Engineered Science
DOI: https://dx.doi.org/10.30919/es1779

A Comparative Study of Multi-Model Prediction of Deep Learning Ability-taking Middle School Mathematics Learning as an Example

Yang Yan,[1] Hao Huang,[2] Zengyi Yan,[3] Yujun Hu,[2] Bowen Zhang,[2] Wenxuan Duan,[2] Yiwei Huang[2] and Fei Li[1,*]

Abstract

Deep learning is increasingly being applied in educational projects. Predicting and evaluating deep learning in middle school mathematics is crucial for optimizing educational project practice. This study, based on Huberman's 5C deep learning model and the NESS-China scale, designed a questionnaire to investigate the current status of deep learning in mathematics among middle school students in Yantai, China. A prediction model for deep learning ability in mathematics was constructed using linear regression, decision tree, and neural network algorithms. The scores on each dimension of deep learning served as independent variables, and the students' overall learning ability level served as the dependent variable. Model performance was evaluated using mean squared error, mean absolute error, and coefficient of determination. The results showed that the linear regression model achieved the best predictive performance, indicating a significant linear correlation between students' deep learning ability and performance on each dimension. This finding suggests that teachers can focus on cultivating students' learning abilities in key dimensions to maximize their overall learning ability. This study provides new insights and empirical evidence for tailoring teaching to individual students and promoting deep learning in mathematics, and offers a fundamental research basis for the application of deep learning in educational projects.

Keywords: Deep learning; Academic performance prediction; Machine learning; Educational big data.
Received: 21 August 2025; Revised: 29 September 2025; Accepted: 06 October 2025
Article type: Research article.

1. Introduction

Deep learning, an innovative machine learning algorithm, was proposed by Professor Hinton of the University of Toronto in 2006.[1] Because its advocacy of active learning and the cultivation of higher-order thinking skills aligns closely with the core needs of education and teaching, the concept of deep learning has been adopted not only in the field of artificial intelligence (AI), but also in the field of education.[2] In 2010, while investigating the application of deep learning technology in educational data mining, Okoli *et al.* determined that deep learning could help solve research challenges in educational tasks and data mining.[3] In 2017, the New Media Consortium released its "Horizon Report", which expressed a consensus among global experts that deep learning algorithms will have a profound impact on educational transformation.[4] The U.S. Defense Advanced Research Projects Agency (DARPA) and the UK Engineering and Physical Sciences Research Council (EPSRC) have launched or funded a series of deep learning research projects to support basic research and talent development. The Chinese education community is also actively exploring the theory and practice of deep learning. The Ministry of Education launched the "Deep Learning" project in 2013, marking a new chapter in the theoretical exploration of deep learning in Chinese education. The following year, the Curriculum and Textbook Research Center of the Ministry of Education further established a "Deep Learning" teaching reform project team to provide theoretical guidance and practical cases for the promotion and application of deep learning in Chinese education.

[1]*School of Mathematics and Statistical Science, Ludong University, Yantai, Shandong, 264025, China*
[2]*School of Economics, Management and Law, Shandong University of Petrochemical Technology, Dongying, Shandong, 257061, China*
[3]*School of Physical Education, Ludong University, Yantai, Shandong, 264025, China*
[*]*Email:* miss_life@163.com (Fei Li)

Based on this background, deep learning theory has garnered widespread attention in educational research and educational engineering practice. However, the biggest obstacle currently facing this field is the lack of "data-driven thinking".[5] The "Blue Book on the Development of Big Data in Basic Education in China" points out that mining educational big data is a key approach to achieving smart teaching.[6-8] With the rapid development of educational big data, machine learning technology has provided new methods and perspectives for quantifying student learning behavior and predicting academic performance. For example, linear regression models have been used to analyze the relationship between student homework completion rates and mathematics scores;[9] neural network models have revealed the significant impact of course engagement on academic performance using student online learning behavior data;[10] and decision tree models have been used to classify student learning behavior characteristics, helping to identify the impact of different learning strategies on academic performance.[11] The application of these models provides powerful quantitative tools for educational research.

Although significant progress has been made in educational data analysis, research on predicting middle school students' deep learning ability in mathematics remains insufficient. Scientific and effective methods for evaluating deep learning ability are urgently needed to provide practical guidance for the application of deep learning in educational projects. The Statistical Product and Service Solutions (SPSS) Modeler platform offers nearly 20 predictive modeling algorithms, with linear regression, decision trees, and neural networks being commonly used data-driven modeling methods.[12] This study selected linear regression, decision tree, and neural network models to predict students' deep learning ability and compared the differences and applicable conditions of these models.[13] This study aims to apply data mining and machine learning techniques to provide educational policymakers with empirical evidence for cultivating students' deep learning ability, while also offering new perspectives and approaches for deepening educational reform and improving educational quality.

