
3. Zarzosa, P.S.; Herraiz, A.D.; Olmo, M.; Ruiz-Benito, P.; Barrón, V.; Bastias, C.C.; De La Riva, E.G.; Villar, R. (2021). Linking functional traits with tree growth and forest productivity in Quercus ilex forests along a climatic gradient. *Science of The Total Environment*, 786, 147468.
4. Larson, J.; Vigren, C.; Wallerman, J.; Ågren, A.M.; Appiah Mensah, A.; Laudon, H. Tree growth potential and its relationship with soil moisture conditions across a heterogeneous boreal forest landscape[J]. *Scientific Reports*, **2024**, *14(1)*: 10611.
5. Weis, A.E.; Simms, E.L.; Hochberg, M.E. Will plant vigor and tolerance be genetically correlated? Effects of intrinsic growth rate and self-limitation on regrowth. *Evolutionary Ecology 14*, 331–352 (**2000**). https://doi.org/10.1023/A:1010950932468.
6. Maes, S.; Perring, M.; Vanhellemont, M.; Depauw, L.; Bulcke, J.; Brūmelis, G.; Brunet, J.; Decocq, G.; Ouden, J.; Härdtle, W.; Hédl, R.; Heinken, T.; Heinrichs, S.; Jaroszewicz, B.; Kopecký, M.; Máliš, F.; Wulf, M.; Verheyen, K. (2018). Environmental drivers interactively affect individual tree growth across temperate European forests. *Global Change Biology*, *25*, 201 - 217. https://doi.org/10.1111/gcb.14493.
7. Steckel, M.; Moser, W.K.; del Río, M.; Pretzsch, H. Implications of Reduced Stand Density on Tree Growth and Drought Susceptibility: A Study of Three Species under Varying Climate. *Forests* **2020**, *11*, 627. https://doi.org/10.3390/f11060627
8. Trouvé, R.; Bontemps, J.; Seynave, I.; Collet, C.; Lebourgeois, F. (2015). Stand density, tree social status and water stress influence allocation in height and diameter growth of Quercus petraea (Liebl.). *Tree physiology*, *35* 10, 1035-46. https://doi.org/10.1093/treephys/tpv067.
9. Coomes, D.; Allen, R. (2007). Effects of size, competition and altitude on tree growth. *Journal of Ecology*, *95*. https://doi.org/10.1111/j.1365-2745.2007.01280.x.
10. Yang, H.; He, H.; Huang, L.; You, W.; Tang, W.; Zhu, G. (2024). Natural canopy model of Quercus hunanensis based on competition and site effect. *Journal of Central South University of Forestry & Technology*(06),92-101+155.doi:10.14067/j.cnki.1673-923x.2024.06.009.
11. Chaturvedi, R.; Raghubanshi, A.; Singh, J. (2011). Leaf attributes and tree growth in a tropical dry forest. *Journal of Vegetation Science*, *22*, 917-931. https://doi.org/10.1111/J.1654-1103.2011.01299.X.
12. Zhao, T. Effects of Shengguan jujube scion on root growth and physiological characteristics of sour jujube rootstock[D]. *Central South University of Forestry and Technology*, **2022**. DOI: 10.27662/d.cnki.gznlc.2022.000223.
13. Mao, J.; Zhao, H.; Yao, J.; Development and application of artificial neural networks[J]. *Electronic Design Engineering*, **2011**, *19(24)*: 62-65.
14. Yang, R.; Yang, Q.; Zeng L.; Chen, Y. Evaluation of rural land ecological security and analysis of influencing factors based on BP-ANN model--taking Fengdu County of Chongqing as an example[J]. *Research on Soil and Water Conservation*,**2017**,*24(03)*:206-213. DOI:10.13869/j.cnki.rswc.2017.03.037.
15. Wang, H.; LI, W.; Niu, J.; Liu D. An improved land price evaluation method based on artificial neural network and fuzzy mathematics[J]. *Journal of Xinyang Normal University (Natural Science Edition)*, **2020**, *33(1)*: 76-82.
16. Ren, S.; Chen, Z.; Deng, X.; Fan, Y.; Sun, A. A preliminary study of vascular sclerosis identification based on wavelet scattering neural network[J]. *Journal of Biomedical Engineering*,**2023**,*40(02)*:244-248.

17. Deng, X.; Liu, Q.; Deng, Y.; Mahadevan, S. An improved method to construct basic probability assignment based on the confusion matrix for classification problem[J]. *Information Sciences*,**2016**,340-341250-261.
18. Zhang, S.; Xie, X.; Xu, Y. Intrusion detection method based on dCNN. *Journal of Tsinghua University (Natural Science Edition)*, **2019**, *59(1)*: 44-52.
