



OPEN ACCESS

EDITED BY

Mostafa Aboulmour Salem,
King Faisal University, Saudi Arabia

REVIEWED BY

Pooja Saigal,
Vivekananda Institute of Professional Studies,
India
Gloria Concepcion Tenorio-Sepulveda,
Tecnologico Nacional de Mexico / TES de
Chalco, Mexico
Aleixandre Brian Duche-Pérez,
Catholic University of Santa María, Peru

*CORRESPONDENCE

Yan Yang
✉ 360226919@qq.com

RECEIVED 21 August 2024

ACCEPTED 05 August 2025

PUBLISHED 29 August 2025

CITATION

Yan L and Yang Y (2025) Assessing the
learning potential of freshmen in labor
education courses using ordinal features and
support vector machine.
Front. Educ. 10:1483964.
doi: 10.3389/feduc.2025.1483964

COPYRIGHT

© 2025 Yan and Yang. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Assessing the learning potential of freshmen in labor education courses using ordinal features and support vector machine

Long Yan and Yan Yang*

The College of Health Humanities, Jinzhou Medical University, Jinzhou, China

Introduction: Artificial intelligence (AI) marks a new wave of the information technology revolution and permeates various sectors as an indispensable tool. Despite its widespread adoption, its application in enhancing college students' labor education remains scantily explored. Conventional teaching approaches often fail to assess students' foundational knowledge accurately, impeding personalized learning. Hence, the current environment underscores the pressing necessity for a robust AI framework capable of reliably predicting individual students' learning aptitude.

Methods: In this study we constructed a multidimensional feature vector model, leveraging data on students' academic performance during their middle school years and their willingness to participate in college-level labor education. Through the usage of Support Vector Machines (SVM), we aim to assess students' learning potential effectively. To validate the efficacy of our predictive model, we conducted jackknife cross-validation testing.

Results: Results indicate a remarkable overall accuracy rate of 97.75%, with an average sensitivity of 93.90% and an average specificity of 95.12%.

Discussion: The proposed method can play a role in enhancing teaching efficiency and tailoring interventions to individual students.

KEYWORDS

AI, college labor, feature vector, support vector machine (SVM), jackknife cross-validation testing

1 Introduction

As a driving force of the latest technological revolution and industrial transformation, artificial intelligence (AI) stands as one of humanity's most remarkable and profound inventions (Han et al., 2022; Sun, 2017). It is reshaping production methodologies and societal dynamics, propelling us into an era of intelligent collaboration between humans and machines, characterized by cross-disciplinary integration, co-creation, and resource-sharing (Zhan and Yang, 2017; Fu and Zhou, 2020). Concurrently, the ubiquity of online course platforms has surged due to rapid information technology advancements (Oliveira et al., 2021; Aldowah et al., 2017; Bryson and Andres, 2020; Haleem et al., 2022; Chen et al., 2020). These platforms meticulously document student activities and performance metrics, accumulating vast repositories of educational data. However, despite this abundance, a significant portion of this data remains underutilized, lacking comprehensive mining to extract its latent value.

Data mining within the realm of artificial intelligence involves deploying algorithms to uncover implicit correlations, trends, and patterns from extensive datasets (Hand et al., 2001; Górriz et al., 2020; Moreno and Redondo, 2016; Beale, 2007). The concept of 'knowledge discovery in database' (KDD) was first proposed at the 11th International Joint

Conference on Artificial Intelligence in 1989, and the term ‘data mining’ was introduced at the 1st International Conference on Knowledge Discovery and Data Mining held in Canada in 1995. Data mining as an interdisciplinary approach utilizes machine learning, pattern recognition, statistics, databases, and visualization techniques to extract valuable information from large datasets. Despite its complexity, data mining equips decision-makers with profound insights, facilitating well-informed and precise decision-making processes (Rui et al., 2022; Hoogerwerf et al., 2013; Pournaras, 2017). Data mining is tasked with handling large-scale data that is usually incomplete, ambiguous, and randomly structured. Therefore, the following requirements should be satisfied to ensure an effective data mining process (Han and Kamber, 2006; Li et al., 2014; Li et al., 2015; Yang et al., 2022; Yang and Wang, 2013; Yu et al., 2012): (i) The data must be of sufficiently large sizes and real; (ii) The discovered knowledge must align with the user’s needs and be interpretable; and (iii) The discovered knowledge must be applicable for addressing specific problems.