2. Research methods

2.1 Data collection

The sample data for this study was derived from a questionnaire survey on the current status of deep learning in mathematics among middle school students. This questionnaire integrates the five fundamental characteristics of deep learning with the core content of the Chinese College Student Survey (CCSS) questionnaire. It covers five primary dimensions: attitude and motivation, questioning and inquiry, integration and construction, transfer and application, and value and evaluation. It is further subdivided into 11 secondary dimensions, with a total of 25 questions.

2.2 Model construction

2.2.1 Linear regression model construction

The linear regression model is a statistical method used to explore the relationship between independent and dependent variables.[14] It is widely used in educational data analysis. It is one of the most commonly used regression analysis methods in the field of artificial intelligence. Its principle is to discover the linear relationship between feature variables and target variables by learning regression coefficients.[15,16] For example, Huang *et al.* compared the predictive capabilities of four mathematical models in their study of predicting academic performance in a mechanical engineering course, concluding that regression models can achieve a high degree of prediction accuracy.[17] Zhu *et al.* demonstrated the advantages of linear regression in processing complex survey data, providing a reference for complex data research in the field of education.[18] Shu *et al.* used regression analysis, neural networks, and other methods to conduct an in-depth analysis of the factors influencing university students' learning outcomes and satisfaction.[19] Adejo *et al.* integrated student academic records and behavioral data and found that the linear regression model had an accuracy rate of up to 97% in predicting academic performance.[20] Wang *et al.* proposed a hybrid modeling method combining linear regression with deep learning to predict multivariate relationships in complex systems.[21] Xiong *et al.* used linear regression to analyze the current status and influencing factors of graduate course construction, providing empirical support for the China educational reform of "Double First-Class" universities.[22] These studies have demonstrated that linear regression models are not only applicable to general educational data but can also provide a scientific basis for educational decision-making in complex scenarios.

Therefore, the linear regression model of this study uses the indicator data of each dimension of deep learning (X_1, X_2, X_3, X_4, X_5) as the independent variable and the student's comprehensive learning ability score (Y) as the dependent variable. The prediction formula is:[23]

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \epsilon \tag{1}$$

In Eq. (1), β_0 is the intercept, β_i is the regression coefficient, and ϵ is the error term, which is used to assess the accuracy of the simulated prediction results.

2.2.2 Decision tree model

A decision tree is a tree-like structure generated through training and can be used to solve classification and regression problems.[24-25] Although some scholars have pointed out the overfitting problem of decision tree models, this problem can be effectively addressed through the application of algorithms such as bagging, boosting, random forests, support vector machines, and neural networks.[26-28] For example, Altujjar *et al.* used a decision tree algorithm to predict student course grades;[29] Sharma *et al.* used the decision tree algorithm ID3 to construct an academic prediction model;[30] Ding *et al.* further pointed out that decision tree models excel in recognizing complex patterns in educational data and classifying student behavior;[31] Tang *et al.* used the classification and regression tree (CART) algorithm to analyze the efficiency of scientific and technological achievement transformation, revealing key factors such as the degree of technology demand matching and demonstrating its practicality in analyzing complex variable relationships;[32] Li *et al.* combined tree models with qualitative analysis to propose a hybrid modeling framework suitable for multivariate analysis.[33]

In this study, the decision tree model used the CART algorithm, selected dimensional data features such as "initiative level" as decision nodes, and predicted students' comprehensive ability scores by hierarchically partitioning the data features. Node division is based on the Gini Index, which is calculated as follows:[34]

$$\mathrm{Gini(D)} = 1 - \sum_{i=1}^{n} p_i^2 \tag{2}$$

In Eq. (2), p_i is the probability of the i class. The study selects split features by maximizing information gain and constructs a tree structure layer by layer.

2.2.3 Neural network model construction

Neural network models, with their powerful nonlinear fitting capabilities, can capture complex patterns in data. For example, Usman *et al.* constructed an artificial neural network model to predict graduation scores based on students' past grades;[35] Shi *et al.* used a Back Propagation Neural Network (BPNN) to achieve dynamic early warning of unsafe behavior risks, demonstrating its predictive power in complex scenarios;[36] Ding *et al.* constructed an AI literacy evaluation system for college students based on neural networks, expanding student ability assessment methods;[37] Wang *et al.* revealed the critical impact of online course participation and content access on academic achievement.[38]

This neural network model adopts a three-layer structure: the input layer receives data from various dimensions for deep learning, the hidden layer neurons process and transform the data through activation functions, and the output layer generates the final prediction results. During model training, the weight parameters are dynamically adjusted through the backpropagation algorithm to continuously optimize model performance.

