International Journal for Housing Science and Its Applications
Publish August 3, 2025. Volume 46, Issue 3 Pages 727-740

https://doi.org/10.70517/ijhsa46354

Research on Differentiated Cultivation Path Planning of College Students' Innovation and Entrepreneurship Education Based on Decision Tree Algorithm in the Environment of Industrial Digital Transformation

Fake Ma[1,*]
[1] Henan Institute of Economics and Trade, Zhengzhou, Henan, 450000, China
Corresponding authors: (e-mail: mafake@henetc.edu.cn).

Abstract In order to permanently improve the level of national economic and social development and international competitiveness, talent cultivation is one of the four basic functions of colleges and universities, and the cultivation of innovative and entrepreneurial talents adapted to the requirements of the times will become a major mission for Chinese colleges and universities to promote the realization of the innovation-driven strategy. This paper uses the information gain method, optimizes and improves the decision tree model, uses the model to analyze the characteristics of the collected student data, and calculates the factors affecting the innovation and entrepreneurship effect of differentiated teaching. For the influence factor of whether it is a high-tech field or not, the first influence is the national scholarship, with a Gini coefficient of 0.589; the second influence is the participation in scientific research activities and academic performance, with Gini indexes of 0.545 and 0.546, respectively; and the third influence is the academic qualification and entrepreneurial field, with Gini indexes of 0.469 and 0.436, respectively. assessing the innovation and entrepreneurship and scientific research ability of the college students level, a total of 408 people are interested in scientific research, accounting for 69.74%, the number of people who have been authorized by the patent is 128, accounting for 21.88%, which is closely related to better academic performance, and a few of the subject students have scientific research ability.

Index Terms Information Gain, Decision Tree Model, Data Characterization, Gini Index, Innovation and Entrepreneurship

I. Introduction

In today's digital era, industrial digital transformation has become an important trend for enterprise survival and development [1]. Industrial digital transformation refers to the deep integration of traditional industries with the real economy through a new generation of information technology, reconfiguring the production process, business model and organizational structure to form a new economic form with data as the core driving force [2]-[4]. The essence of industrial digital transformation is not simply the superposition of technology, but to use digital technology as a lever to pry the collaborative innovation of the various links of the industry chain, to promote the extension of the industrial value chain to the high-end, to create new growth momentum [5], [6]. Under the background of digital transformation, enterprises are facing highly complex and variable technological challenges, and there is an urgent need for high-quality technological innovation talents to adapt to the needs of the new era, which requires colleges and universities to carry out innovation and entrepreneurship targeted and differentiated training for college students to adapt to the needs of the digital industry [7]-[10].

Nowadays, innovation and entrepreneurship education has become the focus of more and more people, and this education mode aims to cultivate students' innovative thinking and entrepreneurial ability, injecting new vitality into the development of the country and society, and its importance is self-evident [11]-[13]. However, the current innovation and entrepreneurship education faces challenges such as imperfect education model, insufficient faculty, and shortage of funds, which cannot realize the effective development of innovation and entrepreneurship, and the differentiated cultivation is even impossible to talk about [14], [15]. In order to cope with the challenges of digital transformation, innovation and entrepreneurship education talent cultivation in colleges and universities needs to make the necessary changes [16]. Educational institutions should strengthen cooperation with enterprises, increase the investment of resources such as funds and human resources, while the curriculum focuses on cultivating students' problem solving ability and teamwork ability, and carry out differentiated innovation and entrepreneurship talent cultivation in response to the talent needs of enterprise digital transformation [17]-[20].

Literature [21] aims to examine the enablers of digital transformation and their impact on performance outcomes, and the results show that leadership, structure and culture are important enablers of enterprise digital transformation, which help industrial enterprises achieve performance outcomes. Literature [22] discusses the transformation and development of traditional industries in the context of the digital economy and examines the dynamic decision-making process between traditional firms, digital technology providers and the government, emphasizing that the government subsidy coefficient can increase the optimal level of effort of traditional firms and digital technology providers. Literature [23] describes the digital transformation of enterprises and emphasizes that digital transformation effectively facilitates business processes, decision-making, etc., and plays an important role in the growth and development of enterprises, and that the management of an enterprise should decide on the consistency, scope, and scale of the digital changes to be made. Literature [24] describes the digital transformation strategy options for industrial firms, identifies multiple key tensions that managers need to address, and provides corresponding guidelines for a digital transformation strategy to overcome each identified tension. The above study explores the digital transformation of enterprises and its implications, expressing the inevitability of digital transformation in that it not only effectively promotes the optimization of business processes, but also contributes to the improvement of business performance. In this context, higher education can only improve students' competition in the workplace and adapt to the market demand for talents by strengthening innovation and entrepreneurship education. Based on the current difficulties in the evaluation of innovation and entrepreneurship teaching ability, literature [25] designed an evaluation model of college students' innovation and entrepreneurship education based on BP neural network, revealing that students' professional knowledge and skills, participation in entrepreneurship training camps and other ways have an important impact on innovation and entrepreneurship ability. Literature [26] emphasized the importance of innovation and entrepreneurship, and pointed out the shortcomings of the current innovation and entrepreneurship education, and carried out practical research in terms of strengthening the awareness, creating the atmosphere, improving the policy and curriculum, and choosing the mentor, strengthening the practice, and building the platform. Literature [27] aims to analyze students' attitudes, self-perceptions, values and intentions towards innovation and entrepreneurship in order to find important variables affecting the relationship between innovation and entrepreneurship education and students' entrepreneurial attitudes, perceptions and intentions, which adds value to the existing knowledge of entrepreneurship. Literature [28] explored innovation and entrepreneurship education for university students and emphasized that innovation and entrepreneurship education significantly improves the quality of innovation and entrepreneurship, as well as the motivation and efficiency of students' participation in innovation and entrepreneurship activities.

This paper addresses the ecological construction of innovation and entrepreneurship education in colleges and universities, and puts forward the corresponding construction direction from three aspects: school-enterprise cooperation, resource allocation, and practical talents. Using the information gain to represent the value of information entropy reduction of the data set after the function of features, designing a questionnaire to collect the data of innovation and entrepreneurship education in colleges and universities, and transforming the questionnaire data into the format suitable for the decision tree model, and using the decision tree to calculate the Gini coefficient by iterating through the remaining features for each feature in the node. Based on the decision tree model, the CFS algorithm is used to measure the correlation between features and discretize continuous variables. We set up a data collection library for innovation and entrepreneurship cases, pre-processed the collected data, analyzed the data using the decision tree model in three aspects: experimental characteristic data, factors affecting innovation and entrepreneurship effects, and differentiated training effects, and planned a reliable differentiated training path based on the results of the analysis.

