



Construction of a Multidimensional Assessment Model for Enhancing Students' Innovative and Entrepreneurial Abilities Driven by Digital Transformation

Tingting Du^{1,*}

¹ Personnel Office, Shanghai Customs College, Shanghai, 201204, China

SUMMARY: *In the current landscape of digital transformation, objectively measuring and fostering students' innovative and entrepreneurial competencies has emerged as a central mission of the educational community. Aiming at the defects of extreme subjectivity and lack of dynamics of the existing assessment methods, this paper constructs a comprehensive assessment model integrating AHP hierarchical analysis and BP neural network, and introduces the SHAP interpretation framework to reveal its decision-making mechanism. First, the multidimensional innovation and entrepreneurship ability evaluation indexes of students are established through feature engineering, and their weights are determined using AHP. The pre-processed samples are then fed into the BP neural network to obtain an accurate measure and evaluation of students' innovation and entrepreneurship potential. SHAP post hoc interpretation is used to uncover the role played by each assessment index. Experimental data demonstrate the results show that the average prediction error between model outputs and measured values of the model in this paper is as small as 0.815% which is an indication of high precision. It also confirms that the innovation and entrepreneurial competence is among the major dimensions that determine the level of students' overall innovation and entrepreneurship proficiency. It implements the dynamism behind individual innovation and entrepreneurial abilities of each student, thereby supporting the smooth operation of the learners' careers.*

KEYWORDS: *AHP hierarchical analysis; BP neural network; SHAP interpretation framework; innovation and entrepreneurial capacity*

1 Introduction

Students' innovation and entrepreneurial potential have become one of the key elements of the competitive environment for modern university students, as these capabilities are not only associated with individual career development, but also serve as a vital engine driving social progress and economic prosperity [1, 2]. Consequently, accurately and rigorously evaluating university students' innovation and entrepreneurial potential carries significant importance across university education, corporate recruitment, and even policy-making domains [3, 4]. Nevertheless, establishing a systematic and comprehensive assessment framework for innovation and entrepreneurial capabilities among students remains a pressing issue in today's modern education system.

In recent years, researchers worldwide have devoted considerable efforts to studies concerning the evaluation of students' innovation and entrepreneurship capacity. Numerous scholars have proposed diverse assessment frameworks and index systems based on various

*Duting070302@163.com

<https://doi.org/10.65102/is2026445>

theoretical foundations, including models grounded in innovation and entrepreneurship capability theories and approaches built upon multiple intelligence theory [5-7]. These studies furnish a vital theoretical foundation and hands-on reference for the development of multi-dimensional evaluation models. However, there are still some problems in the existing studies. On the one hand, most of the existing evaluation models concentrate on students' knowledge mastery and problem solving ability, overlooking other essential competencies of students throughout the process of innovation and entrepreneurship, such as innovative thinking, psychological quality, self-knowledge, teamwork, and practical operation [8-11]. Conversely, the current evaluation model is not dynamic or flexible and thus cannot be used to meet the testing requirements of the skills related to innovation and entrepreneurship in different fields and at various levels. [12, 13]. Under the circumstances of the modern digital transformation period, the evaluation of innovation and entrepreneurship skills should not be restricted to one-dimensional traditional approach, but rather, an assessment framework that is multi-dimensional using the dimensions of technological literacy, innovative thinking, practical ability and risk response should be developed through quantitative measures combined with qualitative assessment [14-17]. The proposed framework offers two core advantages in this regard: first, it delivers a comprehensive depiction of students' innovation and entrepreneurial competencies within the evolving digital landscape, and second, it supports the implementation of evidence-based interventions to strengthen these skills. Most notably, the model enables the development of targeted remedial strategies to address specific gaps in students' innovation and entrepreneurship capabilities — a capability that is critical to advancing the overall quality of innovation and entrepreneurship education at the university level [18-21].

To address the growing demand for accurate and interpretable assessment tools in the digital era, this work develops a comprehensive evaluation model for students' innovation and entrepreneurial competencies using state-of-the-art machine learning methodologies. The modeling workflow begins with the Analytic Hierarchy Process (AHP), which is applied to objectively assign initial weights to each indicator in the multidimensional evaluation system. Next, a Back Propagation (BP) neural network, renowned for its powerful nonlinear fitting and self-learning properties, is trained on the prepared dataset to predict students' innovation and entrepreneurship capabilities. To address the "black box" nature of the neural network, the SHAP interpretability framework is adopted to analyze the optimized model, quantifying the contribution of each evaluation indicator to the final assessment outcomes. Empirical evaluations, encompassing both error analysis and interpretability validation, provide conclusive evidence of the model's high predictive accuracy and the transparency of its decision-making mechanism.

2 Multi-dimensional assessment model construction of students' innovation and entrepreneurship ability

2.1 Calculation method of weights of indicators of students' innovation and entrepreneurship ability

(1) In this paper, following the principle of odd number, a total of six experts are judged according to the relative importance comparison degree, and the judgment matrix of relative importance comparison degree is constructed.

(2) Calculate the weight.

For n elements, two-by-two comparison judgment matrix can be obtained.

$A = (a_{ij})_{n \times n}$, the row vectors of the constructed matrix undergo geometric averaging (square-

root method), and are subsequently normalized to derive the weights of each evaluating factor along with the eigenvectors W :

$$w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \left(\frac{a_{11}}{w_0}, \frac{a_{21}}{w_0}, \dots, \frac{a_{n1}}{w_0} \right)^T \quad (1)$$

Among them:

$$w_0 = \sum_{i=1}^n a_{i1} \quad (2)$$

(3) Consistency test.

To determine the logical consistency of thinking, first calculate the principal characteristic root:

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^n \frac{(AW)_i}{w_i} \quad (3)$$

where W denotes the weight value, A represents the matrix, and n signifies the order.

Then the consistency factor CI is computed:

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (4)$$

Finally, the concordance ratio CR is derived in accordance with Eq:

$$CR = \frac{CI}{RI} \quad (5)$$

Generally speaking,, when $CR < 0.1$, the matrix is considered having good consistency or alternatively the judgement matrix needs to be recalibrated. The weights of the system of factors affecting the innovation and entrepreneurship of students are derived using the judgement matrix and the calculation formula.