Data mining is a crucial research field related to artificial intelligence, statistics, and databases, enabling the development of intelligent and data-driven information technology systems (Losiewicz et al., 2000; Wang and Fu, 2005; Han and Chang, 2002). The pivotal roles of data mining include elucidating extracted information from data in forms of concepts, rules, patterns, and constraints. This harvested intelligence serves to assist decision-making processes or refine existing knowledge paradigms, thereby enhancing utilization of resources within extensive databases. Presently, data mining occupies a prominent position in both academic and industrial spheres, gathering widespread international attention and interest. In the educational domain, data mining technologies hold considerable potential to collect, analyze, and report on students’ learning behaviors and outcomes, thereby improving learning environments and providing pertinent guidance to educators.

Labor education for university students bears significant importance, carrying the mission of nurturing practical and innovative talents essential for national and societal advancement (Huan, 2019; Jiang and Pan, 2019). Under the current circumstances, the enactment of labor education for college students holds multifaceted significance: From a theoretical standpoint, it fosters the strengthening and advancement of values rooted in labor; In the context of our current era, it aids in the development of modern labor skills essential for adapting to advancements in science and technology; In terms of cultural values, it supports the preservation and promotion of the rich tradition of valuing hard work and dedication; Pragmatically, it serves as a driving force behind societal development and advancement.

The traditional teaching approach tends to prioritize the teacher’s authority and rely heavily on textbooks and classroom lectures, focusing mainly on transferring knowledge (Wang, 2007; Yao, 2003; De Lorenzis et al., 2023). Typically, teachers lecture while students passively receive information. However, this method often struggles to assess whether students have truly grasped concepts learned in class, let alone whether they have enhanced their overall skills. Consequently, it becomes challenging to provide timely and objective evaluations of teaching effectiveness and student learning. While this traditional method may streamline teaching for instructors, it fails to accommodate the diverse backgrounds and academic capabilities of students. Particularly, the students with weaker foundations or

learning difficulties may not receive adequate support. To enhance teaching quality, educators must tailor their instructional approaches to suit individual student needs. For college students, traditional educational practices often fall short in evaluating the latent potential cultivated during middle and high school education, thus failing to meet the objective of precise instruction.

With the continuous advancement of artificial intelligence (AI) algorithms, it has become increasingly feasible to analyze and utilize large-scale datasets in greater depth (Callaway, 2024), thereby enabling accurate prediction and classification. This technological leap has introduced transformative changes to the field of education (Wang, 2024; Hu et al., 2024). AI not only assists educators in identifying students’ learning patterns but also facilitates the prediction of their future academic and professional potential. Accurately assessing students’ latent abilities developed during secondary and high school education remains a longstanding challenge for traditional educational approaches, while the integration of AI algorithms may offer an effective solution. Consequently, there is an urgent need for reliable and effective AI-based methodologies capable of accurately evaluating students’ fundamental learning potential.

In this study, we constructed a multidimensional feature vector integrating students’ middle school grades and their inclination towards learning. The Support Vector Machine (SVM) algorithm is then employed to evaluate this learning potential metric. Jackknife cross-validation testing is utilized to validate the method’s performance. The results demonstrate that our proposed technique achieves a high success rate in predicting student learning potential.

2 Materials and methods

2.1 Dataset

Between February 13th and February 20th, 2024, we collected data from all regions of China using random sampling method. This involved an online survey via self-filled questionnaires on the Internet. The inclusion criteria were voluntary participation of current and graduated college students in course evaluations. Through the use of questionnaires, we obtained grade data from the respondents.

1. General demographic information: including location and gender.
2. Grades in Chinese language, politics, and history courses during middle school and high school.
3. Willingness to participate in college labor education courses: categorized as general (60) or enthusiastic (90).
4. Based on the classification method of GPA for major (degree) courses in higher education institutions of China, college labor education course grades are categorized as follows: A – Excellent (≥ 91), B – Good (86–90), C – Average (76–85), and D – Fail (≤ 75).

We utilized the online survey tool “Questionnaire Star”¹ for data collection due to its efficiency, affordability, and user-friendly

¹ <https://www.wjx.cn/>

interface. This platform has been widely used across various survey-related fields. The data were obtained via self-administered questionnaires completed by participants with informed consent, thereby eliminating concerns regarding ethical violations or infringement of rights (Zhou et al., 2020; Zhu et al., 2020; Wu et al., 2020).