II. Construction of an ecosystem for innovation and entrepreneurship education under the digital transformation of industries

II. A. Analysis on the Construction of Ecological System of Innovation and Entrepreneurship Education in Colleges and Universities

II. A. 1) Establishment of a close partnership between industry and business

Under the background of digital transformation, universities need to establish closer cooperation with society, industry and enterprises in order to build an ecosystem of innovation and entrepreneurship education. Through diverse modes of cooperation, including signing cooperation agreements, establishing entrepreneurial mentorship, developing practical projects, and co-building laboratories, universities can provide students with richer practical opportunities and improve their innovation and entrepreneurship abilities in the era of digital economy. At the same time, this deep integration also provides enterprises with more ways to cooperate with universities and acquire new knowledge, which promotes the common development of industry and education.

II. A. 2) Digital empowerment and flexible adjustment of teaching resources supporting settings

By gaining a deeper understanding of the digital needs of the industry, universities can flexibly adjust their specialization in accordance with market changes to ensure that students are equipped with the latest skills and knowledge when they graduate. Updating course content involves not only the updating of theoretical knowledge, but also the design of practical courses in order to cultivate students' innovative and practical abilities needed in the digital economy.

In a nutshell, the digital era has put forward a higher level of talent cultivation requirements for colleges and universities, requiring them to flexibly adjust their professional settings and update their course contents to meet the needs of the digital development of industries. Through in-depth understanding of the digitalization needs of the industry, the establishment of a flexible mechanism and close industry-university-research cooperation, colleges and universities can better cultivate new talents adapted to the era of digital economy.

II. A. 3) Introduction of professionals with practical experience

The construction of a university-enterprise information communication platform improves the efficiency of communication between teachers and industry, and promotes the close integration of practical projects and industrial research. This helps to maintain the efficiency of the faculty in obtaining the latest information about the industry, so that they can better understand the industrial dynamics and incorporate the latest practical experience into their teaching.

Building a university-enterprise information communication platform is another important means to improve the efficiency of communication between faculty and industry. By establishing online platforms or social networks, teachers can obtain timely information on industrial dynamics, market demand, etc., so as to better adjust course content and teaching methods. This also provides a convenient communication platform for industry-university-research cooperation, which promotes both sides to carry out practical projects, subject research and other activities more efficiently. The construction of the information platform helps to establish an information-sharing mechanism between schools and industries, providing teachers with more forward-looking knowledge for teaching.

Colleges and universities can encourage teachers to participate in industrial exchange activities, visits, etc., to feel the pulse of the development of the industry firsthand. This helps teachers understand more deeply the actual operation and technology trends of the industry, provides more specific references for curriculum updating and improvement of teaching methods, and prompts teachers to keep an eye on the dynamics of the industry, so that they can better convey the latest practical experience and industry knowledge in their teaching.

Strengthening the construction of faculty, introducing professionals with rich practical experience, establishing the mechanism of industry-university-research cooperation, and improving the communication efficiency between teachers and industries through the construction of the university-enterprise information communication platform are a series of measures that will help the faculty of colleges and universities to better meet the needs of the industry in the era of the digital economy, and to provide students with a more practical and prospective innovation and entrepreneurship education.

II. B. Algorithms related to the construction of a database of cases of innovation and entrepreneurship education

II. B. 1) Information gain

The information entropy represents the uncertainty while the information gain can measure the magnitude of the impact of a feature on the classification result [29].

The information entropy formula is as follows:

$$Info(D) = -\sum_{i=1}^{c} p_i \log(p_i) Info(D) \tag{1}$$

where D denotes the training dataset, c denotes the number of data categories, and p_i denotes the ratio of the number of samples of category i to all samples.

Corresponding to the dataset D, the information entropy after the action of feature A is $Info(D)$ when feature A is selected as the decision tree judgment node, which is calculated as follows:

$$Info_A(D) = -\sum_{i=1}^{k} p_i \frac{|D_j|}{|D|} \times Info(D_j) \tag{2}$$

where k denotes that the sample D is divided into k parts.

The information gain denotes the value of information entropy reduction of the dataset D after the action of the feature A [30]. The formula is as follows:

$$Gain(A) = Info(D) - Info_A(D_j) \tag{3}$$

The most appropriate feature selection for a decision tree node is the feature with the largest value of $Gain(A) = Info(D) - Info_A(D_j)$.

II. B. 2) Decision Tree Modeling and Evaluation

(1) Decision Tree Modeling

The questionnaire data of a university are collected, covering various aspects of innovation and entrepreneurship education, including the basic teaching function keys, the utilization rate of knowledge resources, the function of interpersonal interaction, and technical support. These data will be used to analyze the problems and reasons of innovation and entrepreneurship ability and innovation and entrepreneurship education. Next, the data will be analyzed using the decision tree model to understand the impact of different factors on student satisfaction and to reveal the root causes of the problems.

First, the questionnaire data were converted into a format suitable for decision tree modeling [31]. The binary and continuous variables were used as the characteristic variables, where the binary variable indicates the existence or non-existence of problems in the construction of online learning space, and the continuous variable indicates student satisfaction. Correspondingly organizing their data forms are:

$$X = \{x_1, x_2, \cdots, x_n\} \tag{4}$$

In Eq. (4), X is the set of elements of all features, x_1 is the first feature, x_n is the last feature, n is the number of features, and the feature n is 2 in binary variables and 7 in continuous variables, i.e., the article uses a Likert isometric seven-level scale

Following this, the data is divided into training and test sets, and a decision tree regression model is used to train the data and predict student satisfaction in the test set. This process uses a decision tree, i.e., the Gini coefficient is computed by traversing the remaining features for each feature within a node, which is computed as:

$$Gini(X) = 1 - \sum_n \left[p(x_i)\right]^2 \tag{5}$$

In equation (5), X is the set of features of the current node, $Gini(X)$ is the Gini coefficient of X, and $p(x_i)$ is the proportion of a specific feature x_i among the remaining features.