2.2 BP Neural Network Model

Neural networks represent a specialized category of machine learning models, designed to mathematically simulate the structural and behavioral patterns of biological nerve cells. The Back Propagation (BP) neural network, often regarded as the cornerstone of modern neural network architectures, is built on a robust theoretical foundation with proven computational reliability, and continues to serve as a practical solution for handling ill-defined and fuzzy problems across multiple domains.

The operational mechanism of the BP neural network can be summarized as follows: For a given set of input data, the model generates an output, which is then compared against the expected target value. The resulting prediction error is calculated iteratively, and the weights of all network parameters are continuously adjusted until the deviation between the actual and target outputs meets a predefined acceptable threshold. In essence, the model's logic revolves

around error assessment and parameter optimization, with the goal of minimizing the gap between predicted and expected outcomes to a specified level.

Fundamentally, the BP neural network relies on the gradient descent algorithm to align the model's outputs with the target function through the process of error backpropagation. Thanks to its inherent strengths in nonlinear mapping, generalization, and fault tolerance, coupled with its status as a mature, well-understood architecture, the BP neural network has found extensive application in research across a wide range of academic and industrial fields.

(1) Nonlinear mapping ability

Deploying the BP neural network model does not require comprehending or investigating the correspondence between input variables and output variables, and its application merely necessitates securing a sufficient volume of authentic and dependable sample data to underpin model training, whereupon it becomes capable of achieving the nonlinear mapping between input variables and output variables, which can eliminate a large number of causality derivation, and it has a certain degree of sophistication and automation.

(2) Generalization ability

Upon completion of BP neural network model training, newly introduced non-sample data fed into the trained model can still be correctly mapped in terms of the correspondence between input variables and output variables within non-sample scenarios, because the perceptron inside the neural network model has a certain degree of generalization ability, and it can apply the laws of sample data to non-sample data.

(3) Fault tolerance

The BP neural network model has considerable fault tolerance, which can screen out some abnormal samples in the training process, thereby safeguarding that the neural network model throughout the training procedure remains unaffected by irregular data inputs that could otherwise precipitate diminished accuracy or system instability.

The theoretical BP neural network, namely the error backpropagation algorithm, originates from the bionic simulation of biological neuron structures in natural nervous systems. Nodes distributed in different layers mimic biological nerve cells. With massive neuron units interconnected with one another, this network realizes bidirectional information interaction and feedback transmission, endowing the model with outstanding self-organization characteristics. Although the fundamental structure of the BP neural network is concise and its standard framework only contains three core layers, it possesses favorable fault tolerance and powerful generalization performance. Such merits enable it to be widely applied to various complex scenarios and fuzzy evaluation tasks. In its typical structural design, full connection is adopted among the input layer, hidden layer and output layer. As the core functional carrier of the whole system, the hidden layer determines the operational mechanism of the model and serves as a critical support for network training and optimization. Accordingly, the quantity of hidden layer neurons must be determined through repeated debugging and experimental verification in model construction. An insufficient number of hidden neurons will restrict the system's feature extraction capacity, resulting in inadequate training outcomes and poor model fitting. Conversely, excessive neuron settings will increase structural redundancy and further trigger the classic overfitting problem, which weakens the model's generalization performance in practical application.

Within the BP neural network model, the input vector associated with the input layer $A = (a_1, a_2, \dots, a_n)^T$, and the output vector linked to the hidden layer is $B = (b_1, b_2, \dots, b_n)^T$. The corresponding output vector of the output layer is $C = (c_1, c_2, \dots, c_m)^T$. Herein, the weights connecting the hidden layer to the output layer and the input layer to the hidden layer are denoted as w_i and v_i , respectively, the input reaching the k -th neuron in the input layer is

ins_k , and the input arriving at the j -th neuron in the hidden layer is ins_j . Based on the above assumptions, The computational formulas of the hidden layer nodes and the nodal output expressions of the output layer of the BP neural network model are given by the following equations respectively, and the output error can be minimized by modifying the corresponding weight values within the BP neural network through weight adjustment as illustrated in Eq. (9) and Eq. (10), respectively.

$$y_j = f\left(\sum_{i=0}^n v_{ij}x_i\right) = f(ins_j) \quad (6)$$

$$o_k = f\left(\sum_{j=0}^m w_{jk}y_j\right) = f(ins_k) \quad (7)$$

$$E = \frac{1}{2} \sum_{k=1}^l \left\{ e_k - f\left[\sum_{j=0}^m w_{jk}f\left(\sum_{i=0}^n v_{ij}x_i\right)\right] \right\}^2 \quad (8)$$

$$\delta_k^c = -\frac{\partial E}{\partial ins_k} = -\frac{\partial E}{\partial o_k} f'(ins_k) \quad (9)$$

$$\delta_j^y = -\frac{\partial E}{\partial ins_j} = -\frac{\partial E}{\partial o_j} f'(ins_j) \quad (10)$$

According to Eq. (9) and Eq. (10), the weights of the output and hidden layers of the BP neural network model are adjusted as Eq. (11) and Eq. (12), respectively.

$$\Delta w_{jk} = \eta (e_k - o_k) o_k (1 - o_k) y_j \quad (11)$$

$$\Delta v_{ij} = \eta \sum_{k=1}^l \delta_k^c w_{jk} y_j (1 - y_j) x_i \quad (12)$$

In order to further enhance the objectivity of the assessment framework for students' innovation and entrepreneurship competencies and introduce additional refinements to AHP, this paper integrates the pertinent characteristics and operational principles of the BP neural network, and employs the BP neural network as a quantitative scoring instrument for appraising students' innovation and entrepreneurship ability.

2.3 SHAP Interpretation Framework

To address the “black box” nature of complex machine learning models, Shapley Additive Explanations (SHAP) offers a theoretically grounded interpretability framework rooted in coalitional game theory. It quantifies each input variable's contribution to a model's predictions by assigning SHAP values, which encode both the direction and strength of a feature's impact on the final output. This framework operates on the principle that every input feature contributes to the model's prediction. SHAP values provide a standardized metric to assess how each feature influences the prediction outcome, enabling cross-validation with domain knowledge and real-world experience to evaluate the model's outputs for consistency and credibility. In

practical terms, a positive SHAP value indicates that the corresponding feature increases the predicted value, while a negative value suggests a suppressive effect that lowers the prediction. The magnitude of the SHAP value (absolute value) directly corresponds to the feature's relative importance in the model. Mathematically, the predicted value for any sample is defined as the sum of the model's base value (typically the mean of the target variable across the entire dataset) and the sum of the SHAP values of all input features. Within this framework, the base value represents the model's prediction in the absence of feature-specific effects, while the output refers to the final predicted score for a given sample. The input features in this work include PBA, OBT, FP, and V_{s20} , whose values collectively determine the model's output. SHAP decomposes the model's output into the additive contributions of these four features, with the output equal to the base value plus the SHAP values of all features. In this study, PBA and FP exhibit positive SHAP values, driving predictions upward, while OBT and V_{s20} show negative values, reducing the predicted outcome. Additionally, SHAP provides a robust way to rank feature importance: features with larger absolute SHAP values exert a more substantial influence on the model's behavior.