Eq (3) is the fundamental formula of the neural network model.[39] where f(x) is the activation function, g(x) is the output function, w_{ij}, v_j is the weight, and b_j, c is the bias. The model optimizes the weight parameters through the backpropagation algorithm.

$$h_j = f(\sum_{b_j} w_{ij} x_j + b_j) \tag{3}$$

The basic model formula is as follows (Fig. 1).

This diagram illustrates the structure and computational flow of a multi-layer neural network. After receiving the signal x_j, the input layer transmits it to the hidden layer, the hidden layer is adjusted by weighting with weights w_{ij} and adjusting the bias b_j, the activation function $f(\cdot)$ then produces the output h_j. Subsequently, the hidden layer output h_j is weighted by the weights v_j and adjusted by the bias c, ultimately yielding the network output $Y = \sum v_j h_j + c$. This demonstrates the signal propagation and nonlinear transformation process from input to output.

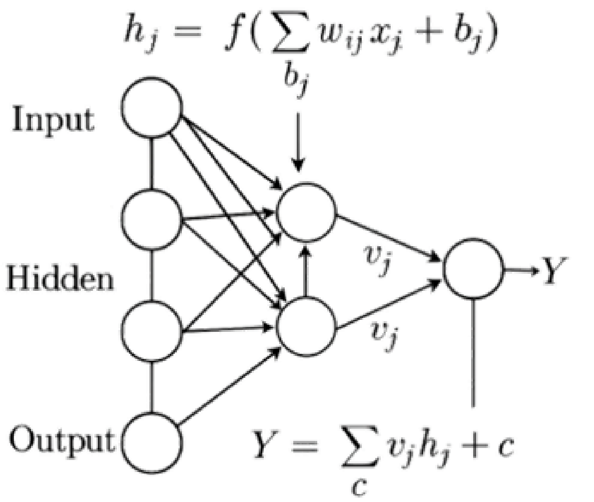

Fig. 1: The basic model formula

2.3 Model evaluation metrics

This study used mean squared error (MSE), mean absolute error (MAE), and sample coefficient of determination (R^2) as metrics for evaluating model performance.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{4}$$

In Eq (4), MSE measures the degree of model prediction bias by calculating the average of the sum of squared errors between the predicted and true values.[34] A smaller MSE value indicates a better model fit.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{5}$$

In Eq. (5), MAE calculates the average of the absolute values of the errors between the predicted value and the true value, which can intuitively reflect the average error.[40] The lower the value, the more accurate the prediction.

$$R^2 = 1 - \frac{1}{n}\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{6}$$

In Eq. (6), R^2 evaluates the proportion of the model's correct prediction results in the total prediction samples.[34] The higher the accuracy, the more reliable the model's predictions.

2.4 Data preprocessing and model training process

The collected questionnaire data was preprocessed, including cleaning, missing value handling, and standardization, to ensure data quality and provide reliable input data for model training.

The dataset was divided into a training set and a test set in a 7:3 ratio. The linear regression, decision tree, and neural network models were trained separately. During training, model parameters were continuously adjusted. For example, the linear regression model used the least squares method to fit the optimal regression coefficients, the decision tree model used criteria such as information gain to determine node partitioning rules, and the neural network model used the backpropagation algorithm to dynamically adjust weights to optimize model performance.

3. Results analysis

3.1 Linear regression model results

The linear regression model achieved a sample determination coefficient of 1.00, with MSE of 4.01e-28 and MAE of 1.73e-14 on the test set, indicating a significant linear relationship between the data and giving the model strong fitting ability.

3.2 Decision tree model results

The sample determination coefficient of the decision tree model is 0.54, with MSE of 14.90 and MAE of 3.03 While decision tree models can handle nonlinear data and, to some extent, compensate for the shortcomings of linear regression, they are prone to overfitting. Especially when the tree depth is too high, the model can overlearn details and noise in the training data, resulting in reduced generalization ability on the test set. Furthermore, the structure of the decision tree significantly influences prediction results. The order of node partitioning and standard deviation can lead to significant deviations in prediction results.