Based on this Gini coefficient, the current node splits to form a new fork in the decision tree, which is computed as:

$$G(X, A) = \frac{|X_L|}{|X|} Gini(X_L) + \frac{|X_R|}{|X|} Gini(X_R) \tag{6}$$

In Eq. (6), X denotes the dataset of the current node, A is the attribute to be classified, and $|X|$ is the total number of samples classified by the current node.

By calculating the mean square error between the prediction result and the actual result, the accuracy of the model can be evaluated and the MSE can be calculated accordingly.Based on this, this paper namely seeks for model optimization as an objective function and thus iterates to obtain a good model performance. After fully iterating the optimized model, the results of the decision tree model are interpreted in order to analyze which factors have a significant impact on student satisfaction and the correlation between these factors. Visualizing the decision tree model provides an intuitive understanding of the relationships between different issues and provides a basis for making targeted recommendations for improvement.

Through the application of data analysis and decision tree modeling, the problems of innovation and entrepreneurship education in a university will be understood in depth and will provide strong support for the development of differentiated and customized strategies. Based on the results of this analysis, the prediction ability based on various types of features in its optimized model is organized.

(2) Decision tree model evaluation

Where P denotes the number of samples of positive cases, N denotes the number of samples of negative cases, TP denotes the number of correctly predicted positive cases, FP denotes the number of negative cases predicted as positive cases, FN denotes the number of positive cases predicted as negative cases, and TN denotes the number of correctly predicted negative cases.

The following metrics are obtained.

Probability of classification accuracy, calculated as:

$$Accuracy = \frac{TP + TN}{P + N} \tag{7}$$

Recall, calculated as:

$$\mathrm{Re}\,call = \frac{TP}{P} \tag{8}$$

The false alarm rate, calculated as:

$$Fprate = \frac{FP}{N} \tag{9}$$

The accuracy is calculated by the formula:

$$Precision = \frac{TP}{TP+FP} \tag{10}$$

Positive Example Coverage is calculated by the formula:

$$Sensitivity = \frac{TP}{TP+FN} \tag{11}$$

Negative example coverage is calculated by the formula:

$$Specificity = \frac{TN}{TN+FP} \tag{12}$$

II. B. 3) Feature selection

Feature selection is a description of an application domain that combines features that are common or relevant to the thing itself. Feature selection can select a minimal subset of N original features, including $M(M \leq N)$, so that the probability distribution values of different categories in the subset including M features are close to the N original features. If F_N is the original feature set and F_M is the selected feature subset, the possible category C conditional probability $P(C \mid F_M = f_m)$ should be close to $P(C \mid F_N = f_N)$, where f_M and f_N are the value vectors of the corresponding feature vectors F_M and F_N.

In general, feature selection algorithms can simplify the data description and simplify the data collection task, which in turn can solve the problem and improve the quality of the dataset. When the dataset has a large number of features, it needs to be cooled down.CFS algorithm, the correlation-based feature selection algorithm. For continuous variables, the CFS method measures the correlation with the score $(Merit_s)$ of the feature subset, as shown in Equation (13):

$$Merit_s = \frac{k\overline{r_{cf}}}{\sqrt{k+k(k-1)\overline{r_{ff}}}} \tag{13}$$

where: k is the number of variables in the subset, $\overline{r_{cf}}$ is the mean value of the correlation between all independent variables and the target variable in the feature subset, and $\overline{r_{ff}}$ is the mean value of the correlation between two independent variables in the feature subset.

For continuous-discrete variables, the continuous variables need to be discretized. If the variables after discretization are X and Y respectively, the calculation formula is shown in Eq. (14), Eq. (15):

$$H(Y) = -\Sigma p(y)\log_2 p(y) \tag{14}$$

$$H(Y \mid X) = -\sum_{x\in X} p(x) \sum_{y \in Y} p(y \mid x)\log_2 p(y \mid x) \tag{15}$$

where: $p(y)$ is the probability of y, $p(y \mid x)$ is the probability of Y under the condition of X, $H(Y)$ is the information entropy of Y, and $H(Y \mid X)$ is the information entropy of Y under the condition of X.

Then the information gain, which is the difference between the a priori information entropy and the a posteriori information entropy, is calculated as shown in Equation (16):

$$gain = H(Y) - H(YX) = H(Y) + H(X) - H(X,Y) \tag{16}$$

The uniform uncertainty between the variables is then calculated as shown in equation (17):

$$symmetrical\ uncerta\,\mathrm{int}\,y = 2.0 \times \frac{gain}{H(Y)+H(X)} \tag{17}$$

If the uncertainty is higher, the correlation is lower. Evaluating the correlation between the variables in each subset of features enables data categorization based on the features and is more conducive to data processing.

II. C. Construction of a data collection bank for innovation and entrepreneurship cases

II. C. 1) Processing framework

The overall framework structure of the system is shown in Figure 1. The processing framework of the system's data collection library is mainly divided into data collection and storage. Combined with the content of Figure 1, among the data collection and storage, it is necessary to collect data from the platform collection data, innovation and entrepreneurship mathematical data, etc., to provide a data source for the innovation and entrepreneurship big data platform, and analyze the data to establish an effective database, i.e., to achieve the visualization of the data through the content of data mining, data screening, and so on.

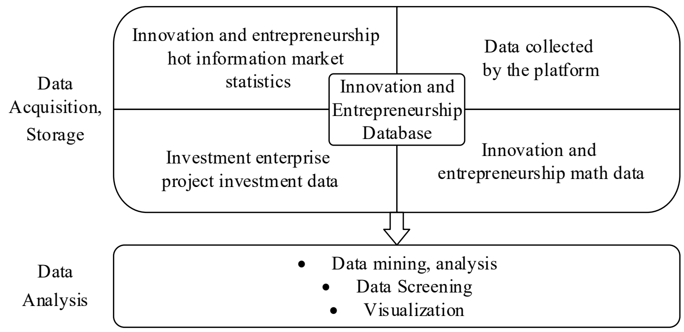

Figure 1: System of data acquisition library on the processing framework

II. C. 2) External library calls

Combine external library calls, use pandas itertools and self time for data processing, combine data such as youth entrepreneurship network, enter “import pandasas pd” for data import, import itertools as it # subset generation, and perform processing to complete the dataset.

