J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 6439--6455

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

The Design of Decision Tree Algorithm Assisted Teaching Activities in Preschool Education

Chunyan Xu[1,✉], Peifeng Sun[2]

[1] School of Preschool Education, Zhengzhou Preschool Education College, Zhengzhou, Henan, 450000, China

[2] School of Science and Technology, Zhengzhou Preschool Education College, Zhengzhou, Henan, 450000, China

ABSTRACT

This paper introduces the decision tree algorithm into the field of preschool education to categorize the styles of children in preschool education. The learning activities of children with different styles are deeply analyzed by the total number of detections, the task score and the total game time in the small train by counting activity. Decision tree algorithm is utilized to integrate online preschool education resources and used in practice so as to assist teaching. The teaching experiment method is used to test its educational effect. Kindergarten children were categorized into 3 types: extroverted, negative emotional and effortful control children. Effort-control style children performed well in play detection behavior, play task score and total play time. In the teaching experiment, children in the experimental group obtained very significant improvements in small muscle activity, art, music and rhythm, blocks, natural science and mathematical thinking, while the control group also improved, but their changes were not significant. Decision tree algorithm has better results in assisting preschool education.

Keywords: decision tree algorithm, style classification, preschool education, resource integration, assisted instruction

1. Introduction

The key to preschool education is to realize goal-oriented learning activities through children's interests and experiences. Information technology provides many types of educational and teaching resources for preschool education, stimulates children's interest in learning and motivates them to take the initiative in learning [1-2]. In this process, information technology has become a cognitive tool for preschool education, or an important tool to assist teachers in the teaching process [3-4]. At present, preschool education information technology-related resources are relatively poor, and the construction of digital teaching is also hindered [5]. Although there are some preschool education

✉ Corresponding author.
E-mail address: xuchunyan@zzpec.edu.cn (C. Xu).

Received 01 February 2024; Revised 20 April 2024; Accepted 10 November 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-353

© 2025 The Author(s). Published by Combinatorial Press. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

resource libraries, ECE theme forums, classroom materials in various fields, and preschool education websites constructed by provincial and municipal education departments, there are fewer resources for special teaching of young children, and the related learning platform for young children is not mature enough in China, and there is a lack of software or platforms for integrating the various resources for teaching young children [6-9]. In this context, the establishment of an online education and resource integration teaching platform for preschool children has become a landmark event in preschool education.

Educated preschoolers can most directly reflect whether the online learning activities are effective or not, teachers are mainly responsible for the development of teaching goals and plans, a series of implementation and supervision of teaching activities [10-12]. The technical support staff is mainly the organizer and builder of the online teaching resource base, to ensure that the teaching process to achieve a good interaction between students and teachers [13-14]. Based on the background of "Internet +", the introduction of intelligent algorithms to assist in the design of teaching activities, taking into account the diversified integration and optimization process of learning resources, making the curriculum resources of the entire teaching resource platform more scientific and reasonable [15-17]. By classifying the resources digitized in preschool education, the scope of development and integration of curriculum resources is broadened, and the resource integration model is designed [18-19]. In this link, the construction of resource environment emphasizing preschool children's learning needs, the content search of teaching resources completed in the teaching platform according to the target course, and the feedback about the needs of multimedia production of teaching resources, so that the integrated teaching resources can practically meet the learning needs of preschool children [20-23].

Practical research on information technology-assisted preschool education teaching activity design at home and abroad is relatively rich, and it is generally believed that information technology has a positive impact on preschool children's language, socialization, and artistic development. Literature [24] proposes a knowledge adaptive education network based on adaptive learning path strategy, which is applied to the preschool education resource contribution and recommendation platform, and its learning path generation and teaching resource integration ability fully meets the highly personalized learning needs of preschool education, and literature [25] shows that the intelligent recommendation system of learning resources can assign matching learning content according to children's cognitive ability, and proposes a learning path generation and teaching resource integration system based on the Hadoop big data framework and machine learning algorithms for preschool education resource recommendation system, which has high practical value in forming educational activities for preschool children. Literature [26] emphasizes the need for personalized innovation in teaching strategies for preschool children to provide a good learning experience, and proposes a fuzzy decision support system to evaluate students' abilities and teaching models to make preschool education more adaptable and feasible. Literature [27] constructed an ontology model based on deep learning algorithms to describe the semantic information of preschool curriculum resources, and improved the efficiency of preschool curriculum resources classification and semantic query through semantic annotation, which fully meets the children's independent learning needs. Literature [28] examined the application of digital education technology in the field of preschool education, and developed a convolutional neural network-driven educational robot, which provides preschool children with scientific and reasonable teaching methods and resources, and promotes the rapid improvement of their learning ability. Literature [29] explored a preschool education resource allocation scheme supported by a resource allocation optimization algorithm, whose automated decision-making effectively improves the resource management efficiency and scale of operation, and realizes the high-quality development of preschool education activities. Literature [30] designed an educational network resource recommendation system based on the decision tree algorithm, which realizes the accurate recommendation of early childhood education network resources by adjusting the effective information of the network resources and modeling the user's interest, which improves the effectiveness of the educational process.

The author tries to embed the decision tree algorithm in preschool education, and categorize preschool children through the decision tree algorithm. Children are categorized into different styles based on their learning activity data. The characteristics shown by children of different styles in learning activities were explored through the game of Little Train Fetch by Number. Then the decision tree algorithm is used to integrate online preschool education resources, build a digital education resources integration model, and put it into practical application. Through teaching experiments, we compare the differences between the experimental group and the control group in terms of small muscle activities, art, music and rhythm, blocks, natural science and mathematical thinking, so as to verify the effectiveness of the decision tree algorithm in assisting preschool education.