3.3 Neural network model results

The neural network model achieved a high sample coefficient of determination of 0.99, with MSE of 0.28, and MAE of 0.35. With its powerful nonlinear fitting capabilities, it excels at capturing complex data features and uncovering complex relationships hidden within the data. However, neural network training is time-consuming and prone to falling into local optima, potentially failing to find the globally optimal weight configuration, which in turn affects the final prediction results.

3.4 Model comparison and optimal model selection

Comparing the prediction results and evaluation metrics of the three models (Table 1) shows that the linear regression model performed significantly better than the other models in this research data scenario. The linear regression model's sample R^2 reached 1.00, MSE was as low as 4.01e-28, and MAE was 1.74e-14. All indicators were far superior to those of the decision tree and neural network models. This shows that the linear regression model can fit the data almost perfectly, with a high degree of fit between the predicted values and the true values. Its prediction accuracy and stability are irreplaceable in this research scenario. This model can more accurately capture the complex relationships between the various factors affecting students' deep learning ability, truly reflect the changing trends in students' learning ability, and provide a more solid and reliable basis for teaching decisions.

Table 1: Evaluation metrics for the three models.

Model Name	R^2 Score	MSE	MAE
Linear regression model	1.00	4.01e-28	1.731e-14
Decision tree model	0.54	14.90	3.03
Neural network model	0.99	0.28	0.35

3.5 Visualization of model prediction results

To more intuitively compare the prediction performance of the three models, this study employed various visualization techniques.

3.5.1 Comparison of predicted and true values

Scatter plots are used to present the distribution characteristics of the predicted and true values for the linear regression, decision tree, and neural network models, respectively. The scatter plot for the linear regression model in Fig. 2a shows that the data points are closely clustered near the diagonal, indicating a high degree of agreement between the predicted and true values, visually demonstrating its superiority in predicting deep learning ability in middle school mathematics. The scatter plot for the decision tree model in Fig. 2b shows a relatively dispersed distribution of data points, scattered around the fitted line, indicating a significant deviation between the predicted and true values. The scatter plot for the neural network model in Fig. 2c shows a relatively concentrated distribution of data points, but still contains a small number of discrete points. This indicates that while its prediction performance has significantly improved, its accuracy is still slightly inferior to that of the linear regression model.

3.5.2 Error distribution boxplots

Boxplots present the distribution of prediction errors for the three models (Fig. 3). The linear regression model's boxplot shows a very small error range, with the median very close to 0, indicating extremely low error dispersion and a very small overall error. The decision tree model's boxplot shows the largest error range, with the median deviating significantly from 0, indicating high error dispersion and a relatively large overall error. The neural network model's boxplot shows a significantly smaller error range than the decision tree model, with the median close to 0, indicating a relatively stable and small error. This further validates the linear regression model's prediction accuracy.

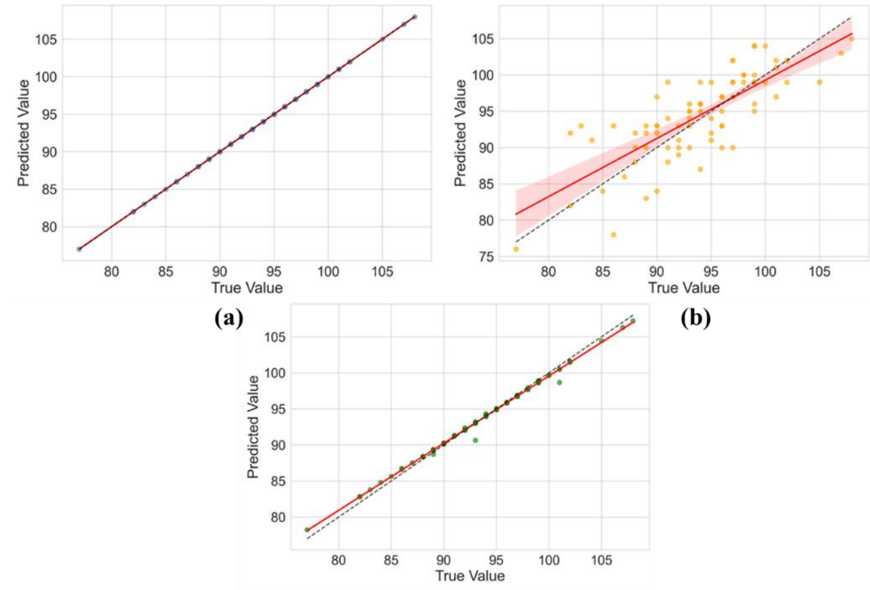

Fig. 2: Scatter plots of three models (a) Scatter plot of linear regression model (b) Scatter plot of the decision tree model (c) Scatter plot of the neural network model.