Formally, let x_i represent the i th sample, x_{ij} its j -th feature, y_i the model's prediction for x_i , y_{base} the global baseline (mean target value of all samples), and $f(x_{ij})$ the SHAP value of the j th feature for the i th sample. The SHAP values are governed by the following equation:

$$y_i = y_{base} + f(x_{i1}) + f(x_{i2}) + \dots + f(x_{ij}) \quad (13)$$

y_i is the model output value of 0.4, y_{base} is the model baseline value of 0.1, and $f(x_{i1})$, $f(x_{i2})$, $f(x_{i3})$ and $f(x_{i4})$ are SHAP values for PBA, OBT, FP and V_{s20} in that order.

The formula for $f(x_{ij})$ is as follows:

$$f(x_{ij}) = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|(|F| - |S| - 1)!}{|F|!} [v(S \cup \{j\}) - v(S)] \quad (14)$$

Here, F refers to the complete set of all influencing features in sample x_i , while S denotes an arbitrary subset of these features. The function $v(S)$ captures the total contribution from the combined influence of all features within subset S , and $v(S \cup \{j\}) - v(S)$ measures the marginal contribution of feature j when added to subset S ,

For Impact Factor j its SHAP value $f(x_j)$ is derived by computing the mean contribution of this feature across the entire dataset, as expressed in the formula below:

$$f(x_j) = \sum_{i=1}^M f(x_{ij}) \quad (15)$$

To interpret SHAP values intuitively, researchers can leverage visualization techniques at both the variable and sample levels, including density scatter plots, feature dependency plots, and force plots. The feature density scatter plot is particularly useful for visualizing the distribution of feature contributions across the entire dataset, with features sorted in descending

order of importance from top to bottom. Each data point on the plot represents the SHAP value of a single sample, encoding both the sign and magnitude of the feature's contribution to that specific prediction. Positive values correspond to features that increase the predicted outcome, while negative values correspond to those that decrease it. The absolute value of each SHAP value quantifies the strength of the feature's influence, and the overall pattern of the point distribution reveals the direction and magnitude of the feature's average effect on the model's predictions.

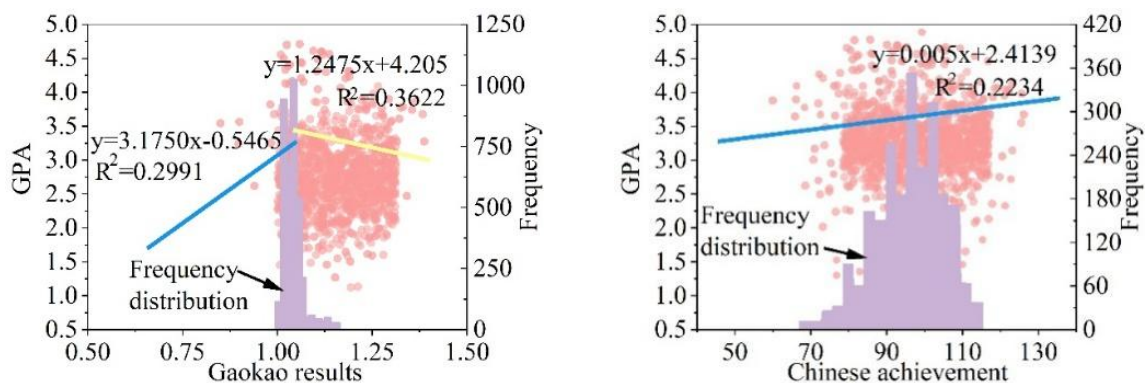
3 Construction of assessment system for students' innovation and entrepreneurship ability

The data for the study were information and data on innovative and entrepreneurial competencies of a total of 4,000 students enrolled in 2022 and 2023 in a local institution in Province A. The data of the students used were anonymized and not related to personal privacy.

3.1 Partial Characterization Engineering Description

The combined scores of the college entrance examination and the GPA have a two-stage transition, the first stage where the score control ratio is 1.04 or less and the GPA value exhibits a positive trend due to the growth in the score control ratio; the second stage where the score control ratio is more than 1.04 and the GPA value tends to decrease as the score control ratio increases. The majority of the GPA values are within the interval of 1.03 -1.05 whereas there are many students whose GPA is below 1.04, which is high, 69.1%.

The average values of the corresponding GPAs of language, mathematics and foreign language scores obtained with 5-point intervals are shown in Fig. 1, and Figs. (a)~(d)) show the distributions of the GPAs of the GCE point-control ratio, the GCE language score, the GCE mathematics score and the GCE English score, respectively. The slopes of the curves fitted to the gaokao language ($R^2=0.2234$, Pearson=0.45), mathematics ($R^2=0.5724$, Pearson=0.74) and foreign language ($R^2=0.6335$, Pearson=0.82) scores are positive, and all of them show a positive effect on the GPA performance, among which the English scores are the most obvious, with the largest slope of the curve The best fit was found in English.