2. Decision tree algorithm based classification of preschool children

2.1. Pre-school education

Pre-school education is one of the important parts of the education field [31], and there are two main definitions of pre-school education as follows. The Law of the People's Republic of China on Pre-school Education (Draft) defines pre-school education as: pre-school education refers to the care and education of pre-school children between the age of three and the age of entry into elementary school by pre-school education institutions such as kindergartens, and a three-year pre-school education system is practiced in China. The Dictionary of Education defines preschool education as the education of children from birth until they enter school. It can be seen that the two definitions do not cover the same age groups; one believes that preschool education covers 0-6 years old, while the other viewpoint believes that preschool education only covers 3-6 years old. In this study, the education received by children aged 3-6 in preschool institutions such as kindergartens is selected as preschool education, which is mainly a planned, organized, and purposeful formal educational activity provided by the government, the market, social organizations, and other main bodies for children aged 3-6. Pre-school education provides children with opportunities for all-round development in language, socialization, logical thinking, cognition, emotion and thinking. This not only helps to improve the overall quality of children, but also lays a solid foundation for their future learning and life.

2.2. Decision Tree Algorithm

The data mining process is often accompanied by corresponding data mining methods, which is an important basis for carrying out data mining work. With the development of data mining, data mining methods are also continuously improved and optimized, at present, the methods commonly used in data mining work are mainly clustering analysis, association rule analysis, regression analysis, classification methods of the four major categories, and at the same time in the use of a certain type of data mining methods at the same time, the researcher will be based on the data mining scenarios and purposes to select the appropriate algorithm as a technical support, through the algorithm to dig deep into the data hidden information value and construct the corresponding data model in the data to provide users with the use of the data. The algorithm digs deep into the hidden information value of the data and builds the corresponding data model in the data to provide users with the use of the data.

Decision tree belongs to a branch of categorical data mining methods [32], which has a simple structure, is easy to understand and interpret, can visualize and analyze the data, and can make feasible and effective results for large data sources in a relatively short time. Therefore, this study combines the research scenario and the research purpose into a comprehensive consideration, and finally the decision tree analysis method is used as the main method for the data mining process in this study.

2.2.1. Overview of decision trees. Decision tree is a method in machine learning with a tree-like structure that approximates a discrete-value structure, in which each internal node represents a judgment on an attribute, each branch represents the output of the result of each judgment, and finally, each leaf node represents the result of a classification. The conceptual diagram of the decision tree is shown in Fig. 1, where rectangles are used to represent the internal nodes and the root node of the decision tree, and ellipses are represented as the leaf nodes of the decision tree, and the structure as a whole is presented as a tree-like structure, where A and B represent different judgmental attributes, a1, a2, b1, and b2 represent the decision-making conditions, and d1, d2, and d3 represent the results of the classification. The root node has only outgoing edges and no incoming edges. Internal nodes have one incoming edge and at least two outgoing edges; leaf nodes have only one incoming edge and no outgoing edges.

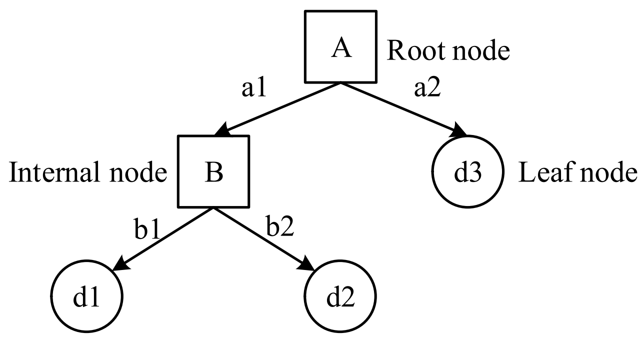


Fig. 1. Decision tree concept diagram

Decision tree is a data mining method based on real cases, he will be a set of messy, irregular data in accordance with certain attributes and conditions for the construction of the decision tree model, in the whole process, the use of top-down recursion, starting from the root node, down the branch to reach the leaf nodes, in which each path is a classification rule, the combination of all the rules, constitutes a classifier.

2.2.2. Decision Tree Correlation Algorithm:

1) Information Entropy. Information entropy refers to the measure of uncertainty of the random variable Y = {c1, c2, c3, ..., ck}, the larger the information entropy, the larger its uncertainty, in the categorization category, that is to say, the more information contained, the larger the entropy, the way to calculate is shown in the following equation (1), X represents a certain category, p(x) represents the probability of choosing the category, H(X) represents the sum of all categories, if there is only one category, p(x)=1, the entropy is 0, the smaller the entropy means the higher the certainty:

$$(X) = -\sum_{i=1}^{n} p(x_i)\log p\left(x_i\right) \tag{1}$$

2) Information Gain. Information gain indicates the difference in information entropy before and after the data set is divided according to the feature T attribute. When dividing the dataset, the method of obtaining the maximum information gain should be selected, and the larger the information gain, the stronger the classification ability.

The specific calculation is shown in equation (2) below:

$$InformationGain(T) = Entropy(S) - Entropy(S/T) \tag{2}$$

3) Information gain rate. C4.5 adopts the information gain rate as a branching criterion for selecting branching attributes for the decision tree, which indicates the ratio of useful information generated by the branch, the larger the value, the more useful information the branch contains, and its expression

As shown in equation (3) below, Gain is information gain and H(S, A) is information entropy:

$$ainRatio(S,A)=\frac{Gain(S,A)}{H(S,A)} \tag{3}$$

4) Gini coefficient. The Gini coefficient is used to express the uncertainty of the data, the larger the Gini coefficient, the greater the uncertainty of the sample set. Specifically in the classification problem, set there are K categories, the probability of the Kth category is pk, its Gini coefficient expression is shown in the following equation (4).

$$Gini(p)=\sum_{k=1}^{k} pk(1-pk)=1-\sum_{k=1}^{k} p^2k \tag{4}$$

If it is a two-class classification problem, the calculation will be relatively simple, if the probability of belonging to the first sample output is p, the Gini coefficient is expressed as the following equation (5):

$$Gini(p)=2p(1-p) \tag{5}$$

2.2.3. Decision Tree Generation Process. Most of the decision tree induction algorithms use a greedy algorithm, which is utilized to recursively classify the initially processed data, and can generate rules and tree-structured model diagrams that are easy for the user to interpret, thus enabling the prediction of new data and the analysis of data classification. The essential process of decision tree is to utilize the known data, generate the decision tree model and related rules, and use the generated rules to parse the test data for prediction and classification.