A total of 135 college students and graduates responded to the questionnaire. After filtering out incomplete responses, we obtained 123 valid questionnaires, resulting in a response rate of 91.11% (Table 1). Respondents were distributed across all regions of China (Figure 1A), with females comprising 57% of the sample (Figure 1B). The number of respondents reflects regional differences in educational development and population mobility. A relatively larger proportion of students were sampled from the more developed eastern regions, while fewer were sampled from the less developed western regions. The sampling distribution across other regions remained generally consistent. Furthermore, the gender and academic performance of students within the sample were relatively balanced. These characteristics support the reliability of our sampling methodology.

The grades were categorized into four classes: A, B, C, and D. Therefore, the dataset can be formalized as follows:

$$\Omega = S^A \cup S^B \cup S^C \cup S^D$$

Where \cup represents the symbol for “union” in the set theory.

2.2 Creating vectors with students' grades and ordinal information

Based on the features that may be associated with students' learning potential, as outlined in Section 2.1, we construct the following vector:

$$R_k = (r_1, r_2, r_3, r_4, r_5, r_6, r_7).$$

Instead of considering the elements within a vector as separate units, we incorporated the order relationship among students' secondary school results and their willingness. By sorting these results in ascending order, we created a new vector denoted as R_s . This process ensures that the order relationships among the data are preserved.

$$R_s = (r_{i_1}, r_{i_2}, r_{i_3}, r_{i_4}, r_{i_5}, r_{i_6}, r_{i_7}).$$

TABLE 1 Grade distribution for university students in labor education course.

| Score | Grade | Number | Percentage |
|------------------------------|--------------|--------|------------|
| ≥ 91 | A: Excellent | 28 | 22.76 |
| $86 \leq \text{and} \leq 90$ | B: Good | 33 | 26.83 |
| $76 \leq \text{and} \leq 85$ | C: Average | 28 | 22.76 |
| $75 \leq$ | D: Fail | 34 | 27.64 |

When two r_i s are equal, they are sorted by lexicographic order of the words, ensuring each element's ordinal number is unique. It implies that the following order

$$r_{i_1} \leq r_{i_2} \leq r_{i_3} \leq r_{i_4} \leq r_{i_5} \leq r_{i_6} \leq r_{i_7}.$$

Each r_i has a unique “position” in R_s , denoted by g_i . The student's secondary school performance and their willingness are represented by an n -dimensional vector V , which includes both position information and the original values.

$$V = (g_1 \times r_1, g_2 \times r_2, g_3 \times r_3, g_4 \times r_4, g_5 \times r_5, g_6 \times r_6, g_7 \times r_7).$$

For instance, suppose

$$R_k = (75, 101, 90, 86, 133, 82, 100),$$

and we obtain R_s where the values are ranked from 1 to 7:

$$R_s = (75, 82, 86, 90, 100, 101, 133).$$

Therefore,

$$\begin{aligned} V &= (1 \times 75, 6 \times 101, 4 \times 90, 3 \times 86, 7 \times 133, 2 \times 82, 5 \times 100) \\ &= (75, 606, 360, 258, 931, 164, 500). \end{aligned}$$

2.3 Support vector machine

Support Vector Machine (SVM) is a widely used supervised learning model in machine learning. SVM offer a principled approach to classification by explicitly optimizing the decision boundary, aiming to identify the hyperplane that maximizes the margin between different classes. It analyzes data to identify patterns and has been applied in various fields. The core principle of SVM involves representing input samples as points in a high-dimensional space and classifying them via an optimal separating hyperplane. SVM training aims for globally optimized solutions, which mitigate overfitting and enable handling a large number of features effectively. More detailed descriptions of the SVM method can be found in publications (Yang et al., 2022; Vapnik, 1998; Steinwart and Christmann, 2008; Chang and Lin, 2011).

The package LIBSVM v3.17 (Yang et al., 2022; Chang and Lin, 2011) which is an implementation of SVM classifier was used in this study. The Radial Basis Function (RBF) kernel, defined below, was selected as the kernel function for our model:

$$k(x, y) = \exp(-\gamma |x - y|^2).$$

The RBF kernel maps the input space into an infinite-dimensional feature space, enabling it to capture highly complex nonlinear relationships. As a local kernel, RBF assigns greater influence to similar samples, while the impact of dissimilar ones