(a) College entrance examination control ratio

(b) Chinese score

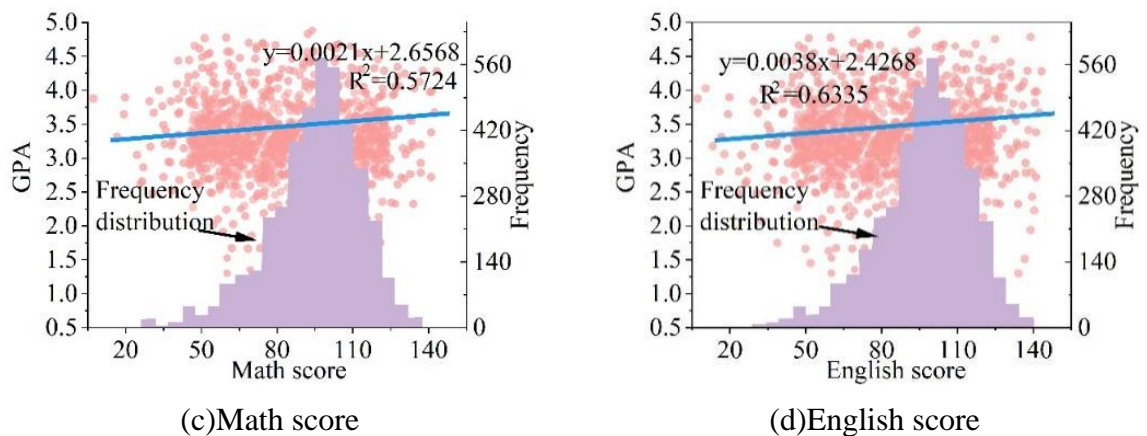


Figure 1: Distribution of grades and GPA

Candidates are divided into four categories, namely, urban freshmen, rural freshmen, former urban freshmen and former rural freshmen, to further analyze the distribution of GPA of the four categories of students. The distribution of GPA is shown in Fig.2, the mean and median of GPA of the rural freshmen-urban freshmen-rural former students-former urban freshmen show a slightly decreasing trend, suggesting that a discernible association exists between the rural and urban freshmen and former students and their academic performance. A small negative tendency, indicating that a significant correlation can be made out between rural and urban freshmen and former categories and academic performance. It is in line with what other researchers have found, namely, that the academic performance of urban and freshmen and former categories and academic performance differs. It is in line with the findings of other researchers that there are disparities in the academic performance of urban and rural children.

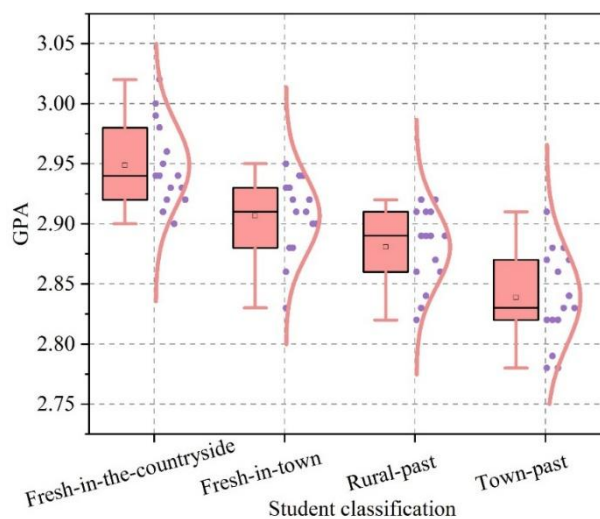


Figure 2: GPA distribution

The students' initial goals for graduation include six goals, which are numbered as A1~A6, such as self-employment, no idea or no consideration of employment, freelancing, direct employment, enlisting in the army and continuing to higher education, etc. Table 1 and Fig. 3 show the distribution of the students' graduation goal expectations and the relationship between the frequency distribution of the expectations of graduation goals, respectively. According to Table 1 and Figure 3, there is a certain connection between graduation goals and students' academic performance, in which the students who want to continue their studies have the best

academic performance, which is obviously better than the students with other goals, while the students who want to start their own business have the worst academic performance, and the proportion of this group of students in the low GPA group is as high as 72%.

Table 1: The students' graduation goals are expected to be distributed

Group	A1	A2	A3	A4	A5	A6	Spearman
Low GPA	72	58	65	60	53	51	-0.82
Middle GPA	26	39	30	36	41	44	0.85
High GPA	2	2	5	4	6	5	0.92

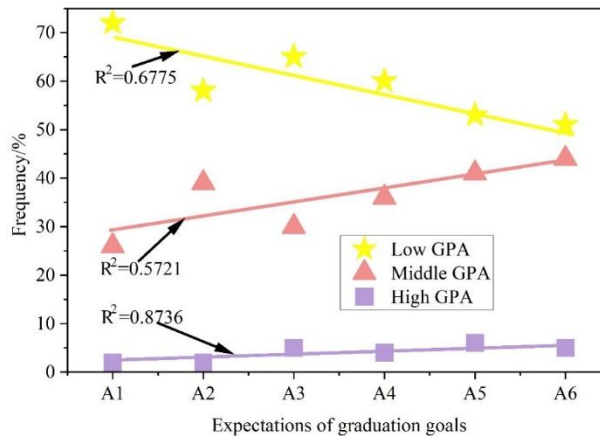


Figure 3: Graduation target expected frequency distribution

3.2 Establishment of the evaluation index system of innovation and entrepreneurship capacity

The evaluation index system of students' innovation and entrepreneurship ability that has been established in this paper comprises of five main indicators and sixteen secondary indicators indicated in Table 2.

Table 2: Student innovation and entrepreneurship ability rating system

Primary indicator	Secondary indicator
Innovation innovation A1	Innovation entrepreneurship A11
	Creative entrepreneurship A12
	Innovative entrepreneurial thinking A13
Innovation and entrepreneurship A2	Professional knowledge reserve A21
	Knowledge mobility A22
	Comprehensive knowledge management A23
Innovative ability A3	Scientific ability A31
	The ability to resolve the problem A32
	Knowledge and information acquisition force A33
Entrepreneurial ability A4	Opportunity recognition A41
	Entrepreneurial ability A42
	Integrated resource ability A43
	Risk control capability A44
Practical ability A5	Innovative entrepreneurship training A51
	Professional skills practice A52
	Results of innovation and entrepreneurship practice A53

4 Results of the assessment of students' innovative and entrepreneurial abilities

4.1 Calculation of weights for assessment indicators

Leveraging the data of primary-level indicators within the students' innovation and entrepreneurship ability evaluation framework, this study employs the Analytic Hierarchy Process (AHP) to derive the weights of each assessment index, as detailed in Table 3. Among all first-level indicators, "innovation and entrepreneurship quality" emerges as the most influential dimension with a weight of 0.3425. At the second level under this category, "innovation and entrepreneurship thinking" stands out with the highest weighting coefficient of 0.3694.