2.3. Classification of children in pre-school education

The construction process of decision tree includes two phases: growth and pruning. The growth process of the decision tree is the building process, which divides the attribute values according to certain criteria and gradually builds the decision tree from top to bottom. The essence of pruning is to eliminate noise or noise-like abnormal data. The two most commonly used pruning methods for decision trees are first pruning and second pruning. In the first pruning method, the tree can be "pruned" by establishing rules to limit the growth of the tree and stopping the construction of the tree in advance. Post-pruning method is more commonly used, which is to prune the decision tree after it is fully grown, and there are two ways of pruning: replacing the subtree with a new leaf node, whose class label is determined by the majority of the classes recorded under the subtree; and replacing the subtree with the most frequently used branch in the subtree. The specific construction process of the decision tree mainly consists of five steps: determining the mining target and object, data collection, data preprocessing, data classification mining, and generating classification rules.

2.3.1. Determination of excavation objectives and targets. This study was conducted on the second semester of Kindergarten T's teaching and learning activities in the year 2023-2024, and the study was based on the data and information of all the students who participated in the learning activities. The study proposes research presuppositions based on the analysis of the collected data and information that the gender, achievement data, and learning styles of the students will have an impact on the students' performance and further explores what kind of impact it will have, and then proposes to suggest rationalizations for teaching to improve the quality of teaching and learning.

2.3.2. Data pre-processing. The collection and data preprocessing of students' basic information and learning activity data. The purpose of this step of data preprocessing is to preprocess some of the incomplete, inconsistent, and noise-containing data in the collected data, so as to improve the quality of the data mining object, and also to improve the accuracy of the research results. Data preprocessing includes four parts: data integration, data cleaning, data conversion, and data normalization.

2.3.3. Data classification mining:

1) Selection of algorithm. Combined with the research object of this study, i.e., the characteristic that students' performance data contain both continuous and discrete data, the C4.5 algorithm in the decision tree algorithm is selected as the method of this study. As an improved algorithm of ID3 algorithm, C4.5 algorithm uses the information gain rate to select attributes [33], avoiding the situation of tendency to select the attributes with the most values when selecting attributes with the information gain in ID3 algorithm, and at the same time it can deal with the residual data, and in the pruning of the decision tree C4.5 algorithm also makes corresponding improvements, and it can be pruned during the process of building the tree.

2) Establish decision tree model. This study uses the C4.5 algorithm in the decision tree algorithm, and the flow of the C4.5 algorithm to establish the model is shown in Figure 2.

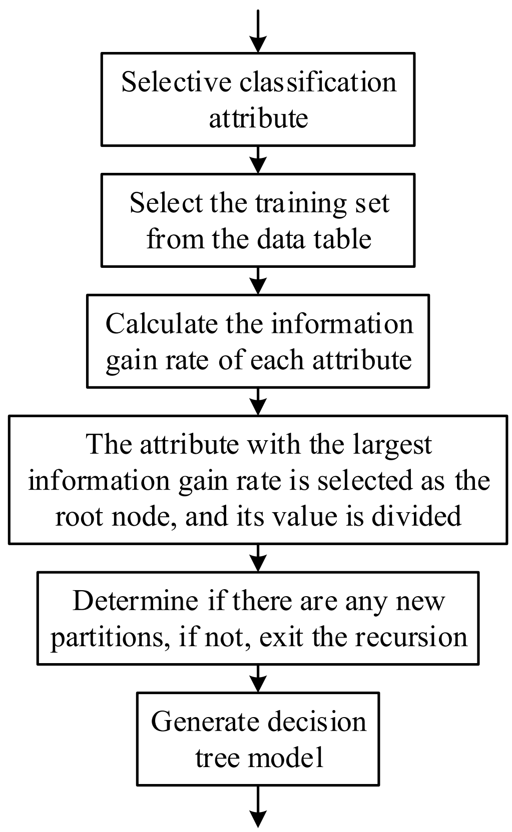


Fig. 2. Flow chart of C4.5 model

As an improvement of ID3 algorithm, C4.5 algorithm adopts the information gain rate as a new metric to overcome the deficiency of ID3 algorithm in selecting the metric by information gain as an attribute, and at the same time, “Split Information” has been added in C4.5 algorithm to optimize the algorithm. The specific operation procedure of the information gain rate is as follows:

Let data partition D be the training set of labeled class tuples. Assume that the class labeling attribute has m distinct values, and thus define m distinct classes $C_i(i=1,2,3,\ldots,m)$. Let $C_{i.D}$ be the set of tuples of class C_i in D. $|D|$ and $|C_{i.D}|$ are the number of tuples in D and $C_{i,D}$, respectively. The amount of information required to classify a given sample is given by the following equation:

$$Info(D) = -\sum pi \log 2(p_i) \tag{6}$$

where the non-zero probability that any tuple in the sample set D belongs to the class C_i is denoted by p_i, which is computed as $|C_{i.D}|/|D|$. Since information is represented in binary encoding therefore a logarithmic function with base 2 is used in the computation. Info(D) denotes the average amount of information required to identify the class labeling of the tuples in the sample set D. Info(D) is also referred to as the entropy of D.

Suppose that the tuples in the sample set D are divided according to a certain attribute A, where there are a total of v different values $\{a_1, a_2, \ldots a_n\}$ with attribute A in the training set. If the data are discrete values, these values correspond to the v outputs of the test on A. The attribute A is used to partition the sample set D into a total of v subsets of $\{D_1, D_2, \ldots, Dv\}$, where D_j contains a certain set of tuples in D that have the value of a_j for A. These partitions correspond to the branches growing from node N. The information entropy of partitioning into subsets according to attribute A at this point is measured by the following equation:

$$Info_A(D) = \sum_{j=1}^{v} \frac{|D_j|}{|D|} \times Info(D_j) \tag{7}$$

The term $\frac{|D_i|}{|D|}$ serves as the weight of the j th partition. The $Info_A(D)$ is the expected information needed to categorize the tuple of D based on partitioning by A attributes. The smaller the desired information needed, the higher the purity of the partition.