II. C. 3) Data import

Full function written by “if __name__ == ‘__main__’: #” import data, data = data.applymap (lambda x: str(x).strip()) # Remove any space values that may be present in the import, for example “education” becomes ‘education’. The final data import is at the end of the code, and the support and confidence levels are set based on the data import.

II. C. 4) Data processing

Output all Goodlist, in fact, is the alternative frequent 1 item set, this function is not too important, the subsequent call can be replaced by other ways, the purpose of writing is to facilitate the logic of the process of a way only.

III. Analysis of the results of college students' innovation and entrepreneurship prediction based on the decision tree algorithm

III. A. ***Experimental data***

III. A. 1) Analysis of data on student characteristics

The purpose of this experiment is to automatically assess students' innovation and entrepreneurship skill level by constructing a classification model based on ID3 decision tree. Specific objectives include.

(1) Determine the performance of the ID3 model on a given dataset.

(2) Evaluate the effects of different features on the classification of students' innovative and entrepreneurial skill levels.

(3) To provide an automated tool to assist teachers in assessing students' innovative entrepreneurial skills.

The experimental data consists of 585 student records, each record contains the characteristics of students' innovative skills,entrepreneurial skills and research skills.20 of them are selected.The characteristics data are shown in Table 1.These data will be used to train and test the ID3 decision tree model. Combining the critical thinking scores and problem solving ability, the combined average score of some sample students is 4.1175, of which ID10, ID16 and ID18 have higher scores, 4.65, 4.55 and 4.5 respectively.

III. A. 2) Decision Tree Algorithm Performance Analysis

The preprocessed 585 data are divided into training and test sets by 7:3, i.e., 410 data in the training set and 175 data in the test set. Where for the C4.5 algorithm and CART algorithm, unlike the ID3 algorithm which uses the training set to generate the tree and then prunes the tree using the training set, a separate pruning set is needed to prune the tree, so the divided training set is again divided into a growing set and a pruning set according to 7:3, i.e., the growing set is 287 pieces of data and the pruning set is 123 pieces of data. The final stage grade prediction model is a grade prediction model constructed on the basis of all the characteristics of learning behaviors after the end of the course and before the final exam. After the end of the course and before the final exam, students are more concerned about whether they will fail the final course, so the final grade is divided into two categories, the final grade is greater than or equal to 60 as passing, and the final grade is less than 60 as failing.

Table 1: Student characteristics data

ID	Age	Innovative ability	Entrepreneurial ability	Scientific ability	Attendance rate /%	Job completion/%	Critical thinking score	Problem-solving ability
1	20	86	92	78	94	90	4.3	4.5
2	21	73	83	92	96	84	4.2	4.1
3	21	90	86	83	95	85	4.5	4
4	22	79	89	76	95	78	3.6	3.9
5	21	81	76	86	90	90	4.1	4.5
6	21	93	87	68	93	86	4.5	4
7	21	75	80	80	95	73	3.8	3.8
8	19	88	75	70	92	87	4.1	4.3
9	21	80	85	67	94	60	4	3.8
10	21	90	80	83	85	93	4.8	4.5
11	21	75	80	74	90	64	3.8	3.7
12	19	90	85	65	84	79	4	4.3
13	22	73	80	90	88	83	3.3	3.8
14	20	85	75	70	93	75	4.5	4.3
15	21	80	86	68	86	96	4.1	4
16	20	90	82	85	94	76	5	4.1
17	21	75	80	70	85	63	3.8	3.7
18	20	90	84	61	93	80	4.1	4.9
19	19	70	83	90	88	77	4	3.3
20	22	85	70	75	92	80	4.4	4.3

The three decision tree algorithms ID3, C4.5 and CART are applied to construct the final stage grade prediction model through the same training dataset, and then the same test dataset is used to evaluate the performance of the three different models, and the algorithm with the optimal performance for this dataset is selected based on the evaluation indexes of Accuracy, Precision, Recall and F1-Measure. Figure 2 shows the performance evaluation comparison of the three algorithms. The accuracy rate of the three algorithms is calculated according to the formula (7) of the accuracy evaluation index in Subsection 2.2.2, and it can be seen that the accuracy rate of the ID3 algorithm is higher than that of the other two algorithms. Calculating the precision rate of the three algorithms according to the formula (10) of the precision rate evaluation metrics in subsection 2.2.2. The decision tree algorithm of ID3 has an accuracy rate of 91.348%, a precision rate of 91.926%, a precision rate of 91.893%, and a F1-Measure of 88.926%, which are better than the other two algorithms.

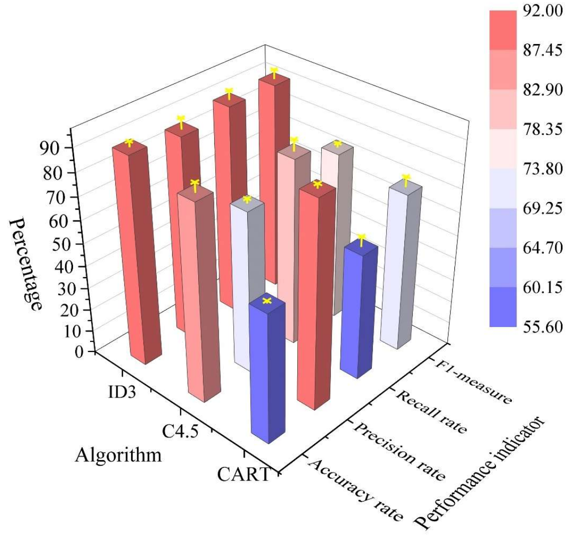

Figure 2: Performance assessment comparison of three algorithms

III. B. *Characterization of Factors Influencing the Creative and Entrepreneurial Effectiveness of Differentiated Instruction*