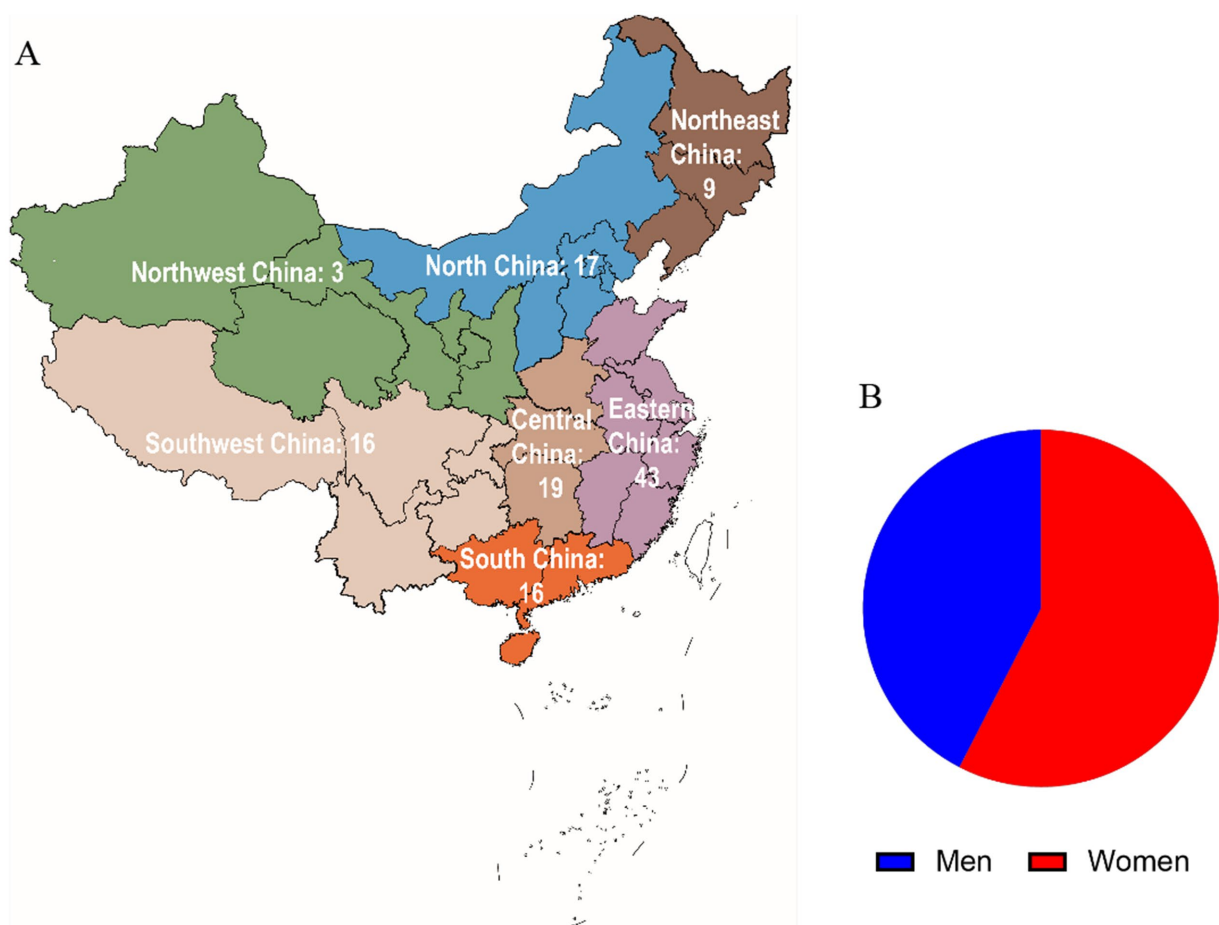


FIGURE 1
Geographical distribution and sex ratio of the participants.

approaches zero. Its inherent smoothness prevents overly fluctuating decision boundaries, thereby contributing to strong generalization capabilities. These characteristics make the RBF kernel particularly well-suited for modeling many types of natural data.

Two parameters, the penalty parameter C and the kernel width parameter γ , were determined via an optimization procedure using a grid search strategy provided by LIBSVM.

2.4 Assessment of the prediction performance

In statistical prediction, researchers commonly use three cross-validation methods to assess the effectiveness of a predictor in real-world scenarios: independent dataset test, subsampling test, and jackknife test. Among these, the jackknife method evaluates model performance by iteratively leaving one sample out as the test set while using the remaining samples for training. This process is repeated for every sample in the dataset, ensuring that all data points are used for both training and evaluation. The method maximizes data utilization and provides a nearly unbiased estimate of model performance (Qiu et al., 2014; Li et al., 2022; Zhao et al., 2021). Therefore, we employed the jackknife test in our study to evaluate the anticipated success rates of our predictor. This test

involves leaving out one sample at a time from the dataset Ω and evaluating it using the predictor trained on the remaining samples.