The information gain is defined as the difference between the original information requirement (based on class proportions only) and the new information requirement (after partitioning of A). Its calculation formula is as follows:

$$Gain(A) = Info(D) - Info_A(D) \tag{8}$$

The C4.5 algorithm employs splitting information similar to Info(D), which is operationally formulated and defined as follows:

$$SplitInfo_A(D) = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} \times \log_2\left(\frac{|D_j|}{|D|}\right) \tag{9}$$

The value of split information volume represents the information generated by dividing the training dataset D into v partitions corresponding to v outputs of the attribute A test.

The operational formula for the information gain rate is shown below:

$$Gainratio(A) = \frac{Gain(A)}{SplitInfo_A(D)} \tag{10}$$

2.3.4. Generating classification rules. The decision tree algorithm uses rules of the if-then form, and the expression for an if-then rule is if condition then conclusion. The practice of extracting if-then rules is as follows: each path in the decision tree from the root node to the leaf node produces a classification rule, the test conditions in the path form the ensemble of the rule antecedent, i.e., the if part of the rule, and the class labeling of the leaf node is assigned to the rule antecedent, i.e., the then part of the rule.

3. Analysis of learning activities of children with different styles

A total of 50 children, 32 boys and 18 girls, were selected to participate in the learning activity (the train counting activity) in the primary and secondary classes of Kindergarten T. The average age of all the children was about 42.58±2.48 months. The evaluation of the children's behavioral performance was done by the teacher through the “train counting” activity, in order to facilitate the acquisition and analysis of observational data and to ensure the accuracy and reasonableness of the data extraction and analysis perspectives. The study combined the children's style scores (sample selection criteria: children who scored the highest in the three dimensions of learning styles were considered to be children with high style scores for that dimension.) As well as the classroom teacher's conversation, 20 children (numbered A1~A20) with more distinctive characteristics of the three dimensions of style scores were identified as the research subjects. The decision tree-based preschool children classification model of this paper was utilized to classify the study subjects. The distribution of children with different styles is shown in Table 1.

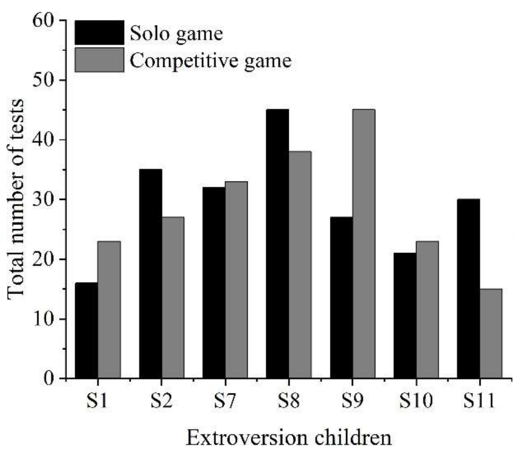
Table 1. Children gender, monthly age and style score (N=20)

Type	Number	Serial
Extroversion	7	A1, A2, A7, A8, A9, A10, A11
Negative emotion	6	A3, A4, A12, A13, A14, A15
Effort control	7	A5, A6, A16, A17, A18, A19, A20

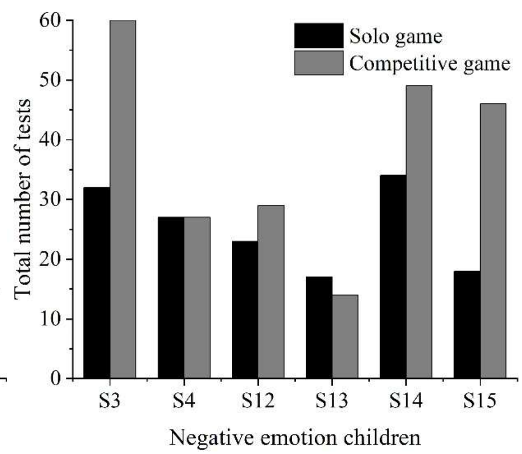
Firstly, descriptive statistics were analyzed for four indicators: the total number of adjustments, the total number of detections, the small train task score (including the selection of card types), and the total length of the game for 20 children of three styles in solo and competitive games; secondly, two representative samples of each of the three styles were selected for a total of six cases to be observed and analyzed for the behavioral performance in the two game situations to extract the differences of the children's different behavioral performances. The three indicators of small train task score, total number of tests, and total game length were analyzed.

3.1. Total number of detections

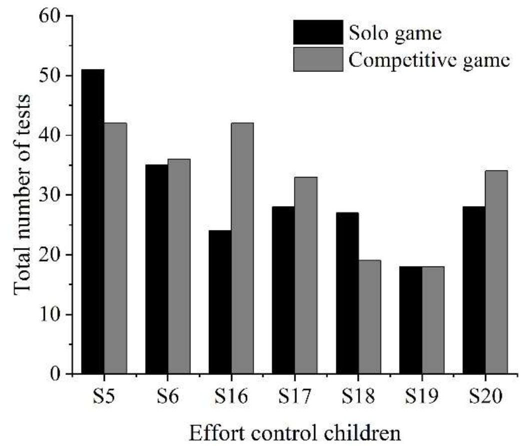
The total number of detections by children in different game situations is shown in Figure 3, (a) to (c) are the total number of detections by extroverted, negative emotional, and effortful control children in separate and competitive games, respectively. From Figure 3, it can be seen that the number of detections in separate and competitive games was different for different styles, with the number of detections in separate games (206) being comparable to the number of detections in competitive games (204) for extroverted children, but the number of detections in separate games (151) was significantly less than the number of detections in competitive games (225) for children with negative emotional style, and the number of detections in competitive games (225) was significantly less for children with effortful control style. significantly more detections (224) than in separate play (211), suggesting that children with negative mood and children with effortful control styles had more competitive play detections than separate play detections, but the magnitude of change in the number of detections for children with effortful control styles was smaller than the magnitude of change in the number of detections for children with negative moods. The researchers found that children with negative emotions often had the situation of "pressing the detection button multiple times without wanting to make adjustments" in competitive games, while children with control styles were able to control their own detection behaviors in the game, reduce the number of invalid "detections", and improve the efficiency of the game.