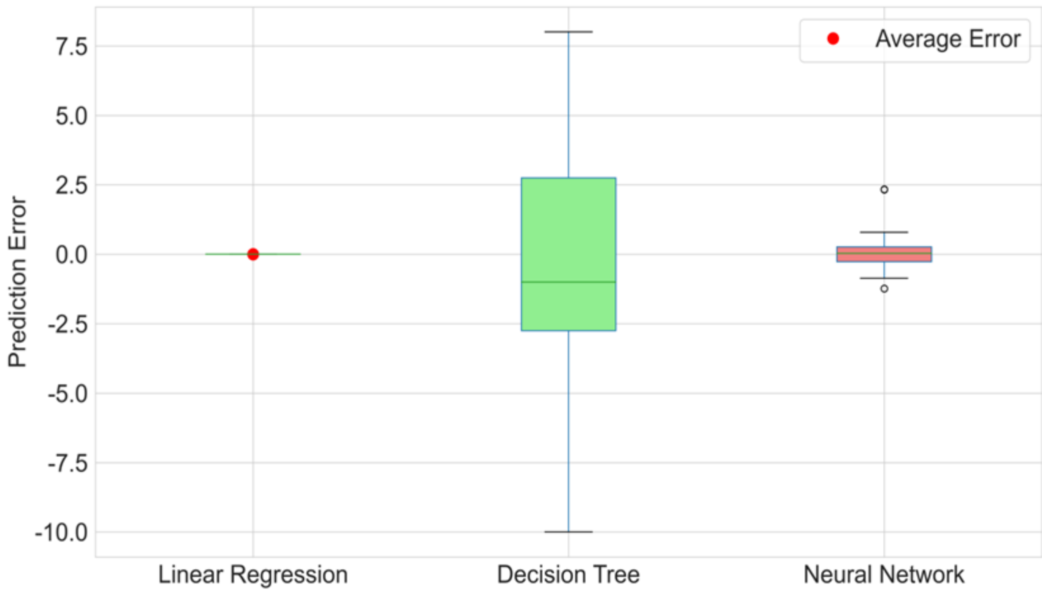

Fig. 3: Distribution of prediction errors of the three models.

3.5.3 MSE and MAE Comparison Charts

The mean squared error and mean absolute error of the three models are compared using box plots as shown in Fig. 4. The figure clearly shows that the linear regression model's box plot has the smallest average error between the predicted and true values, with the median very close to 0. This indicates that the linear regression model fits the data best, has extremely low error dispersion, and offers the most accurate and reliable predictions. The decision tree model's box plot has the largest error range, with the median farther from 0, indicating high error dispersion and a relatively large overall error, reflecting the decision tree model's poor performance, low prediction accuracy, and unreliability. The neural network model's box plot shows that its error range is significantly smaller than that of the decision tree model, with the median close to 0, indicating that its error is relatively stable and small, its predictions are more accurate, and its model is more reliable, but its prediction accuracy and reliability are inferior to those of the linear regression model.

The visualization results further verify the accuracy of the linear regression model's predictions. The predicted values and true value scatter points of the linear regression model are closely clustered near the diagonal line. The error distribution box plot shows that the error range is extremely small and the median is close to 0, which confirms its advantage in capturing data patterns. This result is consistent with the research conclusion of Adejo *et al.*[20] that "linear regression shows high accuracy in predicting academic performance", indicating that

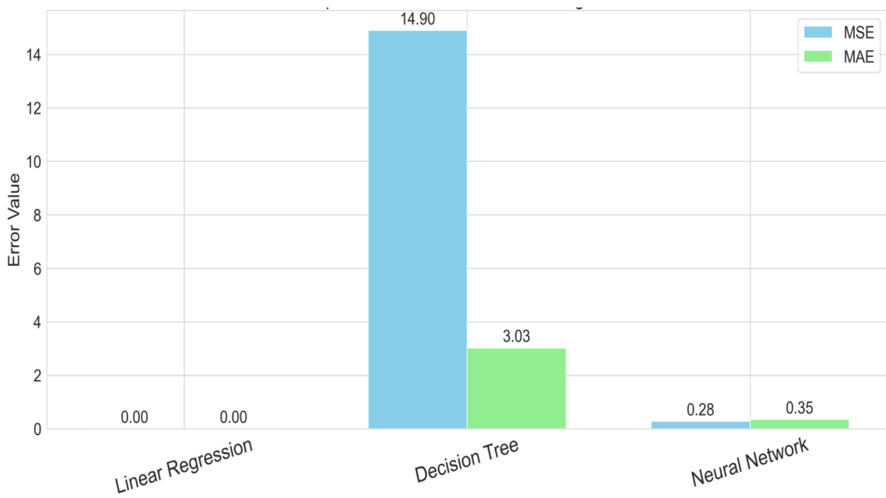

Fig. 4: Comparison of MSE and MAE of three models.

under data characteristics that conform to linear relationships, the linear regression model is an efficient tool for predicting small sample questionnaire data in the education field.