Table 3: Innovation and entrepreneurship ability assessment index weight

Primary indicator	Secondary indicator
Innovation and entrepreneurship A1(0.3425)	Innovation entrepreneurship A11(0.3411)
	Creative entrepreneurship A12(0.2895)
	Innovative entrepreneurial thinking A13(0.3694)
Innovation and entrepreneurship A2(0.3254)	Professional knowledge reserve A21(0.3112)
	Knowledge mobility A22(0.1985)
	Comprehensive knowledge management A23(0.4903)
Innovative ability A3(0.1953)	Scientific ability A31(0.2645)
	The ability to resolve the problem A32(0.6543)
	Knowledge and information acquisition force A33(0.0812)
Entrepreneurial ability A4(0.0789)	Opportunity recognition A41(0.1425)
	Entrepreneurial ability A42(0.2488)
	Integrated resource ability A43(0.1467)
	Risk control capability A44(0.4620)
Practical ability A5(0.0579)	Innovative entrepreneurship training A51(0.4213)
	Professional skills practice A52(0.3541)
	Results of innovation and entrepreneurship practice A53(0.2246)

4.2 Pre-processing of assessment sample data

Pre-processing of the collected data of students' innovation and entrepreneurship competency evaluation, the weighting of evaluation indexes for each data entry in the original dataset is inconsistent with this paper, and to resolve the inconsistency issue, the ratio of the original score to the weight of the original evaluation indexes is adopted as the processed score. This step enables the construction of a comprehensive dataset of students' innovation and entrepreneurship ability scores. Following data cleaning and validation, 955 valid samples are obtained, which are then divided into a training set of 700 samples and a test set of 255 samples. A portion of the processed dataset is illustrated in Table 4.

The range of data magnitudes in each indicator is very diverse, and when the raw data values are used in analysis, the indicators with the least values would be underweighted and those with the greatest values would be overemphasized. To improve the accuracy of the assessment, the samples relating to the students' ability to innovate and entrepreneurship are subjected to normalization treatment. There are two main methods of data normalization: the first one is the MaxAbsScaler normalization model, which is largely based on the ratio of the values in the column to the largest value in the column, that is, in case of an initial data value of d , the

normalized data becomes $d/|d_{max}|$, and the data is limited to the interval of [-1,1] once normalized. Another type of MinMaxScaler normalization is computing the ratio of the difference between the data and the minima of the data in the column over the range of the data in the column and controlling the outcome between [0,1] as given by the following equation:

$$d^* = \frac{d - d_{min}}{d_{max} - d_{min}} \tag{16}$$

The formula is defined by letting d be the original data, d_{max} be the maximum value of the data, d_{min} be the minimum value of the data, and d^* be the normalized output. As all the available information regarding the scores attained in the area of student innovation and entrepreneurship is positive, the paper uses the MinMaxScaler method of normalization to prepare the data.

Table 4: College students' ability to innovate and entrepreneurship (part)

N	A11	A12	A13	A21	A22	A23	A31	A32	A33	A41	A42	A43	A44	A51	A52	A53	Total
1	62	60	77	73	37	80	50	92	33	26	72	24	72	38	98	31	64.33
2	60	60	78	98	41	91	50	80	35	57	93	35	96	56	71	41	71.87
3	63	60	81	76	43	83	54	79	35	16	82	45	84	65	86	26	66.8
4	60	62	84	79	40	81	49	81	41	45	89	39	85	75	75	39	67.94
5	61	59	76	78	40	81	49	83	32	0	79	43	66	36	79	34	63.04
6	60	61	82	77	43	91	50	80	37	35	76	48	54	43	79	40	67.73
7	65	58	80	75	40	80	46	90	34	0	79	31	84	43	97	25	63.25
8	61	61	78	72	38	80	50	87	38	30	71	44	80	39	91	32	66.02
9	58	61	74	75	42	78	48	84	32	0	83	39	81	42	84	29	61.46
10	75	78	79	74	61	82	99	82	46	48	76	76	88	99	94	49	71.36
11	56	62	78	91	41	78	50	80	35	21	54	25	77	43	77	23	68.31
12	58	63	85	98	40	76	97	93	33	78	58	51	65	93	89	30	75.03
13	59	77	82	79	40	90	98	92	48	35	89	33	91	98	91	52	69.67
14	60	61	82	76	43	91	52	90	34	11	64	29	95	46	96	32	65.61
15	63	64	90	104	41	81	99	89	33	47	97	49	84	96	84	20	74.6
16	64	60	78	64	42	82	98	82	33	0	61	44	91	94	85	32	62.26

4.3 Assessment model testing and analysis of assessment results

To evaluate the performance of the proposed assessment model, the test dataset is used for validation, with key results documented in Table 5. In this analysis, the actual value denotes the true innovation and entrepreneurship ability score of each student, which objectively reflects their capability level, while the predicted value is the corresponding output generated by the model based on the test data. The model's generalization ability is considered robust when the relative error between predicted and actual values is controlled within an acceptable threshold.

The agreement between predicted and actual scores is further visualized in Figure 4, where the solid line plots the model's predictions and the dotted line shows the ground-truth values. The strong overlap between the two curves provides clear evidence of the model's reliable predictive performance.

The deviation between the forecasted and observed values across 255 test data entries is depicted in Fig. 5, wherein the peak relative error stands at 1.7947%, the lowest relative error is 0.0264%, and the mean relative error amounts to 0.878%. The assessment model has a certain

degree of universality and fault tolerance, and it carries a meaningful degree of referential value for the evaluation of students' innovative and entrepreneurial ability.