(a) Extroversion children



(b) Negative emotion children

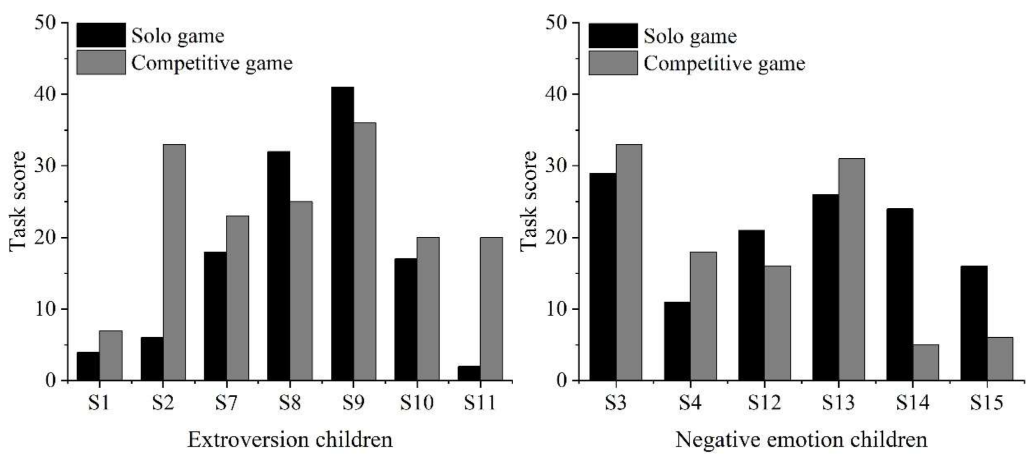


(c) Effort control children

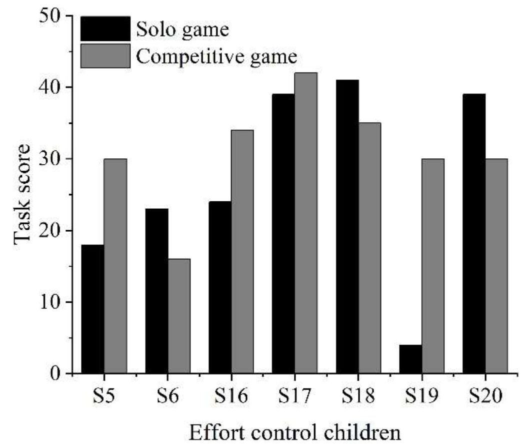
Fig. 3. The test number of children in different game scenarios

3.2. Small Train Task Score

The task scores of the Little Train game in different game situations are shown in Figure 4, (a) to (c) show the task scores of Extraversion, Negative Emotionality and Effortful Control children in solo and competitive games, respectively. From Fig. 4, we can see the changes in the task scores of the miniature train game for children with different styles in different game situations. Comparing the changes in the task scores of children with different styles in solo and competitive games, it was found that the task scores of solo game for children with extroverted styles (120) < the task scores of competitive games (164), the task scores of solo game for children with negative emotional styles (127) > the task scores of competitive games (109), and effortful control style children's separate game task scores (188) < competitive game task scores (217), which shows that extroverted children's task scores are higher than separate game task scores in competitive games, and the opposite is true for negative emotional children, suggesting that competitive game situations can increase extroverted style and effortful control children's game task scores.



(a) Extroversion children (b) Negative emotion children

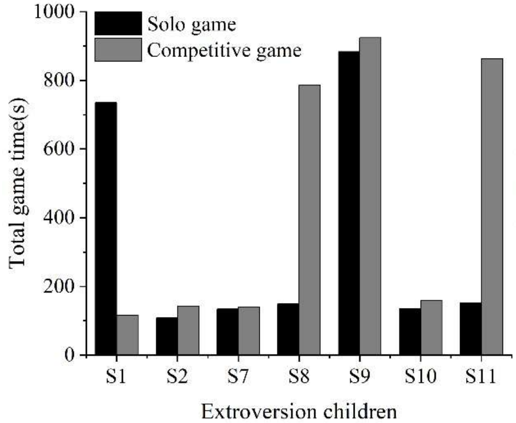


(c) Effort control children

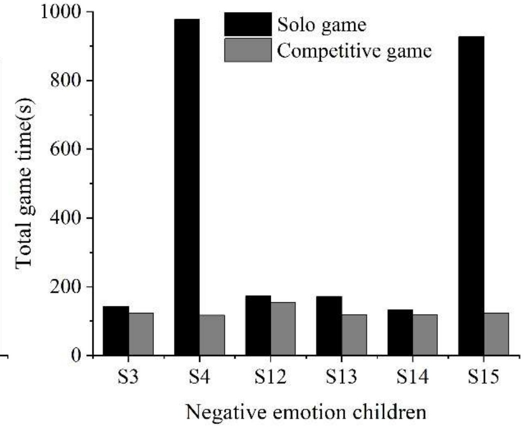
Fig. 4. The task score of children in different game scenarios