To assess accuracy of our model, we adopted sensitivity (S_n), specificity (S_p), accuracy (Acc), and the Matthew's correlation coefficient (MCC), which are widely used for measuring the quality of binary classifications. They are defined as follows:

$$S_n = \frac{TP}{TP + FN}$$

$$S_p = \frac{TN}{TN + FP}$$

$$Acc = \frac{(TP + TN)}{TP + TN + FP + FN}$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

where TP , TN , FP , and FN stand for the number of true positive, true negative, false positive, and false negative obtained from the prediction, respectively.

3 Results

There are two hyperparameters for the SVM model with RBF kernel: the regularization parameter denoted as C and influence of a training example on the decision boundary denoted as γ . A hyperparameter tuning process is required to retrieve the optimal pair of C and γ , which is not predetermined for a specific problem. In this study, we employed a grid search approach from the LIBSVM package to determine these parameters. By testing various values of C and γ (Yang et al., 2022; Chang and Lin, 2011), we found the optimal values to be $C = 2$ and $\gamma = 3.0517578123e-05$. Initially, we converted the prior grades and class attendance willingness into a 7-dimensional feature vector for each of the 123 respondents, forming the LE123 dataset. These feature vectors were then scaled and inputted into the Support Vector Machine (SVM). Through the jackknife test, we calculated the sensitivity, specificity, overall accuracy and Matthew's correlation coefficient, presented in Table 2. The method we used achieved an overall accuracy of 97.75%, with only four incorrect predictions. Notably, the sensitivity and specificity of each category exceeded 90%, averaging over 93%. Furthermore, the Matthews Correlation Coefficient (MCC) reached a value of 0.9573, indicating a high level of agreement between the predicted and actual classifications. These results indicate the potential effectiveness of our approach in predicting students' interest in university labor courses.

4 Discussion

The ongoing research on the application of artificial intelligence technology has garnered significant attention from individuals across diverse sectors. This has led to the production and deployment of smart robots and similar technologies. Scholars worldwide have conducted research on artificial intelligence technology in healthcare, computer information technology, education, biological sciences and other fields (Mavrych et al., 2025; Choudhary et al., 2023; Aljuaid, 2024; Rajabi et al., 2024; Baig and Yadegaridehkordi, 2024; Farrokhnia et al., 2024; Jemmy et al., 2024). This technological innovation has presented both opportunities and challenges for labor education, particularly for college students, manifesting in issues related to the recognition of labor values, the transformation of labor content, and the evolution of labor methodologies. There is an urgent need to explore and promote the innovative and robust development of labor education within the framework of AI technology.

This study proposes a novel machine learning model based on the academic performance of high school students and their interest in labor education courses, utilizing Support Vector Machine (SVM) for

predictive classification. The model yielded promising outcomes and demonstrated the feasibility of applying machine learning methods to the assessment of labor education.

Compared with most previous labor education studies that focused on qualitative analysis or relied on teachers' subjective scoring methods (Black and Wiliam, 1998; Jesse et al., 2025), this study employed SVM to facilitate an objective and data-driven evaluation of students' performance in labor education. This approach not only improves the efficiency and consistency of scoring but also mitigates cognitive bias, addressing recent concerns about "technical trustworthiness" and "fairness" in educational assessment (Williamson and Piattoeva, 2019; Selbst et al., 2019). Compared with earlier studies that utilized linear regression or other simple statistical models (Romero and Ventura, 2010), SVM demonstrates greater robustness in handling small-sample, non-linear classification problems. Its application in this study has shown its potential in the design of intelligent education systems.

This method could be integrated into college course management systems to automatically analyze students' classroom participation, homework completion, and related behavioral data, thereby assisting instructors in conducting comprehensive assessments in labor education. In addition, the model could be used for the early identification of students encountering difficulties in labor education, enabling timely intervention to improve the effectiveness of teaching.

Although the model has achieved promising improvement in educational assessment, the relatively small sample size in this study may limit its generalizability. In future research, we plan to expand the dataset by incorporating more schools and students from diverse disciplinary backgrounds. Multimodal data and longitudinal tracking will also be integrated to enhance the reliability and external validity of the predictive model. We will also integrate AI ethics and educational equity principles to ensure the interpretability, fairness, and transparency of the intelligent scoring system.