Innovative and entrepreneurial universities, namely, provincial (municipal) ordinary key universities and general ordinary colleges and universities, this category of institutions in general is the theme of undergraduate students, and there are a few key undergraduate colleges and universities are developing in the direction of research universities, but on the whole, whether it is the structure of the weight of the students (bachelor's degree, master's degree, doctoral degree), or scientific research and creativity, will be classified as a teaching and research characteristics more in line with the real situation. It is to focus on subject-specialized knowledge, to cultivate undergraduate students' talents, the proportion of undergraduates in school is higher than the proportion of college students, and diversified characteristics have been formed on undergraduate education: firstly, the higher quality of running school subject specialties to the direction of college students' research, i.e., the tendency of running school in a research-oriented university, secondly, the characteristics of undergraduates who are mainly purely knowledge-based education, and thirdly, the education mode of running school mainly in the form of skill-based education, i.e. tends to the innovative and entrepreneurial education goals and characteristics of universities. The development orientation of educational research universities is between research universities and skill-based universities, i.e., to varying degrees, they simultaneously present the characteristics of these two levels of institutions, but instead, their own characteristics are not prominent, and they have the same performance in terms of the effect of undergraduate entrepreneurship. In the following, we will take the factors affecting entrepreneurship as input variables and entrepreneurship effect as output variables in teaching research universities, observe the correlation between input indicators and output indicators to obtain the decision tree model, and analyze the basic characteristics of research and skill oriented in their schooling objectives, which also show a diversified tendency in terms of the characteristics of innovation and entrepreneurship effect of their students.

III. B. 1) Business prospects

Figure 3 shows the output 1-enterprise outlook, the first data of each node is the classification rule of the node, according to the input index to set up the classification rule, e.g.: National Scholarship<0.5, if a sample to be tested in the National Scholarship of the input value of <0.5, then it meets the classification rule, True, split to the left, and vice versa, False, split to the right. Gini: Gini index, the optimal feature selection Indicator. By selecting the input index with the largest Gini index as the classification rule for the node. For example: National Scholarship, gini-0.648, i.e. among the existing samples, the maximum Gini index is 0.648, then select National Scholarship as the classification rule of the node. Sample: Sample, the percentage of the samples to be predicted to the node among all the samples, e.g.: sample=100%, then it means that the amount of samples to be predicted to the node accounts for 100% of the total number of samples. Value: the value of the sample to be tested in the classification rules, the number of output of a certain option as a percentage of the number of all output options. For example: Value= [0.248,0.5], it means that under the classification rule, 24.8% choose option 1 and 50% choose option 2. Analysis of data results: For the enterprise prospect of innovation and entrepreneurship, the most influential is the national scholarship, that is, students who have been awarded the national scholarship are more likely to achieve higher net profit in entrepreneurship, in the second influential position is the entrepreneurial field and academic performance, and in the third influential position is the match between the academic qualifications and the majors they have studied. The following decision tree analyzes the same idea and method, and will not be repeated.