3.3. Total game time

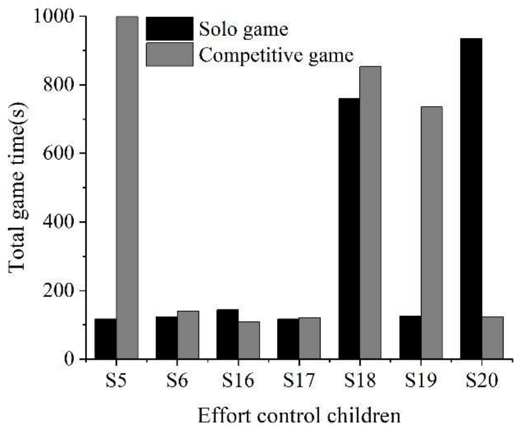
Changes in total play time of children with different styles in different play situations are shown in Figure 5. (a)~(c) shows the total time of play in solitary and competitive play situations for Extraversion, Negative Emotion, and Effortful Control children, respectively. From Figure 5, it can be seen that the three different styles of children in both solo and competitive game situations changed as follows, the total time of solo game of extrovert style children (2297 seconds) was lower than the total time of competitive game (3131 seconds), which indicates that extrovert style children used more total time in competitive game, the researcher found during game observation that extrovert children increased the researcher found during the game observation that extroverted children increased the number and time of communication and interaction with their peers, which may be one of the reasons that caused the total time spent in the competitive game of extroverted children to be more than the total time spent in the solo game. When comparing the duration of the negative emotional style children's games, it was found that the total duration of the solitary game (2525 seconds) was higher than that of the competitive game (754 seconds). The researcher found that the negative emotional style children showed refusal behaviors of “not wanting to continue with the game”, which was reflected in the time spent on selecting the task cards, adjusting the number of fruits, and taking more time in choosing the task cards, Adjusting the number of fruits took more time, so children with negative affective styles may have experienced longer game lengths in competitive games as a result. Comparison of children with effort control style showed that the total time spent playing alone (2204 seconds) was higher than the total time spent playing competitively (2083 seconds), which indicates that children with effort control style showed more concentration and speed in completing the task cards of the train in competitive games, which suggests that children with effort control style can stimulate children's motivation and interest in playing in competitive games.



(a) Extroversion children



(b) Negative emotion children



(c) Effort control children

Fig. 5. Total game time of children in different game scenarios

4. Decision Tree Algorithm Based Online Preschool Resource Integration

4.1. Framework Design of Digital Educational Resources Integration Model for Preschool Education

The process of diversified integration of online teaching resources for preschool education is mainly based on four elements: preschool children, teachers, parents and the content of educational resources. From the perspective of distance learning, the purpose of effectively integrating digital teaching resources should be achieved by integrating remote digital educational resources and original teaching resources in the teaching resource platform. The main behavioral participants in the online learning model are students and teachers, and the main focus is on the diversified integration of teaching resources, so the relative openness of resources is the key to integration.

The educated preschool children can most directly reflect whether the online learning activities are effective or not, the teachers are mainly responsible for formulating teaching goals and plans, and carrying out a series of implementation and supervision of teaching activities; while the technical support staff are mainly the organizers and builders of the online teaching resource library, ensuring good interaction between students and teachers.

In the method of integrating digital educational resources in preschool education designed in this paper, it is mainly based on the background of “Internet+”, taking the online learning model as the basis, classifying the digital resources in preschool education, broadening the scope of development and integration of curriculum resources, designing the resource integration model, emphasizing the construction of the resource environment for preschool children's learning needs, and building the resource environment according to the target curriculum in the teaching platform. The model is based on the teaching platform, which is used to search for teaching resources according to the target courses and provide feedback on the demand for multimedia production of teaching resources, so that the integrated teaching resources can effectively meet the learning needs of preschool children. In addition, the model can also take into account the diversified integration and optimization process of learning resources, making the curriculum resources of the whole teaching resource platform more scientific and reasonable.

4.2. Introduction of a decision tree to complete the integration of resources

The categorized real resource data may have some fields mutilated or incorrect when noise is not taken into account. In the state of high noise, it will interfere with the analysis effect, so that it is necessary to use a decision tree and prune it to make it more easy to use.

Firstly, a decision tree is constructed according to the requirements of resource integration, and then measurement is performed in a classification node to complete the functional evaluation of the node. In the process of measurement and evaluation, the obtained test result with strong function is taken as the optimal node for resource integration, through which classification is carried out to integrate similar resources. In this process, the information gain rate needs to be calculated. In this paper, the information gain rate function is used in the evaluation process, which has the following form:

$$\begin{cases} gainratio = \dfrac{gain(A)}{spliti(A)} \\ spliti(A) = \sum_{1}^{v} \dfrac{p_i}{m} \log\left(\dfrac{p_i}{m}\right) \end{cases} \tag{11}$$

Where $gain(A)$ denotes the information gain, utilizing A as the attribute of the decision tree, v denotes the number of discrete values of A, and p_i denotes the number of positive examples. For the digital educational resources of preschool education, all of them have their fixed IP locations. Considering the factors such as running cost and using download speed, it is necessary to select the resource platforms in the same LAN or within a certain range as the integration environment.

5. Analysis of the effectiveness of online preschool education

5.1. Between-group variation in the effectiveness of children's preschool education

Incorporating the decision tree teaching resource integration method into the preschool education of 20 children selected from kindergarten T. These children were set up as an experimental group and compared with the children who were taught according to the conventional education method (control group). The effectiveness of this paper's online preschool education resource integration method based on decision trees is analyzed. The experiments were conducted in six areas: small muscle activities, art, music and rhythm, blocks, natural science and mathematical thinking, and the teachers rated the students as a way of examining the effectiveness of education.

The results of the comparison between the experimental group and the control group before the teaching experiment are shown in Table 2. According to Table 2, it can be concluded that the difference between the scores of the experimental group and the control group in the six dimensions of small muscle activity, art, music and rhythm, blocks, natural science and mathematical thinking is not more than 0.5 points, and the P-value of the data of these six groups is greater than 0.05, which is not statistically significant, indicating that the level of the two groups is the same, and it will not have an effect on the results of the experiment.