Table 5: Test outcomes of the evaluation model (part)

Sample number	Predictive value	Actual value	Relative error
1	0.7989	0.8135	1.7947%
2	0.7962	0.8069	1.3261%
3	0.7340	0.7281	0.8103%
4	0.7566	0.7568	0.0264%
5	0.7735	0.7759	0.3093%
6	0.7340	0.7273	0.9212%
7	0.7630	0.7624	0.0787%
8	0.7726	0.7691	0.4551%
9	0.7999	0.8130	1.6113%
10	0.8438	0.8562	1.4483%

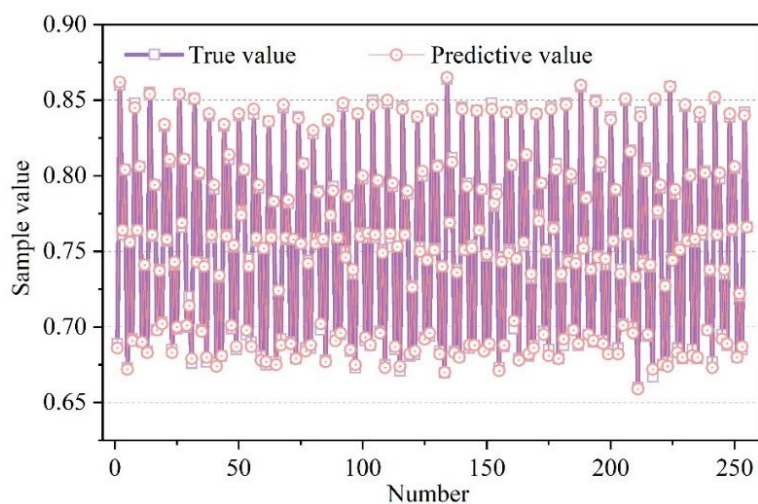


Figure 4: The comparison between the predicted score and the actual score

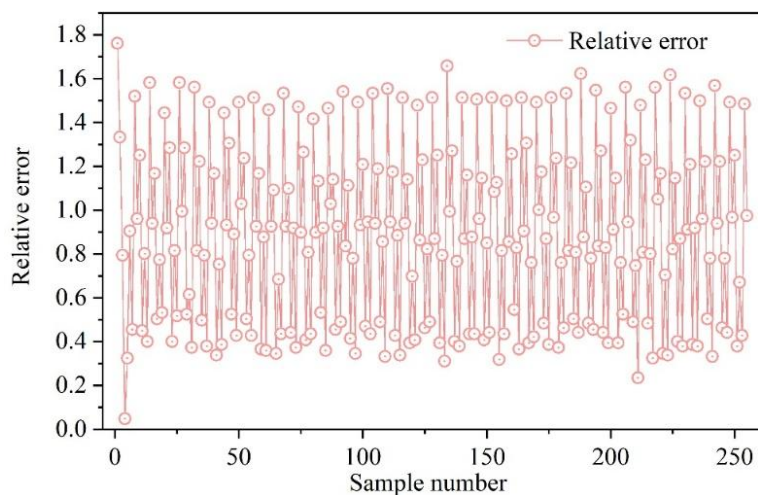


Figure 5: The relative error of the predicted value and the actual value

4.4 Ex post facto interpretability analysis

To realize global interpretation of variables, the SHAP framework is utilized, with the results presented in Fig. 6. The process involves calculating the SHAP value for each sample within each variable, and the variable's contribution magnitude is quantified as the mean of the absolute SHAP values. The analysis is conducted on a dataset of 4000 students, who were selected from three classes taught by the same instructor. Among all variables in the dataset, three key dimensions—innovation and entrepreneurship quality, innovation and entrepreneurship knowledge, and innovation ability—exhibit the strongest influence on students' overall innovation and entrepreneurship ability. Of note, the average absolute SHAP value for innovation and entrepreneurship quality is 1.37. The primary indicators under these three core dimensions—such as innovation and entrepreneurship spirit, awareness, and thinking; professional knowledge reserve; knowledge transfer and comprehensive management capabilities; scientific research ability; problem discovery and resolution skills; and knowledge and information accessibility—play a more prominent role in shaping students' innovation and entrepreneurship ability. This method enables the precise quantification of each feature's contribution weight in the model and the assessment of the extent to which each indicator affects different aspects of innovation and entrepreneurship ability.

Furthermore, feature contribution patterns can be further explored via the summary plot, with the analytical results depicted in Figure 7. Features are ordered along the y-axis in a top-to-bottom manner, corresponding to a descending sequence of their contribution magnitudes. Each point in the summary plot is associated with a single sample: the x-axis reflects the SHAP value of a specific feature for that sample, while the purple color intensity of the point indicates the feature's value (darker hues represent higher feature values). This plot facilitates the determination of each feature's positive or negative impact on innovation and entrepreneurship ability by integrating color cues and SHAP value distributions. As shown in Figure 7, a horizontal examination reveals that an increase in the innovation and entrepreneurship quality principal component's value is accompanied by an increase in its SHAP value—confirming that this principal component exerts a positive influence on high-level innovation and entrepreneurship competence. Notably, the three original indicators (innovation and entrepreneurship spirit, awareness, and thinking), which account for higher weights within the innovation and entrepreneurship quality principal component, also play a positive role in promoting high-level innovation and entrepreneurship competence.

The larger the eigenvalue of entrepreneurial competence principal component, the smaller the SHAP value, indicating that the eigenvalue of entrepreneurial competence principal component makes negative contribution to high-level innovation and entrepreneurship ability, meaning that the two original indexes, entrepreneurial implementation ability and risk control ability, have more significant impact on entrepreneurial competence principal component, negatively contribute to high-level innovation and entrepreneurship capacity. From the stated above, it can be hypothesized that innovation and entrepreneurship spirit has three attributes, namely, innovation and entrepreneurship awareness and innovation and entrepreneurship thinking, which are positive attributes whose values in terms of innovation and entrepreneurship will be projected to be high. Entrepreneurial implementation ability and risk control ability are both negative attributes projected to have high levels of innovation and entrepreneurship. It is obvious that the first factor of quality of innovation and entrepreneurship has a positive impact on the level of innovation and entrepreneurship ability, and the first factor of entrepreneurship ability has a negative impact on the level of innovation and entrepreneurship ability according to the principal components.