3.6 Other factors to consider

Through model comparison, this study clearly demonstrates the advantages of linear regression models in predicting middle school students' deep learning ability in mathematics, but there is still room for improvement. We are currently working to build a larger sample covering a wider range of regions and student groups, optimize the questionnaire design to incorporate more potentially influencing factors, and collaborate with schools to obtain data on students' learning processes. This is not yet fully resolved, and we are actively communicating with all parties to explore feasible data collection solutions. Furthermore, we will conduct in-depth optimization of different models. For decision tree models, ensemble learning methods can be used to reduce the risk of overfitting and improve predictive stability. For neural network models, attention mechanisms or deep learning frameworks such as Long Short Term Memory (LSTM) can be introduced to enhance the ability to capture nonlinear relationships. Interpretability tools such as SHapley Additive Explanations (SHAP) values can be combined to reveal the influence paths of various dimensions and crack the "black box" problem. We will also explore the adaptability of "simple models + complex scenarios", such as combining linear regression models with the needs of stratified student instruction to develop sub-models tailored to students at different learning levels.

Going forward, we will continue this work, striving to reveal the influencing mechanisms of deep learning ability from multiple perspectives and comprehensively, providing a more solid empirical foundation for the application of deep learning theory in middle school mathematics instruction. This will help promote the scientific, precise and personalized development of education, and provide new perspectives and paths for deepening educational and teaching reforms and improving the quality of education.

4. Conclusion

This study utilized questionnaire data on deep learning in mathematics among secondary school students to construct linear regression, decision tree, and neural network models. It systematically compared the predictive performance of these three models regarding mathematical deep learning ability. Results indicate that the linear regression model significantly outperformed the others in prediction, with visualization findings further validating its superiority in capturing data patterns. In contrast, while the decision tree model can handle nonlinear data, its predictive effectiveness was weaker due to overfitting issues. The neural network model excelled at capturing complex relationships but suffered from lengthy training times and susceptibility to local optima. Integrating these findings, educators can prioritize linear analysis methods in practice to precisely identify key factors influencing students' deep learning abilities. Decision trees or neural networks can then be employed to assist in analyzing potential patterns within complex data scenarios, thereby providing personalized learning support for students. This study enriches research on predicting deep learning abilities in education, validating the hypothesis that “the formation of deep learning abilities is a linear cumulative process involving the synergistic interaction of multiple dimensions.” The findings not only provide theoretical foundations for optimizing teaching strategies but also advance big data applications in education. They offer crucial references for educational policy formulation and teaching practices while laying the groundwork for future research exploring the adaptability of educational data and models.

Acknowledgments

The authors are very grateful for financial support from National Statistical Science Research Project (2024LY013).

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

CRediT Statement

Yang Yan: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing - Original draft. **Hao Huang:** Supervision, Methodology, Writing - Review and editing. **Zengyi Yan:** Writing - Review and editing, Validation. **Yujun Hu:** Writing - Review and editing, Validation. **Bowen Zhang:** Writing - Review and editing, Visualization. **Wenxuan Duan:** Writing - Review and editing, Visualization. **Yiwei Huang:** Writing – Review and editing, Resources. **Fei Li:** Funding acquisition, Project administration, Writing - Review and editing, Correspondence.

References

[1] G. E. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets, *Neural Computation*, 2006, **18**, 1527-1554, doi: 10.1162/neco.2006.18.7.1527.

[2] D. X. Chen, Y. Y. Zhan, B. Yang, Application analysis of deep learning technology in educational big data mining, *Journal of Educational Research*, 2019, **40**(2), 68-76, doi: 10.13811/j.cnki.eer. 2019. 02. 009.

[3] C. Okoli, K. Schabram, A guide to conducting a systematic literature review of information systems research, *SSRN Electronic Journal*, 2010, **10**, 1-49, doi: 10.2139/ssrn.1954824.

[4] S. A. Becker, M. Cummins, A Davis, *et al.*, NMC horizon report: 2017 library edition, Austin: *The New Media Consortium*, 2017, 60, ISBN: 978-0-9977215-7-7.