5 Conclusion

In this study, we constructed a multidimensional feature vector that integrates students' secondary school academic performance and their learning inclination. We then employed the SVM algorithm to assess students' learning potential with this feature set. The performance of the method was validated using jackknife cross-validation, and the results showed that sensitivity, specificity, overall accuracy, and the MCC all exceeded 90%, indicating that the proposed technique is effective in predicting students' learning potential.

Through our study, we aim to foster innovative perspectives on labor education reform, leveraging the application of artificial intelligence to support its evolution and progress. Additionally, our

TABLE 2 Prediction results on the dataset LE123 in jackknife test.

| Dataset | Grade | Sensitivity | Specificity | Overall Accuracy | MCC |
|---------|---------|---------------|---------------|------------------|--------|
| LE123 | A | 94.12% | 94.94% | 97.75% | 0.9573 |
| | B | 90.91% | 95.16% | | |
| | C | 96.43% | 96.43% | | |
| | D | 94.12% | 93.94% | | |
| | Average | 93.90 ± 1.96% | 95.12 ± 0.89% | | |

research introduces a novel approach by integrating data mining techniques into the teaching methodologies of college labor education courses, offering valuable insights for educational management entities seeking to tailor targeted teaching strategies for this demographic within the contemporary landscape of educational reform and technological advancement. The dataset in our current study is relatively limited. In future work, we plan to collect a larger and more diverse dataset in order to improve model accuracy and generalizability.

Data availability statement

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions

LY: Software, Writing – original draft. YY: Resources, Writing – original draft, Writing – review & editing.