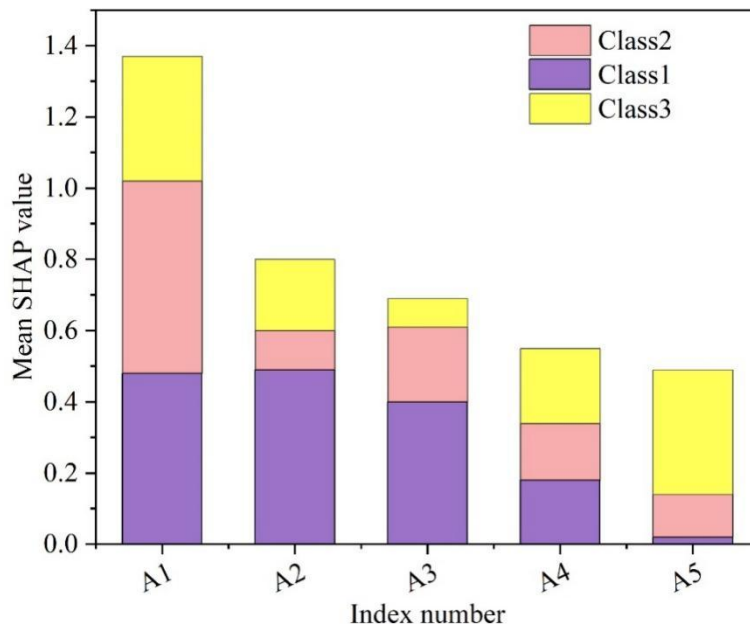


Figure 6: SHAP provides a global explanation for variables

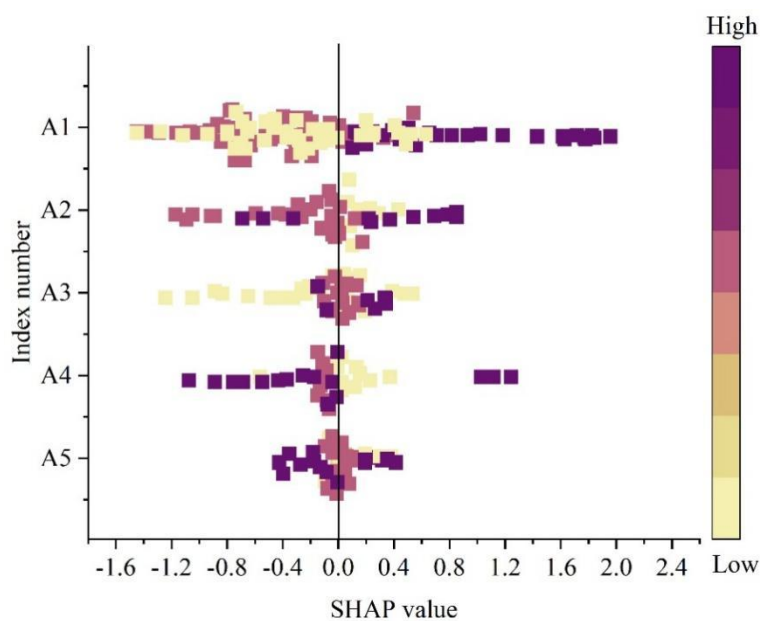


Figure 7: Feature summary results

SHAP further supports local interpretation, which generates conclusions about the significance of individual students' behavioral traits—information that can guide decision-making practices. Local interpretation begins with an individual sample, elucidating how SHAP values affect the predictive performance of the complex model. After completing model training, several samples were randomly chosen to showcase the SHAP interpretation mechanism. The interpretability analysis results for the model's predictions of Sample 1 and Sample 2 are depicted in Figure 8 and Figure 9, respectively. In these figures, purple regions correspond to features with positive SHAP values, while yellow regions represent features with negative SHAP values. The area of each region directly reflects the absolute value of the SHAP value: a larger area signifies a more pronounced impact on model predictions and a more influential role in the prediction process. This relationship can be calculated using the SHAP

value formula: the sum of the SHAP values of all features and the sample’s baseline equals the model’s predicted probability, which serves as the model’s output. As shown in the visualizations, Sample 1 has a SHAP value of 0.082, and its entrepreneurial ability feature is colored red—indicating this principal component acts as a positive factor for the model’s predicted value, leading to an increase in the prediction probability. In contrast, the features of innovation and entrepreneurship quality, innovation and entrepreneurship knowledge, and innovation ability are colored blue, meaning they are negative factors that reduce the prediction probability. Since these three principal components are more significant drivers of the model’s predictive value, decision-makers should focus on these components when evaluating Sample 1.

Similarly, the SHAP value of sample 2 is -0.85, the innovation ability principal component feature of sample 2 is the positive feature of the model predicted value, the entrepreneurial ability and innovation ability feature is the negative feature of the model predicted value, and the entrepreneurial ability and innovation ability principal component feature has a large area, which indicates that it is the key feature of the model predicted value, so for the decision makers for sample 2, it is necessary to focus on the original indicators of the Scientific research capability, discovery and problem solving capability, knowledge and information accessibility, opportunity recognition capability and resource integration capability.

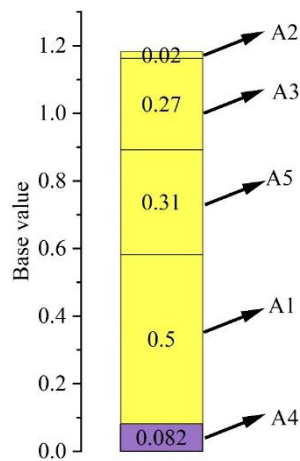


Figure 8: Interpretive analysis of sample 1

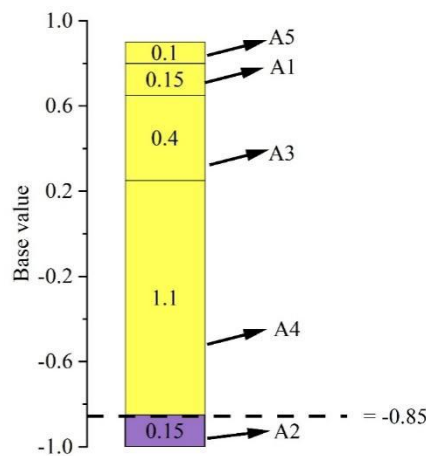


Figure 9: Interpretive analysis of sample 2

5 Conclusion

The present paper is based on the real survey findings regarding the degree of student innovation and entrepreneurship performance, consisting of 16 secondary indicators and the AHP algorithm to evaluate the importance of the factors and select 700 instances of valid data as training data and 255 data points as testing data, and in the case of the same metrics, the data of the evaluation samples are standardized, and the prediction of students innovation and entrepreneurship ability level is done with the help of the BP neural network model.