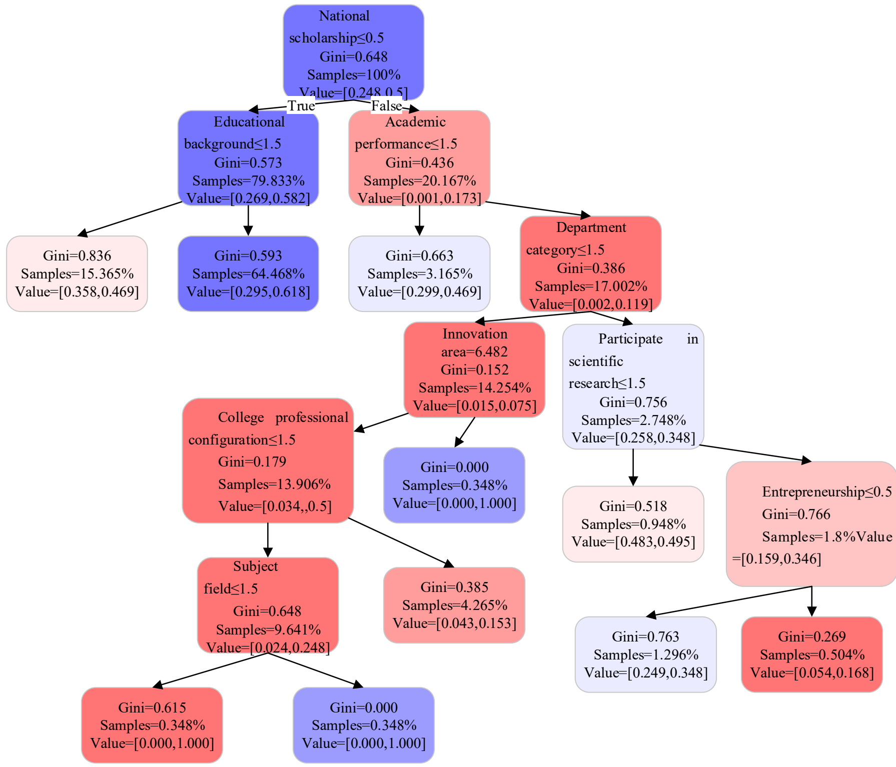

Figure 3: Output 1-enterprise prospects

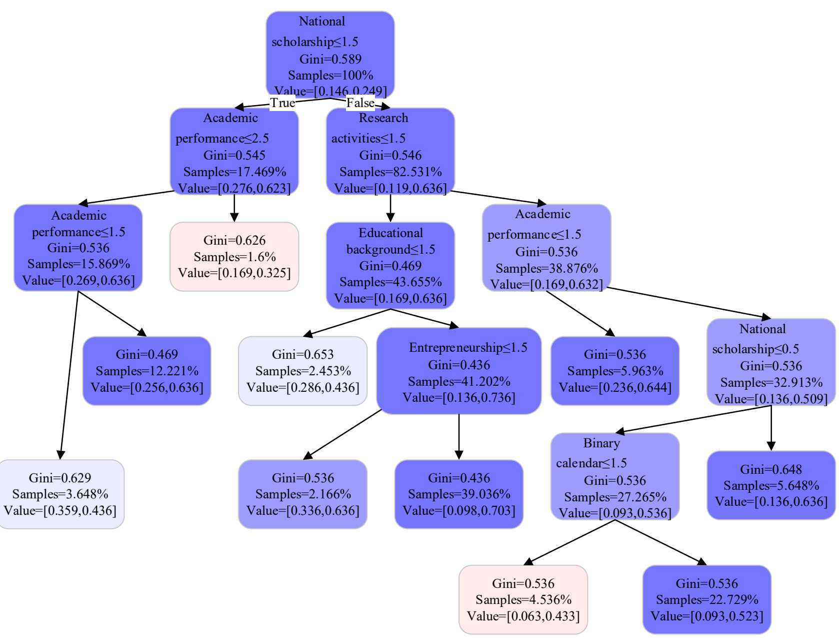

Figure 4: Output 2- whether high-tech domain decision tree

III. B. 2) Whether it is a high-tech field

Figure 4 shows the decision tree of output 2-whether high-tech field or not, and the data results are analyzed:For the question of whether high-tech field or not, the first influential position is the national scholarship (Gini index is 0.589), the second influential position is the participation in scientific research activities and academic performance, with Gini indexes of 0.545 and 0.546, respectively, and the third influential position is the academic qualification and entrepreneurship field, with Gini indexes of 0.469, 0.436, i.e., the synthesis found that in the question of whether high-tech field, ordinary colleges and universities have the same influence factors on college students' entrepreneurship.

III. C. Effectiveness analysis based on differentiated training

Explore the root causes of learning differences among college students and analyze and sort out these causes in order to find the best model of differentiated training for different college students.

III. C. 1) Self-directed learning capacity

In order to get a better access to the current situation of college students' cultivation, state of mind, innovation ability, etc., all the indicators for evaluating college students' independent learning ability are made into a questionnaire for investigation, and the statistical results are analyzed.

Figure 5 shows the statistics of the results, and the independent learning ability is evaluated through the six aspects of innovative entrepreneurship professional course results, learning purpose, self-restraint, satisfaction, learning summary and learning plan. Innovation and entrepreneurship professional course learning results imported by the university student management system, participating in the questionnaire of the university students, innovation and entrepreneurship professional course results of the top 20% of the number of people accounted for 52.137%, the top 50% of the total accounted for 86.325%, so the questionnaire filled in the group of college students with good results, with a variety of development potential.

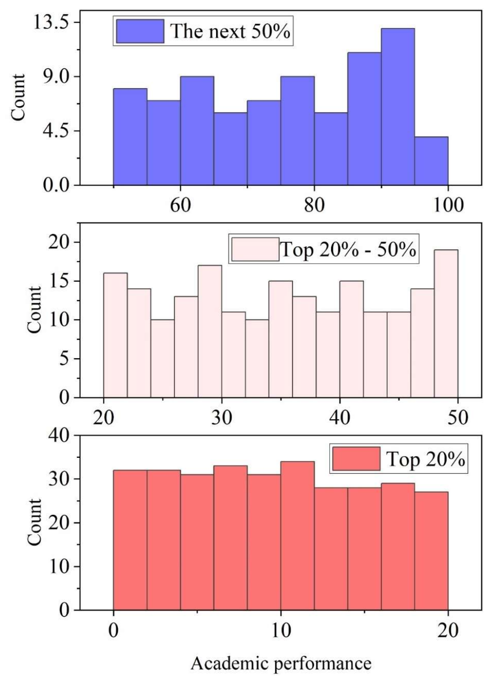

Figure 5: Achievement statistics

The 585 people in the questionnaire have basically the same purpose of learning, basically to improve self-competence, and Table 2 shows the evaluation of self-learning ability, and the proportion of those who learn for improving self-competence reaches 87.69%. In terms of self-restraint, only 52.14% of the students were able to adhere to independent study. The proportion of college students who adhere to independent study is much smaller than the proportion of college students who study in the hope of improving self-competence.

Under the premise of independent study, only 16.24% of the students can do frequent study summarization, and 76.92% of the students only summarize occasionally. Although 52.14% of the students are able to insist on independent study, the proportion of college students who are able to summarize their study is less than 20%. 76.92% of the students feel average or unsatisfied with their study plan, and only 11.62% of the college students have a strong plan.

Table 2: Evaluation of autonomous learning ability

Learning purpose	Microscope	Proportion
Self-improvement	513	87.69%
For graduation	56	9.57%
Not how to study	16	2.74%
Self-discipline	Microscope	Proportion
Stick to your own learning	305	52.14%
Autolearning	183	31.28%
Rarely study independently	49	8.38%
Basic learning	48	8.21%
Learning summary	Microscope	Proportion
Regular summary	95	16.24%
Occasional summary	450	76.92%
Almost no conclusion	40	6.84%
Learning plan	Microscope	Proportion
Well planned	68	11.62%
Satisfaction	150	25.64%
General	300	51.28%
Discontent	30	5.13%
Without plan	37	6.32%

III. C. 2) Practical skills

Science and technology competition is an important reference standard to measure the practical hands-on ability of college students, and the professional competitions of innovation and entrepreneurship category mainly include professional design competition, simulation design competition, enterprise management challenge, Internet + Express Innovation and Entrepreneurship Competition and so on. The statistics of science and technology competitions are shown in Figure 6. The analysis shows that most of the students have not participated in science and technology competitions, 60% of the students are those who have not participated in national and above level competitions, 57.265% of the students have not participated in provincial competitions and above level competitions, and 64.957% of the students have not participated in municipal and school level competitions and above. Those who have participated in 1-3 competitions are in the minority, and those who have participated in 5 or more competitions are 2.393%, 3.419% and 16.239% respectively. It can be concluded that the practical ability of these 585 university students is average, and more students got the awards of science and technology competitions at the municipal and university level.

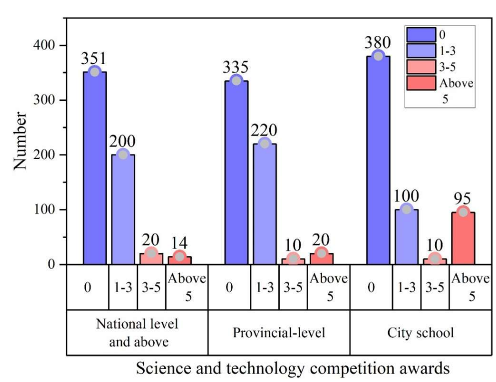

Figure 6: Science and technology competition statistics

The distribution of internship gains is shown in Figure 7. Among the college students who have participated in corporate internships, those who are very satisfied with their realized gains account for 30.77%, and those who are satisfied account for 25.641%, so it can be obtained that college students who have participated in internships have basically gained more.