[5] J. Manyika, M. Chui, B. Brown, *et al.*, Big data: The next frontier for Innovation, Competition and Productivity, *Analytics*, 2011.

[6] Q. Zhou, C. Mou, D. Yang, A review of research progress in educational data mining, *Journal of Software*, 2015, **26**(11), 3026-3042, doi: 10.13328/j.cnki.jos.004887.

[7] R. S. Baker, P. S. Inventado, Educational data mining and learning analytics, *Learning Analytics*, Springer, New York, 2014, 61-75, doi: 10.1007/978-1-4614-3305-7_4.

[8] Z. M. Shu, X. D. Xu, Data mining and analysis of college students' satisfaction education under the perspective of learning analytics, *Research on Educational Technology*, 2014, **5**, 39-44, doi: CNKI: SUN: DHJY.0.2014-05-009.

[9] F. C. Zhu, The relationship between mathematics performance and mathematics homework of students in East Asia, *Journal of Mathematics Education*, 2019, **28**(02), 64-69, doi: CNKI: SUN: SXYB.0.2019-02-012.

[10] F. Q. Sun, R. Feng, Research on factors affecting online academic achievement based on learning analytics, *China Educational Technology*, 2019, **2**, doi: 10.3969/j.issn.1006-9860.2019.03.007.

[11] P. j. Zhao, Y. S. Cao, Method for identifying subway passengers' travelpurpose based on multi-source geographic big data and machine learning, *Journal of Geo-Information Science*, 2020, **22**(9), 1753-1765, doi: CNKI:SUN:DQXX.0.2020-09-002.

[12] S. P. Li, G. Q. Wei, Z. Y. Zhao, Comprehensive performance prediction method of college students based on multi-model fusion, *Control Engineering*, 1-8, doi: 10.14107/j.cnki.kzgc.20240049.

[13] G. Y. Ding, J. K. Cheng, Research on academic performance modeling of college students based on educational data mining, *Heilongjiang Higher Education Research*, 2020, **38**(2), 76-81, doi: 10.19903/j.cnki.cn23-1074/g.2020.02.015.

[14] J. X. You, Z. Sun, Research on Prediction and Intervention of College Students' Academic Performance on Cloud Learning Platform, *China Distance Education*, 2016, **9**, 14-20+79, doi: 10.13541/j.cnki.chinade.20160929.002.

[15] X. Su, X. Yan, C.-L. Tsai, Linear regression, *WIREs Computational Statistics*, 2012, **4**, 275-294, doi: 10.1002/wics.1198.

[16] G. James, D. Witten, T. Hastle, R. Tibshirani, J. Taylor, *An introduction to statistical learning: With applications in python*, Cham: Springer International Publishing, 2023, 69-134.

[17] S. Huang, N. Fang, Predicting student academic performance in an engineering dynamics course: a comparison of four types of predictive mathematical models, *Computers and Education*, 2013, **61**, 133-145, doi: 10.1016/j.compedu.2012.08.015.

[18] D. H. Zhu, K. Y. Wang, Application of Generalized Linear Mixed Model in Customer Satisfaction Research - Analysis of Customer Satisfaction of Bank Wealth Management Products in a Certain Region, *Statistics and Information Forum*, 2014, **29**(1), 94-99, doi: 10.3969/j.issn.1007-3116.2014.01.016.

[19] Z. M. Shu, Q. F. Qu, Analysis of College Students' LearningOutcomes Based on Educational Data Mining, *Journal of Northeastern University: Social Sciences Edition*, 2014, **3**, 312, doi: 10.15936/j.cnki.1008-3758.2014.03.014.

[20] O. W. Adejo, T. Connolly, predicting student academic performance using multi-model heterogeneous ensemble approach, *International Journal of Computer Science, Engineering and Information Technology*, 2017, **10**, 1-12, doi:10.1108/JARHE-09-2017-0113.

[21] Z. Wang, J. Sun, X. Li, A hybrid linear regression and deep learning approach for predictive modeling of smart manufacturing processes, *Engineered Science*, 2019, **6**, 20-26, doi:10.18488/97.1002/es.2019.6.20-26.

[22] L. Xiong, K. C. Lu, Y. Xu, *et al.*, Current status and path optimization of graduate course construction from the perspective of classification development - an empirical analysis based on a

double first-class construction university, *Graduate Education Research*, 2025, **2**, 51-60, doi: 10.19834/j.cnki.yjsjy2011.2025.02.06.019, (03): 48-54.