19. Liu, S.; Cao, J.; Sun, T.; Hu, J.; Fu, Y.; Zhang, S.; Li, S. Inverse kinematics analysis of redundant manipulator based on BP neural network[J]. *China Mechanical Engineering*, **2019**, *30(24)*: 2974-2977.
20. Yao, J.; Wu, Z.; Hu, X.; Sun, Y.; Tian, W.; Lu, Y.; Li, X. Multi-feature backpropagation artificial neural network method for micro-drill resistance tree-ring identification. *Journal of Xinyang Normal University (Natural Science Edition)*, **2024**, *37(4)*: 460-469.
21. Fu, S.; Tang, F.; Hou, L.; Shi, N.; Jin, Y.; Zhang, X. Research on compressive strength prediction of fly ash concrete based on grid search and support vector regression. *Journal of Jiangsu University of Science and Technology (Natural Science Edition)*, **2024**, *38(02)*: 73-79. DOI:10.20061/j.issn.1673-4807.2024.02.012.
22. Wu, J.; Chen, S.; Chen, X.; Zhou, R. Model selection and hyperparameter optimization based on reinforcement learning. *Journal of University of Electronic Science and Technology of China*, **2020**, *49(02)*: 255-261.
23. Waring , R.H.; Thies, W.G.; Muscafo, D.; Wang, Y.M. Trunk growth per unit of leaf area - a measure of tree growth potential[J]. *Shaanxi Forestry Science and Technology*,**1980**,*(05)*:76-79.
24. Hatch, C.R; Gerrard, D.J; Tappeiner II, J.C. 1975. Exposed Crown Surface Area: A Mathematical Index of Individual Tree Growth Potential. *Canadian Journal of Forest Research*. *5(2)*: 224-228. https://doi.org/10.1139/x75-030
25. Wang, H.; Fu, T.; Du, Y; Gao, W.; Huang, K.; Liu, Z.; Chandak, P.; Liu, S.; Van Katwyk, P.; Deac, A.; Anandkumar, A.; Bergen, K.; Gomes, C. P.; Ho, S.; Kohli, P.; Lasenby, J.; Leskovec, J.; Liu, T.; Manrai, A.; Marks, D.; Ramsundar, B.; Song, L.; Sun, J.; Tang, J.; Veličković, P.; Welling, M.; Zhang, L.; Coley, C. W.; Bengio, Y.; Zitnik, M. Scientific discovery in the age of artificial intelligence. *Nature* **620**, 47–60 (2023). https://doi.org/10.1038/s41586-023-06221-2.
26. Ofosu-Ampong, K. Artificial intelligence research: A review on dominant themes, methods, frameworks and future research directions[J]. *Telematics and Informatics Reports*, **2024**: 100127.
27. Liu, Y.; Jiang, C.; Lu, C.; Wang, Z.; Che, W. Increasing the Accuracy of Soil Nutrient Prediction by Improving Genetic Algorithm Backpropagation Neural Networks. *Symmetry* **2023**, *15*, 151. https://doi.org/10.3390/sym15010151.
28. Xiao, M.; Luo, R.; Chen, Y.; Chen, Y.; Ge, X. Prediction model of asphalt pavement functional and structural performance using PSO-BPNN algorithm[J]. *Construction and Building Materials*, **2023**, *407*: 133534.
29. Wang, K.; Huang, H.; Deng, J.; Zhang, Y.; Wang, Q. (2024). A spatio-temporal temperature prediction model for coal spontaneous combustion based on back propagation neural network. *Energy*, *294*, 130824.
30. Anwaier, G.; You, L.; Ye, G.; Nie, S.;Xu, Z.; Chen, F. Construction of Aboveground Biomass Estimation Model for Different Age Groups of Casuarina Equisetifolia [J/OL]. *Journal of Forest and Environmental Science*, 1-11 [2024-11-14]. http://kns.cnki.net/kcms/detail/35.1327.S.20241014.1412.004.html.
31. Gong, R.; Zhang, H.; Lu, X.; Wan, H.; Zhang, Y.; Luo, X.; Zhang, J.; Xie, R. An integrated UAV growth monitoring model of Cinnamomum camphora based on whale optimization algorithm.[J].*PloS one*,**2024**,*19(6)*:e0299362-e0299362.
32. Zheng, S.; Gao, P.; Zhou, Y.; Wu, Z.; Wan, L.; Hu, F.; Wang, W.; Zou, X.; Chen, S. An Accurate Forest Fire Recognition Method Based on Improved BPNN and IoT. *Remote Sens.* **2023**, *15*, 2365. https://doi.org/10.3390/rs15092365.