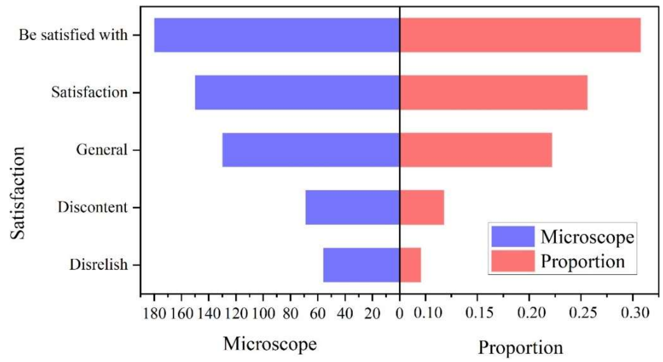

Figure 7: Internship distribution

III. C. 3) Evaluation of innovation, entrepreneurship and research capacity

Table 3 shows the evaluation of innovation and entrepreneurship and scientific research ability, college students' innovation and entrepreneurship is a direction strongly supported by the state, college students' academic literacy and professional skills are stronger than those of college students, and the probability of success in entrepreneurship is relatively high, of the 585 questionnaires, the percentage of those who have actual entrepreneurial experience is 15.04%, and 84.96% of the students do not have actual entrepreneurial experience. The proportion of college students with entrepreneurial ideas is low in the college student group, on the one hand, it is the shackles of academic thinking, on the other hand, it is the increase in age that leads to the decline of entrepreneurial passion. The entrepreneurial ideas are not really put into practice, and this situation is in line with the actual situation of the majority of college students nowadays.

Scientific research ability is a basic ability of college students, and scientific research ability should be an important reference for evaluating college students' learning ability. A total of 408 people, accounting for 69.74%, are interested in scientific research, and 177 people, accounting for 30.26%, are not interested in scientific research. The majority of university students are interested in scientific research, which is closely related to better academic performance. The number of those who have been granted patents is 128, accounting for 21.88%, and those who have not applied for scientific research projects account for 69.57%, which is the majority. To summarize, among these 585 college students, those who really have scientific research ability are a minority, accounting for only about 20% or even less.

Table 3: Evaluation of innovation and research ability

Learning purpose	Have	Nothing
Entrepreneurial analogy	110(18.8%)	475(81.2%)
Innovation and entrepreneurship project declaration	207(35.38%)	378(64.62%)
Entrepreneurship training	180(30.77%)	405(69.23%)
Actual entrepreneurial experience	88(15.04%)	497(84.96%)
Patent application	Microscope	Proportion
Authorized	128	21.88%
Under review	130	22.22%
In conception	150	25.64%
Disinterest	177	30.26%
Scientific research project	Microscope	Proportion
Applied to the national level	28	4.79%
Apply to the provincial level	50	8.55%
Apply for the city school level	100	17.09%
No application	407	69.57%

IV. Research on the differentiated training of university students in innovation and entrepreneurship and pathway planning

IV. A. Combination of practical and simulation platforms

The prerequisite of innovation ability is the ability to transform professional and comprehensive theoretical knowledge into practical activities. Based on the many constraints that the off-campus internship base must fully respect the normal production operation and management of the original unit in the arrangement of the internship time, internship projects and number of interns, the on-campus experimental simulation platform is an important supplement to the off-campus practice base.

IV. B. Combination of course practice and innovation and entrepreneurship competition

Innovation and Entrepreneurship Competition is an authoritative platform to test the innovation and practice ability of students. In order to create a good ecological environment for innovation and entrepreneurship for students in colleges and universities, every year, education authorities, local governments at all levels, societies and associations, and colleges and universities have been organizing innovation and entrepreneurship competitions in various forms around the frontiers of social and economic development. For example, “Internet+” Innovation and Entrepreneurship Competition for College Students, with the theme of “‘Internet+’ Achievement of Dreams, Innovation and Entrepreneurship Open Up the Future”, is aimed at leading entrepreneurship through innovation and promoting employment through entrepreneurship. The cultivation of innovation and practice ability of postgraduates should give full play to the advantages of product innovation and management innovation, and postgraduates of different majors should be formed into teams to participate in the competition, fully communicate with others through the Innovation and Entrepreneurship Competition, transform innovative ideas into innovative achievements, and actively look for entrepreneurial opportunities.

IV. C. Industry-University-Research-Political Collaboration Fostering Platforms

The collaborative innovation practice platform based on industry-university-research has been widely promoted in universities. With the development of university entrepreneurship parks, the cultivation of graduate students' innovation practice ability should make full use of the platforms of university science and technology parks, students' entrepreneurship practice bases, and entrepreneurship nurseries, to support entrepreneurial students to move into entrepreneurship bases and give them guidance.

V. Conclusion

This paper constructs an ecosystem model of innovation and entrepreneurship education in colleges and universities, establishes a decision tree assessment model based on the information gain algorithm, and selects relevant features to collect and preprocess data on college students' innovation and entrepreneurship education data. The factors affecting the innovation and entrepreneurship effect of differentiated teaching are divided into two aspects: enterprise prospect and whether it belongs to the high-tech field. In the enterprise prospect, the Gini index of the national scholarship is the largest at 0.648, Value= [0.248,0.5], which indicates that under this classification rule, 24.8% of the students choose the 1st option, and 50% of the students choose the 2nd option. Under the field of high and new technology, the second influential position is participation in scientific research activities and academic performance, with Gini indexes of 0.545 and 0.546, respectively, indicating that the general universities have the same influence factors on college students' entrepreneurship. At the same time, to assess the effect of differentiated training based on the top 20% of the number of people in the top 20% of the number of grades in the innovation and entrepreneurship professional courses accounted for 52.137%, and the top 50% of the top 50% of the total accounted for 86.325% of the students surveyed had better grades in innovation and entrepreneurship skills, with a variety of potential for development.

References

[1] Abiodun, T., Rampersad, G., & Brinkworth, R. (2023). Driving industrial digital transformation. Journal of Computer Information Systems, 63(6), 1345-1361.

[2] Savastano, M., Amendola, C., Bellini, F., & D’Ascenzo, F. (2019). Contextual impacts on industrial processes brought by the digital transformation of manufacturing: A systematic review. Sustainability, 11(3), 891.