References

- Aldowah, H., Rehman, S. U., Ghazal, S., and Umar, I. N. (2017). Internet of things in higher education: a study on future learning. *J. Phys. Conf. Ser.* 892:012017. doi: 10.1088/1742-6596/892/1/012017
- Aljuaid, H. (2024). The impact of artificial intelligence tools on academic writing instruction in higher education: a systematic review. *Arab World English J.* 1, 26–55. doi: 10.24093/awej/ChatGPT2
- Baig, M. I., and Yadegaridehkordi, E. (2024). ChatGPT in the higher education: a systematic literature review and research challenges. *Int. J. Educ. Res.* 127:102411. doi: 10.1016/j.ijer.2024.102411
- Beale, R. (2007). Supporting serendipity: using ambient intelligence to augment user exploration for data mining and web browsing. *Int. J. Hum.-Comput. Stud.* 65, 421–433. doi: 10.1016/j.ijhcs.2006.11.012
- Black, P., and Wiliam, D. (1998). Assessment and classroom learning. *Assess. Educ. Princ. Policy Pract.* 1998, 7–74.
- Bryson, J. R., and Andres, L. (2020). Covid-19 and rapid adoption and improvisation of online teaching: curating resources for extensive versus intensive online learning experiences. *J. Geogr. High. Educ.* 44, 608–623. doi: 10.1080/03098265.2020.1807478
- Callaway, E. (2024). AI protein-prediction tool AlphaFold3 is now more open. *Nature* 635, 531–532. doi: 10.1038/d41586-024-03708-4
- Chang, C., and Lin, C. J. (2011). Libsvm: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 1–27. doi: 10.1145/1961189.1961199
- Chen, L., Chen, P., and Lin, Z. (2020). Artificial intelligence in education: a review. *IEEE Access* 8, 75264–75278. doi: 10.1109/ACCESS.2020.2988510
- Choudhary, O. P., Saini, J., and Challana, A. (2023). ChatGPT for veterinary anatomy education: an overview of the prospects and drawbacks. *Int. J. Morphol.* 41, 1198–1202. doi: 10.4067/S0717-95022023000401198
- De Lorenzis, F., Prattico, F. G., Repetto, M., Pons, E., and Lamberti, F. (2023). Immersive virtual reality for procedural training: comparing traditional and learning by teaching approaches. *Comput. Ind.* 144:103785. doi: 10.1016/j.compind.2022.103785
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., and Wals, A. (2024). A SWOT analysis of ChatGPT: implications for educational practice and research. *Innov. Educ. Teach. Int.* 61, 460–474. doi: 10.1080/14703297.2023.2195846
- Fu, Z., and Zhou, Y. (2020). Research on human-AI co-creation based on reflective design practice. *CCF Trans. Pervasive Comput. Interact.* 2, 33–41. doi: 10.1007/s42486-020-00028-0
- Górriz, J. M., Ramírez, J., Ortiz, A., Martínez-Murcia, F. J., Segovia, F., Suckling, J., et al. (2020). Artificial intelligence within the interplay between natural and artificial computation: advances in data science, trends and applications. *Neurocomputing* 410, 237–270. doi: 10.1016/j.neucom.2020.05.078
- Haleem, A., Javaid, M., Qadri, M. A., and Suman, R. (2022). Understanding the role of digital technologies in education: a review. *Sustain. Oper. Comput.* 3, 275–285. doi: 10.1016/j.susoc.2022.05.004
- Han, J., and Chang, K. C. C. (2002). Data mining for web intelligence. *Computer* 35, 64–70. doi: 10.1109/MC.2002.1046977
- Han, J., and Kamber, M. (2006). Data mining: Concepts and techniques. 2nd ed. Burlington, Massachusetts: Morgan Kaufmann.
- Han, S., Kelly, E., Nikou, S., and Svec, E. O. (2022). Aligning artificial intelligence with human values: reflections from a phenomenological perspective. *AI & Soc.* 37, 1383–1395. doi: 10.1007/s00146-021-01247-4
- Hand, D., Mannila, H., and Smyth, P. (2001). Principles of data mining. Cambridge: MIT Press, 2.
- Hoogerwerf, M., Lösch, M., Schirrwagen, J., Callaghan, S., Manghi, P., Iatropoulou, K., et al. (2013). Linking data and publications: towards a cross-disciplinary approach. *Int. J. Digit. Curation* 8, 244–254. doi: 10.2218/ijdc.v8i1.257
- Hu, F. Y., Sun, Z. H., and Zhou, J. G. (2024). Application and evaluation of artificial intelligence technology in medical imaging teaching. *Inf. Med.* 21, 152–155. doi: 10.16659/j.cnki.1672-5654.2024.12.152
- Huan, W. (2019). The four new paths of labor education for college students in the new era. *Central China Normal University Journal of Postgraduates*.
- Jemmy, J., Aina, M., Wahdah, W., Joshua, W., and Sabri, S. (2024). Impact of ChatGPT in higher education learning. *J. Int. Lingua Technol.* 3, 43–57. doi: 10.55849/jiltech.v3i1.505
- Jesse, W., Lily, F., and Janell, W. (2025). Making plants matter: modeling photosynthesis. *Sci. Teach.* 92, 26–37. doi: 10.1080/00368555.2024.2445521
- Jiang, B. Z., and Pan, H. Z. (2019). On the practice of labor education for college students in the new era. *J. Heze Univ.* 41, 69–72. doi: 10.16393/j.cnki.37-1436/z.2019.03.016
- Li, C., Yang, Y., Fei, W. C., He, P. A., Yu, X., Zhang, D., et al. (2015). Prediction of success for polymerase chain reactions using the Markov maximal order model and support vector machine. *J. Theor. Biol.* 369, 51–58. doi: 10.1016/j.jtbi.2015.01.017
- Li, C., Yang, Y., Jia, M. D., Zhang, Y., Yu, X., and Wang, C. (2014). Phylogenetic analysis of DNA sequences based on k-word and rough set theory. *Physica A* 398, 162–171. doi: 10.1016/j.physa.2013.12.025
- Li, X. J., Zhang, S. L., and Shi, H. Y. (2022). An improved residual network using deep fusion for identifying RNA 5-methylcytosine sites. *Bioinformatics* 38, 4271–4277. doi: 10.1093/bioinformatics/btac532
- Losiewicz, P., Oard, D. W., and Kostoff, R. N. (2000). Textual data mining to support science and technology management. *J. Intell. Inf. Syst.* 15, 99–119. doi: 10.1023/A:1008777222412
- Mavrych, V., Ganguly, P., and Bolgova, O. (2025). Using large language models (ChatGPT, copilot, PaLM, bard, and Gemini) in gross anatomy course: comparative analysis. *Clin. Anat.* 38, 200–210. doi: 10.1002/ca.24244
- Moreno, A., and Redondo, T. (2016). Text analytics: the convergence of big data and artificial intelligence. *IJIMAI* 3, 57–64. doi: 10.9781/ijimai.2016.369
- Oliveira, G., Grenha Teixeira, J., Torres, A., and Morais, C. (2021). An exploratory study on the emergency remote education experience of higher education students and

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

teachers during the COVID-19 pandemic. *Br. J. Educ. Technol.* 52, 1357–1376. doi: 10.1111/bjet.13112

Pournaras, E. (2017). Cross-disciplinary higher education of data science—beyond the computer science student. *Data Sci.* 1, 101–117. doi: 10.3233/DS-170005