The integrated model of AHP - BP has shown a better quality of assessment and practical utility through the provision of scientific and rigorous values of the indicators and the results of the experiment on the accuracy test of the BP neural network in solving complicated nonlinear relationships show that the average relative error between the value predicted by the model and the actual value is only 0.878 percent and the maximum relative error is held at 1.8 percent and hence it can be concluded that it can act as a measure of evaluating the level of innovativeness and entrepreneurship skills of students It has been found to be useful as a measure of the degree of innovation and entrepreneurship skills of the students, and it is one of the most significant technical tools at the disposal of university faculty to choose and assess the level of innovation and entrepreneurship skills of their students.

SHAP interpretation framework has proven to be effective in visualizing and quantifying the evaluation process. SHAP interpretation framework produces a result that in the developed index system, the SHAP average score of quality of innovation and entrepreneurship is 1.37, which is the most important dimension behind the growth of the students' innovation and entrepreneurship skills and this revelation changes theoretical information into particular and intervenible quality aspects, which gives an orientation towards the practice of innovation and entrepreneurship education under the influence of digitalization.

About the Author

Tingting Du was born in Shandong, China, P.R. China, in March 1987. She obtained a Master's degree from East China University of Science and Technology in China. She is currently employed as an Assistant Research Fellow in Shanghai Customs College. I am currently employed in the Personnel Office, Shanghai Customs College. My main research direction is innovation and entrepreneurship among college students.

References

- [1] Wang, Y. (2022). Research on the Cultivation of College Students' Innovation and Entrepreneurship Ability Based on Subject Competition—Taking Logistics Management as an Example. *International Journal of New Developments in Education* (1.0).
- [2] Zhu, H. B., Zhang, K., & Ogbodo, U. S. (2017). Review on innovation and entrepreneurship education in Chinese universities during 2010-2015. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(8), 5939-5948.
- [3] Pan, H., Jiang, Y., Huang, Y., Zhang, Z., Du, J., Song, X., ... & Li, Y. (2025). A study on the mechanism affecting the Innovation and Entrepreneurship Ability of medical students based on Constructivist Theory: mediating role of Innovation and Entrepreneurship Willingness. *Frontiers in Medicine*, 12, 1630168.

- [4] Yingying, Z., Mei, W., & Dwen, S. (2023). Analysis of Innovation and Entrepreneurship Education in Chinese Universities. *Kuram ve Uygulamada Egitim Bilimleri*, 23(1), 15-27
- [5] Xu, X. (2022). The structural VPR model for the evaluation of innovation and entrepreneurship education. *Chinese Education & Society*, 55(4-5), 307-321.
- [6] Aykan, E. B., Fidanci, B. E., & Yildiz, D. (2022). Evaluation of entrepreneurship, individual innovation and obstacles to innovation in nursing students. *International Journal of Caring Sciences*, 15(3), 1792.
- [7] Ding, L., Chai, X., & Zeng, F. (2021). Evaluation of innovation and entrepreneurship ability of computer majors based on neural network optimized by particle swarm optimization. *International Journal of Emerging Technologies in Learning (iJET)*, 16(20), 19-34.
- [8] Purzer, S., Fila, N., & Nataraja, K. (2016). Evaluation of Current Assessment Methods in Engineering Entrepreneurship Education. *Advances in Engineering Education*, 5(1), n1.
- [9] Ye, P., & Yee, C. M. (2024, July). Design of college students' innovation and entrepreneurship education quality evaluation model based on grey relational degree algorithm. In *International Conference on Image Processing and Artificial Intelligence (ICIPAI 2024)* (Vol. 13213, pp. 797-802). SPIE
- [10] Yuan, Y., Zhang, H., Wang, J., & Li, X. (2018). A research on the quality evaluation of the innovation and entrepreneurship education at chinese universities based on linguistic operators. *Educational Sciences-Theory & Practice*, 18(5).
- [11] Sheng, Y., Yu, Q., & Chen, L. (2016). A Study on the Process Oriented Evaluation System of Undergraduate Training Programs for Innovation and Entrepreneurship. *Creative Education*, 7(15), 2330-2337.
- [12] Liang, G., Alghazzawi, D. M., & Joseph, N. R. (2022). The evaluation of college students' innovation and entrepreneurship ability based on nonlinear model. *Applied Mathematics and Nonlinear Sciences*, 7(1), 791-802.
- [13] Meng, Q., Xu, C., Cai, L., & Zhang, X. (2025, February). Design of College Students' Innovation and Entrepreneurship Competence Evaluation System in the Machine Translation Industry Based on the Onion Model. In *2025 6th International Conference on Computer Science, Engineering, and Education (CSEE)* (pp. 23-29). IEEE Computer Society.
- [14] Ma, X., & Pan, H. (2023). Improving entrepreneurial skills and professional association of college students using big data analysis and IoT. *Soft Computing*, 27(19), 14253-14267.
- [15] Huang, Y., Zhang, J., Xu, Y., & Bu, Y. (2024). How can digital innovation capabilities of Chinese college students be cultivated?. *The International Journal of Management Education*, 22(3), 101085.
- [16] Luo, Y., & Yang, H. (2024, August). Construction of Evaluation Model for College Students' Innovation and Entrepreneurship Ability Under the Background of Digital

Economy. In International Conference on Computational Vision and Robotics (pp. 311-324). Cham: Springer Nature Switzerland.

- [17] Lesinskis, K., Mavlutova, I., Spilbergs, A., & Hermanis, J. (2023). Digital transformation in entrepreneurship education: The use of a digital tool KABADA and entrepreneurial intention of generation Z. *Sustainability*, 15(13), 10135.
- [18] Yiyun, W. A. N., SOMJAI, S., & SOMJAI, A. (2024). THE APPLICATION OF DIGITAL TRANSFORMATION TO THE INTELLIGENT FINANCE STUDENT INNOVATION CAPABILITY. *Procedia of Multidisciplinary Research*, 2(11), 51-51.
- [19] Liu, Z., & Cui, A. (2025). An effective evaluation method of college students' innovation and entrepreneurship education based on data mining algorithm. *International Journal of High Speed Electronics and Systems*, 2540405.
- [20] Zuolong, Y., Hernandez, A. A., & Yiwen, W. (2025). The influence of innovation and entrepreneurship education in higher education institutions on the development of students' innovative thinking. *Edelweiss Applied Science and Technology*, 9(9), 155-171.
- [21] Ni, H., & Zhang, K. (2025). Structural Dimensions and Current Status Assessment of College Students' Innovation and Entrepreneurship Capabilities Based on Grounded Theory. *Critical Humanistic Social Theory*, 2(4).