[23] C. Q. Li, W. T. Liu, Z. Y. Yan, *et al.*, Prediction Method and Application of Pressure Drop in Fiber Filter Rods Based on Multivariate Linear Regression and Fluid Mechanics, Science, *Technology and Engineering*, 2025, **25**(14), 5983-5991, doi: 10.12404/j.issn.1671-1815.2402603.

[24] X. D. Wu, V. Kumar, *The Top Ten Algorithms in Data Mining: Classification and Regression Trees*, Boca Raton, CRC Press, 2009, **9**, 179.

[25] C. Mauro Assis Gomes, G. C. Lemos, E. G. Jelihovschi, 87Avaliação psicológica, 2020, 19(1), pp. 87-96Comparing the predictive power of the CART and CTREE algorithms, *Revista Avaliação Psicológica*, 2020, **19**, 87-96, doi: 10.15689/ap.2020.1901.17737.10.

[26] A. N. Hu, X. G. Wu, Y. S. Chen, Analysis of treatment effect heterogeneity - opportunities and challenges brought by machine learning methods, *Sociological Research*, 2021, **1**, 91-114, 228.

[27] J. Franklin, The elements of statistical learning: data mining, inference and prediction, *The Mathematical Intelligencer*, 2005, **27**, 83-85, doi: 10.1007/BF02985802.

[28] P. Cortez, using data mining to predict secondary school student performance, *EUROSIS*, 2008, 5-12, https://api.semanticscholar.org/CorpusID:16621299.

[29] Y. Altujjar, W. Altamimi, I. Al-Turaiki, M. Al-Razgan, Predicting critical courses affecting students' performance: a case study, *Procedia Computer Science*, 2016, **82**, 65-71, doi: 10.1016/j.procs.2016.04.010.

[30] G. Sharma, K. Santosh, Analysis and prediction of student's academic performance in university courses, *International Journal of Computer Applications*, 2017, **160**, 40-44, doi: 10.5120/ijca2017913045.

[31] W. Ding, A. S. Abdel-Salam, M. Almasri, S. T. Waller, A machine learning approach to predict student performance in engineering education, *Engineered Science*, 2021, **15**, 103-112, doi: 10.18488/97.1002/es.2021.15.103-112.

[32] L. Y. Tang, S. Y. Xie, S. Y. Hu, Research on the factors affecting the transformation of scientific and technological achievements oriented by technology demand: A case study of 101 high-tech enterprises, *China Science and Technology Forum*, 2023, **4**, 16-24, doi: 10.13580/j.cnki.fstc.2023.04.013.

[33] X. Y. Li, J. W. Lian, X. Xue, *et al.*, Fusion tree model Exploration and Practice of Qualitative Comparative Analysis Method, *Journal of Library and Information Technology*, 2025, **69**(13), 43-55, doi: 10.13266/j.issn.0252-3116.2025.13.004.

[34] Y. Y. Zhao, Y. Y. Chen, B. Sang, Student Characteristics, School Characteristics and Student Academic Performance - An Empirical Study Based on Machine Learning Methods, *Education and Economy*, 2023, **39**(1), 47-58.

[35] O. L. Usman, A. O. Adenubi, Artificial Neural Network (ANN) Model for Predicting Students' Academic Performance, *Journal of Science and Information Technology*, 2013, **2**, 36, doi: 10.1016/j.caeai.2021.100018.

[36] J. Shi, D. Y. Chang, P. Zheng, Early warning model of unsafe behavior of construction workers based on BP neural network, *Journal of Safety Science*, 2022, **32**(1), 27-33, doi: 10.16265/j.cnki.issn1003-3033.2022.01.004.

[37] J. H. Ding, L. Y. Guo, W. X. Zhang, *et al.*, Construction of artificial intelligence literacy index system for college students based on AHP-BPNN method, *Journal of Distance Education*, 2025, **43**(1), 46-56, doi: 10.15881/j.cnki.cn33-1304/g4.2025.01.005.

[38] L. Wang, Y. Hao, S. Wang, Achievement prediction and analysis based on neural network for smart education, *Discover Education*, 2025, **4**, 161, doi: 10.1007/s44217-025-00606-3.

[39] N. Michael, *Neural Networks and Deep Learning*, Beijing: Determination Press, 2015.

[40] N. Wang, Q. Wang, Application Research of Deep Learning-Based Student Comprehension Prediction Models in Online Education Environments, *Data Analysis and Knowledge Discovery*, 2025, 1-17, https://link.cnki.net/urlid/10.1478.G2.20250219.1242.004.