[3] Yakymova, L., Novotná, A., Kuz, V., & Tamándl, L. (2022). Measuring industry digital transformation with a composite indicator: A case study of the utility industry. Journal of International Studies, 15(1).

[4] Ghosh, S., Hughes, M., Hodgkinson, I., & Hughes, P. (2022). Digital transformation of industrial businesses: A dynamic capability approach. Technovation, 113, 102414.

[5] Tashenova, L. V., Babkin, A. V., & Mamrayeva, D. G. (2019). Digital transformation of industrial production in the context of Industry 4.0. Bulletin of the Karaganda university Economy series, 96(4), 154-162.

[6] Ghobakhloo, M., & Iranmanesh, M. (2021). Digital transformation success under Industry 4.0: a strategic guideline for manufacturing SMEs. Journal of Manufacturing Technology Management, 32(8), 1533-1556.

[7] Zhang, J., & Chen, Z. (2024). Exploring human resource management digital transformation in the digital age. Journal of the Knowledge Economy, 15(1), 1482-1498.

[8] Teng, Y., Zheng, J., Li, Y., & Wu, D. (2024). Optimizing digital transformation paths for industrial clusters: Insights from a simulation. Technological Forecasting and Social Change, 200, 123170.

[9] Li, X., Chen, Z., & Chen, Y. (2024). The Impact of Digital Talent Inflow on the Co-Agglomeration of the Digital Economy Industry and Manufacturing. Systems, 12(8), 317.

[10] Bai, X., & Wan, Y. (2025). Evaluation method of innovation and entrepreneurship education quality in colleges and universities based on entropy weight method. International Journal of Innovation and Sustainable Development, 19(1), 81-93.

[11] Chen, Q. (2025). Evaluation method of service quality of college students' innovation and entrepreneurship education based on AHP-DEA method. International Journal of Innovation and Sustainable Development, 19(1), 26-42.

[12] Zhang, X., & Li, W. (2025). University Innovation and Entrepreneurship Education. Neutrosophic Sets and Systems, 355.

[13] Iddris, F. (2025). Entrepreneurship education on international entrepreneurship intention: the role of entrepreneurship alertness, proactive personality, innovative behaviour and global mindset. Journal of Applied Research in Higher Education, 17(2), 640-662.

[14] Wang, Z. (2025). Research on Integration of Innovation and Entrepreneurship Education Resources Based on Embedded Neural Network Algorithm. International Journal of High Speed Electronics and Systems, 2540337.

[15] Liu, Q., An, X., & Chen, W. (2025). College students' entrepreneurship education path and management strategy of start-up enterprises using causal attribution theory. Scientific Reports, 15(1), 2706.

[16] Khalid, A. (2025). An Analysis of Teaching Practice Courses in Entrepreneurship Education for Economics and Management. European Journal of Contemporary Education and E-Learning, 3(1), 41-54.

[17] Liu, Z., Zhang, M., Guo, Y., Mao, T., Deng, S., & Li, Y. (2025). Entrepreneurship education stimulates entrepreneurial intention of college students in China: A dual-pathway model. The International Journal of Management Education, 23(2), 101107.

[18] Zhang, X. (2025). Decision-Making in Efficiency Evaluation of Resource Allocation in University Innovation and Entrepreneurship Education. Neutrosophic Sets and Systems, vol. 77/2025: An International Journal in Information Science and Engineering, 355.

[19] Xue, S. (2025). An evaluation method of innovation and entrepreneurship ability of college students based on implicit knowledge mining. International Journal of Business Intelligence and Data Mining, 26(1-2), 174-189.

[20] Yu, Y., Gu, H., Liang, B., Chen, X., Chen, Z., & Wang, L. (2024). Performance Evaluation of School Enterprise Collaborative Innovation and Practice of Innovation and Entrepreneurship Education Based on the Improved AHP Fuzzy Comprehensive Evaluation Method. Discrete Dynamics in Nature and Society, 2024(1), 5583728.

[21] Imran, F., Shahzad, K., Butt, A., & Kantola, J. (2021). Digital transformation of industrial organizations: Toward an integrated framework. Journal of change management, 21(4), 451-479.

[22] Zhang, W., Zhao, S., & Wan, X. (2021). Industrial digital transformation strategies based on differential games. Applied Mathematical Modelling, 98, 90-108.

[23] Gigova, T., Valeva, K., & Nikolova-Alexieva, V. (2019, March). Digital transformation–opportunity for industrial growth. In 2019 International Conference on Creative Business for Smart and Sustainable Growth (CREBUS) (pp. 1-4). IEEE.

[24] Danuso, A., Giones, F., & da Silva, E. R. (2022). The digital transformation of industrial players. Business Horizons, 65(3), 341-349.

[25] Wang, W., Qiu, D., Chen, X., & Yu, Z. (2023). An empirical study on the evaluation system of innovation and entrepreneurship education in applied universities. Computer Applications in Engineering Education, 31(1), 100-116.

[26] Wang, X. (2020). Research on the path of college students' innovation and entrepreneurship education. Open Journal of Social Sciences, 8(3), 298-305.

[27] Saji, B. S., & Nair, A. R. (2018). Effectiveness of innovation and entrepreneurship education in UAE higher education. Academy of Strategic Management Journal, 17(4), 1-12.

[28] Wang, L. S., Li, Y., Sun, H. F., & WANG, L. (2018). The construction and operation of "Five-oriented" cultivation model for innovative and entrepreneurial quality of college students. US-China Educ. Rev, 8, 248-259.

[29] Nan Zheng,Yurong Li,Wuxiang Shi & Qiurong Xie.(2025).Sparse Bayesian based NARX modeling of cortical response: Introducing information entropy for enhancing the stability. Neurocomputing,626,129569-129569.

[30] Lukas Magnaguagno,Stephan Zahno & Ernst Joachim Hossner. (2025). The information gain of explicitly provided over self-generated contextual knowledge for behavioral control. PloS one,20(2),e0318994.

[31] Benjamín Luna Benoso,José Cruz Martínez Perales,Úrsula Samantha Morales Rodríguez,Rolando Flores Carapia & Víctor Manuel Silva García. (2024). A New Classification Model Using a Decision Tree Generated from Hyperplanes in Dimensional Space. Applied Artificial Intelligence,38(1).