Table 2. Pre-test comparison of experimental and control group

Dimension	Experimental		Control		T	P
	M	SD	M	SD		
Small muscle activity	23.25	4.85	23.68	4.69	-0.842	0.623
Art	27.69	5.06	27.85	5.07	-0.426	0.538
Music and rhythm	28.15	5.84	27.96	4.63	0.674	0.498
Building block	25.48	4.89	25.32	5.42	0.368	0.721
Nature science	23.42	5.32	23.62	5.33	-0.296	0.634
Mathematical thinking	24.86	5.42	25.06	5.16	-0.614	0.598

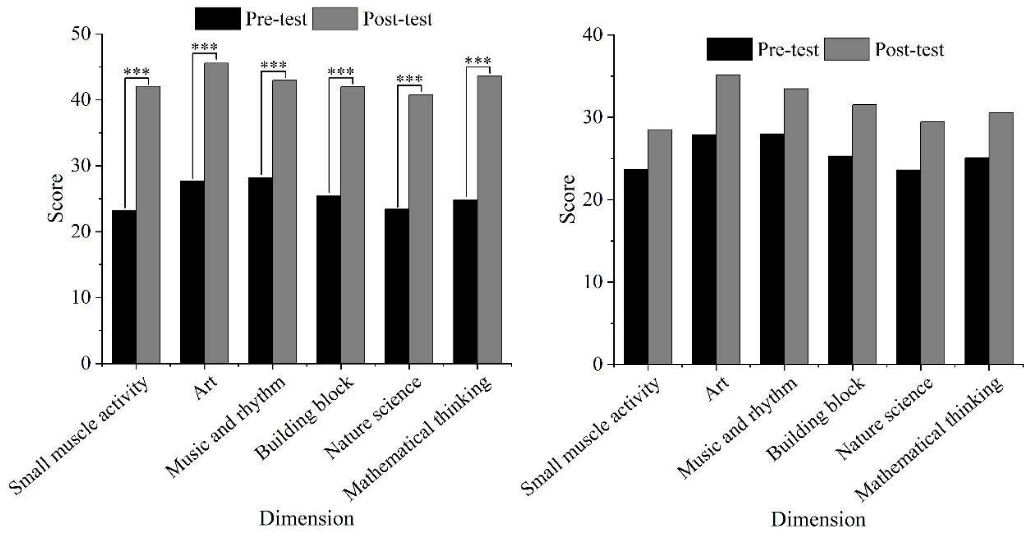
The comparison results between the experimental group and the control group after the teaching experiment are shown in Table 3. According to Table 3, the experimental group achieved the average scores of 42.06, 45.55, 42.96, 41.95, 40.75 and 43.62 in the six dimensions of fine motor activity, art, music and rhythm, building blocks, natural science and mathematical thinking, respectively, while the control group obtained the average scores of 28.49, 35.14, 33.45, 31.52, 29.44 and 30.54, respectively, and the difference between the two groups widened from less than 0.5 points to [9.51, 13.57] points. And all dimensions showed very significant differences.

Table 3. Post-test comparison of experimental and control group

Dimension	Experimental		Control		T	P
	M	SD	M	SD		
Small muscle activity	42.06	6.84	28.49	4.69	25.165	0.000
Art	45.55	7.48	35.14	5.07	19.452	0.000
Music and rhythm	42.96	8.06	33.45	4.63	15.496	0.000
Building block	41.95	7.58	31.52	5.42	21.342	0.000
Nature science	40.75	6.98	29.44	5.33	20.546	0.000
Mathematical thinking	43.62	7.46	30.54	5.16	24.765	0.000

5.2. Between-group variation in the effectiveness of children's pre-school education

The test results of children in the experimental and control groups on small muscle activity, art, music and rhythm, blocks, natural science and mathematical thinking were compared before and after the experiment, and the results are shown in Fig. 6, where (a) and (b) are the between-group changes in the experimental and control groups, respectively. As can be seen in Figure 6, children in the experimental group produced highly significant differences before and after the experiment in small muscle activity, art, music and rhythm, blocks, natural science, and mathematical thinking, with mean scores in the six dimensions increasing by 18.81, 17.86, 14.81, 16.47, 17.33, and 18.76 points, respectively. The control group also improved in all 6 dimensions, but none of them were significant, with a difference of 4.81, 7.29, 5.49, 6.20, 5.82, and 5.48 points in each dimension.



(a) Experimental group (b) Control group

Fig. 6. Intergroup change of experimental and control group

6. Conclusion

The article embeds the decision tree algorithm into the teaching of preschool education, and uses the decision tree algorithm to classify the learning styles of children in the preschool stage and analyze their learning activities. Online preschool education resources are integrated through the decision tree algorithm, which are imported into teaching to test the effect of its use.

The kindergarten children were categorized into 3 types of extroverted, negative emotion and effortful control children through the decision tree. Effort-control style children were able to control their detection behaviors in the game and improve the efficiency of the game. Competitive game situations were able to improve the game task scores of Extraversion and Effort Control children. Effort-control children showed greater concentration during competitive play.

After integrating the decision tree algorithm with preschool educational resources and using it in actual teaching, children in the experimental group showed very significant improvement in six areas: small muscle activity, art, music and rhythm, blocks, natural science, and mathematical thinking, while the changes in the control group were not significant. After the experiment, the children in the experimental group were 13.57, 10.41, 9.51, 10.43, 11.31, and 13.08 points higher than the control group in the six dimensions.

References

[1] Wang, X., Sun, H., & Li, L. (2019). An innovative preschool education method based on computer multimedia technology. International Journal of Emerging Technologies in Learning, 14(14).

[2] Lazareva, M., & Zvezda, L. (2021, June). Digital technologies in preschool education: Risks and prospect. In 2021 1st International Conference on Technology Enhanced Learning in Higher Education (TELE) (pp. 93-95). IEEE.

[3] Lupu, D., & Laurenţiu, A. R. (2015). Using new communication and information technologies in preschool education. Procedia-Social and Behavioral Sciences, 187, 206-210.