Qiu, W. R., Xiao, X., and Chou, K. C. (2014). iRSpot-TNCPseAAC: identify recombination spots with rinucleotide composition and pseudo amino acid components. *Int. J. Mol. Sci.* 15, 1746–1766. doi: 10.3390/ijms15021746

Rajabi, P., Taghipour, P., Cukierman, D., and Doleck, T. (2024). Unleashing ChatGPT's impact in higher education: student and faculty perspectives. *Microelectron. J.* 2:100090. doi: 10.1016/j.chbah.2024.100090

Romero, C., and Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* 40, 601–618. doi: 10.1109/TSMCC.2010.2053532

Rui, Y., Carmona, V. I. S., Pourvali, M., Xing, Y., Yi, W. W., Ruan, H. B., et al. (2022). Knowledge mining: a cross-disciplinary survey. *Mach. Intell. Res.* 19, 89–114. doi: 10.1007/s11633-022-1323-6

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., and Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. *Proceedings of the Conference on Fairness, Accountability, and Transparency 2019; (FAT* '19)*: 59–68.

Steinwart, I., and Christmann, A. (2008). Support vector machines. New York, NY: Springer-Verlag.

Sun, W. P. (2017). Reflection on the value of artificial intelligence. *Philos. Res.* 63, 120–126.

Vapnik, V. (1998). Statistical learning theory. New York, NY: Wiley.

Wang, T. (2007). The comparison of the difficulties between cooperative learning and traditional teaching methods in college English teachers. *J. Hum. Resour. Adult Learn.* 3, 21–30.

Wang, S. S. (2024). Artificial intelligence technology in education. *High-Tech. Commer.* 7, 87–97. doi: 10.36079/lamintang.ijeste-0702.784

Wang, L., and Fu, X. (2005). *Data mining with computational intelligence*. Berlin: Springer Science & Business Media.

Williamson, B., and Piattoeva, N. (2019). Objectivity as standardization in data-scientific education policy, technology and governance. *Routledge.* 44, 64–76. doi: 10.1080/17439884.2018.1556215

Wu, W., Zhang, Y., Wang, P., Zhang, L., Wang, G., Lei, G., et al. (2020). Psychological stress of medical staffs during outbreak of COVID-19 and adjustment strategy. *J. Med. Virol.* 92, 1962–1970. doi: 10.1002/jmv.25914

Yang, X. W., and Wang, T. M. (2013). A novel statistical measure for sequence comparison on the basis of k-word counts. *J. Theor. Biol.* 318, 91–100. doi: 10.1016/j.jtbi.2012.10.035

Yang, B., Yang, L., Yu, S., Yang, M. Y., Zhang, W. T., Bin, C. (2022). AUTODC: an automatic machine learning framework for disease classification. *Bioinformatics* 38, 3415–3421. doi: 10.1093/bioinformatics/btac334

Yao, X. J. (2003). The teaching of the traditional sports of the minority people. *J. Wuhan Inst. Phys. Educ.* 2, 112–114. doi: 10.15930/j.cnki.wtxb.2003.02.036

Yu, X. Q., Zheng, X. Q., Meng, L. Y., Li, C., and Wang, J. (2012). A support vector machines based method to predict success for polymerase chain reactions. *Comb. Chem. High Throughput Screen.* 15, 486–491. doi: 10.2174/138620712800563936

Zhan, Q., and Yang, M. A model of maker education in China universities by smart technologies: the perspective of innovation ecosystem. *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*: 628–633.

Zhao, H. C., Li, Y. H., and Wang, J. X. (2021). A convolutional neural network and graph convolutional network-based method for predicting the classification of anatomical therapeutic chemicals. *Bioinformatics* 37, 2841–2847. doi: 10.1093/bioinformatics/btab204

Zhou, S. J., Zhang, L. G., Wang, L. L., Guo, Z. C., Wang, J. Q., Chen, J. C., et al. (2020). Prevalence and socio-demographic correlates of psychological health problems in Chinese adolescents during the outbreak of COVID-19. *Eur. Child Adolesc. Psychiatry* 29, 749–758. doi: 10.1007/s00787-020-01541-4

Zhu, J., Sun, L., Zhang, L., Wang, H., Fan, A., Yang, B., et al. (2020). Prevalence and influencing factors of anxiety and depression symptoms in the first-line medical staff fighting against COVID-19 in Gansu. *Front. Psych.* 11:386. doi: 10.3389/fpsy.2020.00386