[4] Zhao, Y., Gao, W., & Ku, S. (2023). Optimization of the game improvement and data analysis model for the early childhood education major via deep learning. Scientific reports, 13(1), 20273.

[5] Xu, B. (2022). Application of edge computing and data mining processing system in preschool education

courses. Mobile Information Systems, 2022(1), 7872897.

[6] Sun, J., Wu, H., & Shi, S. (2023). A research of the evaluation of preschool education resource allocation level and spatio-temporal differences: Based on repeated indicators method and Theil index. Heliyon, 9(6).

[7] Huang, J. (2022). Preschool education resource allocation model for index system evaluation based on nonlinear random matrix. Mathematical Problems in Engineering, 2022(1), 6271690.

[8] Chen, C. (2022, April). Application design in preschool education resource system platform based on deep learning technology. In International Conference on Multi-modal Information Analytics (pp. 729-735). Cham: Springer International Publishing.

[9] Batrakova, I., Ushanov, A., & Ioseliani, A. (2024). Information technology and online resources in the professional activities of preschool teachers. Journal of Applied Research in Higher Education, 16(2), 675-689.

[10] Liu, H. (2019). The research of theoretical construction and effect of preschool wisdom education system in the background of big data. Cluster Computing, 22(Suppl 6), 13813-13819.

[11] Qing, S. (2022). Design and application of preschool education system based on mobile application. Mathematical Problems in Engineering, 2022(1), 8556824.

[12] Qi, F. (2021, May). Application of new media technology in the information teaching of preschool education. In International Conference on Society Science (pp. 157-163).

[13] Otterborn, A., Schönborn, K., & Hultén, M. (2019). Surveying preschool teachers' use of digital tablets: general and technology education related findings. International journal of technology and design education, 29(4), 717-737.

[14] Zhang, Q. (2023). Secure preschool education using machine learning and metaverse technologies. Applied Artificial Intelligence, 37(1), 2222496.

[15] Zhang, Z., Zhang, X., & Liu, X. (2025). A novel algorithm for personalized learning in preschool education using artificial intelligence. Journal of Computational Methods in Sciences and Engineering, 14727978251314562.

[16] Pang, Z., & Chen, J. (2023). Construction of Preschool Education Resources under the Internet+ Background. Journal of Testing and Evaluation, 51(1), 27-38.

[17] FUNTIKOVA, O. (2021). Use of electronic educational resources in preparation of future educator of preschool education. EUROPEAN HUMANITIES STUDIES: State and Society, (4), 122-143.

[18] Tuo, M., & Long, B. (2022). Construction and application of a human - computer collaborative multimodal practice teaching model for preschool education. Computational Intelligence and Neuroscience, 2022(1), 2973954.

[19] Khasanova, M. (2024). Improving the Professional Knowledge and Skills of Preschool Education Teachers Using Innovative Educational Technologies. Modern Science and Research, 3(1), 1-2.

[20] Weiwei, S. (2022). Design of auxiliary teaching system for preschool education specialty courses based on artificial intelligence. Mathematical Problems in Engineering, 2022(1), 4504707.

[21] Li, J. (2021). Research on the reform and innovation of preschool education informatization under the background of wireless communication and virtual reality. Wireless Communications and Mobile Computing, 2021(1), 3176309.

[22] Fedina, N. V., Burmykina, I. V., Zvezda, L. M., Pikalova, O. S., Skudnev, D. M., & Voronin, I. V. (2017). Use of distance learning technologies in the course of implementing educational programs in preschool education. Eurasia journal of mathematics, science and technology education, 13(11), 7561-7571.

[23] Georgiev, V., & Nikolova, A. (2020). Tools for creating and presenting online learning resources for preschool kids. TEM Journal, 9(4), 1692.

[24] Su, Y., Dong, Z., & Li, S. (2025). Distributed Sharing and Personalized Recommendation System of College Preschool Education Resources Under the Intelligent Education Cloud Platform Environment. International Journal of High Speed Electronics and Systems, 2540430.

[25] Yang, Y. (2023, March). Intelligent Recommendation System for Early Childhood Learning Platform Based on Big Data and Machine Learning Algorithm. In The International Conference on Cyber Security Intelligence and Analytics (pp. 140-150). Cham: Springer Nature Switzerland.

[26] Hou, Y. (2025). Design and Implementation Evaluation of Personalized and Differentiated Teaching Strategies for Preschool Children Based on Fuzzy Decision Support Systems. International Journal of Computational Intelligence Systems, 18(1), 42.

[27] Huang, J. (2024). The Sharing Model of Kindergarten Art Enlightenment Education Curriculum Resources Based on Deep Learning Algorithm. International Journal of High Speed Electronics and Systems, 2440123.

[28] Prasad, P. Y., Prasad, D., Malleswari, D. N., Shetty, M. N., & Gupta, N. (2022). Implementation of Machine Learning Based Google Teachable Machine in Early Childhood Education. International Journal of early childhood special education, 14(3).

[29] Wang, L., & Liu, T. (2022). Optimization algorithm of preschool education resource allocation based on data envelopment analysis model. Scientific Programming, 2022(1), 3858664.

[30] Zhang, L. (2022, October). Accurate Recommendation Algorithm of Preschool Education Network Resources Based on Improved Decision Tree. In 2022 Global Reliability and Prognostics and Health Management (PHM-Yantai) (pp. 1-8). IEEE.

[31] Ilaria Pietropoli & Pablo Gracia. (2025). Social inequalities in children's cognitive and socioemotional development: The role of home learning environments and early childhood education. Research in Social Stratification and Mobility,97,101034-101034.

[32] Zhuo Li & Xin Guo. (2025). Insights into PAHs oxidation assisted by CO2 and H2O at elevated pressures: A combined ReaxFF MD simulations and decision trees model. Chemical Engineering Journal,509,161259-161259.

[33] Chun Pan. (2024). Construction of Risk Prediction Models for Enterprise Finance Sharing Operations Using K-Means and C4.5 Algorithms. International Journal of Computational Intelligence Systems,17(1),208-208.